

Aus der Klinik und Poliklinik für Diagnostische und Interventionelle
Radiologie der Universitätsmedizin der Johannes Gutenberg-Universität
Mainz

Software für die automatische Tumordiagnose in der funktionalen MRT der weiblichen Brust

Inauguraldissertation
zur Erlangung des Doktorgrades der physiologischen Wissenschaften

der Universitätsmedizin
der Johannes Gutenberg-Universität Mainz
vorgelegt von

Dirk Mayer
aus Soest

Mainz, 2009

The lack of information can't be remedied by any mathematical trickery.
Cornelius Lanczos, 1961

Inhaltsverzeichnis

Abstract	v
Abkürzungsverzeichnis	vii
1 Einleitung	17
1.1 Motivation	17
1.2 Zielsetzung	18
1.3 Aufbau der Arbeit	18
2 Grundlagen	19
2.1 Brustkrebs	19
2.1.1 Morphologie der weiblichen Brust	19
2.1.2 Tumoreigenschaften	20
2.1.3 Pathologische Veränderungen	21
2.1.4 Risikofaktoren, Prognose und Behandlung	24
2.2 Mammadiagnostik	25
2.2.1 Palpation	25
2.2.2 Mammographie	26
2.2.3 Sonographie	27
2.2.4 Magnetresonanzmammographie	28
2.3 Befundung	31
2.3.1 Bildsubtraktion	31
2.3.2 Kinetik	32
2.3.3 Morphologie	34
2.3.4 T_2 Wichtung	35
2.3.5 Diagnose	35
2.3.6 BI-RADS TM	36
2.4 Bildverarbeitung	39
2.4.1 Wissensrepräsentation	41
2.4.2 Low-Level	42
2.4.3 Intermediate-Level	42
2.4.4 High-Level	43

3	Stand der Forschung	45
3.1	Kategorien von CAD-Systemen	45
3.2	Kommerzielle CAD-Systeme	47
3.2.1	Low-Level-Visualisierungssysteme	47
3.2.2	Low-Level-Diagnosesysteme	47
3.2.3	High-Level-Diagnosesysteme	48
3.3	Wissenschaftliche Publikationen	49
3.3.1	Registrierung	49
3.3.2	Segmentierung	50
3.3.3	Merkmalsextraktion	51
3.3.4	Klassifikation	52
3.4	Fortschritt durch die vorliegende Arbeit	53
4	Material und Methoden	55
4.1	Bilderhebung	57
4.1.1	MRT-Protokoll	57
4.1.2	Import	60
4.2	Registrierung	60
4.2.1	Sequenzielle Teilschritte	62
4.3	Segmentierung	64
4.3.1	Einzelne Bildserie	64
4.3.2	Funktionale Bildserien	66
4.4	Merkmalsextraktion	67
4.4.1	Form und Begrenzung	67
4.4.2	Textur	70
4.4.3	Kinetik	72
4.4.4	Sonstige	74
4.4.5	Übersicht	75
4.5	Klassifikation	75
4.5.1	Goldstandard	75
4.5.2	Lerndaten	78
4.5.3	Neuronale Netze	79
4.5.4	Training	80
4.5.5	Export	81
4.6	Evaluierung	81
4.6.1	Abstandsmaße	82
4.6.2	Goldstandardvergleich	83
4.6.3	Multivariate Analyse	84
4.6.4	Kreuzvalidierung	86
4.6.5	Alternative Klassifikatoren	88

5	Ergebnisse	91
5.1	Registrierung	91
5.2	Segmentierung	96
5.3	Merkmalsextraktion	101
5.4	Klassifikation	102
6	Diskussion	105
6.1	Registrierung	105
6.1.1	Qualitätsabschätzung des Algorithmus	107
6.1.2	Evaluation der Bildverbesserung	112
6.1.3	Einfluss auf Kontrastmittel anreichernde Strukturen	114
6.1.4	Anwendung auf fremde Bildinhalte	116
6.1.5	Erkenntnisse und Grenzen	120
6.2	Segmentierung	121
6.2.1	Qualitätsabschätzung des Algorithmus	123
6.2.2	Zusätzlich segmentierte Objekte	129
6.2.3	Vergleich zum Goldstandard	131
6.2.4	Einfluss der Registrierung auf die Segmentierung	133
6.2.5	Erkenntnisse und Grenzen	136
6.3	Merkmalsextraktion	137
6.3.1	Merkmalsvergleich mit BI-RADS TM	138
6.3.2	Mensch versus Maschine	141
6.3.3	Robustheit der Merkmalsberechnung	143
6.3.4	Clusteranalyse	149
6.3.5	Hauptkomponentenanalyse	151
6.3.6	Erkenntnisse und Grenzen	154
6.4	Klassifikation	156
6.4.1	Merkmalskombinationen	158
6.4.2	Merkmalsselektion	160
6.4.3	Vorhersage des ungesicherter Goldstandards	164
6.4.4	Klassifikatorenvergleich	166
6.4.5	Beurteilung zusätzlicher Funde	169
6.4.6	Problemfall	169
6.4.7	Erkenntnisse und Grenzen	172
6.5	Vergleich zu kommerziellen Systeme	173
6.5.1	Registrierung	174
6.5.2	Segmentierung	174
6.5.3	Merkmalsextraktion	175
6.5.4	Klassifikation	175
6.6	Nutzen als Zweitmeinungssystem	177
6.7	Ausblick	179
7	Zusammenfassung	181

A Koordinatensystem nach DICOM	183
A.1 Patientenkoordinatensystem	183
A.2 Modellkoordinatensystem	185
A.3 Koordinatentransformation	185
B KLT Feature Tracking	189
C Vektorinterpolation	193
Stichwortverzeichnis	195
Abbildungsverzeichnis	201
Tabellenverzeichnis	205
Literaturverzeichnis	207
Danksagung	223

Abstract Breast cancer is the most common cancer among women in the Western industrialized nations, worldwide its part of all cancers among women adds up to 21 %. About one in nine women develops breast cancer during her lifetime; the mortality rate in Germany comes up to 27 %.

Breast carcinomas mostly have a relatively slow growth rate. Detecting and removing all breast cancers 5 to 10 *mm* in size could almost eliminate breast cancer death, since the 20 years survival rate is with over 95 % very high [Kai07].

Magnetic resonance mammography is a very sensitive diagnostic method giving the chance to detect breast tumors down to a size of about 3 *mm* in diameter. The drawback of this relative new method is its complex and time consuming diagnostic demand; a numerous pitfalls exist in interpretation of the acquired images, the long diagnostic learning curve lead to a high inter-observer variability.

A computer aided diagnostic system can increase the quality of such complex diagnostics or at least speed up the process. Therefore the aim of this thesis is to develop a full automatic diagnosis software which is usable as an automatic reader for second opinion [MVF⁺06]. To my knowledge, such complete software does not exist today.

The software executes a pipe of different image processing operations that have been developed following the diagnostic approach of a physician. It results in an autonomous diagnosis for each object found: First, as a pre-processing step, a 3d image registration eliminates motion within the acquired breast images advancing the image quality for the subsequent steps. Then, each contrast-enhancing lesion is detected as a single object by a rule based segmentation applying individual thresholds. Thereafter, kinetic and morphological features are calculated describing the characteristics of contrast-agent uptake as well as the shape, margin and texture properties of each segmented object. Finally, based on the extracted feature vector, two trained neuronal networks rate each object to be an additional finding or to be a benign or malignant lesion.

The performance of the software was tested on the image data of 101 female patients inclosing 141 histological proven lesions. The prediction of healthiness on these objects results in 88 % sensitivity and 72 % specificity, which is similar to known prediction values of expert radiologists in literature. The predictions also include on average 2.5 additional malignant findings per patient, which turned out to be misclassified artifacts.

Abkürzungsverzeichnis

T_2fs	fettgesättigte T_2 -gewichtete Sequenz
2D	zweidimensional
3D	dreidimensional
4D	vierdimensional
ACR	American College of Radiology
AUC	Area Under Curve
BI-RADS TM	Breast Imaging Reporting and Data System
CAD	Computer Aided Detection
CAD	Computer Aided Diagnosis
CD	Computer Diagnosis
CT	Computertomographie
DCIS	duktales Karzinom in-situ
DICOM	Digital Imaging and Communications in Medicine
et al.	und andere
FDA	flexible Diskriminanzanalyse
fMRT	funktionale MRT
FoV	Field of View
GS	Goldstandard
GUI	Graphical User Interface – graphische Benutzerschnittstelle
IOP	Image Orientation Patient
IPP	Image Position Patient
KLT	Kanade-Lucas-Tomasi
KM	Kontrastmittel
knn	k-nearest Neighbor
KOS	Koordinatensystem
LCIS	Lobuläres Karzinom in-situ
LDA	lineare Diskriminanzanalyse
MI	Mutual Information
MIP	Maximum Intensitäts Projektion
MPR	Multiplanare Rekonstruktion
MRM	Magnetresonanzmammographie
MRT	Magnetresonanztomographie
NEMA	National Electrical Manufacturers Association

NN	Neurales Netz
QDA	quadratische Diskriminanzanalyse
ROI	Region of Interest
SI	Signalintensität
SN	Seriennummer
SVM	Support Vektor Maschinen
TP	Zeitpunkt
TTP	Time To Peak
VEGF	vascular epithelial growth factor

Kapitel 1

Einleitung

1.1 Motivation

In den westlichen Industrieländern ist das Mammakarzinom (s. Kap. 2.1.3) der häufigste bösartige Tumor der Frau. Sein weltweiter Anteil an allen Krebserkrankungen der Frau beläuft sich auf etwa 21 % [PPF99].

Das Robert Koch Institut gibt die jährliche Neuerkrankung in Deutschland mit über 57.000 Frauen an. Das Mammakarzinom ist damit die häufigste Krebsneuerkrankung mit steigender Inzidenzrate der letzten drei Jahrzehnte. Es ist für mehr als 25 % der gesamten Krebserkrankungen verantwortlich. Das mittlere Erkrankungsalter bei Frauen von 63 Jahren liegt um sechs Jahren unter dem Durchschnittserkrankungsalter aller Krebserkrankungen.

Inzwischen ist jede neunte Frau bedroht, während ihres Lebens an Brustkrebs zu erkranken. Insgesamt hat die Häufigkeit dieser Krebsform in den letzten 20 Jahren in ganz Europa zugenommen. Die Mortalität nimmt seit Mitte der 1990er Jahre wieder leicht ab, die alterstandardisierte Rate liegt derzeit bei knapp 27 % [Rob04].

Kaiser [Kai07] legt dar, dass die Existenz eines diagnostischen Verfahrens, mit dem alle Mammakarzinome unter 1 *cm* Durchmesser erkannt und entfernt werden, den Tod durch Brustkrebs praktisch beseitigen könnte. Denn die 20-Jahres-Überlebensrate bei Erkrankung durch initiale Karzinome kleiner als 1 *cm* liege mit über 95 % sehr hoch.

Mit der Kontrastmittel gestützten Bildgebung durch die MRT (s. Kap. 2.2.4) steht eine relativ junge Untersuchungsmethode zur Verfügung, die sensitiv genug zur Erkennung von Karzinomen ab einer Größe von 3 *mm* Durchmesser ist. Die diagnostische Methodik (s. Kap. 2.3) ist jedoch komplex, fehleranfällig, erfordert eine lange Einarbeitungszeit und somit viel Erfahrung des Radiologen.

1.2 Zielsetzung

Ziel der Arbeit ist die Entwicklung einer Prototypsoftware [MVF⁺06], die dem Radiologen als Zweitmeinungssystem bei der zeitaufwendigen und komplexen Diagnose des Mammakarzinoms auf der Kontrastmittel gestützten Magnetresonanztomographie (MRT) Untersuchung zur Seite steht. Wie auch der Radiologe muss die Software verdächtige Veränderungen selbstständig erkennen, diese analysieren und eine Diagnose pro Befund ableiten. Für eine Akzeptanz als Zweitmeinungssystem sollte die Software dabei möglichst autonom ohne Interaktion durch den Anwender arbeiten.

1.3 Aufbau der Arbeit

Als interdisziplinäre Arbeit werden zuerst die medizinischen (s. Kap. 2.1 - 2.3) und mit der Bildverarbeitung sodann die informatischen Grundlagen (s. Kap. 2.4) vermittelt, die zum Gesamtverständnis der Arbeit beitragen.

Der Stand der Forschung (s. Kap. 3.1) kategorisiert bekannte Softwareprodukte (s. Kap. 3.2), gibt einen ausführlichen Überblick über bereits bekannte Bildverarbeitungstechniken (s. Kap. 3.3) und grenzt die Leistung dieser Arbeit gegenüber den bekannten Produkten ab (s. Kap. 3.4).

Jede gewonnene Teilkomponente aus der algorithmischen Umsetzung der Vorgehensweise des Radiologen bei der Diagnostik wird im Kapitel Material und Methoden beschrieben. Es wird auf die Erzeugung der Bilddaten (s. Kap. 4.1), deren qualitative Verbesserung durch eine Bewegungskorrektur als Vorverarbeitung (s. Kap. 4.2), das Erkennen verdächtiger Erkrankungen (s. Kap. 4.3), die Charakterisierung der Befunde (s. Kap. 4.4) und die Ableitung der Diagnose (s. Kap. 4.5) eingegangen. Das Kapitel schließt mit der Erläuterung verschiedener Evaluationsverfahren (s. Kap. 4.6) ab. Durch sie wird die Qualität der Ergebnisse der funktional sehr umfangreichen Software in der Diskussion bewertet.

Die Ergebnisse des Prototyps werden in Kapitel 5 durch Darlegung der Ausgaben jeder Teilkomponente vorgestellt. Anhand der Bilddatenbank mit dem histologisch gesicherten Goldstandard der zugrunde liegenden MRT-Untersuchungen (s. Kap. 4.5.1) wird die Leistungsfähigkeit der Software bei der Erkennung von Karzinomen beurteilt.

Die Diskussion (s. Kap. 6) prüft die Qualität der Ergebnisse aller Teilkomponenten durch Anwendung verschiedener mathematischer Evaluationsverfahren. Zuvor werden jeweils die Notwendigkeit und die Wahl der verwendeten Algorithmen begründet und dabei ein Vergleich mit der Literatur gezogen. Abschließend wird der Frage nachgegangen, inwieweit der Prototyp den Radiologen als Zweitmeinungssystem unterstützen kann (s. Kap. 6.6) und welche Verbesserungen vor dem Beginn einer Multizentrischen Klinischen (s. Kap. 6.7) Studie umgesetzt werden müssen.

In der Zusammenfassung in Kapitel 7 wird ein Querschnitt aller vorherigen Kapitel gegeben und geprüft, inwieweit die Zielsetzung erfüllt wurde.

Das Glossar (Seite vii) und Stichwortverzeichnis (Seite 195) vereinfachen den Umgang mit Abkürzungen und den definierten Fachbegriffen.

Kapitel 2

Grundlagen

Folgend werden sowohl grundlegendes medizinisches als auch informatisches Wissen der interdisziplinäre Arbeit vermittelt. Im sich anschließenden Kapitel 2.1 werden Informationen über die Entstehung und Behandlung von Brustkrebs gegeben. Kapitel 2.2 stellt heute eingesetzte Verfahren für die Diagnostik dieser Krebserkrankung vor, während sich das darauf folgende Kapitel 2.3 mit der Befundung nur durch die Magnetresonanztomographie als eine dieser Untersuchungsmethoden beschäftigt. Abschließend werden die Wurzeln der Bildverarbeitung erläutert (s. Kap. 2.4), die zum Verständnis der angewendeten Bildverarbeitungsmethoden (s. Kap. 4) einen Beitrag leisten.

2.1 Brustkrebs

2.1.1 Morphologie der weiblichen Brust

Die weibliche Brust (*Mamma*, s. Abb. 2.1) kann funktional in zwei Bereiche unterteilt werden:

- die *Brustdrüse* (*Glandula mammaria*), bestehend aus ca. 15 bis 25 einzelnen *Milchdrüsen* in Form von Läppchen (*Lobi glandulae mammariae*), deren *Milchgänge* sich zu *Milchsäckchen* (*Ductus lactiferi colligentes*) erweitern und strahlenförmig zur Brustwarze laufen; dort münden sie nach außen. Die Zellen der Milchgänge und -drüsen sind von einer *Basalmembran* umgeben.
- das stützende und formgebende *Binde- und Fettgewebe*, das auch die Milchdrüsen voneinander abgrenzt.

Mit fortschreitendem Alter der Frau wandelt sich auch die Morphologie der Mamma. Die Ausbildung der Drüsenläppchen beginnt am Ende der Pubertät und setzt sich etwa bis zum 35. Lebensjahr fort. Danach wird mit beginnendem Rückbildungsprozess das Drüsen-gewebe allmählich durch Fettgewebe ersetzt. Hormonell bedingt wird die Brust während der Menstruation stärker durchblutet und lagert Flüssigkeit ein, was mit Spannungen einhergehen kann [Lip06, Fis00, TMV91].

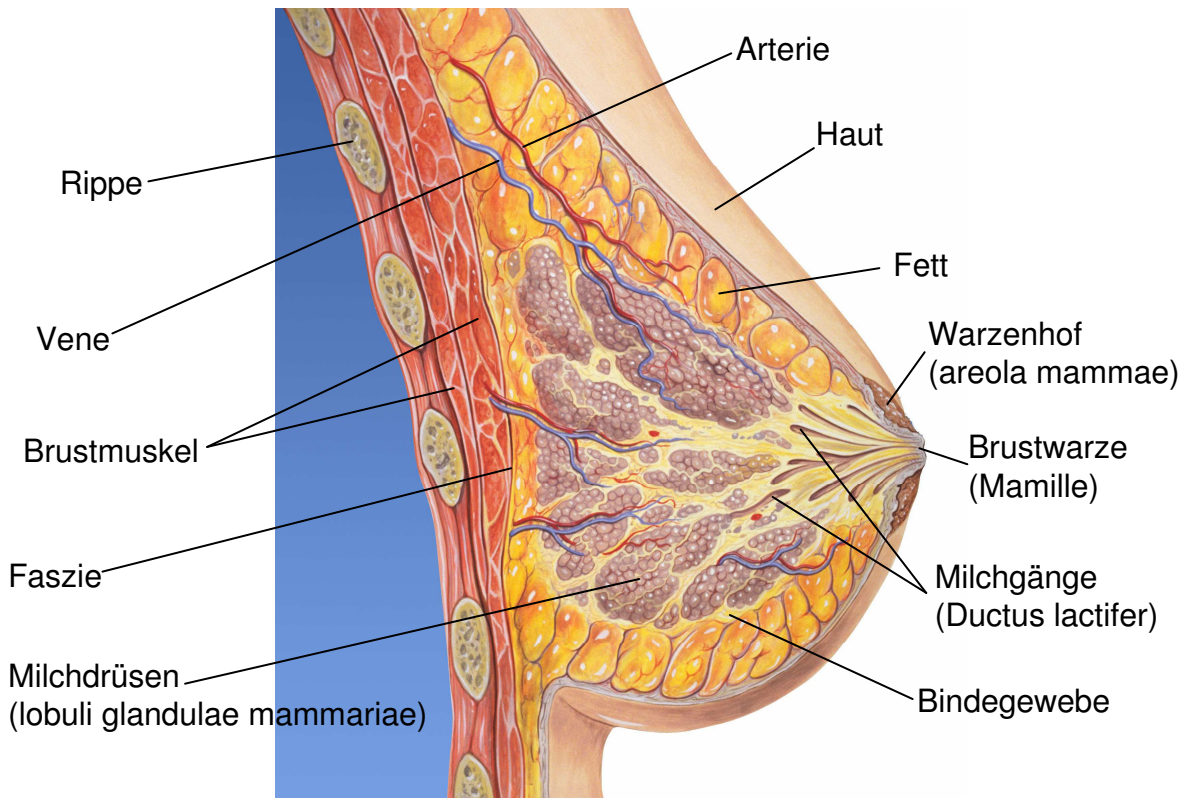


Abbildung 2.1: Anatomie einer gesunden weiblichen Brust in der Sagittalansicht (von Patrick J. Lynch)

2.1.2 Tumoreigenschaften

Tumore entstehen durch die Fehlregulation des Wachstums körpereigener Zellen, wodurch es zu autonomen und progressiven Neubildungen von neuen Zellen kommt. Unterschieden wird der Grad ihrer Bösartigkeit (*Dignität*) in gutartig oder bösartig (s. Tab. 2.1).

Gutartige (benigne) Tumore zeichnen sich überwiegend durch langsames Wachstum aufgrund einer niedrigen Zellteilungsrate aus, ihr Gewebe ist gut differenziert. Dabei verdrängt die Geschwulst lediglich das umliegende Gewebe, infiltriert es aber nicht und grenzt sich dadurch scharf ab.

Dagegen können *bösartige (maligne)* Tumore unbehandelt zum Tode des Betroffenen führen. Ihr Wachstum ist für gewöhnlich schneller aufgrund der höheren Zellteilungsrate. Im Gegensatz zu benignen Geschwülsten verlässt ein maligner Tumor den Ort seines Entstehens und wächst invasiv und destruierend in das angrenzende Gewebe. Je bösartiger der Tumor, desto weniger ist sein Gewebe differenziert. Maligne Tumore können Tochtergeschwüre (*Metastasen*) in andere Bereiche des Organismus bilden.

Als weitere Form nehmen *semimaligne* Tumore in ihrem Verhalten eine Stellung zwischen malignen und benignen Tumoren ein. Sie bilden zwar keine Metastasen, infiltrieren das umliegende Gewebe aber wie ein bösartiger Tumor. [TMV91].

Merkmal	Benigne	Maligne
Wachstum	langsam, verdrängend	schnell, infiltrierend, invasiv, destruierend
Abgrenzung	scharf, glatt	unscharf, irregulär
Differenzierung	weitgehend ausdifferenziert, homologes Gewebe	wenig differenziert, heterologes Gewebe
Zellgehalt	niedrig	hoch
Verlauf	länger dauernd, symptomarm, keine Metastasen, selten Rezidive	kürzer, häufig letal, Metastasen, häufig Rezidive

Tabelle 2.1: Dignität von Tumoren und unterscheidbare Merkmale [TMV91]

2.1.3 Pathologische Veränderungen

Pathologische Veränderungen der Mamma können in gut- und bösartige Neubildungen unterschieden werden. Die Grenzen sind jedoch fließend, da es zudem auch intermediäre Neubildungen und Vorstufen (*Präkanzerosen*) gibt. Alle vorkommenden pathologischen Veränderungen der Beschaffenheit, Form oder Konsistenz der Brust werden folgend als *Läsionen* (Schädigung, Verletzung oder Störung anatomischer Strukturen oder Funktionen) bezeichnet. Sie werden häufig von der Frau selbst durch Ertasten von verdichteten Gewebeanteilen in der Brust oder Sekretaustritt aus der Mamille erkannt.

Benigne In etwa drei Viertel der selbst ertasteten Fälle liegt ein gutartiger Zustand vor. Meist sind dies keine Tumore, sondern eine Vermehrung des Bindegewebes (*Mastopathie*) zwischen den Milchdrüsen oder die Bildung flüssigkeitsgefüllter Zysten. Weitere gutartige, nicht tumoröse Veränderungen stellen Entzündungen der Brust (*Mastitis*) dar, meist verursacht durch Bakterien, die von außen in die Milchgänge gelangen. Diese Gefahr ist während der Stillzeit am größten. Liegt jedoch ein gutartiger Tumor vor (s. vorheriges Kapitel), geht er häufig vom Bindegewebe (*Fibroadenom*) oder von den Milchgängen (*Papillom*) aus, selten vom Fettgewebe (*Lipom*). Papillome können in seltenen Fällen entarten und zählen entsprechend als *semimaligne*.

Maligne Die häufigste bösartige pathologische Veränderung der Mamma ist der Brustkrebs oder auch das so genannte *Mammakarzinom*. Bösartige Tumore (s. vorheriges Kapitel) werden allgemein als Karzinome bezeichnet, wenn sie vom Epithelgewebe ausgehen, das innere und äußere Oberflächen des Körpers bedeckt. Das Mammakarzinom entsteht im Drüsenepithel (*Adenokarzinom*).

Histologisch werden Mammakarzinome in invasive und nicht invasive Formen unterschieden (s. Abb. 2.2). Die Hauptgruppe der invasiven Tumore bildet das von den Oberflächenzellen der Milchgänge ausgehende *duktale Karzinom* (60-70 %) und das von den Milchdrüsenläppchen ausgehende seltenere *lobuläre Karzinom* (20 %) [FDTC+99]. Invasive Karzinome wachsen infiltrierend, die lobuläre Form zeigt dabei typischerweise ein diffuses

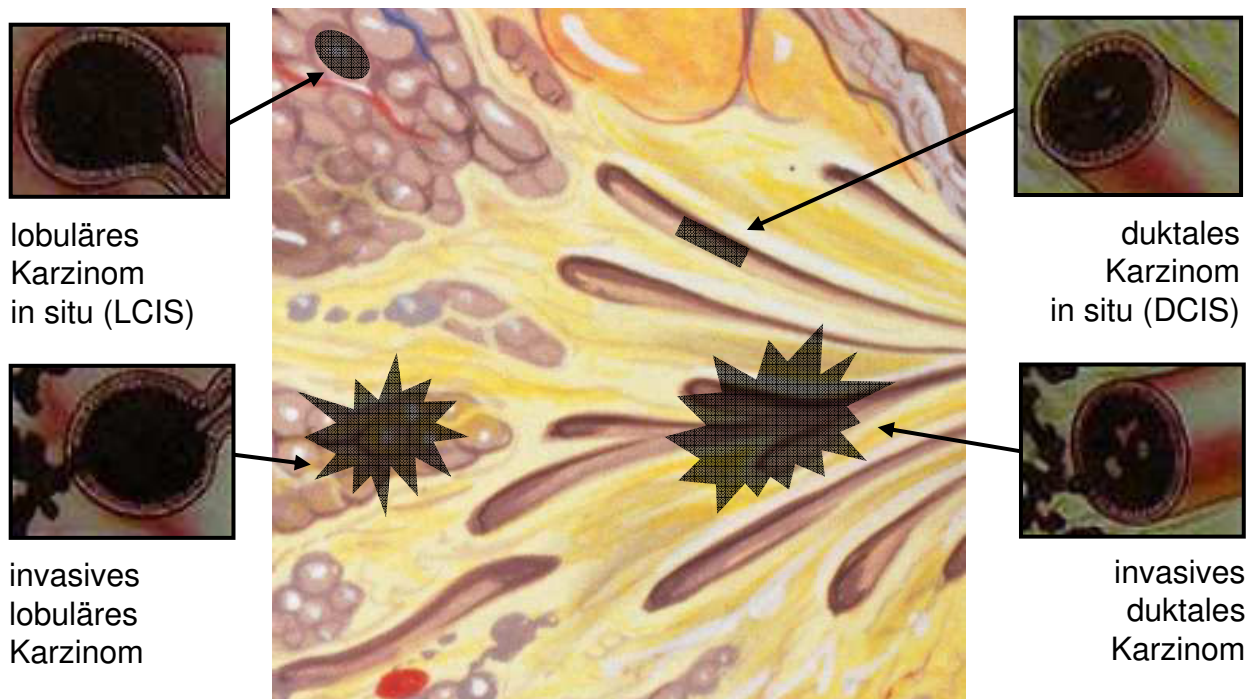


Abbildung 2.2: Unterscheidung invasiver und in-situ Varianten duktaler und lobulärer Karzinome

Wachstum, bildet also selten ertastbare Knoten. Beide Karzinomarten können auch nur örtlich begrenzt innerhalb der Läppchen oder Gänge (*in-situ*) wachsen (*DCIS = duktales Karzinom in-situ*, *LCIS = lobuläre Karzinom in-situ*). Hier konnte der Tumor noch nicht die *Basalmembran* durchbrechen. Daher kann er auch keine Metastasen bilden. In diesem Zustand ist die betroffene Frau durch die alleinige komplette Entfernung des Karzinoms heilbar.

Weitere Tumore wie das *medulläre* und *muzinöse* Karzinom sind Sonderformen des invasiven lobulären Karzinoms, die sich glatt und somit ähnlich wie die gutartigen Tumore zum gesunden Gewebe abgrenzen. Tabelle 2.2 fasst die häufigsten Tumorarten der Mamma zusammen, erhebt jedoch keinen Anspruch auf Vollständigkeit weiterer Misch- und Sonderformen [Lip06, Fis00, HKS03, TMV91].

Die Verdopplungszeit des Volumens liegt bei etwa zwei bis fünf Monaten, so dass bis zum Heranwachsen eines 1 cm großen Tumors aus einer initialen Krebszelle 10 Jahre und mehr vergehen können [Pea76]. Aufgrund des langsamen Wachstums von Brustkrebs bietet sich ein ausgiebiger Zeitraum für eine mögliche Früherkennung.

Ab einem Durchmesser von etwa 2 mm wird für das bösartige Wachstum der Tumore und die Fähigkeit zur Metastasierung die Neubildung von Gefäßen induziert (*Neoangiogenese*). Diese sind für die Versorgung mit Nährstoffen und den Abtransport von Stoffwechselprodukten unerlässlich. Eine wichtige Rolle spielt das Wachstumshormon *VEGF* (*vascular epithelial growth factor*), das zur Ausbildung neuer Gefäße und zum Abbau der

Dignität	Tumortyp	Muttergewebe	Häufigkeit	Morphologie
Malignant	Duktales Karzinom	Milchgang	sehr häufig	Sternförmig oder knotig, unregelmäßig begrenzt
	Lobuläres Karzinom	Drüsengewebe	häufig	Diffus, unregelmäßig begrenzt
	Medulläres Karzinom	Drüsengewebe	selten	Rundlich knotig oder polygonal, glatt begrenzt
	Muzinöses Karzinom	Drüsengewebe	selten	Rundlich knotig oder polygonal, glatt begrenzt
	DCIS LCIS	Milchgang Drüsengewebe	selten selten	- -
Benign o. Semi-malignant	Fibroadenom	Bindegewebe	häufig	rundlich/oval glatt begrenzt
	Papillom	Milchgang	selten	-
Benign	Lipome	Fett	extrem selten	scharf begrenzt
	Mastopathie	Bindegewebe	sehr häufig	-

Tabelle 2.2: Ursprungsgewebe und Häufigkeiten nach Biopsie verschiedener Tumorarten [HKS03, Fis00].

Basalmembran führt. Die Wände der heranwachsenden neuen Gefäße haben zunächst eine undichte Interzellularverbindung mit einer erhöhten Durchlässigkeit (*Permeabilität*) für niedermolekulare Stoffe [GLCF72, Fol92].

Die Häufigkeit des Auftretens von Karzinomen innerhalb der Brust ist ungleichmäßig verteilt (s. Abb. 2.3). Etwa die Hälfte der bösartigen Tumore entsteht im oberen axillaren Viertel der Brustdrüse, im unteren inneren Viertel ist sein Auftreten am seltensten. Der Grund hierfür liegt in der regionalen Verteilung des Brustdrüsengewebes. Korrelierend findet man im oberen äußeren Quadranten am meisten Drüsengewebe.

Das Vorliegen einer bösartigen Veränderung kann durch eine Gewebeuntersuchung nach Durchführung einer Biopsie gesichert werden. Dazu wird eine Nadel minimal invasiv durch die Haut der Brust an die verdächtige Stelle geschoben und die entnommene Gewebeprobe anschließend untersucht (*histologischer Befund*). Um die Treffsicherheit zu erhöhen, wird die Steuerung der Nadel durch ein bildgebendes Verfahren (s. Kap. 2.2) überwacht [Lip06].

Auch Männer können in sehr seltenen Fällen an Brustkrebs erkranken (etwa im Verhältnis 1:100 zu Frauen), meist begünstigt durch eine Hormonbehandlung beispielsweise mit weiblichen Geschlechtshormonen. Das Erkrankungsalter liegt jedoch im Schnitt 10 Jahre höher als bei Frauen [Lip06, Fis00].

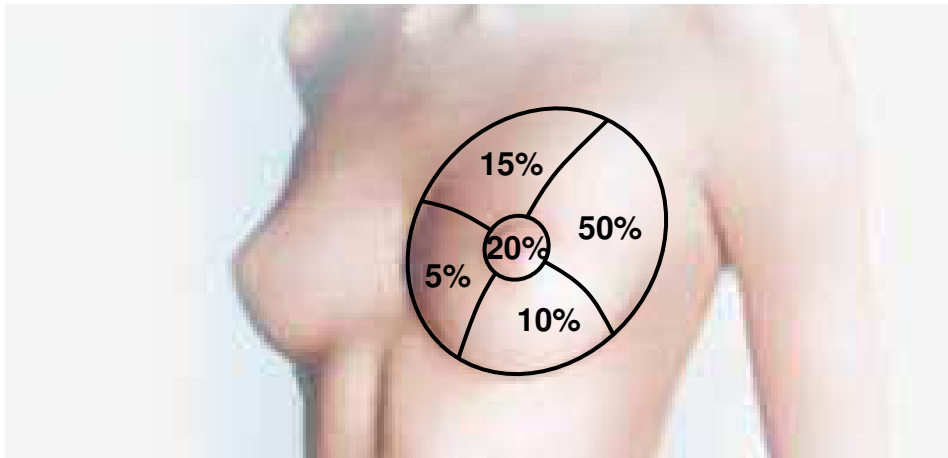


Abbildung 2.3: Häufigkeit des Karzinomauftritts in Abhängigkeit der Verteilung im Drüsengewebe.

2.1.4 Risikofaktoren, Prognose und Behandlung

Verschiedene Faktoren erhöhen das Risiko an einem Mammakarzinom zu erkranken. Besonders starken Einfluss hat die genetische Disposition, beispielsweise durch eine familiäre Belastung. Sie wird vermutet, wenn ein erstgradiger Verwandter, also die Mutter oder eine Schwester, erkrankt sind. Bekannte Mutationen liegen auf den *BRCA 1* oder *BRCA 2* Genen. Liegen im Extremfall Mutationen beider Gene vor, kann das individuelle Risiko, an Brustkrebs zu erkranken, bis auf nahezu 100 % ansteigen. Auch das Alter an sich oder eine zurückliegende Erkrankung der anderen Brust erhöhen das relative Erkrankungsrisiko um ein Vielfaches. Während im Alter unter 20 Jahren das Mammakarzinom extrem selten ist, steigt das Risiko mit zunehmendem Alter an und hat seinen Höhepunkt im Alter zwischen 55 und 80 Jahren. Statistiken belegen zudem, dass übermäßiger Alkoholkonsum, Fettsucht, Kinderlosigkeit bzw. späte Schwangerschaften nach dem 30. Lebensjahr, Wirkung der Hormonersatztherapie oder die Bildung von Mastopathien das Risiko für eine Erkrankung erhöhen [SW07].

Die Häufigkeit von Mammakarzinomen in der Weltbevölkerung ist regional unterschiedlich; in Japan, Südamerika oder der Arktis tritt es wesentlich seltener auf als in Europa oder Amerika. Es wird vermutet, dass regionale Gegebenheiten wie Ernährung bzw. das Stillverhalten (Eskimos stillen 3 Jahre oder länger) das Risiko einer Erkrankung beeinflussen [Lip06].

Während gutartige Veränderungen der Brust für gewöhnlich ohne Folgen bleiben, führt das Mammakarzinom unbehandelt zwangsläufig zum Tod der Erkrankten. Bereits drei Jahre nach Entdeckung des Karzinoms verstirbt die Hälfte der unbehandelten Betroffenen. Die Lebensgefahr entsteht erst durch die Bildung von *Metastasen*. Besonders häufig werden zuerst die Lymphknoten in der Achsenhöhle befallen. Eine Ausbreitung über den Blutweg erfolgt meist in das Knochenmark, in die Lunge, in die Leber oder in die Eierstöcke.

Entscheidend für die Prognose sind die Tumorgröße und das Vorhandensein von Lymph-

knotenmetastasen. Patientinnen, bei denen kleine Karzinome unter 1 *cm* Durchmesser diagnostiziert wurden, haben eine 5-Jahres-Überlebensrate von über 95 %. Die Überlebensrate nimmt mit der Tumorgöße zum Zeitpunkt des Behandlungsbeginns exponentiell ab. Sind bereits mehr als vier Lymphknoten positiv, wird die Prognose weitestgehend durch das Ausmaß der Metastasen bestimmt [FYM93].

Eine Dauerheilung der Erkrankten ist nur möglich, wenn das Karzinom komplett entfernt wird und keine Metastasen im Körper verbleiben. Zur Lokalbehandlung steht die operative Entfernung des Karzinoms (*Ektomie*) oft mit nachfolgender Bestrahlung der Brust und der Lymphabflusswege im Vordergrund. Je nach der lokalen Voraussetzung muss entweder die gesamte Brustdrüse (modifizierte Mastektomie oder sub-cutane Mastektomie) und eventuell der darunterliegende Brustmuskel (radikale Mastektomie) oder als brusterhaltende Operation nur das Karzinom im gesunden Gewebe entfernt werden. Damit brusterhaltend operiert werden kann, muss der Durchmesser des Karzinoms im Verhältnis zur Brust ausreichend klein sein. Das Karzinom darf weder in die Muskelfaszie noch in die Haut eingewachsen sein und es darf kein multizentrisches Karzinom vorliegen. Zusätzlich werden die Lymphknoten der Achselhöhle oder neuerdings zumindest der sogenannte Wächterlymphknoten entfernt. Systemweite Therapien wie Chemotherapie oder Hormontherapie werden unterstützend eingesetzt [Lip06].

2.2 Mammadiagnostik

Neben der Tastuntersuchung (*Palpation*, s. Kap. 2.2.1) stehen mit der Mammographie (s. Kap. 2.2.2), Sonographie (s. Kap. 2.2.3) und kontrastmittelgestützter MRT (s. Kap. 2.2.4) drei bildgebende Verfahren für die Diagnose von Brustkrebs zur Verfügung.

2.2.1 Palpation

Viele Läsionen können als knotige Verhärtungen ertastet werden. In oberflächlicher Lage oder kleiner Brust ist es möglich, bereits kleinste Karzinome zu entdecken. Dagegen können tiefer liegende Tumore von mehr als 2 *cm* Durchmesser in einer großen Brust unentdeckt bleiben. Tatsächlich bleiben jedoch mehr als 50 % der Tumore zwischen 1,0 bis 1,5 *cm* durch die Palpation unauffällig [CTC91, RBC93]. Besonders diffus wachsende Tumore wie z.B. das *lobuläre Karzinom* oder in-situ Karzinome lassen sich kaum ertasten. Eine vorhandene Mastopathie, die ebenfalls als Knoten wahrgenommen wird, erschwert zudem die Palpation. Dennoch kann auf den Tastbefund nicht verzichtet werden.

In 30 - 40 % aller Tastbefunde findet sich ein Mammakarzinom in der Biopsie. Bis zu 10 % aller klinisch auffälligen Tumore bleiben in der Mammographie unauffällig [BLGG91, Fei91, Bak82]. Auch axillare Ausläufer können sich dem Blickwinkel der Mammographie entziehen. Jeder unklare oder karzinomverdächtige Tastbefund muss durch eine weitere diagnostische Untersuchungsmethode (s. Kap. 2.2.2 bis Kap. 2.2.4) abgeklärt werden. Insgesamt senkt die Tastuntersuchung als alleinige diagnostische Methode die Mortalität nicht [HKS03, Lip06].

2.2.2 Mammographie

Die *Mammographie* ist das Standardverfahren sowohl zur Vorsorgeuntersuchung (*Screening*) als auch zur Abklärung symptomatischer Patientinnen.

Technik Die Untersuchung basiert auf Röntgentechnik, bei der ein Objekt von einer Seite mit elektromagnetischen Wellen durchstrahlt wird, während auf der Gegenseite die durch das Objekt hervorgerufene Schwächung gemessen wird. Als Bild entsteht ein zweidimensionaler (*2D*) Schattenwurf des Objektes, auf dem röntgendichtere Strukturen heller dargestellt sind als strahlendurchlässigere. Die für die Mammographie speziell eingesetzten Röntengeräte haben eine sehr hohe örtliche Auflösung (ca. $50 \mu\text{m}$ pro Pixel in der digitalen Mammographie) und erzeugen kontrastreiche Aufnahmen bei geringer Strahlung. Vor der Aufnahme wird die Brust zwischen zwei Plexiglasplatten komprimiert, es folgen für gewöhnlich Aufnahmen aus zwei verschiedenen Ebenen.

Während das Fettgewebe der Brust strahlendurchlässig ist, sind das Drüsengewebe und Tumore röntgendichter. Bei fettreicher Brust ist die Mammographie bei Erkennung von Karzinomen somit sehr sensitiv. Bei Frauen über 65 Jahren liegt die Sensitivität bei 81 bis 94 %, bei jungen Frauen unter 40 Jahren sinkt sie aufgrund des zumeist höheren Anteils an Drüsengewebe und der Neigung zur Bildung von Mastopathien auf 54 bis 58 % [KGB+96, RHW+98].

Besonders sensitiv ist die Mammographie beim Erkennen der sehr röntgendichten Mikroverkalkungen, die von ca. 25 bis 40 % aller invasiven und etwa 40 bis 90 % der nicht invasiven Karzinome gebildet werden [HKS03, SFP88, GFC+92].

Die Mammographie ist jedoch wenig spezifisch, denn sie lässt nur wenige Möglichkeiten zur Differenzierung von gut- und bösartigen Verdichtungen. Wegweisend ist oft ein internes Verkalkungsmuster. Scharf abgegrenzte Tumore, Zysten und Lipome lassen sich nur bei kompletter Lage im Fettgewebe sicher als gutartige befunden. Viele andere Verdichtungen können nicht näher eingeordnet werden. Erschwerend können auch gutartige Veränderungen Mikroverkalkungen bilden.

Da das bildgebende Verfahren nur zweidimensionale Aufnahmen zulässt, können überlagernde Strukturen, wie ein dichter Drüsenkörper und kleine Tumore, kaum unterschieden werden. Alle unklaren Befunde müssen somit im Anschluss an die Mammographie genauer untersucht werden, entweder durch Biopsie oder einem der beiden folgenden bildgebenden Verfahren [HKS03].

Die Mammographie ist eine kostengünstige Untersuchungsmethode und zeichnet sich durch schnelle und einfache Handhabung aus. Eine regelmäßige Früherkennung (Screening-Untersuchung) wird bei Frauen im zweijährigen Turnus ungefähr zwischen dem 50. und 70. Lebensjahr in mehreren Europäischen und Amerikanischen Staaten durchgeführt. Hingegen ist die Mammographie für jüngere Frauen unter 40 Jahren aufgrund ihres dichten und strahlenempfindlichen Drüsengewebes nicht zur Vorsorge geeignet. Hier stellt die Sonographie (s. nächstes Kapitel) das primär eingesetzte Untersuchungsverfahren zur Abklärung von Tastbefunden dar [HKS03, Fri93].

2.2.3 Sonographie

Die *Sonographie* (*Ultraschall*) gilt als additives Untersuchungsverfahren in der Mammadiagnostik, das besonders zur Abklärung unklarer klinischer Befunde oder von Auffälligkeiten in der Mammographie eingesetzt wird.

Technik Die Sonographie verwendet eine Sonde, die kurze, gerichtete Ultraschallwellenimpulse aussendet. Diese breiten sich von der Ultraschallsonde in das darunter liegende Gewebe aus und erfahren je nach vorliegender Gewebeeigenschaft eine Abschwächung durch die Prozesse Schallabsorption, Streuung (an Gewebegrenzflächen) und Reflexion in unterschiedlichem Maße. Die reflektierten Anteile der ursprünglich homogenen Schallwellen werden von der Sonde in der Folge wieder registriert, sie bestimmen die Echogenität, die sich je nach Gewebetyp in der Schalllaufzeit und -amplitude unterscheidet. Aus den Abweichungen zwischen gesendeten und empfangenen Signalen wird ein *2D*-Bild konstruiert, in dem sich Gewebe mit einer geringen Echogenität (echoarm) mit dunkleren und eine hohe Echogenität (echoreich) mit helleren Grauwerten darstellt [Mil87].

Mit modernen Sonographiegeräten lassen sich Tumore mit einem Durchmesser kleiner als 5 mm darstellen. Fragliche Strukturen, die sich in der Mammographie überlagern, können bei Kenntnis der ungefähren Lage mit der Sonographie weiter charakterisiert werden.

Flüssigkeiten sind gering bis gar nicht echogen, somit lassen sich einfache Zysten durch die Sonographie eindeutig diagnostizieren.

Während die Echogenität von Drüsengewebe hoch und die von Fettgewebe geringer ist, kann die Echogenität von Karzinomen sehr variabel sein; oft sind sie echoarm mit gemischten Anteilen. Ein kleines Karzinom im echoarmen Fettgewebe kann kaum zu sehen sein. Ein echoarmes Karzinom im echoreichen Drüsengewebe wird durch die Sonographie dagegen leicht erkannt. Eine wesentliche Rolle für die Diagnostik spielen auch die Schallphänomene, welche aus Richtung der Sonde hinter (dorsal) dem Befund entstehen. Einfache Zysten zeigen hier oft eine dorsale Schallverstärkung, während bösartige Tumore eine Schallauslöschung (Schallschatten) hervorrufen [OVT⁺91, BS91].

Durch die Sonographie lassen sich zudem morphologische Kriterien wie die Schärfe des Randes, runde oder längliche Form und Echotextur des Befundes bestimmen, die Anhaltspunkte für die Differenzierung von gut- oder bösartigen Prozessen geben können (vgl. Tab. 2.1).

Die Sonographie selbst eignet sich jedoch nicht zum Einsatz im Screening. Neben dem hohen zeitlichen Untersuchungsaufwand ist die Sensitivität zu niedrig und die Treffsicherheit von kleinen Karzinomen zu gering [HKS03, JB98].

2.2.4 Magnetresonanzmammographie

Die *Magnetresonanzmammographie* (*MRM*) als relativ junges Verfahren zur Diagnose von Brustkrebs basiert auf der *Magnetresonanztomographie* (*MRT*). Es ist das sensitivste Verfahren zur Erkennung von invasiven Karzinomen ab einer Größe von 3 mm und stellt neben der Morphologie auch die Gefäßversorgung des Tumors mittels Kontrastmittel (*KM*) dar.

Technik Die Funktionsweise der MRT-Geräte basiert auf der Messung magnetischer Momente von Atomen. Durch den Eigendrehimpuls (*Spin*) der Protonen und Neutronen von Atomkernen wird ein magnetisches Moment hervorgerufen. Vom Standpunkt der klassischen Physik kann ein Atomkern vereinfacht auch als magnetischer Kreisel angesehen werden. Mit Hilfe von magnetischen Feldern und hochfrequenten elektromagnetischen Wellen lässt sich dieser physikalische Effekt zur Erzeugung von Bilddaten nutzen. Hierbei spielt nur das Wasserstoffatom aufgrund der hohen Konzentration im biologischen Gewebe des Patienten und des großen magnetischen Moments eine tragende Rolle. Da besonders Weichteilorgane viele Wasserstoffatome pro Volumen aufweisen, werden diese besonders kontrastreich in Bildinformationen umgesetzt.

Der Patient wird in dem starken statischen Magnetfeld des MRT-Geräts mit einer Feldstärke von heute etwa 0,5 T bis 3,0 T (in der Einheit T für *Tesla*¹) gelagert, wodurch die Rotationsachse der Spins in einen Gleichgewichtszustand ausgerichtet werden. Die hervorgerufene magnetische Wirkung lässt sich als vektorielle Größe in eine Längs- und Querkomponente zerlegen, die zunächst einmal parallel zum äußeren Magnetfeld ausgerichtet sind. Um die Kernmagnetisierung nachweisen zu können, wird sie durch die Einstrahlung eines äußeren Hochfrequenzimpulses aus ihrer Rotationsachse herausgelenkt. Das anschließende Bestreben der Magnetisierung, wieder in den Gleichgewichtszustand zurückzukehren, wird als *Relaxation* bezeichnet. Die eingebrachte Energie des Anregungsimpulses wird während des Relaxationsvorgangs freigegeben und durch eine externe Messspule als Signal gemessen. Die Dauer der Rückkehr des Magnetisierungsvektors in sein Ausgangszustand wird durch die zwei Zeitkonstanten T_1 und T_2 beschrieben.

Die Zeitkonstante T_1 gibt die Relaxationszeit der *Längskomponente* an, sie ist je nach Gewebe sehr unterschiedlich: Fettgewebe relaxiert wesentlich kürzer als Gewebe wie Hirn, Leber, Niere oder Muskel, während Flüssigkeiten mit hohen T_1 Werten deutlich langsamere Zeiten aufweisen. Die Zeitkonstante T_2 beschreibt das Relaxationsverhalten der *Querkomponente*. Flüssigkeiten haben sehr hohe T_2 Zeiten, vitale Tumore und Entzündungen zeichnen sich häufig durch verlängerte Zeiten im Vergleich zum umliegenden gesunden Gewebe aus. Strukturen mit geringer Anzahl an Wassermolekülen wie Knochen oder Sehnen haben so kurze Relaxationszeiten, dass sie kaum messbare Signale freigeben.

Die durch die Magnetisierungsvektoren induzierte Spannung in der Empfangsspule kann zu einem Zeitpunkt nur als eindimensionales Signal für das gesamte Untersuchungsgebiet gemessen werden. Erst durch den Einsatz von orthogonal zueinanderstehenden Gradientenfeldern erfolgt für jedes gemessene Signal eine räumliche Zuordnung. Dazu wird eine Kombination von Hochfrequenzimpulsen und Magnetfeldern bestimmter Frequenzen bzw.

¹1,0 T entsprechen etwa der 20.000 fachen Feldstärke des Erdmagnetfeldes

Stärken mehrfach pro Sekunde ein- und ausgeschaltet. Erst nach Ablauf einer solchen *Sequenz* können alle gemessenen Signale zu einem Bild rekonstruiert werden. Die Technik der MRT erlaubt es dem Anwender heute, dreidimensionale (*3D*) Sequenzen zu verwenden und dabei ein beliebig orientiertes rechtwinkliges Volumen ($FoV = \text{Field of View}$) im Bereich der Empfangsspule zu definieren und auf das zu messende Körperteil anzuwenden.

Auf den erzeugten *3D*-Bilddaten einer T_1 -gewichteten Sequenz erscheinen Gewebe mit langer T_1 Relaxationszeit wie beispielsweise Luft, Flüssigkeiten oder entzündliche bzw. tumoröse Gewebe signalarm (*hypointens*), während beispielsweise Fettgewebe mit einer kurzen T_1 Relaxationszeit signalreich (*hyperintens*) abgebildet wird. Dagegen werden auf T_2 -gewichteten Aufnahmen Gewebe mit langer T_2 Relaxationszeit wie beispielsweise Flüssigkeiten und entzündliche bzw. tumoröse Gewebe hyperintens und Gewebe mit einer kurzen T_2 Relaxationszeit wie beispielsweise Muskulatur, Knochen oder auch fibröses Gewebe hypointens abgebildet. Hyperintense *Signalintensität* (*SI*) werden im rekonstruierten Bild durch helle Grauwerte abgebildet, hypointense *SI* durch dunkle Grauwerte dargestellt.

Sowohl für T_1 - als auch T_2 -gewichtete Messungen kann durch geschaltete Präparationsimpulse vor dem eigentlichen Anregungsimpuls die spätere Darstellung von Fett im Bild unterdrückt werden. Dazu existieren verschiedene Verfahren, beispielsweise die Verwendung eines Inversionsimpuls in Abhängigkeit der Relaxionszeit des Fettgewebes (*Fettunterdrückung*) oder, basierend auf der Verschiebung der Resonanzfrequenzen der Protonen von Wassermolekülen und Fettsäuren, der Einsatz eines CHESS-Impuls zur Absättigung der Resonanzen der Fettsäurenprotonen (*Fettsättigung*). Diese Techniken erfordern eine gute Magnetfeldhomogenität im gesamten FoV und führen auch wegen eines ungünstigeren Signal-Rausch-Verhältnisses meist zu schlechterer Bildqualität als ohne die Verwendung einer *Fettsuppression* [Sch05, Fis00].

Der Durchbruch in der Untersuchung der Mamma durch die MRT kam erst in den Achtziger Jahren durch die Möglichkeit zum Erhalt dynamischer Aufnahmen unter Verwendung von paramagnetischen MR-Kontrastmittels und dem Einsatz spezieller Oberflächenspulen [HHS⁺86, WLM84, KK89]. Die klassische MRT wird dabei um einen funktionalen (dynamischen) Anteil erweitert (*fMRT*), denn durch die zusätzliche Messung der *Kinetik* wird die Verteilung des Kontrastmittels im zeitlichen Verlauf und damit die Gefäßversorgung der Tumore erfasst (s. Kap. 2.3.2). Die technische Umsetzung erfolgt mit der Messung des gleichen Brustbereichs durch mehrere aufeinanderfolgende Sequenzen. Die *Anreicherung* einer Läsion durch das KM pro Zeitserie lässt sich anschließend aus den Bildfolgen als *Zeit/Signalintensitätskurve* zusammenstellen (s. Abb. 4.7).

In der MRT finden paramagnetisches Kontrastmittel basierend auf Gadolinium-DTPA (Gd-DTPA) Anwendung, die ungepaarte Elektronen mit einem hohen magnetischen Moment besitzen. Diese Eigenschaft verkürzt die T_1 und T_2 Relaxationszeiten für KM aufnehmendes Gewebe und führt insbesondere auf den T_1 gewichteten Sequenzen zu erhöhten Signalintensitäten. Somit eignen sich T_1 gewichtete Aufnahmen zur Aufzeichnung der Kinetik und Bewertung der Morphologie.

Überlagerungsprobleme von Strukturen wie in der zweidimensionalen Mammographie existieren nicht, durch das Mitbetrachten der zeitlichen Komponente erzeugt die fMRT

vierdimensionale ($4D$) Bilddaten. $3D$ -Sequenzen heutiger 1,5 Tesla Geräte erzeugen ein Bildvolumen mit einer Ortsauflösung zwischen 0,5 und 8,0 mm^3 (s. Kap. 4.1.1).

Die funktionale MRT bildet derzeit für die Erkennung invasiver Karzinome mit 90 bis 100 % das sensitivste Verfahren [Kai90, HKBL⁺92]. Durch Untersuchung sowohl der Kinetik (s. Kap. 2.3.2) als auch der Morphologie (s. Kap. 2.3.3) auf den T_1 -gewichteten Aufnahmen lassen sich zusammen mit der Beurteilung der T_2 -gewichteten Aufnahmen (s. Kap. 2.3.4) vielseitige Kriterien für eine Diagnose (s. Kap. 2.3.5) gewinnen. Die Spezifität variiert in der Literatur und liegt zwischen 37 % und 97 % [HFH⁺93, FHV⁺93, MBB⁺95, SLQD⁺97, LDC⁺98, HWP⁺89, KZ89]. Ihr Wert wird stark durch die Qualität der Bilddaten und die Erfahrung des Radiologen mitbestimmt. Speziell für in-situ Karzinome liegt die Sensitivität mit 50 - 94 % niedriger als bei der Mammographie, denn Mikroverkalkungen sind in der MRT nicht signalgebend [GML⁺96, FWB⁺96, ORS⁺97, SKH⁺97]

Die Befundung der $4D$ -Bilddaten ist sehr zeitaufwendig und komplex, die Diagnostik ist nicht standardisiert (s. Kap. 2.2). Während der Radiologe in der Mammographie für gewöhnlich zwei Bilder pro Brust befundet, bereitet die fMRT aufgrund der $3D$ -Darstellung und erfassten zeitlichen Komponente eine regelrechte Flut von mehreren hundert Bildschichten. Gerade hier kann eine computergestützte Auswertung zur Erhöhung der Qualität und vor allem der Spezifität beitragen. Die Befundungszeit erhöht sich zudem durch jede falsch-positive Anreicherung (s. Kap. 2.3.1) und dauert – von der Zahl der anreichernenden Strukturen abhängig – insgesamt zwischen 5 bis 40 Minuten. Die Bilderzeugung ist durch Vorbereiten der KM-Applikation und der Akquisitionsdauer des MRT-Geräts sehr zeitaufwendig. Die Kosten einer Untersuchung sind wesentlich höher als für die beiden zuvor genannten bildgebenden Verfahren.

Die MRT ist aufgrund des Zeit- und Kostenfaktors ungeeignet zum Einsatz für das Screening. Sie wird nur für bestimmte Fragestellungen ergänzend eingesetzt:

- Patienten mit nachgewiesenem Brustkrebs und dem Verdacht auf weitere Absiedelungen in derselben Brust oder aber auf der Gegenseite
- Patienten mit Brustimplantaten, multizentrischen Tumoren oder zur Differenzierung einer postoperativen Narbe von einem Karzinom
- Patienten mit Lymphknotenmetastasen in der Achselhöhle bei bislang unbekanntem Ursprungstumor
- erweiterte Früherkennung bei Patienten mit Genmutationen (*BRCA 1*, *BRCA 2*)
- Krankheitsstadienbestimmung (Staging) für die Planung des therapeutischen Vorgehens (beispielsweise neoadjuvante Therapie zur Verkleinerung des Tumors vor einem operativen Eingriff)

2.3 Befundung

Auffällig und durch den Radiologen zu klassifizieren sind alle KM-Anreicherungen (engl. *Enhancement*) im Bildmaterial mit einem deutlich höheren Signal als das umgebende normale Drüsengewebe. Läsionen mit einer solch auffälligen Anreicherung können erst nach Ausblendung des Fettgewebes (s. Kap. 2.3.1) mit hoher Sensitivität erkannt und ihre Morphologie und Kinetik (s. Kap. 2.3.2 - 2.3.4) zur Ableitung der Diagnose (s. Kap. 2.3.5) bestimmt werden. Die Aufzeichnung der Einzelbefunde ist durch BI-RADSTM (s. Kap. 2.3.6) standardisiert.

2.3.1 Bildsubtraktion

Das Erkennen von KM-Anreicherung auf T_1 -gewichteten Aufnahmen wird durch das ebenfalls signalreiche und reichlich vorhandene Fettgewebe der Mamma (s. Kap. 2.1.1) eingeschränkt. Zur eindeutigen Beurteilung der Morphologie müssen zuerst die Fettbereiche ausgeblendet werden.

In den USA werden dazu standardmäßig fettsupprimierte Aufnahmen eingesetzt (s. Kap. 2.2.4). Diese bedürfen jedoch längerer Messzeiten und sind aufgrund von Inhomogenitäten und Bildrauschen qualitativ meist schlechter.

In Europa wird überwiegend ohne Fettsuppression gearbeitet, stattdessen hat sich das Verfahren der Bildsubtraktion etabliert. Vor der Applikation des KMs wird im Allgemeinen eine native Serie gemessen, deren Signalintensitäten von allen nach der KM-Gabe folgenden schichtidentischen Bilddaten subtrahiert werden. Signalintensitäten von gesundem Geweben wie beispielsweise das Fett, welche durch das applizierte KM kaum oder keinen Signalanstieg zeigen (s. Abb. 2.4 a), werden durch die Subtraktion ausgelöscht. Die berechneten Differenzbilder stellen letztlich nur noch die Anreicherungen von Geweben durch das Kontrastmittel dar (s. Abb. 5.2 a).

Die Qualität der Differenzbilder verschlechtert sich, wenn die Bildinhalte der untersuchten Gewebeabschnitte der zu subtrahierenden Sequenzen vor und nach der KM-Applikation nicht mehr übereinstimmen. Gründe für die Veränderungen sind vielfältig, beispielsweise spielen der Herzschlag, das Ansprechen, die Atmung oder Entspannung der Patientin während der langen Untersuchung eine Rolle. Dadurch sind die Strukturen der Brust nicht mehr deckungsgleich und werden durch die Subtraktion nicht mehr vollständig ausgelöscht. Solche *Bewegungsartefakte* zeichnen sich durch negative oder positive Subtraktionswerte ohne Information im Differenzbild ab. Die positiven Werte bergen dabei die Gefahr, als zusätzliche (falsch-positive) KM-Anreicherung wahrgenommen zu werden (s. Abb. 5.2 b). Durch die Bildverschiebung erscheinen auf den Differenzbildern auch die echten Kontrastmittelanreicherungen verfälscht. Damit sind sowohl die morphologischen Informationen als auch die kinetischen Analysen (s. nachfolgende Kapitel) auf Bildpunktebene kompromittiert.

Die KM-Darstellung auf den erzeugten Bildvolumen fettsupprimierter Sequenzen wird durch Bewegungen weniger beeinflusst, da das Fettgewebe zum Zeitpunkt der Messung durch das MRT-Gerät selbst und nicht durch eine Subtraktion mit einer zeitlich zurückliegenden nativen Aufnahme ausgelöscht wird. Die kinetische Analyse über allen Messungen

unterliegt jedoch auch hier eventuellen Bewegungen im Bild und ist somit eine Problemlage unabhängig von der gewählten Ausblendungsvariante [Fis00].

2.3.2 Kinetik

Das KM trägt unterschiedlich zur Bildgebung der Mamma Diagnostik bei. In der Brust-MRT können Läsionen anhand eines schnelleren Anreicherns mit Kontrastmittel von gesundem Gewebe differenziert werden (s. Abb. 2.4 a). Der pathophysiologische Hintergrund ist nicht vollständig geklärt, die Eigenschaft der Neoangiogenese (s. Kap. 2.1.2) von malignen Tumoren spielt dabei jedoch eine entscheidende Rolle.

Für das typische Anreicherungsverhalten von Karzinomen durch das KM stellen sich zwei zeitliche Effekte ein. Sie sind vor allem durch die Permeabilität der neu induzierten Gefäße des Karzinoms geprägt (s. Kap. 2.1.2), die auch dem Kontrastmittel den Austritt in den extrazellulären Raum erlaubt: Die hohe Gefäßdichte und beschleunigte Diffusion (gleichmäßige Verteilung) des KMs in den extrazellulären Raum des Tumors aufgrund der höheren Gefäßpermeabilität führt zu einem intensiveren Anreichern mit Kontrastmittel. Gleichzeitig führen arteriovenöse Shunts und wiederum die höhere Permeabilität später auch zu einem schnelleren Auswaschen des KMs. Somit ergeben sich für Karzinome im Allgemeinen ein starker Anstieg in der initialen Phase, gefolgt von einem *Wash-Out* in der postinitialen Phase (s. Abb. 2.4 b) [KSS95, Kai90, LC69, SY88, SWH+97].

Durch die gemessene Kinetik wird also letztendlich versucht, Rückschlüsse auf eine mögliche Hypervaskulierung eines Karzinoms zu ziehen. Verschiedene Untersuchungen haben anhand von histologischen Auswertungen bestätigt, dass die Zunahme der Gefäßdichte mit dem Grad der Bösartigkeit von Tumoren korreliert [FWIH89, CMCD95, FLS+94, GWB+94, OKC+95, TKT93, WSWF91].

Diese klare Korrelation zeigte sich jedoch leider nicht, wenn man die Gefäßdichte anhand der durch die fMRT dargestellten Kinetik beurteilt. Zahlreiche Vergleichsstudien befassen sich mit der Analyse der KM-Anreicherung von Mammakarzinomen mit uneinheitlichen Ergebnissen [TBS+06, Fis98, HES97, SOS97, BMM97, FGC94, WFP92, MBH97]. Die Kinetik für Karzinome ist nicht zwingend spezifisch, gutartige Läsionen zeigen ein sehr variables Anreicherungsverhalten. So konnten für gutartige Fibroadenome und auch für Papillome teilweise ein karzinomtypisches Kontrastmittelverhalten nachgewiesen werden [FHV+93, FASW92]. Auch zeigten Karzinome eine verzögerte KM-Aufnahme ähnlich den gutartigen Läsionen [KMK+99, HKHP+94].

Die Ursachen hierfür können auch in der Methodik der MRM begründet sein. Beispielsweise wird die Beziehung von KM-Konzentration im Gewebe und Signalanreicherung im Bild nicht linear abgebildet, sondern hängt wesentlich von der Relaxationszeit des Gewebes ab [ABBM05, JTKG00]. Kuhl [Kuh00] sieht den Grund der inkonsistenten Korrelationen auch in der Komplexität des Anreicherungsprozesses, an dem viele Verteilungsfaktoren beteiligt sind. Tabelle 2.3 führt einige Faktoren auf, die den Verlauf der Kinetik beeinflussen können; sie sind nach der für die Untersuchung eingesetzten Hardware, Gewebeeigenschaft und Auswertungsprozess unterteilt. Während die Gefäßdichte und -permeabilität als veränderliche Faktoren bekannt sind und die Diagnosegrundlage bilden, muss jeder weitere

Kategorie	Faktoren
Hardware	MRT-Gerät (Feldstärke, Hersteller, Standort, Feldinhomogenitäten, ...) Protokoll (eingesetzte Sequenzen, Anzahl der Serien, Ortsauflösung, ...) Spule Kontrastmittel (Diffusionseigenschaften, Relaxivität, ...) Bolus (Zeitpunkt, Injektionsgeschwindigkeit, Volumen, ...)
Gewebe	Gefäßdichte und -permeabilität Relaxationszeit KM-Konzentration, Flüssigkeitsmenge (Patientengewicht) Flussgeschwindigkeit des KMs Temperatur hormonale Regulation Ausscheidungsverhalten des KMs durch die Niere
Auswertung	Messregion (Größe, Position, Anzahl der Stützstellen, ...) Bewegungsartefakte

Tabelle 2.3: Einflussfaktoren auf die Bildqualität und der dargestellten Kinetik in der funktionalen MRT

Faktor möglichst konstant gehalten werden, um die aus den Bilddaten abgeleitete Kinetik mit gleichen Maßstäben bewerten zu können.

Die Menge an applizierten Kontrastmittel beeinflusst die spätere Konzentration im Gewebe und damit die gemessene Signalstärke im Bild. Die Höhe der Dosierung des KMs wird daher am Körpergewicht der Patientin bemessen. Für den Einsatz von 3D-Sequenzen wird eine Dosis zwischen 0,1 und 0,2 mmol Gd-DTPA/kg Körpergewicht empfohlen [Fis00, HKHP+94].

Die Durchblutung der Mamma unterliegt der hormonalen Regulation und beeinflusst den Anreicherungsprozess. Kuhl et al. [KSS95] hatten an gesunden Probandinnen gezeigt, dass das Verhalten der Signalanreicherung in der zweiten Zykluswoche am günstigsten für die MRT-Untersuchung ist.

Ein nekrotischer Teil eines Tumors zeigt eine andere Kinetikcharakteristik als der mit Gefäßen versorgte Bereich. Entscheidend ist die Wahl der Messregion zur Bestimmung der Kinetik. Fischer [Fis00] schlägt zur Bestimmung der Kinetik die Wahl einer Messregion nur an der signalstärksten Position der Läsion (*Hotspot*) vor. Innerhalb der Begrenzung wird der durchschnittliche Signalwert als Mittelwert über alle eingeschlossenen Voxel berechnet. Die Region wird an gleicher Position auf allen Sequenzen (nativ und nach KM-Gabe) angewendet und die einzelnen Messwerte zu einer Zeit/Signalintensitätskurve zusammengestellt (s. Abb. 4.7). Führen Verschiebungen der Bildinhalte (s. vorheriges Kapitel) dazu, dass die Anreicherung nicht mehr im Fokus der gewählten Region liegt, muss der Hotspot entsprechend mitverschoben werden. Die Signalintensitäten können sowohl auf den T_1 -gewichteten

Aufnahmen als auch auf den Differenzbildern bestimmt werden (vgl. Kap. 2.3.6).

Trotz der vielen Einflussfaktoren und uneinheitlichen Vergleichsstudien ist der allgemeine Konsens, dass Tumore aufgrund ihrer Gefäßdichte schneller als das gesunde Drüsen- gewebe (s. Abb. 2.4 a) anreichern. Besonders invasiv wachsende Karzinome erreichen üblicherweise eine maximale Anreicherung innerhalb der ersten 3 Minuten. Gesundes Drüsen- gewebe nimmt das KM meist langsam und kontinuierlich auf. Die Deutsche Röntgenge- sellschaft empfiehlt zur Erfassung der Kinetik einen Untersuchungszeitraum von 8 bis 10 Minuten. Weiter befürwortet sie für die fMRT mindestens 5 Messungen mit einer zeitlichen Auflösung von 1 bis maximal 2 Minuten pro Serie nach der KM-Applikation [DRG95].

Für die Entdeckung von Tumoren gilt für die fMRT eine Grenze von etwa 3 mm Durchmesser, denn kleinere Objekte werden aufgrund der fehlenden Angiogenese kaum abgebildet [Fol92].

Zur Erhöhung der Spezifität bei der Differenzierung der Dignität von Tumoren muss die Kinetik zusammen mit der Morphologie (s. nächstes Kapitel) bewertet werden [Fis00, Kuh00].

2.3.3 Morphologie

Das unterschiedliche Wachstumsverhalten von gut- und bösartigen Läsionen führt auch zu einer differenzierbaren Morphologie (s. Tab. 2.1). Sie kann in den Differenzbildern durch die Anreicherung in ihrer äußeren *Form*, der *Begrenzung* des Randbereichs und der *Textur* charakterisiert werden und trägt somit zur Diagnose bei.

Je höher die Ortsauflösung der Bilddaten, desto detailreicher kann die Morphologie der Anreicherung erfasst und entscheidende Kriterien zur Differenzierung noch erkannt werden. Ebenso bestimmt die Ortsauflösung darüber, ab welcher Größe Anreicherungen im Bild erst erkennbar werden [VTK04, KJM+06].

Vor allem das Vorliegen spikulierter (sternförmiger) Formen ist ein sicherer Hinweis auf einen malignen Tumor, der radiäre Ausziehungen gebildet hat. Rundliche Formen deuten dagegen eher auf eine Gutartigkeit der Läsion hin. Eine glatte Begrenzung spricht ebenso für benigne Tumore, während irreguläre Ränder häufig durch Karzinome gebildet werden. Bei zu niedriger Bildauflösung kann eine irreguläre Begrenzung dennoch glatt erscheinen. Ist die endotumorale Anreicherung homogen, weist dies auf einen gutartigen Prozess hin, während ein heterogenes Erscheinen auf das Gegenteil deuten lässt.

Während die Angiogenese vor allem im Randbereich des Tumors ausgeprägt ist, herrschen im Zentrum besonders in fortgeschrittenen Stadien invasiv duktaler Karzinome oft *Nekrosen* und *Fibrosen* vor. Daher geben besonders randständige Anreicherungen (*Ringenhancement*) einen sicheren Hinweis auf ein schnell wachsendes Karzinom [NSO+97, BFV02, LML+02].

Die morphologischen Kriterien sind jedoch nicht immer eindeutig in Bezug auf die Dignität. Die gutartigen Fibroadenome können beispielsweise aufgrund der Bildung von Septen in ihrer Textur inhomogen erscheinen und erfüllen damit ein Malignitätskriterium (vgl. Tab. 2.1).

2.3.4 T_2 Wichtung

Durch das Mitbetrachten einer T_2 -gewichteten Sequenz lässt sich vor allem der Wasser- und Fettgehalt von Gewebe beurteilen. Einfache Zysten und Gefäße lassen sich so von tumorösen Läsionen anhand ihres hohen Signals differenzieren. Auch gutartige Läsionen wie Fibroadenome oder Mastopathien neigen ebenfalls zu erhöhten Signalen. Maligne Veränderungen erscheinen auf der T_2 -gewichteten Aufnahme dagegen häufig mit gering geminderter Signalintensität gegenüber dem gesunden Drüsenparenchym. Karzinome können eine Ödemzone um sich herum bilden (*peritumorales Ödem*), die signalreich um den dunkler dargestellten Tumor erkennbar ist. [KKM⁺99]

2.3.5 Diagnose

Während für die Befundung der MRM ein Standard erarbeitet wurde (s. Kap. 2.3.6), bleibt die Ableitung der Dignität aus den vielen bestimmbaren Einzelbefunden dem individuellen Vorgehen des Radiologen überlassen. Es herrscht Einigung darüber, dass sich erst durch das gleichzeitige Betrachten sowohl der Kinetik als auch der Morphologie die Spezifität der Diagnostik erhöht [Fis00, Kuh00]. Die Diagnostik der MRM gestaltet sich aufgrund der Informationsvielfalt komplex und zeitaufwendig und hängt somit im besonderen Maß von der Erfahrung des Radiologen und der Bildqualität ab.

Als ein häufig angewendetes Diagnoseverfahren wird eine Publikation aus dem Jahr 1999 genannt, in der Fischer et al. [FKG99] verschiedene Kriterien zur Beurteilung der Mamma-MRT zu einem Scoring System zusammenfassten. Das als *Göttinger Score* bekannte Verfahren bewertet die kinetischen und morphologischen Befunde durch Punkte, die zu einem Summenscore addiert werden. Je spezifischer ein Kriterium, desto höher seine Bewertung (1 oder 2 Punkte), unspezifische Merkmale erhalten gar keine Punkte (s. Tab. 2.4).

Zur Beurteilung der Kinetik unterteilt Fischer die Zeit/Signalintensitätskurve in eine initiale (1. bis 3. Minute) und eine postinitiale (3. bis 8. Minute) Phase und berechnet jeweils die Steigung (s. Formel 2.1 u 2.2, vgl. Abb. 2.4) zum Maximalwert:

$$\text{Initiale Phase} = \frac{(SI_{\max 1.-3.\min} - SI_{\text{native}})}{SI_{\text{native}}} \cdot 100\% \quad (2.1)$$

$$\text{Postinitiale Phase} = \frac{(SI_{8.\min} - SI_{\max 1.-3.\min})}{SI_{\max 1.-3.\min}} \cdot 100\% \quad (2.2)$$

Durch Anwendung von Schwellwerten wird das Verhalten der initialen und postinitialen Phase kategorisiert und wiederum mit Punkten bewertet (s. Tab. 2.4). Je stärker der Anstieg des Signals in der frühen Phase, desto höher die Punktbewertung für ein karzinomverdächtiges Verhalten. Ebenso ist ein Signalabfall in der späten Phase ein Zeichen für Bösartigkeit und wird als spezifisch bewertet. Fischer weist darauf hin, dass die Ergebnisse abhängig zur Signalstärke im Bild sind, so dass die Schwellwerte entsprechend der klinischen Handhabung des KM zur Evaluation angepasst werden müssen (vgl. Kap. 6.3.3).

Durch Nachschlagen des berechneten Summenscores aus Kinetik und Morphologie in der Scoringtabelle 2.5 kann die endgültige Dignität pro Läsion abgeleitet werden, zudem wird eine Empfehlung für das weitere Vorgehen gegeben (vgl. nächstes Kapitel).

Fischer et al. wendeten den Göttinger Score auf 463 präoperativ untersuchte Patientinnen mit insgesamt 143 benignen und 405 malignen Läsionen an und stellten die Ergebnisse der klinischen Untersuchung, Mammographie und Sonographie gegenüber (s. Tab. 2.6). Anhand des Summenscores konnte für die Mamma-MRT mit 93 % die höchste Sensitivität aller Methoden mit einer Spezifität von 65 % erzielt werden.

Seither wird versucht, immer mehr Kriterien aus den Bilddaten abzuleiten und für die Beurteilung zu nutzen, um vor allem die Spezifität weiter zu verbessern. Das Verwenden uneinheitlicher Kriterien birgt die Gefahr, dass Befunde und Diagnosen verschiedener Institute nicht mehr vergleichbar oder nachvollziehbar sind. Das nächste Kapitel beschreibt einen Standard zur Herbeiführung einer Vereinheitlichung der Brustuntersuchung.

2.3.6 BI-RADSTM

Um die Mammographiebefundung zu standardisieren und zu vereinheitlichen, wurde vom American College of Radiology (ACR) das *Breast Imaging Reporting and Data System* (BI-RADSTM) entwickelt [Ame03]. Es wurde nach 1992 und 1998 zuletzt im Jahr 2003 aktualisiert und setzt sich in der Übersetzung zunehmend auch im Deutschen Sprachgebrauch durch [FH06]. Es definiert für die Mamma-MRT die technischen Aspekte, Terminologien und Mindestanforderungen für eine einheitliche Befunderstellung.

Zur Bestimmung der Kinetik als Zeit/Signalintensitätskurve gibt BI-RADSTM Empfehlungen sowohl für die Position als auch für die Größe der Messregion, die im Bereich der stärksten Anreicherung liegen und mindestens drei Pixel umfassen sollte. Werden mehrere Regionen untersucht, ist abschließend die auffälligste Kinetik zu bewerten. Signalverläufe der initialen und postinitialen Phase werden durch die Terminologien *langsam*, *mäßig* und *stark* bzw. *Persistent*, *Plateau* und *Wash-Out* unterschieden (s. Abb. 2.4 b).

Zur Bestimmung der Morphologie werden die Anreicherungen in Foki, Herdbefunde und nicht raumfordernd unterschieden.

Als *Fokus* werden kleine punktförmige Anreicherungen bezeichnet, die aufgrund ihrer geringen Größe (kleiner ca. 5 mm) morphologisch nicht weiter charakterisiert werden können.

Herdbefunde sind dagegen raumfordernde Läsionen, die sich morphologisch näher beschreiben lassen. Das BI-RADSTM-Lexikon verwendet dazu definierte Terminologien (vgl. Kap. 2.3.3) für die Form, Begrenzung und interner Charakteristik der Anreicherung (s. Tab. 2.7).

Als *nicht raumfordernde Enhancements* werden alle Anreicherungen bezeichnet, die weder Foki noch Herdbefunde sind. Sie lassen sich durch ihre Verteilungsmuster innerhalb der Mamma, interne Charakteristika und der Symmetrie ihres Auftretens bezüglich beider Brüste beschreiben (s. Tab. 2.7).

Die verschiedenen Terminologien der Morphologie veranschaulicht der Standard durch Bildbeispiele. Das BI-RADSTM-Lexikon vereinheitlicht somit die Begrifflichkeiten und den

Kriterium	0 Punkte	1 Punkt	2 Punkte
Initiale Phase (in %, s. Formel 2.1)	$x < 50$, gering	$50 < x < 100$, mäßig	$x > 100$ stark
Postinitiale Phase (in %, s. Formel 2.2)	$x > +10$, Persistent	$-10 < x < +10$, Plateau	$x < -10$, Wash-Out
Form	rund, oval	irregulär dentritisch	-
Begrenzung	scharf	unscharf	-
KM-Verteilung	homogen, septiert	inhomogen	ringartig

Tabelle 2.4: Bewertung von morphologischen und kinetischen Kriterien durch den Göttinger Score (maximal 8 Punkte).

Summenscore	Kategorie	Bewertung	Konsequenz
0–1	I	sicher benigne	keine
2	II	wahrscheinlich benigne	ggf. Kontrolle
3	III	unklar	unklar, Zusatzinformation nötig
4–5	IV	wahrscheinlich maligne	histologische Abklärung
6–8	V	sicher maligne	histologische Abklärung

Tabelle 2.5: Durch die Scoringtabelle wird der Göttinger Summenscore (s. Tab. 2.4) abschließend bewertet.

Verfahren	Sensitivität	Spezifität	Treffgenauigkeit
Klinische Untersuchung	58 %	76 %	62 %
Mammographie	86 %	32 %	72 %
Sonographie	75 %	80 %	76 %
MRT	93 %	65 %	85 %

Tabelle 2.6: Gegenüberstellung verschiedener Diagnoseverfahren bei der Tumordifferenzierung [FKG99].

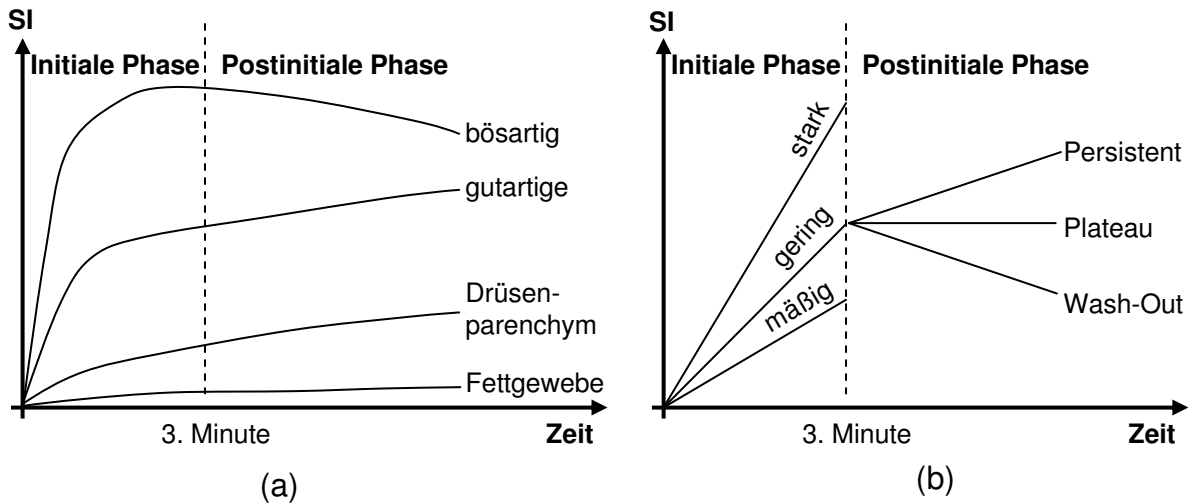


Abbildung 2.4: (a) Typische Signalverläufe gesunder Gewebe gegenüber gut- und bösartigen Veränderungen. (b) Schematischer Verlauf der Kinetik in der initialen und postinitialen Phase nach BI-RADS™

Fund	Kriterien	Terminologien
Fokus	-	-
Herdbefund	Form Begrenzung Endotumorale Anreicherung	rund, oval, lobuliert, irregulär glatt, irregulär, spikuliert homogen, heterogen, randständig, dunkle Septierung, anreichernde Septierung, zentrale Anreicherung
nicht raumforderndes Enhancement	Interne Charakteristika Symmetrie	fokal, linear, duktal, segmental, regional, multipel, diffus, homogen, heterogen, punktiert, gruppiert, dendritisch symmetrisch, asymmetrisch
Kinetik	Initiale Phase Postinitiale Phase	langsam, mäßig, stark Persistent, Plateau, Wash-Out

Tabelle 2.7: Auszug definierter Terminologien des BI-RADS™ -Lexikon.

Kategorie	Befund
0	keine abschließende Beurteilung, Untersuchung unvollständig oder mangelhaft
1	unauffällig / negativ
2	gutartig
3	wahrscheinlich gutartig – kontrollbedürftig
4	verdächtig – Biopsie empfohlen
5	karzinomverdächtig, angemessene Behandlungsmaßnahmen einleiten
6	gesichertes Karzinom

Tabelle 2.8: BI-RADSTM -Kategorien für die Brust-MRT.

Umfang eines Befundes, bietet dem Radiologen aber kein Verfahren zur abschließenden Diagnose (s. vorheriges Kapitel). Jedoch kann die Entscheidung des Radiologen aufgrund des standardisiert aufbereiteten Befundes nachvollzogen bzw. überprüft werden.

Hat der Radiologie eine abschließende Beurteilung anhand der gesammelten Kriterien vorgenommen, kann er diese durch die Vorgaben des BI-RADSTM -Lexikons standardisiert aufzeichnen. Das ACR schlägt hierfür die Einteilung des Befundes in sechs Kategorien vor (s. Tab. 2.8). Dadurch wird die Zuordnung von Befunden zur Dignität erleichtert und es werden klare Richtlinien für das weitere Vorgehen gegeben, beispielsweise ob eine Verlaufskontrolle ausreichend ist oder eine Biopsie erfolgen sollte.

2.4 Bildverarbeitung

In der internationalen Literatur findet man rund um den Begriff Bildverarbeitung eine Fülle von Schlagwörtern: *Machine* bzw. *Computer Vision*, *Image Understanding*, *Pattern Recognition*, *Image Interpretation* oder *Computer Imaging*. Hinter all diesen Begriffen verbirgt sich das generelle Problem der Wahrnehmung und Interpretation von beliebigen Bildszenen. Ein Eingangsbild wird dabei mit Hilfe des Computers in ein Ausgangsformat überführt. Die hierzu verwendeten Verfahren haben das Ziel, die Bildinhalte deutlicher sichtbar zu machen, Informationsgehalte zu erkennen und diese zu beschreiben oder zu interpretieren. In der vorliegenden Arbeit wird der Begriff *digitale Bildverarbeitung* als Synonym für diesen Komplex an Verfahren und Techniken verwendet.

Neben der Medizin und Biologie (Computergestützte Diagnose, Mikroskopie) findet die Bildverarbeitung in der Überwachung bzw. Beobachtung (Astronomie, Gesichtserkennung, Lippenlesen), Robotik (sichtbeeinflusste Maschinen, autonome Navigation), Industrie (Qualitätskontrolle, Bar-Code lesen, biometrische Sicherheitssysteme), Kriminalistik (Fingerabdruck, Schrifterkennung) oder Militär (Aufklärung, Zielerfassung und -verfolgung) Anwendung.

Anhand dieser Aufführung wird deutlich, dass ein hohes Interesse an der Bildverarbeitung besteht. Umso erstaunlicher ist es, dass kein generell einsetzbares Bildverarbeitungs-

tool mit akzeptabler Performanz existiert. Dafür gibt es verschiedene Gründe [CL97]:

Der menschliche Sehapparat ist funktionell nicht vollständig bekannt. Da die Informatik immer bemüht ist, ein Stück der realen Welt auf ein Computersystem abzubilden, scheitert sie hier bei der Realisierung des kompletten Leistungsspektrums unseres Sehvermögens. Mit der Terminologie der Informatik ausgedrückt, produziert unser visuelles System hoch aufgelöste Bilder, verarbeitet und interpretiert diese in Echtzeit und liefert dabei nahezu fehlerfreie Ergebnisse auf Eingangsdaten uneingeschränkter Inhalte. Eine Vorgabe, die schon alleine einen zu hohen Anspruch an die Leistungsfähigkeit heutiger Hardware stellt [Mor01]. Solange der Sehapparat und die Verarbeitung der Informationen durch das Gehirn nicht komplett entschlüsselt sind, wird es auch kein vereinheitlichtes Bildverarbeitungssystem geben.

Signifikant erfolgreiche Bildverarbeitungssysteme arbeiten daher immer auf Bildern mit begrenzten Inhalten (Handschriften, Gesichter, Satellitenaufnahmen, Fingerabdrücke, medizinische Bilder eines bestimmten Organs, ...). Diese spezialisierten Systeme haben meist akzeptable Laufzeiten, sind aber kaum erweiterbar für Bilder anderer Inhalte. Systeme, die für eine allgemeine Anwendung angedacht sind, erfordern einen höheren Entwicklungsaufwand und die resultierenden Endanwendungen sind zumeist langsamer als ein spezialisiertes System. Aber selbst solche allgemeinen Toolkits haben immer noch interne Beschränkungen bezüglich des Eingangs- und Ausgangsformats der Verarbeitung.

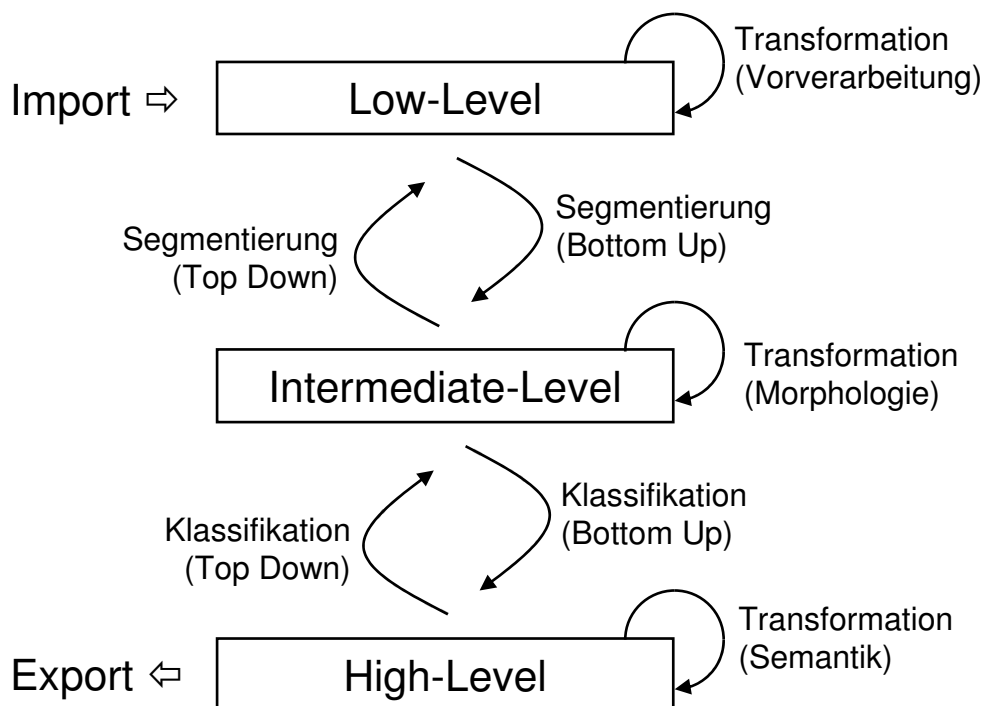


Abbildung 2.5: Verfahren der Bildverarbeitung in Zuordnung der Wissensrepräsentation

2.4.1 Wissensrepräsentation

Die Frage nach der Darstellung von Bildverarbeitungsergebnissen wurde in der Literatur ausgiebig diskutiert. Sie wird traditionell in drei Wissensrepräsentationen unterteilt [BB82]:

- Die *Low-Level-Repräsentation* (s. Kap. 2.4.2) bezeichnet die Haltung der Daten der Bildgebung und bildet nach deren Import den Ausgangspunkt jeder Bildverarbeitungsanwendung. Der Ausdruck Bild ist ein Synonym für die Low-Level-Repräsentation. Allgemein werden die Daten der Low Level-Repräsentation als eine Folge von diskreten Werten gehalten und erfahren durch ein unterliegendes Koordinatensystem (*KOS*) einen räumlichen Bezug zueinander, sie liegen im sogenannten *Ortsraum*. Üblicherweise werden Bilddaten unter der Verwendung eines kartesischen KOS dargestellt. Bildpunkte im 2D-KOS werden als *Pixel*, im 3D als *Voxel* bezeichnet.
- Die *Intermediate-Level-Repräsentation* (s. Kap. 2.4.3) beschreibt die relevanten Inhalte des Low-Level durch einfache Symbole, die folgend als *Tokens* bezeichnet werden. Das Format dieser Objekte kann sehr unterschiedlich sein, beispielsweise eine Menge von Bildpunkten, einfache Primitive wie Linien, Oberflächenmodelle oder andere geometrische Beschreibungen. Die Intermediate-Level-Repräsentation ist gegenüber den Bilddaten eindeutig interpretierbar, denn ihr Inhalt wird auf abstrakterer Ebene gehalten und ist auf die wesentlichen Informationen reduziert. Er ist somit Ressourcen schonender gespeichert, der Zugriff ist folglich leichter und schneller und die Inhalte verschiedener Bilddaten werden durch die Repräsentation vergleichbar.
- Während sich die beiden vorherigen Repräsentationen mit der Darstellung von unabhängigen Informationsblöcken (Bildpunkten oder Tokens) beschäftigen, beinhaltet die *High-Level-Repräsentation* (s. Kap. 2.4.4) deren Semantik. Durch das Benennen der einzelnen Tokens oder der Herstellung von Beziehungen zwischen ihnen wird der Inhalt der Low-Level-Repräsentation nun auf einer abstrakten Ebene beschrieben. Diese Beschreibungen stellen gleichzeitig auch das Ergebnis der Bildverarbeitungsanwendung für den Export dar.

Die Wissensrepräsentation beschreibt die Bildinhalte beim Übergang vom Low- zum High-Level auf eine immer abstraktere Ebene. Der Gehalt an Wissen steigt mit jedem Level genauso wie die Anzahl an benötigten Verfahren und Algorithmen, um dieses Wissen zu extrahieren.

Das Erkennen und Verstehen von Bildinhalten ist eine äußerst anspruchsvolle Aufgabe, eine direkte Beziehung zwischen Bildpunkten und der Semantik des Bildes ist für komplexere Fragestellungen nicht möglich. Die Informatik bedient sich der Intermediate Repräsentation als Bindeglied zwischen der Ein- und Ausgabe. Die verwendete Bildverarbeitung zur Lösung einer anspruchsvollen Problemstellung involviert meist Verfahren aller Wissensrepräsentationen.

In Abb. 2.5 ist die Bildverarbeitung in Beziehung zu der Wissensrepräsentation gesetzt. Die Darstellung und folgenden Beschreibungen haben keinen Anspruch auf Vollständigkeit, vermitteln aber die Konzepte und Grundlagen der digitalen Bildverarbeitung.

2.4.2 Low-Level

Unter *Vorverarbeitung* (s. Abb. 2.5) versteht man die qualitative Aufbereitung von Bilddaten oder das Verändern ihrer Darstellungsform. Allen Vorverarbeitungsalgorithmen ist gemein, dass sowohl die Eingabe als auch die Ausgabe der Operationen ausschließlich der Low-Level-Repräsentation zugehörig sind. Die Vorverarbeitung stellt somit eine Transformation dar, die am Wirkungsbereich unterschieden werden kann:

- *Selbsttransformationen* verändern ausschließlich die Werte der Low-Level-Daten, unterschieden wird die Anzahl an Bildpunkten, die der eingesetzte Operator umfasst. *Punktoperatoren* transformieren die Werte der Bildpunkte einzeln und verändern somit das Histogramm, also die statistische Häufigkeit der einzelnen Bildwerte. Dagegen beziehen lokale Operatoren die Bildwerte der lokalen Umgebung in die Berechnung des Transformationswerts ein. Synonyme für die lokale Umgebung sind die Begriffe *Maske* oder *Fenster*. Lokale Operatoren werden häufig zur Bildfilterung (Glättung bzw. Weichzeichnung beispielsweise durch Mittelwert-, Median- oder Gaußfilter) oder Kantenhervorhebung (beispielsweise nicht lineare Filter wie Roberts-Kreuz oder Gradientfilter wie der Canny Operator) eingesetzt.
- *Affine Transformationen* verändern den Bezug der Daten zum Koordinatensystem beispielsweise durch Anwendung von Rotation, Translation, Skalierung und anderen (s. Kap. A.3). Die neue Position der Bildpunkte nach Anwendung der Transformation verlässt häufig das diskrete Raster, so dass zusätzlich eine Interpolation zum Ermitteln der Bildwerte im Kontinuierlichen erfolgen muss.

Wie die Selbsttransformation verlässt auch die affine Transformation nicht den Raum: das Eingabebild wird zwar verändert, ist aber immer noch deutlich als solches zu erkennen.

- Durch *globale Transformationen* wie Wavelet, Hough oder Radon wird die Bildinformation in einen anderen Raum abgebildet. Die Fourier-Transformation beispielsweise überführt den Ortsraum der Bilddaten in einen *Frequenzraum*. Dort bietet sich die Möglichkeit der effizienten Anwendung von linearen Filtern auf die Frequenz- und Phasenanteile. Der Frequenzraum wird wiederum durch eine Folge diskreter Werte und einem Koordinatensystem durch die Low-Level-Repräsentation dargestellt. Das Eingabebild lässt sich aber durch die Darstellung von Frequenz- und Phasenanteile nicht mehr wiedererkennen [Jäh05, NH04].

2.4.3 Intermediate-Level

Durch die *Segmentierung* werden die Bilddaten des Low-Levels in das Intermediate-Level überführt. Ziel einer Segmentierung ist die Unterteilung der Bilddaten in disjunkte Bereiche ähnlicher Eigenschaften. Zwei prinzipielle Vorgehensweisen lassen sich dabei unterscheiden (s. Abb. 2.5):

Das Bottom-Up Verfahren geht von den Bildpunkten aus und unterteilt das Bild anhand eines Homogenitätskriteriums. Alle Bildpunkte, die das Kriterium erfüllen, werden demselben Token zugeordnet. Das Homogenitätskriterium kann wie bei der Bildvorverarbeitung durch Punktoperatoren bestimmt (z.B. Schwellwertsegmentierung) oder durch lokale Operatoren berechnet werden (z.B. Texturerkennung).

Komplexere Segmentierungsmethoden prüfen neben dem Erfüllen des Kriteriums auch die räumlichen Zusammenhänge – gegeben durch das Koordinatensystem – der Bildpunkte. Ein bekannter Vertreter solcher Regionen basierter Segmentierungen ist das *Region Growing* (s. Kap. 4.3.1).

Kanten- und konturbasierte Segmentierung ermitteln die Token nicht aus der Ähnlichkeit einzelner Bildpunkte, sondern aus Unterteilungen des Bildes durch Erkennen von Grenzbereichen. Alle Bildpunkte eines in sich abgegrenzten Bereichs können einem Token zugewiesen werden.

Vorteil der Bottom-Up Verfahren sind ihre meist einfache Implementierung. Ihre Laufzeit wächst aber mit der Anzahl der zu betrachteten Bildpunkte. Das *Top-Down* Verfahren dagegen arbeitet mit einem A-priori-Wissen über das, was in den Bilddaten erkannt werden soll. Praktische Anwendung haben Formmodelle, die beispielsweise anatomisches Wissen in einem mathematischen Modell mit Freiheitsgraden kodieren. Das Modell kann sich an einer vorgegebenen Position dem Bildinhalt anpassen, das Token wird direkt aus dem endgültig geformten Modell erzeugt.

Sind die relevanten Bildinhalte erst einmal als abstraktes Token erfasst, kann dieses durch morphologische Operationen verändert werden (s. Abb. 2.5). Während die bisher vorgestellten Operatoren auf den Bildwerten arbeiten, betrachten die morphologischen Operationen nur das Token selbst. Bekannte morphologische Transformationen sind *Erosion* und *Dilatation* und die daraus definierten Opening- und Closing-Operatoren. Vereinfacht führt die Erosion zu einer Verkleinerung des Tokens, während es die Dilatation vergrößert [Jäh05, Soi98].

Durch die Prüfung der Zugehörigkeit einzelner Bildpunkte zu einem einzelnen oder mehreren Token kann die Bildinformation binarisiert werden.

2.4.4 High-Level

Die High-Level-Klassifikation unterteilt die Token anhand von Merkmalsvektoren in verschiedene Klassen. Ein Vektor wird für jedes Token durch Extraktion verschiedener Merkmale (*Merkmalsextraktion*, beispielsweise Größe, Form, Texturierung, räumlichen Informationen, ...) kodiert, seine Dimension entspricht der Anzahl an Merkmalen. Jedes Merkmal kann durch Analyse der Informationen der beiden vorherigen Level bestimmt werden. Die Klassifikation der Token durch den Merkmalsvektor kann – wie bei der Segmentierung – anhand zweier Vorgehensrichtungen ermittelt werden (s. Abb. 2.5):

Beim *Bottom-Up* Vorgehen werden unüberwachte (untrainierte) Klassifikationsverfahren eingesetzt (z.B. Clustering). Sie unterteilen den durch die Vektoren gebildeten Merkmalsraum in Klassen, Vektoren mit ähnlichen Eigenschaften werden denselben Klassen

zugewiesen. Daraus erschließt sich jedoch noch nicht zwingend die Semantik der einzelnen Klassen.

Dagegen können überwachte Klassifikationsverfahren, die mit Merkmalsvektoren von Referenzobjekten mit bekannten Zielklassen trainiert wurden (z.B. Neuronale Netze oder Support Vektor Maschine, s. Kap 4.5.3 u. 4.6.5), die für unbekannte Merkmalsvektoren vorhergesagten Klassen semantisch interpretieren (*Top-Down*). Das Verfahren der überwachten Klassifikation gibt durch das antrainierte A-priori-Wissen neben der Unterteilung in Klassen auch gleich deren Bedeutung wieder.

Überwachte und unüberwachte Klassifikatoren lassen sich auch zum Segmentieren verwenden. Dann beinhaltet der Merkmalsvektor nur Bildpunktinformationen des Low-Levels. Dementsprechend liefert das Ergebnis auch nur eine Aussage über die Bildpunkte (*Low-Level-Klassifikation*).

Beim Erschließen der Semantik (s. Abb. 2.5) versucht die Informatik den universellen und leistungsstarken Sehapparat des Menschen nachzuahmen. Hier müssen die Token unter Berücksichtigung der erarbeiteten Information noch zueinander in Beziehung gesetzt werden. Für gewöhnlich können räumliche oder zeitliche Verbindungen zwischen den Token festgelegt werden (z.B. durch Semantische Netze). Durch Auswertung dieser Beziehungen lässt sich schrittweise der Grad an Semantik immer weiter erhöhen. Da häufig A-priori-Wissen über die Bildinhalte vorliegt, kann die Semantik auch durch formulierte Gesetzmäßigkeiten (z.B. Fuzzy Logic) erschlossen werden.

Aufbauend auf der abstrakten Repräsentation des Intermediate-Levels eröffnet sich erst im High-Level die Möglichkeit, die dargestellte Szene des Low-Level zu verstehen.

Kapitel 3

Stand der Forschung

Die komplexe Aufgabenstellungen der Diagnostik von Brusttumoren in der fMRT erfordern die algorithmischen Umsetzungen aller drei Wissensrepräsentationen (s. Kap. 2.4.1). Wesentliche Komponenten eines CAD-Systems (s. nächstes Kapitel) für die Brust-MRT sind die Kompensation von Bewegungen (*Registrierung*) als Low-Level-Funktionalität, das Erkennen von einzelnen Läsionen (*Segmentierung*) als Übergang zum Intermediate-Level und deren morphologische und kinetische Beschreibung (*Merkmalsextraktion*). Abschließend erfolgt eine Dignitätsbeurteilung (*Klassifikation*) der Objekte als High-Level-Funktionalität [DMK⁺99].

In den letzten Jahren haben sich einige CAD-Systeme für die Befundung von Brustkrebs auf fMRM Bilddaten etabliert. Das nächste Kapitel unterteilt ihren Funktionsumfang in drei Kategorien. Für jede Kategorie werden im Kapitel 3.2 kommerzielle Produkte vorgestellt und deren Funktionsweise näher beleuchtet.

Die Hersteller der CAD-Systeme geben nur wenig Informationen zur genauen Funktionsweise ihrer Produkte preis. Auch vereinen sie nicht die Funktionalität aller Komponenten. Daher werden ergänzend zu den existierenden Systemen wissenschaftliche Publikationen vorgestellt, die zwar jede für sich nur einzelne Komponenten eines CAD-System behandeln, aber zusammengenommen einen Überblick heutiger Entwicklungsmöglichkeiten für ein vollständiges System geben (s. Kap. 3.3).

Abschließend wird der Fortschritt der vorliegenden Arbeit gegenüber den existierenden CAD-Systemen hervorgehoben (s. Kap. 3.4).

3.1 Kategorien von CAD-Systemen

Viele Softwareprodukte für die fMRM bezeichnen sich als CAD-Systeme. In der Medizinischen Informatik wird die Abkürzung *CAD* entweder für *Computer Aided Detection* oder *Computer Aided Diagnosis* verwendet [Vli01]. Während in beiden Fällen das *CA* für die gleiche Aussage steht, verbirgt sich hinter dem *D* ein gewaltiger funktionaler Unterschied, den man mit Hilfe der Wissensrepräsentation (s. Kap. 2.4.1) am Beispiel der Brust-MRT Befundung (s. Kap. 2.4.1) verdeutlichen kann:

- *CADetection* hebt beispielsweise verdächtige Anreicherungen im Bild als Farbkodierung hervor oder vereinfacht die Auswertung durch eine graphische Darstellung der Kinetik. Je nach Komplexität gehören solche Funktionen fast ausschließlich der Low-Level-Ebene an; sie treffen daher nur Aussagen für einzelne Bildpunkte. Sie verfügen über kein Wissen, welcher Bereich des Bildes eine einzelne Läsion darstellt bzw. wo deren Begrenzung zum gesunden Gewebe liegt. Solche Systeme können keine eigene Einschätzung bezüglich der Dignitäten von Auffälligkeiten geben; diese Entscheidung ist alleine dem Benutzer überlassen.
- Dagegen gibt ein *CADiagnosis*-System eine eigene Einschätzung über Bös- oder Gutartigkeit. Aber gerade bei Diagnosesystemen gibt es eklatante Unterschiede, auf welcher Ebene der Wissensrepräsentation diese Aussage erfolgt:

Viele Systeme klassifizieren nur einzelne Bildpunkte anhand eines Merkmalsvektors, der die kinetische Informationen aus den fortlaufenden Serien enthält (Low-Level-Klassifikation, s. Kap 2.4.4). Das Ergebnis wird wiederum meist als Farbkodierung pro Bildpunkt eingeblendet, wobei verschiedene Farben den Grad der Bösartigkeit der einzelnen Bildpunkte hervorheben. Dies ist gegenüber einem *CADetection* System nur eine geringe Leistungssteigerung, denn es wird auch hier kein Wissen über die Existenz einzelner Läsionen bzw. deren Abgrenzung im Bild erzeugt. Die verwendete Klassifikation stellt lediglich eine Segmentierung dar, die die Bildpunkte in Klassen mit ähnlicher Eigenschaft wie beispielsweise *Steigung der initialen Phase* einteilt. Das System unterstützt den Benutzer auch hier ohne Funktionalitäten der High-Level Ebene.

Funktional stark weiterentwickelte Systeme segmentieren zuerst verdächtige Läsionen (Intermediate-Level-Funktionalität). Sie erzeugen dadurch Informationen über die Anzahl an Läsionen im Bild und grenzen umgebenes Gewebe ab. Auf den Token können wesentlich informationsreichere Merkmalsvektoren berechnet werden, die sowohl kinetische als auch morphologische Informationen umfassen. Es folgt somit eine Klassifikation als High-Level-Funktionalität. Solche Systeme sind in der Lage, Dignitätsaussagen auf Läsionen statt auf Bildpunkten zu treffen.

Nur anhand des Etiketts *CAD* kann also nicht erkannt werden, ob die Software auch wirklich endgültige Aussagen treffen kann oder lediglich das Bildmaterial für den Betrachter aufbereitet. Folgend werden CAD-Systeme funktional in drei Kategorien unterschieden, das nächste Kapitel ordnet ihnen bekannte Produkte zu:

Low-Level-Visualisierung visuelle Aufbereitung hilfreicher Informationen für den Benutzer, keine diagnostische Aussage

Low-Level-Diagnose Klassifikation auf Grundlage der Kinetik, diagnostische Aussage für Bildpunkte

High-Level-Diagnose Klassifikation als High-Level-Funktionalität durch Bewertung der Kinetik und Morphologie, Dignitätsaussagen für einzelne Läsionen

3.2 Kommerzielle CAD-Systeme

Die Hersteller von CAD-Systemen geben nur wenig Anhaltspunkte zu der Funktionsweise ihrer Systeme. Auskünfte über deren Produkte wurden durch die Firmenwebseiten, Informationsbroschüren, Einsicht in die Patentanmeldungen und durch persönliche Gespräche und Besuche von Messen in Erfahrung gebracht. Wissenschaftliche Publikationen (s. Kap. 3.3) lagen nur selten vor. Die folgenden Beschreibungen (s. Kap. 3.2.1 bis 3.2.3) kommerzieller Software beschränken sich nur auf die Funktionen, die einen direkten Zusammenhang mit dem Prozess der Diagnoseerstellung durch den Benutzer haben.

3.2.1 Low-Level-Visualisierungssysteme

*DynaCAD*TM wurde von *Mevis* [Mev] in Bremen entwickelt und ist ein reines Visualisierungstool. Der Benutzer erhält die Möglichkeit, eine Registrierung der Bilddaten durchzuführen (2D oder 3D) und kann diese durch Parameter qualitativ beeinflussen. Nach Anwahl einer Mappingfunktion werden die Signalintensitäten als zusätzliche Information auf Bildpunktebene durch eine Farbkodierung eingeblendet. Neben den Standardansichten einschließlich *MPR* lässt sich die Farbkodierung auch zusammen mit einer 3D *MIP* darstellen. Der Benutzer kann sich für jeden Bereich des Bildes die Kinetik als Zeit/Signalintensitätskurve anzeigen lassen.

3.2.2 Low-Level-Diagnosesysteme

Die Firma *CADsciences* [CAD] setzt als Mindestanforderung für ihr Produkt *Breast fTP* ein 1,0 T Gerät und DICOM-3-Kompatibilität (s. Kap. A) voraus. Herzstück des Systems ist das als *fTP* (Full-Time-Point) patentierte Verfahren zur Klassifikation der Kinetik.

Full-Time-Point basiert auf einem pharmakokinetischen Modell (s. Kap. 3.3.3) von Tofts [TBS95], welches für die Verwendung von KM mittels schneller Bolus Injektion anwendbar gemacht wurde. Aus den Gleichungen der einzelnen Kompartimente können drei Zustände ermittelt werden: Ein maligner Zustand wird rot dargestellt, blau wird für benignes Verhalten verwendet und grün für Mehrdeutigkeit. Diese Heuristik wurde durch die Auswertung von Bildern mit bekannten Histologien erarbeitet.

Einfacher geht das als *3TP* (Three-Time-Point) bezeichnete Verfahren vor, das *CADsciences* zu Beginn der Produktentwicklung angeboten hat [DGW97, FHD02, KFHD02]. Es benötigt zur Berechnung der gleichen Farbkodierung lediglich die Signalintensitäten von drei Zeitpunkten.

Das Produkt *CADstream*TM von *Confirma*TM [Con] mit Hauptsitz in den USA und Zweigstelle in Berlin hat seinen Schwerpunkt in der Optimierung des Befundungsablaufs. Es kann beliebig orientierte 3D-Sequenzen variabler Serienanzahl verarbeiten. Die graphische Benutzerschnittstelle (*GUI*) erlaubt alle Standardansichten einschließlich *MIP* und *MPR*. Um Bewegungsartefakte auf den Subtraktionen zu mindern, führt die Software zuerst eine automatische Registrierung der Bilder durch. Auf den registrierten Bilddaten wird

durch eine Farbkodierung angezeigt, welcher Bereich der Brust verdächtig ist. Grundlage hierfür bildet die Kinetik, die auf den Bildpunkten berechnet wird und je nach Verlauf der postinitialen Phase (s. Kap. 2.3.5) den Grad der Bösartigkeit (ansteigend, gleichbleibend, auswaschend) durch einstellbare Schwellwerte in einen von drei Farbwerten kodiert.

Der Benutzer kann Teilbereiche der Farbmarkierungen manuell selektieren und definiert dadurch eine einzelne Läsion. Daraufhin generiert das System automatisch einen Report, der mit Bild-, Größen-, Kinetik- und Ortsangaben der Selektion übersichtlich gestaltet ist. Der Anwender kann durch das BI-RADSTM-Lexikon (s. Kap. 2.3.6) unterstützt seine Diagnose in den Report eintragen.

Für die Ortsangaben im Report führt CADstreamTM eine weitere Segmentierung aus, um die Lage der Brust im Bild zu erkennen. Auf deren Ergebnisse kann durch Abstandsberechnungen auch die Position des Herzens geschätzt werden. Da es ebenfalls Anreicherungen zeigt, wird im geschätzten Bildbereich die Farbkodierung ausgeblendet, andernfalls würde das Herz vor allem auf Bilddaten mit transversaler Schichtführung zu zusätzlichen unerwünschten Einfärbungen führen.

3.2.3 High-Level-Diagnosesysteme

Ein Team um den Mathematiker Alan Penn forscht in den USA seit über einem Jahrzehnt an der Entwicklung eines CAD-Systems [PKT⁺01]. Das mittlerweile als *ONCAD*TM [Pen] benannte Produkt kann Bilder von primär fettsupprimierten Sequenzen verarbeiten (1,5 T - GE Medical Systems), die mit einer Auflösung von 512x512x32 Voxeln die Brust durch hohe Schichtdicken erfassen. Über eine GUI muss der Benutzer jede verdächtige Läsion, für die er eine Diagnose benötigt, manuell grob vorsegmentieren. Für diese wird anschließend für eine morphologische Aussage automatisch eine patentierte fraktale Dimension [PL97] berechnet und durch eine Heuristik in eine dreiwertige Aussage – schwaches, mittleres oder hohes Anzeichen von Bösartigkeit – überführt. Da dieses einzelne Merkmal allein nicht aussagekräftig ist, um eine abschließende Diagnose zu treffen, muss der Benutzer für jede Läsion – unterstützt durch vorgegebene Bewertungskategorien – weitere morphologische Merkmale selbst beurteilen. Gleiches gilt auch für die Kinetik, die als Zeit/Signalintensitätskurve dargestellt, aber durch den Benutzer selbst einzuschätzen ist. Die vielen manuellen und wenig automatisch ermittelten morphologischen Merkmale entsprechen in etwa dem Umfang der Kriterien des Göttinger Scores (vgl. Kap. 2.3.5). Das System summiert die Bewertungspunkte zu einem normierten AIF-Wert zwischen 0,0 und 1,0 auf. Durch den Vergleich des AIF-Werts mit Werten bekannter Histologien kann die Software eine Aussage darüber treffen, ob die markierten Läsionen maligne oder benigne sind.

3.3 Wissenschaftliche Publikationen

Das vorherige Kapitel hat bereits gezeigt, dass heutige CAD-Systeme nicht über eine Bildverarbeitung aller Ebenen der Wissenrepräsentation (s. Kap. 2.4.1) für die Brust MRT-Untersuchung verfügen. Fehlende Funktionalitäten verlangen zwangsweise nach Eingaben des Anwenders. Auch fehlt zu der vorhandenen Funktionalität meist der genaue wissenschaftliche Hintergrund.

Um dennoch einen kompletten Überblick über die funktionalen Bausteine eines CAD-Systems geben zu können, wird folgend auf verschiedene wissenschaftliche Publikationen zurückgegriffen, die meist jedoch nur eine Teilkomponente wie Registrierung (s. Kap. 3.3.1), Segmentierung (s. Kap. 3.3.2), Merkmalsextraktion (s. Kap. 3.3.3) oder Klassifikation (s. Kap. 3.3.4) behandeln.

3.3.1 Registrierung

In der Informatik werden Algorithmen, die Bilder mit veränderten Inhalten wieder in räumliche Übereinstimmung bringen, als *Registrierung* bezeichnet. Registrierungsprobleme entstehen immer dann, wenn Bilder durch unterschiedliche Geräte (Modalitäten) oder zu unterschiedlichen Zeitpunkten akquiriert wurden. Die Entwicklung von Registrierungsalgorithmen begann bereits in den achtziger Jahren für eine Vielzahl von Modalitäten. Sie stellen eine Selbsttransformation im Low-Level dar (s. Kap. 2.4.2). Eine Übersicht allgemeiner [MV98] und speziell für die Brust-MRT entwickelter Registrierungen [Siv05] ist in der Literatur vorhanden. Die Registrierungsalgorithmen können aufgrund ihrer vielfältigen Methodik und Zielsetzung durch bis zu neun Grundkriterien unterschieden werden [EMPV95]. Auf einige soll folgend näher eingegangen werden.

Der *Grundcharakter* eines Registrierungsalgorithmus ist dadurch bestimmt, dass die Bildveränderungen zwischen Ausgangs- und Zielbild durch künstliche Vorgaben markiert sind oder selbstständig nur aus den Bildinformationen erarbeitet werden. Die Selbstständigen Verfahren können in Bezug auf die Erkennung der Korrespondenzen zwischen Ausgangs- und Zielbild in drei Gruppen unterteilt werden:

- Erkennung von Korrespondenzen durch sogenannte *Landmarken* wie beispielsweise eindeutige anatomische Strukturen oder manuelle Vorgaben.
- Durch vorgeschaltete *Segmentierungen* (s. Kap. 3.3.2) werden anatomisch gleiche Strukturen erkannt und dienen als Intermediate Token (z.B. triangulierte Oberflächenmodelle) als Eingabe für den Angleichungsprozess.
- Vergleich der Eigenschaften von *Bildpunktemengen* zur Erkennung von Korrespondenzen.

Verfahren basierend auf Landmarken oder Segmentierungen leiten aus den Korrespondenzen eine Mappingfunktion zur Formulierung der erkannten Bewegung ab.

Bei den auf Bildpunkten basierten Verfahren kann mit Gesamt- oder Teilbild der Umfang der Punktmengen unterschieden werden. Sie müssen die Ähnlichkeit des untersuchten Bereichs zwischen Ausgangs- und Zielbild auf den Punktmengen bewerten. Dazu werden Verfahren wie Entropien, Korrelationen, Minimierung von Abstandmaßen (s. Kap. 4.6.1), optischer Fluss (s. Anhang B) u.a. eingesetzt.

Anhand des Grundcharakters der Registrierung entscheidet sich die Art der Bildtransformation, also die Neuordnung der Grauwerte als Selbsttransformation aufgrund der erkannten Bewegung. *Rigide Transformationen* führen nur Translationen oder Rotationen aus (s. Anhang A.3). Das Ausgangsbild wird also nur gedreht oder verschoben, um wieder eine Übereinstimmung mit dem Zielbild herzustellen. Affine Transformation sind etwas flexibler, erhalten aber noch die Abstandsverhältnisse paralleler Strecken. Rigide und affine Transformationen sind meist das Ergebnis von Registrierungen mit künstlichen Markern oder auf der Basis erkannter Landmarken. Bildpunkt-basierte und teilweise auch auf Segmentierung basierende Verfahren führen dagegen häufig zu *nicht rigiden* bzw. *elastischen Transformationen*, bei denen die Abstände lokaler Regionen voneinander abweichen können.

Bei der Anwendung der Registrierung können Ziel- und Ausgangsbild gleichen (*mono modal*) oder verschiedenen (*multi modal*) Modalitäten entsprechen. Letzteres verlangt nach anspruchsvolleren Methoden zum Feststellen der Bildkorrespondenzen. Auch kann die Dimension der Bilder variieren, beispielsweise können 2D-Mammographiebilder auf 3D-MRT-Messungen desselben Patienten registriert werden [RSM⁺06]. Dies erfordert zusätzliche affine Transformationen der Bildkoordinaten, um die Bildausschnitte entsprechend zu positionieren.

Die ersten Verfahren zur Registrierung von fMRT Bildern der Brust verwendeten rigide Registrierungsalgorithmen [ZJB⁺96, KCHL99]. Affine Registrierungen wurden von Kumar und Lucht et al. vorgestellt [KAH⁺96, LKB00], neuere Ansätze verwenden häufig elastische Methoden [MV98].

3.3.2 Segmentierung

Während die Registrierung als notwendiger Vorverarbeitungsschritt auch für andere Bildmodalitäten bereits ausgiebig in der Literatur behandelt wurde, befassen sich mit der Segmentierung (s. Kap. 2.4.3) von Läsionen in der Brust nur wenige Publikationen. Ziel der Segmentierung ist es, die relevanten, Läsionen darstellenden Bildbereiche von den übrigen Bildinformationen abzugrenzen und ins Intermediate-Level zu übertragen. Wie bereits bei der Kategorisierung von CAD-Systemen (s. Kap. 3.1) verdeutlicht, muss die Art der Abgrenzung in Segmentierung aller Anreicherungen durch Low-Level-Klassifikation und Segmentierung der Läsionen als einzelne Token unterschieden werden.

Die Mehrheit der Publikationen befasst sich mit *Segmentierungen durch Low-Level-Klassifikationen* (s. Kap. 2.4.4) auf Grundlage der kinetischen Informationen. Meistens unüberwachte Klassifikatoren bewerten die Kinetik aus den Bildfolgen jeweils auf Bildpunkt-basis in eine dreiwertige Aussage. Mit überwachten Klassifikatoren konnten trainierte Muster maligner und benigner Läsionen durch Auswerten lokaler Bildbereiche vorherge-

sagt werden [KFB⁺98, SWBA, CGB⁺05, FPF⁺06]. Gomez et al. [GGVMB⁺04] stellten für die Dignitätsvorhersage auf Bildpunktbasis Neuronale Netze, Support Vektor Maschinen und K-Means Clustering (s. Kap. 4.5.3 u. 4.6.5) gegenüber mit dem erfolgreicherem Abschneiden der Neuronalen Netze. Alternativ zu einer Klassifikation kann der Verlauf der Kinetik durch Modelle parametrisiert werden (s. nächstes Kapitel). All diese Verfahren arbeiten zwar automatisch, erkennen aber nur die Anreicherungen als solche im Bild und bewerten dabei die Bösartigkeit einzelner Bildpunkte. Sie vollziehen aber nicht die Abgrenzung einzelner Läsionen innerhalb der detektierten Anreicherungen. Somit sind derartige Segmentierungen ungeeignet für eine anschließende Morphologiebestimmung.

Bekanntere Verfahren zur Segmentierung der *Läsionen als einzelne Objekte* sind halbautomatische Verfahren, welche nach einer (grob) manuellen Selektion den Rand des Objektes erkennen [CGB06, ZBE⁺07, WSK⁺06]. Durch die Selektionsvorgabe eröffnen sich Möglichkeiten zum Einsatz von Top-Down Verfahren.

In einem einzelnen Ansatz wird die Symmetrie der rechten und linken Brust ausgenutzt, um durch Abstandsmaße die Abnormalitäten im Vergleich zu erkennen [AP03].

Publikationen über automatische Segmentierungen, die einzelne Läsionen erkennen, konnten nicht gefunden werden.

3.3.3 Merkmalsextraktion

Sofern eine manuelle oder semiautomatische Segmentierung vorhandener Läsionen in einzelne Objekte erfolgte (s. vorheriges Kapitel), ist mit der Merkmalsextraktion die Beschreibung der Eigenschaften von Läsionen der nächste Schritt. Die berechneten Merkmale, bestehend aus morphologischen und kinetischen Informationen, dienen in der abschließenden Klassifikation (s. Kap. 3.3.4) zur Vorhersage der Dignität.

Kaiser [Kai07] stellt nicht weniger als 147 kinetische und morphologische Merkmale zur Beschreibung von Läsionen zusammen. Der Autor gibt einen Überblick über alle ihm bekannten aktuellen Analysemöglichkeiten, um eine Grundlage zur Verbesserung der fMRM nachgesagten niedrigen Spezifität zu schaffen. Während in der Literatur nur wenige morphologische Merkmale algorithmisch umgesetzt sind, ist die Auswertung der Kinetik schon länger Bestandteil verschiedener kommerzieller Systeme (s. Kap. 3.2).

Die Parametrisierung der Kinetik in der Literatur kann in zwei Kategorien unterteilt werden. Bei den *Schwellwertmethoden* wird nach Gabe des KMs die absolute oder die relative Intensitätsverstärkung zur nativen Aufnahme berechnet und durch Anwendung von Schwellwerten bewertet [HWP⁺89, Kai93]. Als Erweiterung wird der Schwellwert auf einen bestimmten Zeitpunkt oder ein bestimmtes Zeitintervall nach KM-Applikation angewendet [Fis00, HHS⁺86, Kai93]. Ein Beispiel für eine solche *zeitabhängige Schwellwertinterpretation* wurde bereits in Kapitel 2.3.5 durch Formel 2.1 gegeben.

Wesentlich komplexer sind *pharmakokinetische Modelle*, die den Signalverlauf über die Zeit durch verschiedene Kompartimente abstrahieren. Dadurch können KM beeinflussende Faktoren wie beispielsweise die Permeabilität der Blutgefäße (s. Kap. 2.3.2) parametrisiert werden. Für die pharmakokinetische Analyse T_1 -gewichteter MRT-Bilddaten hat das Modell von Brix [BSP⁺91] weite Verbreitung gefunden, unter anderem auch auf die Anwen-

derung von Mammatumoren. Durch ein 2-Kompartiment modellierten Brix et al. [BMK⁺01] Plasma und Interstitium und charakterisierten die KM-Dynamik durch drei Parameter unter der Annahme, dass das KM mit einer konstanten Rate in das zentrale Plasmakompartiment einfließt. Parametrisiert wird die Kontrastveränderung durch die Amplitude (A) in Abhängigkeit zum Plasmavolumen und die Austausch- (k_{ep}) und Eliminationsrate (k_{el}) des KMs zwischen Plasma und Interstitium. Die beiden Raten charakterisieren letztendlich die initiale und postinitiale Phase der Kinetik und lassen Rückschlüsse auf die Gefäßpermeabilität zu.

Morphologien werden in der Literatur überwiegend durch 3D-Merkmale berechnet, die Größen-, Rand- oder Textureigenschaften der Anreicherung darstellen. Shahar et al. [SSB02] leiteten geometrische Merkmale aus den gerenderten 3D-Oberflächengittern zur Beschreibung des Oberflächen/Volumen Verhältnisses und Indices für die Kugelform ab. Neben kinetischen Merkmalen bestimmten Gilhuijs et al. [GGB98, GDM⁺02] verschiedene Variationen der Randschärfe. Tzacheva et al. [TNB03] berechneten sowohl intensitätsbasierte Merkmale wie Textur und Randeigenschaften als auch auf binarisierten Bildinformationen klassische Formmerkmale wie Exzentrizität und Größe. Weitere Formparameter wie Komplexität, Konvexität, Rundheit und Dehnung wurden von Liney et al. [LSG⁺06] erhoben.

Aufwendigere Einzelmerkmale wurden durch Penn et al. veröffentlicht, die mit der fraktalen Dimension textuelle Aussagen über eine Läsion in 2D machen [PL97] oder in Kombination mit der Kinetik das als Blooming bezeichnete Randschärfeverhalten über die Zeit beschreiben [PTB⁺06].

3.3.4 Klassifikation

Verschiedene Publikationen untersuchen die Möglichkeiten der Diskriminanz maligner und benigner Funde durch überwachte Klassifikatoren (vgl. Kap. 2.4.4). Betrachtet werden folgend nur Publikationen, die Merkmalsvektoren bestehend aus morphologischen oder kinetischen Inhalten verwenden, jedoch keine zur Segmentierung eingesetzten Low-Level-Klassifikationen (vgl. Kap. 3.3.2).

Die Merkmalsvektoren setzten sich entweder aus automatisch berechneten Werten (s. vorheriger Abschnitt) oder aus deskriptiven Bewertungen durch einen Radiologen zusammen. Die Dimension des Vektors variiert dabei von 4 bis 40 Merkmale. Als Goldstandard (s. Kap. 4.5.1) wurden überwiegend histologisch gesicherte Tumore verwendet.

Häufig wurden – wie in der Medizin häufig angewendet – neuronale Netze zur Quantifizierung eingesetzt. Abdolmaleki et al. [ABN01] untersuchten 120 Patientinnen mit 14 überwiegend kinetischen Merkmalen und erreichten eine Treffgenauigkeit (s. Kap. 4.6.4) von 79 %. Auf lediglich 14 Läsionen (10 maligne und 4 benigne) erreichten Tzacheva et al. [TNB03] anhand von 4 Merkmalen eine Treffgenauigkeit von 91,2 %.

Ein Vergleich zwischen Expertenradiologen, statistischem Modell und Neuronalem Netz mit 40 manuell erhobene Merkmalen zeigte bei Vomweg et al. [VBK⁺03] auf 604 Läsionen angewendet eine Dominanz der Netze (Sensitivität/Spezifität: 94 %/92 %, s. a. Kap. 4.6.4) gegenüber dem Radiologen (92 %/86 %) und dem statistischen Modell (91 %/69 %). Sza-

bo et al. [SWBA] quantifizierten 75 maligne und 30 benigne Läsionen und verglichen die Vorhersage auf verschiedenen Kombinationen kinetischer und morphologischer Merkmale mit einem Expertenradiologen. Letzterer erreichte die höchste Treffgenauigkeit (80 %), gefolgt von dem Neuronalen Netz (77 %), das mit Randeigenschaft, initialer und postinitialer Phase über die drei signifikantesten Merkmale trainiert wurde. Das Betrachten von rein kinetischen Merkmalen (74 %) brachte eine bessere Treffgenauigkeit als von rein morphologische Merkmalen (68 %).

Nattkemper et al. [NAL⁺05] führten einen Vergleich verschiedener Klassifikatoren durch. Neben K-Means Clustering und Self-Organizing Maps (eine Art von neuronalen Netzen) als unüberwachte Vertreter setzten sie auch überwachte Verfahren wie k-nearest Neighbor und mit Support Vektor Maschinen (SVM) oder Entscheidungsbäumen auch neuere Methoden zur Tumorklassifikation ein und erkannten dabei ein Potential der SVM (vgl. Kap. 6.4.4).

3.4 Fortschritt durch die vorliegende Arbeit

In der vorliegenden Arbeit werden die Ergebnisse eines Forschungsprojekts beschrieben, das die Entwicklung einer Prototypsoftware zur Erkennung von Brustkrebs auf funktionalen MRT-Bilddaten der weiblichen Brust im Rahmen einer Machbarkeitsstudie zum Ziel hatte. Hierbei sollte die Vorgehensweise des Radiologen bei der Befundung möglichst nachgeahmt werden. Gegenüber den zuvor aufgeführten Systemen (s. Kap. 3.2) unterscheidet sich der Prototyp in zwei wesentlichen Merkmalen:

- Die Software umfasst die Bildverarbeitung *aller Ebenen der Wissensrepräsentation* (s. Kap. 2.4.1), im Einzelnen die Komponenten Registrierung (s. Kap. 4.2), Segmentierung (s. Kap. 4.3), Berechnung von sowohl kinetischer als auch morphologischer Information (s. Kap. 4.4) und überwachte High-Level-Klassifikation (s. Kap. 4.5). Damit kann sie alle Arbeitsschritte des Radiologen nachbilden und *Dignitätsvorhersagen pro Läsion* treffen. Die Prototypsoftware übertrifft dadurch den Funktionsumfang existierender CAD-Systeme.
- Die gesamte Bildverarbeitung aller Komponenten des Prototyps läuft ohne jegliche Benutzerinteraktion *automatisch* ab. Daher benötigt die Software auch keine graphische Benutzerschnittstelle.

Durch die Prototypsoftware der vorliegenden Arbeit wurde nach meinen Erkenntnissen erstmals ein automatisches Diagnosesystem auf funktionaler MRT der weiblichen Brust mit einer eindeutigen Dignitätsvorhersage pro Läsion entwickelt.

Genauso wie der Begriff *CAD* werden auch mit *automatisch* bzw. *automatisiert* sehr unterschiedliche Leistungen bezeichnet. Selbst Systeme, die dem Benutzer mehr als 10 Interaktionen (Parametereinstellungen, Markierungen, Vorgabe oder Ausschluss von Bildbereichen durch die GUI, ...) abverlangen, bevor ein Ergebnis ausgegeben werden kann,

bezeichnen sich oft noch als „automatisiert“. Automatisch meint in der vorliegenden Arbeit, dass ohne jegliche Benutzerinteraktion das System selbstständig Diagnosen pro Läsion berechnet.

Bei der Bezeichnung CAD für die Prototypsoftware steht das *D* eindeutig für *Diagnosis* als High-Level-Funktionalität (vgl. Kap. 3.1). Allerdings müsste, da es sich um eine automatische Software handelt, das *A* für *Aided* in *CAD* ersetzt werden: *Computer Automatic Diagnosis*. In der Literatur wird hierfür der Begriff *Computer Diagnosis (CD)* verwendet.

Kapitel 4

Material und Methoden

Gegenstand dieses Kapitels sind die autonom ausgeführten Bildverarbeitungsschritte (vgl. Kap. 2.4 und 3.3), die die Prototypsoftware zur Gewinnung der Diagnose vollzieht. Dabei ahmt sie im weitesten Sinne die Vorgehensweise des Radiologen nach.

Nach dem Import (s. Kap. 4.1) wird mit der Registrierung als Vorverarbeitung eine Bewegungskorrektur der Bilddaten durchgeführt (s. Kap. 4.2). Die Segmentierung (s. Kap. 4.3) erkennt alle verdächtigen Anreicherungen. Es folgt eine Bewertung jedes einzelnen Befundes sowohl nach kinetischen als auch morphologischen Kriterien (s. Kap. 4.4). Aus den gesammelten Informationen erfolgt abschließend eine Dignitätsaussage pro Läsion durch Anwendung überwachter Klassifikation (s. Kap. 4.5).

Der Diskussionsteil (s. Kap. 6) beschäftigt sich ausführlich mit der Beurteilung der Qualität der durch diese Prototypsoftware erzielten Ergebnisse. Die zur Evaluation eingesetzten überwiegend statistischen Verfahren werden im Kapitel 4.6 zusammengefasst.

Abbildung 4.1 gibt eine Übersicht der verwendeten Bildverarbeitungsschritte (Benennung der linken Spalte) zusammen mit dem Fluss (Linien) der Daten (geometrische Formen). Die verwendeten Abkürzungen, Details zu dem Diagrammablauf und der Bezug der vorgestellten Methoden im Kontext mit der allgemeinen Bildverarbeitung sind in den jeweils folgenden Kapiteln beschrieben.

Die gesamte Funktionalität der Prototypsoftware wurde in der Programmiersprache C++ mit dem Microsoftcompiler Visual Studio V6.0 entwickelt [Mic]. Neben den folgenden Bildverarbeitungsalgorithmen stellen auch die Datenstrukturen für die Low-, Intermediate- und High-Level-Repräsentation eigene Implementierungen dar. Als Fremdbibliotheken wurde die Standardbibliothek von Microsoft gegen das freie Paket STLport [STL] ausgetauscht. Zum Auslesen der DICOM-Tags (s. Anhang A) wurde die DCMTK-Bibliothek von Offis aus Oldenburg verwendet [Off]. Für alle Klassifikationen (s. Kap. 4.5.3 und 4.6.5) und zum Erstellen einer statistischen Aussage wurde das Statistikpaket R [GI] verwendet. Zum Ausführen der R-Pakete wurden eigene Schnittstellen in die Software integriert.

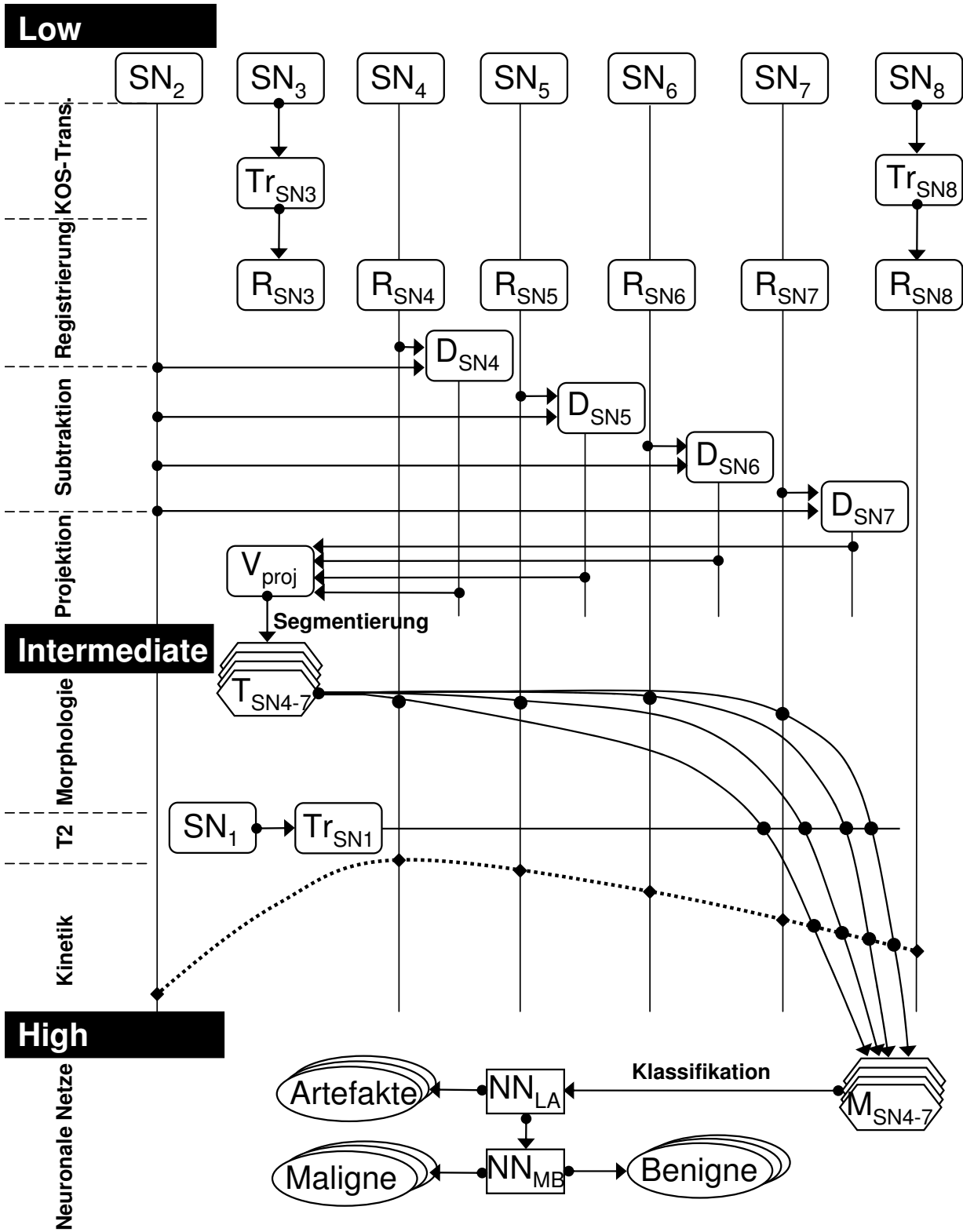


Abbildung 4.1: Ablauf der ausgeführten Bildverarbeitungsschritte der Prototypsoftware

4.1 Bilderhebung

Das gesamte Patientenkollektiv der vorliegenden Arbeit wurde an demselben 1,5-Tesla-MRT (Magnetom Sonata von Siemens) untersucht (s. Abb. 4.2). Als Oberflächenspule wurde eine handelsübliche CP Doppelbrustspule ebenfalls von Siemens verwendet (s. Bildausschnitt). Diese wurde auf dem Tisch des MRT-Gerätes unter der in Bauchlage befindlichen Patientin platziert, deren beide Mammae in die Vertiefung der Spule herabgingen. Um Bewegungen während der Untersuchung zu minimieren, wurden beide Mammae durch die integrierte Kompressionsvorrichtung der Spule fixiert. Die Bildakquisition wurde mit identischem Protokoll (s. nächstes Kapitel) unter Anwendung desselben KMs durchgeführt, das abhängig vom Körpergewicht der Patientin dosiert wurde. Der Injektor für die KM-Applikation ist auf der linken Bildseite an einem Infusionsständer befestigt, ein Zugang zur Armvene der Patientin ist gelegt. Der Injektor wurde nach Messung der nativen Serie durch das MRT-Gerät aktiviert. Der Untersuchungstag der Patientinnen war immer zwischen dem fünften und fünfzehnten Tag ihres Menstruationszyklus datiert (vgl. Kap. 2.3.2).

Bevor der Import von den Bilddaten erläutert wird (s. Kap. 4.1.2), beschreibt das nächste Kapitel das verwendete MRT-Protokoll.

4.1.1 MRT-Protokoll

In der MRT-Technik verhält sich die zeitliche Auflösung antiproportional zur örtlichen. Zur Erhebung der Bilddaten wurde daher ein spezielles Protokoll kombinierter 3D-Sequenzen eingesetzt [VTK04, VTS02]. Es enthält einerseits schnelle Sequenzen innerhalb der ersten Minuten nach KM-Gabe für eine zeitlich hochaufgelöste Darstellung der Kinetik, andererseits ermöglichen drei Sequenzen mit einer hohen Ortsauflösung eine detailreiche Abbildung der Morphologie.

Die Abfolge der gemessenen Sequenzen mit zugewiesenen Seriennummern (SN) kann der Abbildung 4.3 entnommen werden. Kombiniert werden die Orientierung (vgl. Abb. A.1) und die Ortsauflösung. Vor der intravenösen Applikation des Kontrastmittels (unterbrochene Zeitachse) wird zunächst mittels einer fettgesättigten T_2 -gewichteten Sequenz (T_2fs , s. Kap. 2.2.4) untersucht. Als weitere native Aufnahmen folgen eine koronare und eine transversale T_1 -gewichtete Sequenz mit hoher Ortsauflösung (*high-res*). In einer anschließenden 25 Sekunden dauernden Pause wird das paramagnetische Kontrastmittel (Gadolinium DT-PA, Magnavist von Schering) als Bolus in einer Dosis von 0,15 mmol/kg Körpergewicht durch eine Druckinjektionsspritze (MR Injector, Medrad Indianola) intravenös injiziert und mit 20 ml einer 0,9 %igen Kochsalzlösung gespült. Die erste gemessene Nachkontrastserie (durchgezogene Zeitachse) markiert den Zeitpunkt null. Es folgen drei T_1 -gewichtete koronare Aufnahmen mit geringer Ortsauflösung (*low-res*). Durch diese schnellen Sequenzen kann die Kinetik der ersten 213 Sekunden nach KM-Gabe mit hoher zeitlicher Auflösung erfasst werden. Zum Abschluss werden wieder *high-res* Sequenzen gemessen. Es folgen zwei weitere T_1 -gewichtete Aufnahmen in beiden Orientierungen.

Die Tabelle 4.1 fasst die wesentlichen Unterschiede der Protokollsequenzen zusammen.

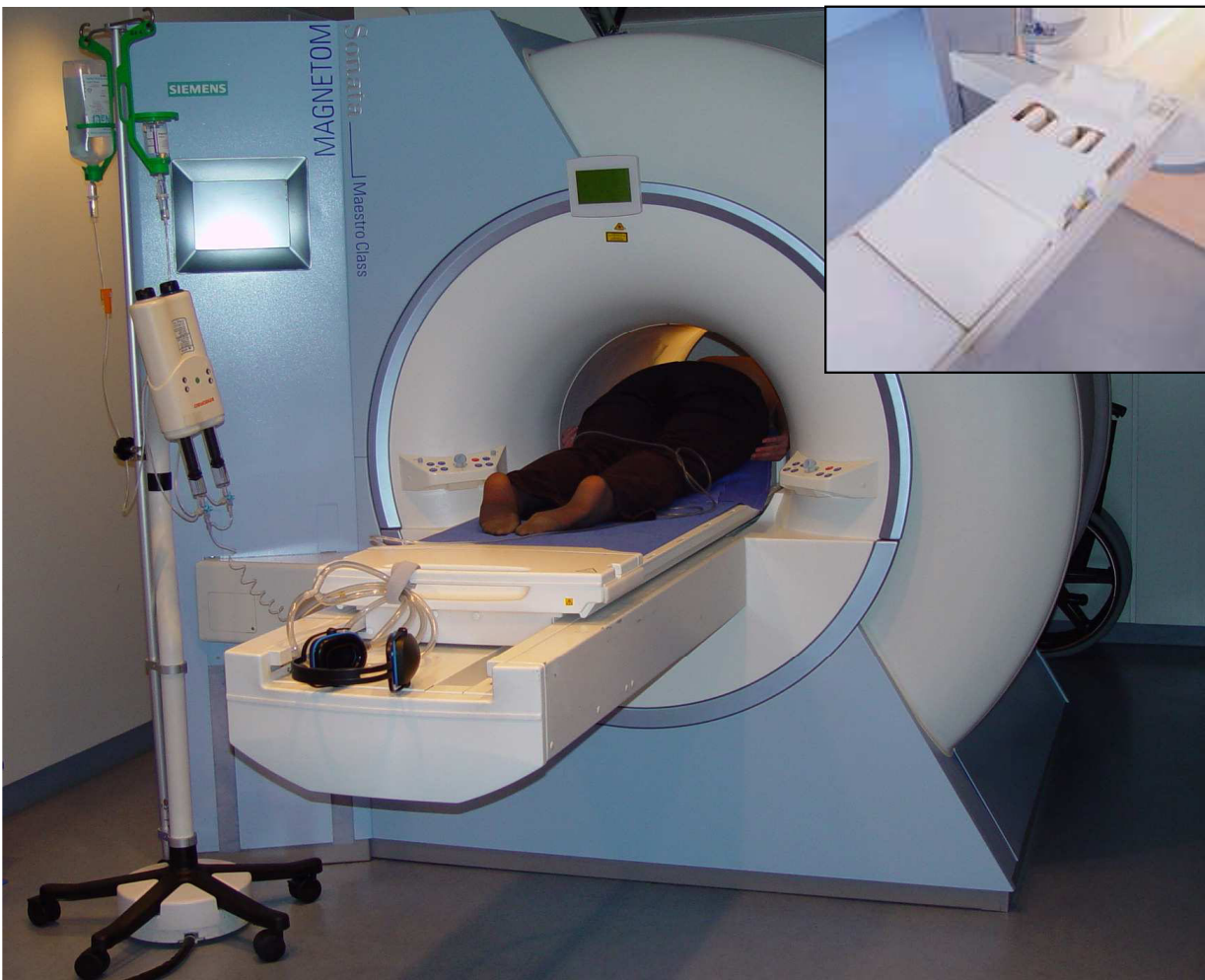


Abbildung 4.2: Dynamische MRT-Untersuchung der Brust, der Bildausschnitt zeigt die verwendete Brustspule

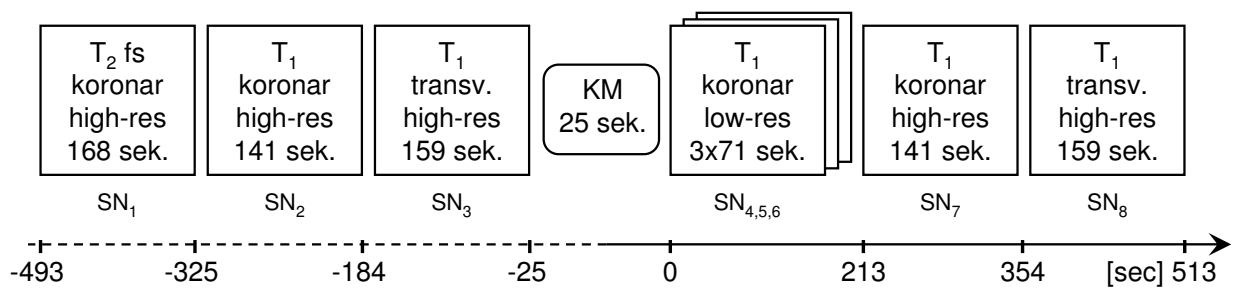


Abbildung 4.3: Sequenzen und zeitlicher Ablauf des eingesetzten Protokolls

Name	T_1w low-res	T_1w high-res	T_1w high-res	T_2w fs
Orientierung	koronar	koronar	transversal	koronar
Anzahl der Schichten	72	72	72	50
Schichtdicke (mm)	1,7	1,7	2,0	2,3
FoV (mm ²)	320x160	320x160	320x160	320x160
Matrix	256x128	512x256	512x256	512x256
Ortsauflösung (mm ²)	1,25x1,25	0,625x0,625	0,625x0,625	0,625x0,625
Messzeit (s)	71	141	159	168

Tabelle 4.1: Das Protokoll der Bilderhebung verwendet vier verschiedene Sequenzen unterschiedlicher Orientierung und Ortsauflösung.

Die 3D-Bilddaten einer T_1 -gewichteten koronaren high-res Sequenz setzen sich aus 72 Bildschichten mit einer Schichtdicke von je 1,7 mm zusammen. Die gemessenen Signale werden in jeder Schicht durch eine Matrix von 512 x 256 Bildpunkten diskretisiert; jeder Bildpunkt hat eine Ortsauflösung von 0,625 x 0,625 mm². Die low-res Sequenz besteht aus gleicher Schichtdicke, verwendet aber mit lediglich 256 x 128 Bildpunkten eine reduzierte Matrix. Jeder diskrete Bildpunkt beschreibt mit einer 1,25 x 1,25 mm² Schichtauflösung ein vierfach größeres Volumen. Die low-res Bilddaten werden nach Abschluss der Messung durch das MRT-Gerät durch Interpolation auf eine 512 x 256 Matrix wieder auflösungsgleich zu den Bilddaten der high-res Sequenzen; sie enthalten aber eine reduzierte Bildinformation. Die Bilddaten sowohl der low- als auch der high-res Sequenzen messen einen identischen Raum mit einem Volumen von 320 x 160 x 122 mm³.

Die transversalen T_1 -gewichteten Sequenzen sind in Bezug auf Matrixgröße und Schichtanzahl identisch zu den koronaren. Jedoch liegt die transversal orientierte Schichtdicke mit 2,0 mm höher. Das Volumen des akquirierten Raumes umfasst 320 x 144 x 160 mm³.

Die T_2 -gewichtete fettgesättigte Aufnahme wird ebenfalls koronar aufgenommen. Sie besteht aus 50 Bildschichten, die jedoch mit 2,3 mm dicker sind. Hierdurch wird der Raum durch ein Volumen von 320 x 160 x 115 mm³ dargestellt.

Das FoV umschloss jeweils beide Mammae einschließlich der Mamillen und der ventralen Thoraxwand, so dass auch die axillar gelegenen Lymphknoten aufgenommen wurden. Die Ausdehnung der Bildvoxel beschreibt eine anisotrope Form. Die Kantenlänge der Voxel innerhalb der Bildschicht beträgt nur etwa ein Drittel gegenüber der Schichtdicke.

Durch die koronar orientierten Aufnahmen kann flexibler auf die Größe der Mamma eingegangen werden. Die Erfassung größerer Brüste kann durch eine Erhöhung der Schichtanzahl oder der Schichtdicke sichergestellt werden. Die koronare Darstellung ist für den Kliniker, also im Wesentlichen den Gynäkologen, sehr gut erfassbar. Die Milchgänge, die sternförmig auf die Brustwarze zulaufen, werden durch die Schichtführung nur im Querschnitt dargestellt. Das ist ein potentieller Nachteil bei der Bildgebung eines DCIS (s. Kap. 2.1.3), das bei sagittaler oder transversaler Orientierung durch einen Längsschnitt besser abgebildet wird.

4.1.2 Import

Alle Aufnahmen des Protokolls werden von dem MRT-Gerät im DICOM-Standard (s. Anhang A) exportiert. Allein durch die T_1 -gewichteten koronaren Sequenzen werden 360 Bildschichten akquiriert, was einer Datenmenge von 90 MB (Megabyte) entspricht.

Beim Import der DICOM-Bildschichten müssen die Daten der einzelnen Dateien wieder zu einem Bildvolumen zusammengefügt werden (s. Anhang A). Neben den Bildinformationen extrahiert die Software für die sich anschließenden Bildverarbeitungen auch Patientendaten, Akquisitionszeitpunkte der einzelnen Bildserien und Informationen über ihre räumliche Orientierung.

Zur effizienteren Datenhaltung legt der Prototyp die Bildvolumen der 8 Serien im Modell-KOS ab (s. Anhang A.2). Diese Darstellung enthält jedoch keine Informationen über den räumlichen Bezug der Bildpunkte zum gelagerten Patienten im MRT-Gerät. Die erzeugten Bilddaten der verschiedenen Sequenzen verfügen aufgrund uneinheitlicher Orientierung, Auflösung oder Schichtabstände über Modellkoordinatensysteme, die auf gleichen Koordinaten einen anderen Bereich des Patienten abbilden. Ein Vergleich der Bildinhalte auf Grundlage der Modellkoordinaten verschiedener Sequenzen ist so meist nicht möglich. Anhang A beschreibt die Erstellung von Transformationsmatrizen, die den Bezug zwischen Modell- und Patienten-KOS herstellen (s. Abb. A.2). Erst nach Anwendung einer KOS-Transformation beziehen sich die Koordinaten nun im Patienten-KOS auf gleiche Bildausschnitte. Bei Algorithmen, die nur einzelne Bildvolumen verarbeiten, genügt die Betrachtung des Bildvolumens im Modell-KOS. Andere Bildverarbeitungsschritte müssen jedoch die Bildinformationen aller Zeitserien gleichzeitig betrachten (z.B. Segmentierung in Kap. 4.3 oder Analyse der Kinetik in Kapitel 4.4) und erfordern zuvor eine Transformation.

Keine Transformation benötigt ein Vergleich der koronaren low- und high-res Sequenzen, die – trotz der unterschiedlichen Ortsauflösung – bereits durch die anschließende Interpolation des MRT-Geräts über identische KOS verfügen.

4.2 Registrierung

In einem Vorverarbeitungsschritt müssen zuerst Bewegungen in den dynamischen Bildfolgen erkannt und eliminiert werden, um die Anzahl an falsch-positiven KM-Anreicherungen in den Differenzbildern zu minimieren (s. Kap. 2.3.1). Erst dadurch wird sichergestellt, dass sowohl die Kinetik (s. Kap. 2.3.2) als auch die Morphologie (s. Kap. 2.3.3) korrekt in den Bilddaten dargestellt sind. Dieser Vorverarbeitungsschritt stellt eine Selbsttransformation der Low-Level-Repräsentation dar (s. Kap. 2.4.2).

Es gibt verschiedene Ursachen für Bewegungsartefakte. Häufig entspannen sich die Patientinnen nach anfänglicher Angst während der langen Untersuchungsphase und sinken tiefer in die Brustspule ein. Die Verschiebung der Brüste lässt sich trotz Verwendung der Kompressionsvorrichtung (s. Kap. 4.1) nicht völlig unterbinden. Auch die Vibration des MRT-Geräts, Herzschlag oder Atmung führen zu geringfügigen Verschiebungen, die durch die heute hoch aufgelösten Bilddaten immer deutlicher erkennbar werden. Zudem können

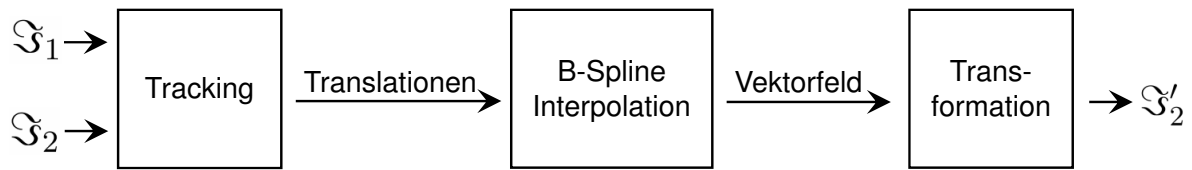


Abbildung 4.4: Teilschritte des Registrierungsalgorithmus

Rückfaltungsartefakte entstehen, wenn Objekte außerhalb des FoV jedoch noch innerhalb der Empfangsspule liegen. Diese sind auf den koronaren Sequenzen besonders im Bereich des Thorax zu erkennen (s. Abb. 6.18).

Die Registrierung der Brust MRT-Bilder ist anspruchsvoll. Es existieren keine steifen Strukturen wie Knochen, die als Orientierung für die Verschiebung dienen könnten (vgl. Kap. 6.1.4). Die Brust verschiebt sich nicht als Ganzes, vielmehr können sich benachbarte Bereiche des weichen Gewebes unabhängig voneinander in unterschiedliche Richtungen verschieben. Die Bewegung der Brust stellt sich als eine Kombination globaler und lokaler Einflüsse dar. Auch sind die Bilddaten allgemein wenig kontrastreich, es zeigen sich wenig charakteristische Strukturen, anhand derer Bewegungen erkannt werden können.

Folgend wird ein 3D-Algorithmus entwickelt, um Bewegungen in allen Raumrichtungen basierend auf der Berechnung des optischen Flusses (s. Anhang B) korrigieren zu können.

Ziel der Registrierung ist es, die durch Bewegung verursachten Bildverschiebungen durch eine Selbsttransformation des Bildes zu eliminieren. Im nächsten Kapitel wird dazu eine Kombination aus drei sequentiell ausgeführten Bildverarbeitungsalgorithmen (s. Abb. 4.4) vorgestellt [VFM07]:

1. Bestimmen von Verschiebungsvektoren (*Translationen*) durch Tracken ausgewählter, charakteristischer Bildpunkte zweier aufeinanderfolgender Bilder \mathfrak{S}_1 und \mathfrak{S}_2 .
2. Erzeugen eines Vektorfelds für die Beschreibung der Bewegung im gesamten Bild durch Interpolation der zuvor bestimmten Einzeltranslationen.
3. Durchführung der elastischen Bildtransformation auf das Bild \mathfrak{S}'_2 durch Übertragung der Grauwerte von \mathfrak{S}_2 an die neue Position in \mathfrak{S}_1 anhand der Bewegungsinformation des Vektorfeldes.

Die Registrierung wird auf allen T_1 -gewichteten Sequenzen SN_2 bis SN_8 angewendet (s. Abb. 4.3 u. 4.1). Da mit koronarer (K) und transversaler (T) Schichtführung Sequenzen unterschiedlicher Orientierung gemessen wurden, müssen vor der Durchführung der drei Schritte die Bildvolumen in ein gemeinsames KOS transformiert werden (s. Anhang A.3). Der geringste Aufwand besteht darin, die beiden transversalen Aufnahmen in ein koronares Modell-KOS durch die Matrix $M_{T \rightarrow K}$ (s. Formel A.2) affin zu transformieren. Es sei $SI_{\mathfrak{S},p}$ der Signalwert des Bildvolumens \mathfrak{S} an der Voxelposition p :

$$\text{Signalwert} = SI_{\mathfrak{S},p}, \text{ mit } p = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathfrak{S} \quad (4.1)$$

Das transformierte Bild $Tr_{\mathfrak{S}}$ ergibt sich aus:

$$\begin{aligned} Tr_{\mathfrak{S}} &= f_{T \rightarrow K}(\mathfrak{S}), \mathfrak{S} \in \{SN_3, SN_8\}, \text{ mit} \\ SI_{Tr_{\mathfrak{S}},p} &= SI_{\mathfrak{S},p'}, \text{ mit } p' = M_{T \rightarrow K} \cdot p, \forall p \in Tr_{\mathfrak{S}} \end{aligned} \quad (4.2)$$

Da die transformierten Bildkoordinaten p' das diskrete Bildraster verlassen, müssen beim Übertragen die Grauwerte des transversalen Bildvolumens interpoliert werden. Hierfür wird mit einer trilinearen Interpolation [Hec94] ein Standardverfahren verwendet.

Tr_{SN_3} und Tr_{SN_8} ersetzen im Folgenden die transversalen Volumen (s. Abb. 4.1), aufgrund des identischen Modell-KOS können alle T_1 -gewichteten Bildvolumen nun auf Voxel-ebene miteinander verglichen werden.

Da die Nachkontrastaufnahmen zwecks Fettauslöschung zur nativen Aufnahme normiert werden (s. Kap. 2.3.1), erfolgt die Registrierung der Bildinhalte der Serien Tr_{SN_3} , SN_4 - SN_7 und Tr_{SN_8} auf den Zeitpunkt der nativen Aufnahme SN_2 .

4.2.1 Sequenzielle Teilschritte

Tracking Im ersten Schritt werden die Translationen einzelner korrespondierender Bildpunkte der nativen zu den nachfolgenden Bildserien ermittelt. Viele Bildbereiche wie das homogene Fettgewebe oder die Luft können keine Bewegungsinformation wiedergeben. Daher werden auf der nativen Aufnahme zuerst nur diejenigen charakteristischen Bildbereiche (*Feature*) ermittelt, die auf den nachfolgenden Bildserien möglichst eindeutig identifiziert (*Tracking*) werden können. Anhang B beschäftigt sich ausgiebig mit den mathematischen Herleitungen des Featuretrackings, die folgend kurz zusammengefasst werden: Für jeden Punkt des Bildvolumens wird innerhalb eines Bildausschnitts – definiert durch das Fenster $WF_{x,y,z}$ – die Gradientenmatrix G berechnet und der minimale Eigenwert als Qualitätsmaß bestimmt. Fenster homogener Bildbereiche führen durch niedrige Qualitätswerte zu schwachen Features. Durch eine untere Schranke, die durch den Skalierungswert $TH_{Feature}$ relativ zum höchsten Qualitätswert des Volumens bestimmt wird, werden solche schwachen Features bei Unterschreitung der Schranke von der weiteren Verarbeitung ausgeschlossen.

In Regionen mit markanten Bildinformationen häufen sich vermehrt starke Features, andere Bereiche enthalten nur wenige und schwächere. Um eine möglichst gleichmäßige Verteilung der Featurepositionen über das gesamte Volumen zu erreichen, wird durch ein Fenster der Größe $WS_{x,y,z}$ die lokale Umgebung aller Features getestet. Beginnend bei den stärksten Qualitätswerten, werden alle ebenfalls im Fenster befindlichen Features ausgeschlossen. Letztendlich verbleiben als eine Teilmenge aller Bildpunkte nur diejenigen Positionen mit den lokal stärksten Features.

Die selektierten Features werden auf die Bildvolumen $Tr_{SN_3}, SN_4 - SN_7$ und Tr_{SN_8} der Reihe nach getrackt. Ein solcher 3D-Trackingalgorithmus basierend auf Kanade-Lucas-Tomasi wird in Anhang B hergeleitet; er legt einen Fensterbereich um die Features und versucht den Inhalt dieses Bildausschnitts auf dem zeitlich späteren Bildvolumen an verschobener Position wiederzuerkennen. Das Tracken der einzelnen Features auf die nächste Bildserie erfolgt iterativ, d.h. die Zielposition der Verschiebung wird schrittweise angenähert. Als erfolgreich beendet gilt das Tracking eines einzelnen Features, wenn der berechnete Translationsvektor während einer Iteration eine durch den Vektor TH_{Trans} vorgegebene Genauigkeit unterschreitet. Zur Qualitätswahrung muss diese innerhalb von TH_{Iter} Iterationsschritten erreicht werden, andernfalls gilt das Tracking eines Features als gescheitert. Ein erfolgreiches Tracken resultiert in einem Translationsvektor, der die Richtung und Länge der Verschiebungen für das selektierte Feature auf das nachfolgende Bildvolumen wiedergibt.

Vektorinterpolation Da die Translationen nur für einige Bildpunktkorrespondenzen im Volumen vorliegen, müssen diese Informationen für die übrigen Bildbereiche interpoliert werden, um den Registrierungsprozess durch die nachfolgende Transformation der Bildgrauwerte abschließen zu können. Die Interpolation wird dadurch erschwert, dass die berechneten 3D-Translationsvektoren abhängig von ihrer Featureposition unregelmäßig über das ganze Bildvolumen verteilt sind. Sie geben durch ihre Länge und Richtung zudem eine Mischung aus globalen und lokalen Bewegungen wieder. Anhang C erläutert die auf B-Splines basierende Interpolation der Vektoren, die folgend kurz zusammengefasst wird: Über ein definiertes Gitter wird iterativ ein Vektorfeld erzeugt. Durch Veränderung der Gitterabstände können globale und lokale Verschiebungen gleichermaßen berücksichtigt werden. Das initial grobe Gitter mit dem höchsten Abstand $G_{Distance}$ wird iterativ immer weiter verkleinert, bis jeder Bildpunkt über einen eigenen Gitterpunkt verfügt, der den interpolierten Translationsvektor wiedergibt.

Bildtransformation Das erzeugte Vektorfeld jeder Nachkontrastserie enthält die Information vorhandener Bewegungen ausgehend vom Zeitpunkt der nativen Aufnahme. Die abschließende elastische Bildtransformation stellt eine Selbsttransformation (s. Formel 4.3) im Low-Level dar, die die registrierten Bildvolumen R_{SN_X} durch die Funktion f_{Reg} erzeugt (s. Abb. 4.1). Hierzu wird für jeden Bildpunkt p der Grauwert aus den Nachkontrastserien \mathfrak{S} an der durch das Vektorfeld (f_{Spline} , s. Formel C.1) vorgegebenen Position übernommen.

$$\begin{aligned} R_{\mathfrak{S}} &= f_{Reg}(\mathfrak{S}), \mathfrak{S} \in \{Tr_{SN_3}, SN_4, SN_5, SN_6, SN_7, Tr_{SN_8}\}, \text{ mit} & (4.3) \\ SI_{R_{\mathfrak{S}},p} &= SI_{\mathfrak{S},p'}, \text{ mit } p = f_{Spline}(p'), \forall p \in R_{\mathfrak{S}} \end{aligned}$$

Wie schon bei der KOS-Transformation verlassen die Translationsvektoren des Vektorfelds das diskrete Bildraster, der Grauwert von p' wird wiederum durch trilineare Interpolation bestimmt.

Der ganze Prozess der Registrierung auf den Zeitpunkt der Serie SN_2 erfordert also ein einmaliges Bestimmen der Features auf der nativen Serie, während das Tracking sequentiell über alle folgenden Nachkontrastserien vorangetrieben wird. Dabei werden die Translationsvektoren der zuvor getrackten Serie als Startpositionen für die nächste übernommen.

Entsprechend der erkannten Bewegung werden die Vektoren auf jeder weiteren Serie kontinuierlich verändert. Vor der Durchführung der elastischen Grauwerttransformation f_{Reg} wird pro Nachkontrastserie ein Vektorfeld interpoliert.

Die zweite native Serie SN_3 wird für die sich anschließende Bildverarbeitung nicht benötigt. Sie wird nur für spätere statistische Evaluationen (s. Kap. 6.1.2) in den Prozess aufgenommen.

4.3 Segmentierung

Die Bildvorverarbeitung durch die Registrierung (s. vorheriges Kapitel) verbessert die Qualität der Differenzbilder (s. Kap. 2.3.1), so dass im Idealfall nur noch die Anreicherung durch das Kontrastmittel und keine Bewegungsartefakte enthalten sind. Zur Erzeugung der für die Segmentierung notwendigen Differenzen D_{SN_4} bis D_{SN_7} wird eine Subtraktion f_{Sub} zur nativen Aufnahme auf Bildpunktebene durchgeführt (s. Abb. 4.1):

$$\begin{aligned} D_{\mathfrak{S}} &= f_{Sub}(\mathfrak{S}), \mathfrak{S} \in \{SN_4, SN_5, SN_6, SN_7\}, \text{ mit} \\ SI_{D_{\mathfrak{S}},p} &= SI_{\mathfrak{S},p} - SI_{SN_2,p}, \forall p \in D_{\mathfrak{S}} \end{aligned} \quad (4.4)$$

Aufgabe der Segmentierung ist es, die anreichernden Läsionen der Low-Level-Repräsentation von den übrigen Bildinformationen abzugrenzen. Jede erkannte Läsion wird als eigenes Token in das abstraktere Intermediate-Level überführt, die nachfolgenden Bildverarbeitungsschritte können durch die Tokenobjekte vereinfacht auf die relevanten Informationen des Low-Levels zugreifen (s. Kap. 2.4.3). Erst mit der Abgrenzung durch die Token lässt sich die Anreicherung detailliert morphologisch beschreiben (s. Kap. 4.4).

In der vorliegenden Arbeit wurde eine neu adaptive Schwellwertsegmentierung entwickelt, die durch eine aus Regeln formulierte Heuristik optimiert wurde. Nach Erläuterung des prinzipiellen Algorithmus anhand eines einzelnen 3D-Bildvolumens im nächsten Kapitel folgt eine Verallgemeinerung zur Segmentierung funktionaler 4D-Bilddaten (s. Kap. 4.3.2).

4.3.1 Einzelne Bildserie

Das entwickelte 3D-Segmentierungsverfahren nutzt adaptive Schwellwerte, um sich dem individuellen Anreicherungsverhalten einzelner Läsionen anzupassen. Das iterative Vorgehen vermag durch Anwenden einer aus drei Regeln formulierten Heuristik die Anreicherungen der Läsionen von denen des umliegenden gesunden Parenchyms zu unterscheiden. Alle potentiellen Voxel einer einzelnen Läsion werden als zusammenhängendes 3D-Tokenobjekt ins Intermediate-Level überführt. Die Heuristik für die Entscheidungsfindung setzt sich

aus folgenden Regeln zusammen:

1. Je höher eine gegebene Signalintensität, desto wahrscheinlicher stellt sie den Teil einer Läsion dar.
2. Die Beziehung zwischen dem signalstärksten Voxel SI einer Läsion und einem individuellen Schwellwert $SI_{Hom} = f_{Thres}(SI)$, der diese Läsion optimal segmentiert, ist streng monoton fallend.
3. Das Bildrauschen kann von den SI der Läsion durch einen unteren Schwellwert SI_{Noise} abgegrenzt werden.

Die Segmentierung basiert auf dem Region Growing, einem Standardverfahren der Bildverarbeitung, das sich als ein durch die Bildeigenschaften gesteuertes Wachstumsverfahren definiert [Han00]. Ausgehend von einer initialen Saatzelle wird jedes angrenzende Voxel aufgenommen, wenn es dasselbe Homogenitätskriterium erfüllt. Der Prozess wird für die Nachbarn eines jeden neu angelagerten Voxels wiederholt. Dadurch wächst die Saatzelle so lange zu einem zusammenhängenden Objekt heran, bis keine angrenzenden Voxel mehr das Homogenitätskriterium erfüllen. Wenn als Saatzelle ein Voxel mit der Intensität SI_{Seed} dient, gilt das Homogenitätskriterium als erfüllt, wenn die SI eines anzulagernden Voxels größer oder gleich dem Schwellwert $SI_{Hom} = f_{Thres}(SI_{Seed})$ der Regel (2) ist.

Die Segmentierung startet nach Analyse des Histogramms aufgrund von Regel (1) auf dem Voxel mit dem höchsten Signalwert SI_{Max} . Das Region Growing überführt die dazugehörige Läsion durch Segmentierung der angrenzenden Voxel als zusammenhängendes 3D-Objekt mit individuellem Schwellwert $f_{Thres}(SI_{Max})$ in das Intermediate-Level. Nach Regel (2) wird die Läsion optimal segmentiert.

Die übrigen Anreicherungen werden durch wiederholtes Anwenden des Region Growings segmentiert, die Saatzellen werden dabei iterativ (SI_{Iter}) anhand des SI bestimmt. Nachdem mit den höchsten Signalwerten ($SI_{Iter} = SI_{Max}$) begonnen wurde, wird der Signalwert mit jeder Iteration dekrementiert ($SI_{Iter} = SI_{Iter} - TH_{Dec}$) und alle diejenigen Voxel als Saatzellen bestimmt, deren SI größer oder gleich SI_{Iter} ist, aber noch nicht ins Intermediate-Level überführt wurden. Erfüllen gleich mehrere Voxel diese Bedingung, werden sie einzeln nacheinander abgearbeitet. Der Signalwert SI_{Iter} wird solange dekrementiert, bis der individuelle Schwellwert $f_{Thres}(SI_{Iter})$ aus Regel (2) die Schranke SI_{Noise} für das Bildrauschen aus Regel (3) unterschreitet.

Durch die iterative Segmentierung wird sichergestellt, dass alle Bildbereiche, die ein vom Bildrauschen deutlich unterscheidbares Signal zeigen, als zusammenhängende 3D-Tokenobjekte vom restlichen Bildinhalt abgegrenzt werden. Nun kann es bei der Segmentierung eines neuen Objektes vorkommen, dass während des Region Growings Voxel eines bereits zuvor segmentierten Objektes angelagert werden. Da ein Voxel jedoch nur einer Läsion zugeteilt werden kann, müssen solche Überlagerungen behandelt werden. Mögliche Konsequenzen wären das Verschmelzen der beteiligten oder Entfernen eines der betroffenen Objekte. Ein Verschmelzen wäre inkonsequent, da durch Regel (2) vorausgesetzt wurde, dass das bereits existierende Objekt aufgrund des individuell berechneten Schwellwertes

optimal segmentiert wurde. Es würden Bereiche vereint, die ein unterschiedliches Homogenitätskriterium erfüllen.

Beim Entfernen stellt sich die Frage, welches der beiden Objekte beibehalten werden soll. Die Segmentierungsheuristik wird durch Überlagerungsregeln erweitert: Durch das Dekrementieren des Signalwertes SI_{Iter} ergibt sich für jedes neue Region Growing ein Schwellwert, der aufgrund der streng monoton fallenden Funktion f_{Thres} niedriger als der aller zuvor segmentierten Objekte ist. Kommt es nun zu einer Überlagerung, muss davon ausgegangen werden, dass der durchschnittliche Signalwert der zuletzt segmentierten Läsion niedriger ist als der zuvor segmentierten. Wiederum durch Regel (1) begründet, behält der Algorithmus bei Überlagerungen bestehende Objekte bei und schließt neue Segmentierungen von der Ergebnisliste aus. Für den Ausschluss ist unerheblich, ob Überlagerungen mit einem Objekt der Ergebnisliste oder mit einem bereits ausgeschlossenen Objekt entstehen.

4.3.2 Funktionale Bildserien

Mit dem beschriebenen Segmentierungsverfahren des vorherigen Kapitels kann jedes 3D-Differenzvolumen D_{SN_X} separat segmentiert werden. Um Aussagen pro Patient statt pro Zeitserie zu leisten, wird es nun zu einem 4D-Verfahren erweitert. Das generierte künstliche 3D-Bildvolumen V_{proj} (s. Abb. 4.1) bildet in einem zusätzlichen Zwischenschritt die 4D-Bildinformation durch die Projektion f_{proj} der Einzelvolumen über die Zeit in sich ab:

$$V_{proj} = f_{proj}(D_{SN_4}, D_{SN_5}, D_{SN_6}, D_{SN_7}) \quad (4.5)$$

Die Funktion f_{proj} muss so gewählt sein, dass die Anreicherung einer Läsion – auch wenn sie nur auf einer Serie anreichert – im Projektionsvolumen erkenn- und abgrenzbar abgebildet wird.

Das Volumen SN_8 geht, da die KOS-Transformation die Bildqualität (s. Kap. 6.1.2) und damit die nachfolgende Morphologieberechnung beeinflusst, nicht in die Segmentierung ein. Sie findet lediglich in der Bestimmung der Kinetik Anwendung (s. Kap. 4.4.3).

Auf dem Projektionsvolumen kann nun die 3D-Segmentierung des vorherigen Abschnitts initiiert werden. Jede Läsion, auch wenn sie zuvor auf mehreren Serien anreicherte, wird so höchstens einmal erkannt.

Da angenommen werden muss, dass sich die Morphologie der Läsion durch die Projektion verändert hat, ist eine Nachverarbeitung der Segmentierungsergebnisse notwendig. Dazu werden die Token anhand ihrer durchschnittlichen Signalintensität absteigend sortiert. Beginnend mit den signalstärksten Token – aufgrund von Regel (1) – wird das Projektionsvolumen wieder verlassen und für jedes einzelne Token diejenige Bildserie D_{SN_X} bestimmt, in der sich die Läsion am charakteristischsten darstellt. Dort wird das endgültige Token durch erneutes Anwenden des Region Growings für die nachfolgende Morphologiebestimmung nachsegmentiert.

Kuhl [Kuh00] schlägt für die Morphologiebestimmung diejenige Bildserie vor, in der die Läsion die stärkste Anreicherung zeigt. Dies ist konform zur Definition in Regel (1). Die

charakteristische Bildserie für die Nachsegmentierung wird also über die Signalintensität bestimmt. In der algorithmischen Umsetzung wird das jeweilige Token des Projektionsvolumens als Schablone benutzt und dasjenige Differenzvolumen (D_{SN_4} bis D_{SN_7}) bestimmt, das über alle Voxel der Schablone den höchsten mittleren Grauwert aufweist. Für die Nachsegmentierung auf dem gewählten Differenzvolumen geht dieser Mittelwert für die individuelle Schwellwertberechnung in die Funktion f_{Thres} der Regel (2) ein. Als Saatzelle wird das Voxel der maximalen Intensität bestimmt. Die Bildserie der Nachsegmentierung bestimmt zudem, in welche Tokenliste T_{SN_4} bis T_{SN_7} das jeweils endgültig segmentierte Objekt zugeordnet wird (s. Abb. 4.1). Diese nach Bildserien sortierte Tokenliste ersetzt die ursprüngliche Liste des Projektionsvolumens für die sich anschließende Merkmalsextraktion.

4.4 Merkmalsextraktion

Durch die Segmentierung wurden die Läsionen als relevante Bildbereiche durch einzelne Objekte erkannt und als Token im Intermediate-Level abgelegt (T_{SN_4} bis T_{SN_7} , s. Abb 4.1). Jedes Token listet seine referenzierten Voxellemente und zugehörige Bildserie auf und ermöglicht dadurch einen schnellen Zugriff auf die entsprechenden Grauwerte. Die registrierten Bilder (R_{SN_x}) des Low-Levels werden ab jetzt nur noch benötigt, um auf die Intensitätswerte der durch das Intermediate-Level maskierten Voxel zuzugreifen.

Während der Radiologe Läsionen durch ihre morphologischen und kinetischen Eigenschaften beschreibt, bezeichnet die Informatik die Gewinnung beschreibender Informationen als Merkmalsextraktion. Anhand der extrahierten Merkmale kann die Klassifikation die Objekte abschließend in maligne oder benigne unterscheiden (vgl. Kap. 2.4.4).

Als *statische Merkmale* werden folgend alle morphologischen Merkmale bezeichnet, die sich innerhalb einer einzelnen Bildserie berechnen lassen. Durch die Segmentierung wurde bereits festgelegt, auf welcher Serie die Morphologie eines Tokens bestimmt werden muss. Wie der Radiologe befundet die Software die Anreicherung anhand ihrer Form, Begrenzung (s. nächstes Kapitel) und Textur (s. Kap. 4.4.2).

Dagegen charakterisieren die *dynamischen Merkmale* (s. Kap. 4.4.3) mit einer zeitlichen Veränderung über alle Bildserien die Kinetik (s. Kap. 2.3.2).

Es werden folgend verschiedene Bildverarbeitungsverfahren zur Merkmalsextraktion vorgestellt. Manche Verfahren bilden Eigenschaften ab, die auch der menschlichen Wahrnehmung entsprechen. Die meisten Merkmale liegen jedoch außerhalb des visuellen Verständnisses.

4.4.1 Form und Begrenzung

Folgende statische Merkmale werden nur anhand der Form der Token, d. h. nur durch Anordnung der Voxelmengen im Raum bestimmt, ohne auf die Signalintensitäten des Low-Levels zugreifen zu müssen.

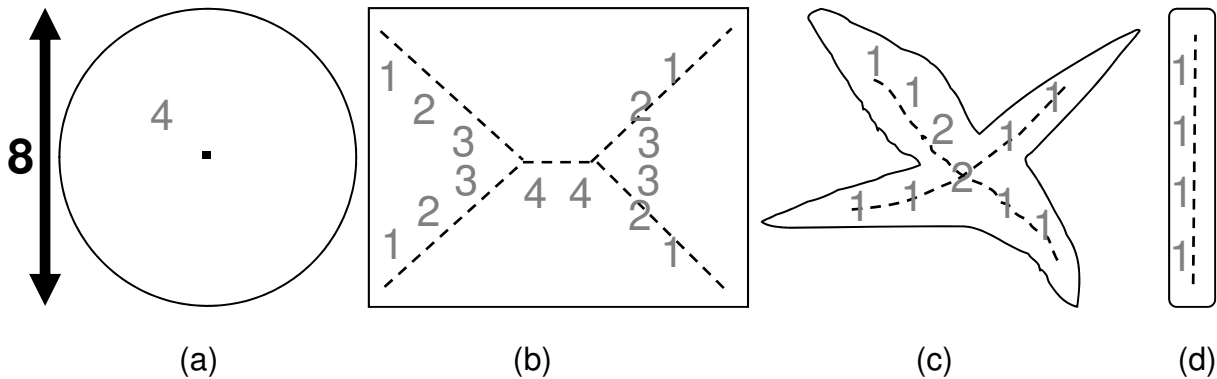


Abbildung 4.5: Beispiele von Skelettierungen im 2D mit den minimalen Randdistanzen

Zentraler Algorithmus zur Formbeschreibung eines Tokens ist die Skelettierung. Dieses Verfahren gehört zur Familie der morphologischen Operationen (s. Kap. 2.4.3) und vereinfacht die Darstellung eines Tokens und damit die Möglichkeit von dessen Beschreibung.

Eine Skelettierung reduziert ein Token auf eine minimale Menge an Voxel, die dennoch alle topologischen und geometrischen Informationen des Ursprungsobjektes wiedergeben können. Die resultierende Voxelmenge bildet dabei eine zusammenhängende Skelettlinie. Abb. 4.5 zeigt Beispiele in 2D. Die Ursprungsformen sind an den durchgehenden Linien zu erkennen, die gestrichelten Skelettlinien sind in ihren Inneren abgebildet.

Das hier verwendete 3D-Skelettierungsverfahren [LVG80, LRC94, Wei07] entfernt iterativ alle einfachen Randvoxel des Tokens, durch deren Wegfall sich die Topologie (räumliche Beziehung) nicht verändert. Diese einfachen Voxel werden anhand der Euler Charakteristik erkannt. Sie gibt eine Kennzahl für geschlossene Flächen und bildet eine topologische Invariante, denn topologisch identische Flächen besitzen gleiche Kennzahlen. Lobregt et al. [LVG80] geben ein effizientes Verfahren, das nur durch die Anzahl der sichtbaren Ecken, Kanten und Flächen der direkt benachbarten Voxel lokal entscheidet, ob sich durch Wegnahme eines Voxel die Kennzahl ändern würde.

Zu jedem Voxel der resultieren Skelettlinie wird durch die iterative Reduzierung die minimale Distanz (SD_{min}) zum Rand der Ursprungsform mitberechnet. Der Algorithmus verarbeitet durch Berücksichtigung des Voxelspacing (s. Anhang A.1) auch Bildvolumen mit anisotropen Voxel, die Randdistanz liegt in der Einheit Millimeter (mm) vor. Würde auf jedes Voxel der Skelettlinie der Mittelpunkt einer Kugel gelegt, deren Durchmesser diesem minimalen Abstand entspräche, könnte durch die Vereinigungen aller Kugeln das ursprüngliche Tokenobjekt wiederhergestellt werden (s. Abb. 4.5).

Die resultierende Skelettlinie zusammen mit den minimalen Randdistanzen ermöglicht statistische Auswertungen zur Formbeschreibung der Ursprungstoken. Zuvor werden noch einige klassische Merkmale aufgeführt.

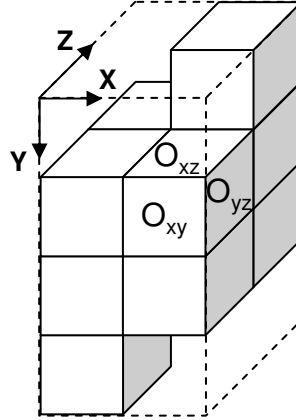


Abbildung 4.6: Voxeltoken mit minimaler Boundingbox

Die Größe des Volumens eines Tokens in der Einheit mm^3 berechnet sich aus der Anzahl der Tokenvoxel ($|Token|$) und dem Voxelspacing (s. Anhang A.1):

$$M_{Volume} = |Token| \cdot S_x \cdot S_y \cdot S_z \quad (4.6)$$

Der maximale Durchmesser wird aus der minimalen, das Token vollständig umfassenden Bounding Box ($BBox$) geschätzt (s. Abb. 4.6). Der Durchmesser entspricht der maximalen Kantenlänge der Bounding Box in der Einheit mm :

$$M_{MaxDiameter} = \max(BBox_x \cdot S_x, BBox_y \cdot S_y, BBox_z \cdot S_z) \quad (4.7)$$

Die größte aller minimaler Distanzen (SD_{min}) der Skelettlinienvoxel (*Skeleton*) entspricht dem minimalen Durchmesser des Tokens in der Einheit mm :

$$M_{MinDiameter} = \max_{i=1}^{|Skeleton|} SD_{min_i} \quad (4.8)$$

Das Verhältnis des minimalen zum maximalen Durchmesser (s. Formel 4.7 und 4.8) beschreibt die Form des Tokens:

$$M_{Form1} = \frac{M_{MinDiameter}}{M_{MaxDiameter}} \quad (4.9)$$

Ebenso leistet das Verhältnis der Oberfläche zum Objektvolumen einen Beitrag zur Formbeschreibung. Die Oberfläche wird im Diskreten durch die Anzahl der freien Voxelflächen in den drei Raumebenen (O_{xy} , O_{xz} , O_{yz} , s. Abb. 4.6) in der Einheit mm^2 bestimmt. Das Objektvolumen wurde bereits durch das Merkmal M_{Volume} berechnet.

$$M_{Form2} = \frac{|O_{xy}| \cdot S_x \cdot S_y + |O_{xz}| \cdot S_x \cdot S_z + |O_{yz}| \cdot S_y \cdot S_z}{M_{Volume}} \quad (4.10)$$

Die Anzahl der Skelettlinienvoxel erhöht sich mit der geometrischen Komplexität der Tokenoberfläche und wird daher ins Verhältnis zu der Anzahl der Tokenvoxel gesetzt:

$$M_{\text{Begrenzung1}} = \frac{|\text{Skeleton}|}{|\text{Token}|} \quad (4.11)$$

Um die Skelettlinie weiter zu charakterisieren, wird der Durchschnitt der minimalen Randdistanzen SD_{\min} zu der maximalen Distanz ins Verhältnis gesetzt:

$$M_{\text{Form3}} = \frac{1}{|\text{Skeleton}| \cdot M_{\text{MinDiameter}}} \sum_{i=1}^{|\text{Skeleton}|} SD_{\min_i} \quad (4.12)$$

Eine weitere Möglichkeit der Formcharakterisierung ist die Aussage der *fraktalen Dimension*. Die Dimension D für geometrische Formen wie Punkte ($0D$), Linien ($1D$), Flächen ($2D$) oder Körper ($3D$) ist im Alltag gegenwärtig und lässt sich durch ganzzahlige Werte beschreiben [Man87]. Doch bereits 1890 gab es Diskussionen über die „flächenfüllende“ *Peano-Kurve*, die durch eine initiale Linie und unendlichem rekursivem Ersetzen der geraden Abschnitte durch eine Schleife letztendlich eine ganze Fläche zu füllen vermochte. Ist die Dimension einer so konstruierten Linie $1D$ oder $2D$?

Die Mathematik entscheidet dies durch die Berechnung fraktaler (gebrochener) Dimensionen und verlässt somit den ganzzahligen Wertebereich. Sie ist durch ihre Selbstähnlichkeit charakterisiert und lässt sich für beliebige im Raum angeordnete Punktmengen bestimmen. Hierzu findet der Boxcounting-Algorithmus auf das Token Anwendung [Han00], der ein $3D$ -Gitter über die Voxelmenge legt und die Anzahl der Überdeckungen (N_δ) mit den Gitterzellen (δ) zählt. Durch immer weiteres Verfeinern der Gitterzellen ($\delta \rightarrow 0$) und erneutes Bestimmen der Überdeckungen kann die Konvergenz zur Fraktalen Dimension D_k berechnet werden:

$$M_{\text{Form4}} = D_k = \lim_{\delta \rightarrow 0} \frac{\log N_\delta}{\log \delta} \quad (4.13)$$

4.4.2 Textur

Nachdem im vorherigen Kapitel Merkmale nur anhand der abstrakten Information der Intermediate Repräsentation berechnet wurden, wird zur Beschreibung der endotumoralen Charakteristik der Läsionen nun zusätzlich auf die Signalwerte der Low-Level Strukturen zugegriffen. Während für die Segmentierung die Signalintensitäten (SI) der Differenzbilder D_{SN_4} bis D_{SN_7} herangezogen wurden, berechnen sich die folgenden Merkmale auf den SI der registrierten Bilddaten R_{SN_4} bis R_{SN_8} (s. Abb. 4.1). Durch die Nachsegmentierung wurde dies bereits für jede Läsion bestimmt.

$$\text{Grauwert} = SI_{\mathfrak{S},p}, \text{ mit } \mathfrak{S} \in \{R_{SN_4}, \dots, R_{SN_8}\}, p = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathfrak{S} \wedge p \in \text{Token} \quad (4.14)$$

Auf der Voxelmenge eines Tokens ($Token$) werden der durchschnittliche Grauwert μ und die Varianz σ berechnet. Letztere wird vom Menschen als Kontrast wahrgenommen:

$$\begin{aligned}\mu &= \frac{1}{|Token|} \sum_{i=1}^{|Token|} SI_{\mathfrak{S},p_i} \\ M_{Kontrast} = \sigma &= \frac{1}{|Token|} \sum_{i=1}^{|Token|} (SI_{\mathfrak{S},p_i} - \mu)^2\end{aligned}\quad (4.15)$$

Bereiche im Inneren einer Läsion, die nicht ausreichend Kontrastmittel aufgenommen haben, erfüllen das Homogenitätskriterium nicht und sind nicht Teil des Segmentierungsergebnisses. Sie zeigen sich als Hohlräume ($Bubbles$) in der zusammenhängenden Voxelmenge eines Tokens. Die Hohlräume können mit Hilfe der Zusammenhangskomponenten [Han00] erkannt werden. Jeder einzelne zusammenhängende Raum bildet dabei ein neues Tokenobjekt ($Bubble_i, i = 1, \dots, |Bubbles|$). Die Anzahl der Bubbletoken entspricht der Anzahl an Hohlräumen. Die Gesamtgröße aller Hohlräume wird ins Verhältnis zur Größe des Tokens gesetzt:

$$M_{Nekrose} = \frac{\sum_{i=1}^{|Bubbles|} |Bubble_i|}{|Token|}\quad (4.16)$$

Im Vergleich zum vorherigen Merkmal wird folgend nur der größte Hohlraum betrachtet:

$$M_{Ringenhancement} = \frac{Max_{i=1}^{|Bubbles|} |Bubbles_i|}{|Token|}\quad (4.17)$$

Der Randbereich einer Läsion wird mit dem angrenzenden gesunden Gewebe verglichen. Während der Rand der Läsion durch das segmentierte Token bereits gegeben ist, muss der umliegende Bereich des gesunden Gewebes erst lokalisiert werden. Die Dilatation [LOPR97] als morphologische Operation auf dem Token (s. Kap. 2.4.3) erzeugt diesen Bereich, indem der Mittelpunkt eines 3D-Strukturelements in Form eine $3 \times 3 \times 3$ Fensters nacheinander auf allen Tokenvoxeln platziert wird. Alle Bereiche des Strukturelements, die ausserhalb des Tokens liegen, bilden eine ein Voxel breite Hülle ($Wrap$) um das Token. Der Randbereich ($Boundary$) des Tokens ergibt sich aus der Teilmenge der Tokenvoxel, die eine direkte Nachbarschaft zur Hülle haben. Der durchschnittliche Signalwert des Tokenrandes wird ins Verhältnis zu dem der Hülle gesetzt:

$$M_{Begrenzung2} = \frac{|Boundary| \sum_{i=1}^{|Wrap|} SI_{\mathfrak{S},p_i}}{|Wrap| \sum_{i=1}^{|Boundary|} SI_{\mathfrak{S},p_i}}\quad (4.18)$$

4.4.3 Kinetik

Während die bisherigen Merkmale die Morphologie einer Läsion anhand einer einzelnen Serie beschreiben, basieren die folgenden kinetischen Merkmale auf den Signalwerten der nativen (SN_2) und aller registrierten Serien ($R_{SN_4}, \dots, R_{SN_8}$, s. Abb. 4.1). Der Verlauf der Zeit/Signalintensitätskurve definiert sich durch die auf diesen sechs Stützstellen bestimmten Signalwerte. Jeder Stützstelle muss also jeweils ein Zeit- und Signalwert zugewiesen werden. Da sich die Akquisitionszeiten der einzelnen Sequenzen unterscheiden, sind die Abstände der Stützstellen unregelmäßig. Abb. 4.3 weist die Akquisitionszeiten aus. Diese Protokollinformationen sind aus dem DICOM-Tag ($0x0008, 0x0032$) hergeleitet (vgl. Anhang A). Der Zeitpunkt der Kontrastmittelgabe muss dagegen zuvor bekannt sein. Diese Information ist nicht im DICOM vermerkt.

Als Stützstellenzeitpunkte TP aller Nachkontrastmessungen wird die Mitte der Akquisitionsspanne gewählt. Die native Serie wird dagegen auf den Zeitpunkt der Kontrastmittelgabe festgelegt. Der Zeitpunkt null entspricht dem Beginn der Messung der ersten Nachkontrastaufnahme (s. Abb. 4.7, vgl. Abb. 4.3). Die Kinetik wird über ein Zeitintervall von 7:39 Minuten bestimmt, beginnend bei dem Zeitpunkt der KM-Gabe t_{min} bis zur Mitte der Messung der 8. Serie t_{max} .

$$\begin{aligned}
 t_{min} = t_{TP_1} &= t(SN_2) = -25s \\
 t_{TP_2} &= t(R_{SN_4}) = 36s \\
 t_{TP_3} &= t(R_{SN_5}) = 107s \\
 t_{TP_4} &= t(R_{SN_6}) = 178s \\
 t_{TP_5} &= t(R_{SN_7}) = 284s \\
 t_{max} = t_{TP_6} &= t(R_{SN_8}) = 434s
 \end{aligned} \tag{4.19}$$

Die Signalintensitäten der Stützstellen werden als Durchschnittswert über die durch eine *ROI* (*Region of Interest* = relevanter Bildbereich) definierte Voxelmenge berechnet (s. Formel 4.20). Es werden pro Läsion immer zwei alternative ROIs angewendet:

Token ROI Die ROI entspricht dem gesamten Token (*Token*).

Hotspot ROI Zur Bestimmung des Hotspots (vgl. Kap. 2.3.2) als eines Teilbereichs der Token ROI wird zuerst der Bereich der stärksten lokalen Anreicherung ermittelt. Er wird über den höchsten durchschnittlichen SI innerhalb eines definierten $3 \times 3 \times 3$ Fensters bestimmt. Auf dieser lokalisierten Position wird weiter ein kugelförmiges Fenster zentriert, das die Voxelteilmenge der Hotspot ROI aus der Schnittmenge zur Token ROI definiert. Die Größe der Kugel wird derart angepasst, dass die Schnittmenge mindestens 20 Voxel umfasst.

Wenn bei kleinen Token keine Schnittmenge von 20 Voxeln erreicht werden kann, wird das ganze Token als Hotspot angenommen. Die beiden ROI Varianten führen dann zu identischen Resultaten.

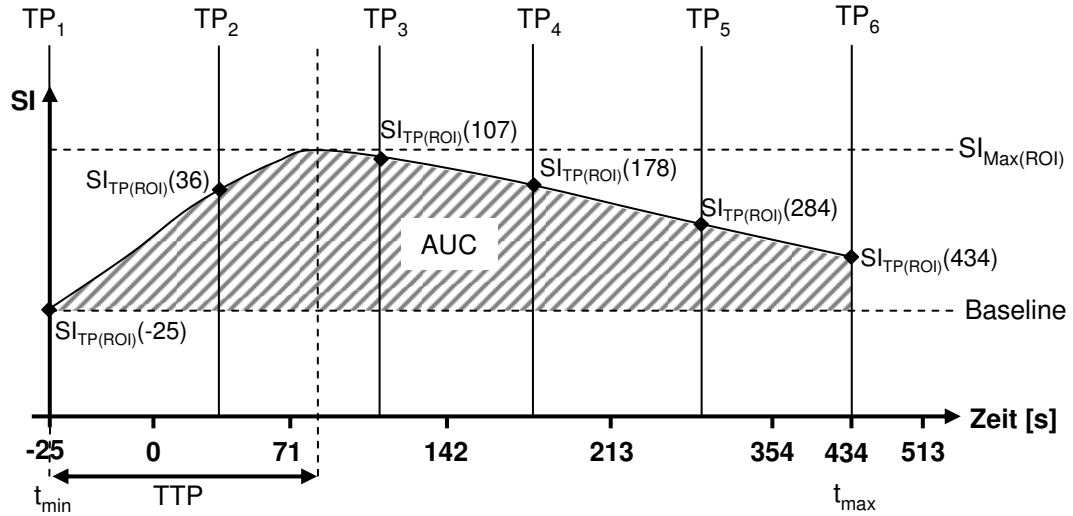


Abbildung 4.7: Interpretation der Kinetik durch Zeit- und Flächenmerkmale

Durch beide ROI Varianten – Hotspot und vollständiges Token – werden folgend immer zwei Signalintensitätskurven berechnet (s. Abb. 4.7). Dazu wird über alle Voxel der ROI der durchschnittliche SI pro Stützstelle berechnet:

$$SI_{TP(ROI)}(t_x) = \frac{1}{|ROI|} \sum_{i=1}^{|ROI|} SI_{\mathfrak{S}, p_i}, p_i \in ROI, ROI \in \{Token, Hotspot\}, t_x = t(\mathfrak{S}) \quad (4.20)$$

Die Signalwerte (SI_{Inter}) zwischen den Stützstellen werden durch das Verfahren der Polynominterpolation bestimmt [Sto04]. Mit Hilfe der Interpolationsfunktion f_P wird die Signalkurve im Sekundenabstand von t_{min} bis t_{max} insgesamt 459 Mal abgetastet:

$$SI_{Inter(ROI)}(t_x) = f_P(t_x), t_x \in \{t_{min}, \dots, t_{max}\}, ROI \in \{Token, Hotspot\} \quad (4.21)$$

Der Vorteil der Polynominterpolation liegt neben der einfachen Berechnung in der exakten Näherung der vorgegebenen Funktionswerte:

$$SI_{Inter(ROI)}(t_x) = SI_{TP(ROI)}(t_x), t_x \in \{t_{TP_1}, t_{TP_2}, t_{TP_3}, t_{TP_4}, t_{TP_5}, t_{TP_6}\} \quad (4.22)$$

Folgend werden zur Charakterisierung Kinetikmerkmale verwendet, die häufig im Zusammenhang mit der Auswertung von Perfusionen (Durchströmung) im biologischen Gewebe Anwendung finden [KTMW⁺05]. Für das erste kinetische Merkmal wird durch die Interpolationsfunktion f_P für beide Kinetikkurven (Token und Hotspot) jeweils der maximale Intensitätswert $SI_{Max(ROI)}$ ermittelt. Jedoch wird nicht der Signalwert selbst, sondern die verstrichene Zeit ($TTP = \text{Time To Peak}$) nach Kontrastmittelgabe in der Einheit Sekunden bestimmt (s. Abb. 4.7):

$$M_{TTP(ROI)} = t_x - t_{min}, t_x \in \{t_{min}, \dots, t_{max}\} \wedge SI_{Inter(ROI)}(t_x) = SI_{Max(ROI)} \quad (4.23)$$

Eine weitere Charakterisierung des Kinetikverlaufs ergibt sich durch die Berechnung der durch die Intensitätskurve gebildeten Fläche ($AUC = \text{Area Under Curve}$). Unberücksichtigt bleibt dabei die Fläche unterhalb der *Baseline*, die durch den Signalwert der nativen Serie bestimmt wird. Die Fläche wird zudem zum maximalen Signalwert normiert (s. Abb. 4.7):

$$M_{AUC(ROI)} = \sum_{t_x=t_{min}}^{t_{max}} \frac{SI_{Inter(ROI)}(t_x) - SI_{Inter(ROI)}(t_{min})}{SI_{Max(ROI)} - SI_{Inter(ROI)}(t_{min})} \quad (4.24)$$

Wie beim vorherigen Merkmal wird die Fläche unter dem interpolierten Signalverlauf gebildet, die Abtastung findet aber nur bis zum Zeitpunkt der maximalen Intensität (s. Formel 4.23) statt. Das Zeitfenster kann sich gegenüber dem vorherigen Merkmal verkürzen:

$$M_{AUCTTP(ROI)} = \sum_{t_x=0}^{M_{TTP(ROI)}} \frac{SI_{Inter(ROI)}(t_x + t_{min}) - SI_{Inter(ROI)}(t_{min})}{SI_{Max(ROI)} - SI_{Inter(ROI)}(t_{min})} \quad (4.25)$$

Die postinitiale Phase wird ähnlich dem Fischer Merkmal (s. Formel 2.2) berechnet, jedoch wird nicht der Maximalwert im Intervall der ersten bis dritten Minute bestimmt, sondern der bereits beim TTP ausschlaggebende Zeitpunkt des Maximalwerts (s. Formel 4.23) verwendet. Als Beschreibung eines möglichen Wash-Out-Verhaltens ergibt sich:

$$M_{WashOut(ROI)} = \frac{SI_{Inter(ROI)}(t_{max}) - SI_{Max(ROI)}}{SI_{Max(ROI)}} \quad (4.26)$$

In den bisherigen Merkmalsberechnungen wurde der Absolutwert der maximalen Anreicherungen nicht berücksichtigt. Dieser wird folgend zum Bildhistogramm der Signalintensitäten der letzten Bildfolge normiert. Als Normierungswert dient das oberste Quartil ($SI_{Quartil}$) des Histogramms der Serie R_{SN_8} :

$$M_{MaxSI(ROI)} = \frac{SI_{Max(ROI)}}{SI_{Quartil}} \quad (4.27)$$

4.4.4 Sonstige

Das T_2 -gewichtete fettgesättigte Bildvolumen (SN_1) gibt Informationen über den Wassergehalt einer Läsion. Sein Modell-KOS unterscheidet sich jedoch von den T_1 -gewichteten Aufnahmen. Es muss vor dem Zugriff eine KOS-Transformation erfolgen (s. Anhang A, vgl. Kap. 4.2).

$$\begin{aligned} Tr_{SN_1} &= f_{T_2 \rightarrow T_1}(SN_1), \text{ mit} \\ SI_{Tr_{SN_1}, p} &= SI_{SN_1, p'}, p' = M_{T_2 \rightarrow T_1} \cdot p, \forall p \in Tr_{SN_1} \end{aligned} \quad (4.28)$$

Nun lässt sich für jedes Tokenvoxel der SI Wert auf der transformierten T_2 -gewichteten Aufnahme Tr_{SN_1} bestimmen. Da die Transformation das diskrete Voxelastraster verlässt, erfolgt wiederum eine trilineare Grauwertinterpolation (vgl. Kap. 4.2.1). Die Menge W der Wasser darstellenden Elemente definiert sich aus allen Tokenvoxeln, die auf der T_2 -gewichteten Bildserie einen Grauwert größer als der Schwellwert TH_{T_2} zeigen. Deren Anzahl wird ins Verhältnis zur Tokengröße gesetzt:

$$\begin{aligned} W &= \{x | x \in Token \wedge SI_{Tr_{SN_1},x} > TH_{T_2}\} \\ M_{T_2} &= \frac{|W|}{|Token|} \end{aligned} \quad (4.29)$$

Das Patientenalter M_{Alter} wird aus den DICOM-Tags gewonnen (vgl. Anhang A). Entweder ist es im Tag (0x0010, 0x0010) direkt angegeben oder lässt sich aus dem Geburtstag (0x0010, 0x0030) des Patienten und dem Untersuchungsdatum (0x0008, 0x0022) herleiten.

4.4.5 Übersicht

Insgesamt extrahiert das System 24 kinetische und morphologische Merkmale (s. Tab. 4.2) pro Läsion, mit denen sich überwachte Klassifikatoren für Dignitätsentscheidung antrainieren lassen (s. nächstes Kapitel). Die Merkmalswerte werden entsprechend der Serie der Nachsegmentierung in vier Listen geführt (M_{SN_4} - M_{SN_7} , s. Abb. 4.1).

4.5 Klassifikation

Nachdem die Läsionen auf der Low-Level-Ebene erkannt und einzeln ins Intermediate-Level übertragen wurden, konnten für jede Läsion durch verschiedene Bildverarbeitungsalgorithmen morphologische und kinetische Merkmale extrahiert werden (s. vorheriges Kapitel). Anhand des Merkmalsvektors soll das System in die Lage versetzt werden, selbstständig die Dignitätsentscheidung zu treffen. Damit führt das System nun auch die Funktionalität der höchsten Wissensrepräsentation aus. Als High-Level-Funktion kommen überwachte Klassifikatoren zum Generalisieren der Problemstellung zum Einsatz (s. Kap. 2.4.4).

Im nächsten Kapitel wird auf die Bereitstellung des Goldstandards und anschließend auf die Erzeugung der Lerndaten aus dem Goldstandard (s. Kap. 4.5.2) für das Training der Klassifikatoren (s. Kap. 4.5.4) eingegangen. Zuvor erläutert Kapitel 4.5.3 die Grundlagen des verwendeten Klassifikators.

4.5.1 Goldstandard

Durch den Goldstandard (GS) wird das Wissen über das Vorliegen gut- oder bösartiger Läsionen im Bild bereitgestellt, mit dem die Klassifikatoren antrainiert werden können. Der gesamte Prozess der Goldstandardfestlegung muss ein hohes Qualitätsmaß erfüllen,

Merkmal	Beschreibung	Formel
$M_{Volumen}$	Morphologie (Größe, Tokengröße)	4.6
$M_{MaxDiameter}$	Morphologie (Größe, BoundingBox)	4.7
$M_{MinDiameter}$	Morphologie (Größe, Skeleton)	4.8
M_{Form1}	Morphologie (Durchmesser Verhältnis)	4.9
M_{Form2}	Morphologie (Oberflächenvolumenverhältnis)	4.10
M_{Form3}	Morphologie (Skeleton)	4.12
M_{Form4}	Morphologie (Fraktale Dimension)	4.13
$M_{Begrenzung1}$	Morphologie (Randkontur, Skeleton)	4.11
$M_{Begrenzung2}$	Morphologie (Randschärfe)	4.18
$M_{Kontrast}$	Morphologie (Varianz)	4.15
$M_{Nekrose}$	Morphologie (Hohlräume)	4.16
$M_{Ringenhancement}$	Morphologie (Hohlräume)	4.17
$M_{TTPToken}$	Kinetik (Zeit)	4.23
$M_{TTPHotspot}$	Kinetik (Zeit)	4.23
$M_{AUCToken}$	Kinetik (Fläche)	4.24
$M_{AUCHHotspot}$	Kinetik (Fläche)	4.24
$M_{AUCTTPToken}$	Kinetik (Fläche)	4.25
$M_{AUCTTPHotspot}$	Kinetik (Fläche)	4.25
$M_{WashOutToken}$	Kinetik (Steigung)	4.26
$M_{WashOutHotspot}$	Kinetik (Steigung)	4.26
$M_{MaxSIToken}$	Kinetik (normierte SI)	4.27
$M_{MaxSIHotspot}$	Kinetik (normierte SI)	4.27
M_{T2}	Morphologie (T_2)	4.29
M_{Alter}	DICOM-Information	

Tabelle 4.2: Übersicht der berechneten kinetischen und morphologischen Merkmale

im Einzelnen bei der Akquisition des Bildmaterials (s. Kap. 4.1), der histologischen Untersuchung und ihrer Zuordnung zum Bildmaterial. Nur dadurch ist ein Lernen korrekter Sachverhalte gewährleistet.

Zur Festlegung der Dignitäten wurden nur histologische Berichte zugelassen, die die genaue Tumordiagnose einschließlich des Tumortyps wiedergaben und eindeutig der Anreicherung im Bild zuzuordnen waren. Alle Histologien wurden aus Biopsien oder Operationen gewonnen, die höchstens 3 Monate vor oder nach der Untersuchung durchgeführt wurden. In der Zwischenzeit durften keine größeren Eingriffe wie Chemotherapie oder andere Operationen erfolgt sein, die eine Veränderung der Brust verursacht haben könnten.

Bilddaten ohne Auffälligkeiten wurden nur dann als negativer (gesunder) Befund aufgenommen, wenn auch eine Wiederholungsuntersuchung, die nach mindestens 12 und höchstens 24 Monaten durchgeführt wurde, ohne Verdacht blieb. Nachsorgepatienten wurden nur dann ins Patientenkollektiv aufgenommen, wenn biopsierte Tumore vorlagen.

Der vorhandene histologische Bericht wurde durch einen Expertenradiologen (> 1500

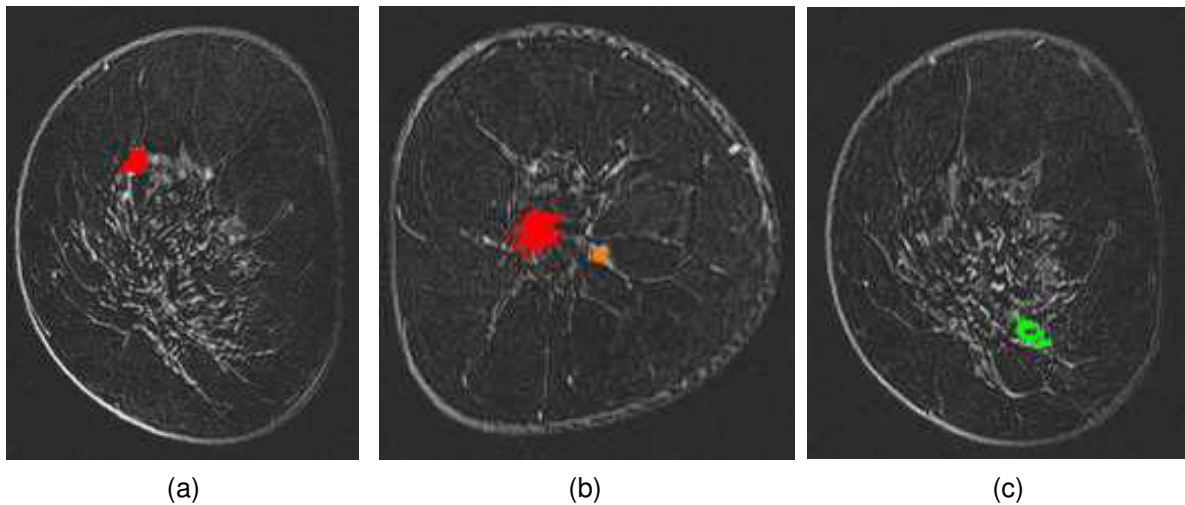


Abbildung 4.8: Beispiele von GS-Vorgaben: (a) maligner Tumor , (b) maligner (rot) und verdächtiger (orange) Tumor, (c) benigner Tumor, jeweils auf unregistrierten Bilddaten

Brust MRT-Befunden) dem Bildmaterial zugewiesen (s. Abb. 4.8 u. Abb. 6.19). Als Grundlage diente das Differenzvolumen der koronaren high-res Serie SN_7 und der nativen Serie SN_2 der unregistrierten Bilddaten. Mit Hilfe eines 3D-Zeichentools markierte der Radiologe alle Bildpunkte der untersuchten Anreicherung und wies jedem der manuell segmentierten Objekte die Dignität zu. Maligne Tumore erhielten die GS-Klasse 1, benigne die GS-Klasse 2 (s. Tab. 4.3). Darüber hinaus wurde der nähere Tumortyp (s. Tab. 5.2) für jede Markierungen festgelegt.

Um später Läsionen von unbedeutenden Segmentierungen unterscheiden zu können, markierte der Radiologe auch alle histologisch ungesicherten Läsionen. Lagen sie in der direkten Nachbarschaft zu histologisch gesicherten Objekten (z.B. bei multizentrischen Tumoren), wurden sie als verdächtige Läsionen bewertet und ihre Dignität von dem biopsierten Objekt übernommen (GS-Klasse 3 für verdächtig maligne oder GS-Klasse 4 für verdächtig benigne, s. Abb. 4.8 b u. Abb. 6.19 b). Konnte die Dignität nicht durch den histologischen Bericht oder aus der Umgebungsinformation übernommen werden, wurde mit der GS-Klasse 5 ein unsicherer Fund markiert. Dazu zählten auch alle anreichernden Lymphknoten. In etwa der Hälfte der Studien zeigten sich KM-Anreicherungen im Bereich der Brustwarzen, die gesondert mit der GS-Klasse 6 erfasst wurden. Die histologisch ungesicherten GS-Klassen 3 bis 6 (s. Tab. 4.3) wurden jedoch nicht für die Trainingsphase (s. Kap. 4.5.4) der Klassifikatoren verwendet.

Bildmaterialien mit starken Bewegungsartefakten wurden uneingeschränkt zugelassen, da vor der Klassifizierung eine Bewegungskorrektur (s. Kap. 4.2) durchgeführt wird.

Die gesamte manuelle Selektion durch den Expertenradiologen legt den Goldstandard fest, durch den die Lerndaten bestimmt (s. nächstes Kapitel), die Klassifikatoren (s. Kap. 4.5.5) trainiert und die Leistungsfähigkeit (s. Kap. 4.6.4) der Prototypsoftware abschließend beurteilt werden können.

Ursprung	GS	Klasse
histologisch gesichert	maligner Tumor	1
	benigner Tumor	2
histologisch ungesichert	maligner verdächtiger Tumor	3
	benigner verdächtiger Tumor	4
	unsicherer Tumor	5
	Brustwarze	6
Software	zusätzliche Funde / Artefakte	7

Tabelle 4.3: Goldstandard Klassen

4.5.2 Lerndaten

Um die Informationen der manuellen GS-Vorgaben für das nachfolgende Training erschließen zu können, müssen die Positionen der Tokenobjekte der automatischen Segmentierung mit den Markierungen des Expertenradiologen verglichen werden. Bei Übereinstimmung kann die GS-Klasse der Markierung übernommen und dem Vektor der 24 Merkmalen (s. Tab. 4.2) als Zielklasse angefügt werden. Die Zusammensetzung aus Merkmalsvektor und Zielklasse wird folgend als *Instanz* bezeichnet.

Während die automatische Segmentierung auf den registrierten Bilddaten ausgeführt wurde, markierte der Expertenradiologe den GS auf den unregistrierten Differenzvolumen. Dadurch bleibt die Möglichkeit erhalten, Klassifikationsergebnisse basierend auf verschiedenen Registrierungsalgorithmen miteinander vergleichen zu können. Eine manuelle Markierung auf den bereits registrierten Bilddaten hätte dagegen für jeden neuen Algorithmus wiederholt werden müssen. Vor dem Positionsvergleich muss der GS daher auf den Zeitpunkt der nativen Messung registriert werden. Dazu wird das bereits während des Registrierungsprozesses (s. Kap. 4.2.1) erzeugte Vektorfeld (s. Formel C.1) verwendet. Im Vergleich zum Transformationsschritt der Registrierung, der beim Verlassen des Voxelrasters die Grauwerte interpoliert, können die binären GS-Markierungen nur als Ganzes im Diskreten verschoben und der nächstgelegenen Rasterzelle zugeordnet werden.

Zur Beurteilung der Übereinstimmung von automatischer Segmentierung (SEG) mit dem Goldstandard (GS) werden die zwei *Überdeckungskriterien* O_{in} und O_{out} erhoben:

$$O(SEG, GS)_{in} = \frac{GS \cap SEG}{GS} \cdot 100\% \quad (4.30)$$

$$O(SEG, GS)_{out} = \frac{SEG \setminus GS}{SEG} \cdot 100\% \quad (4.31)$$

O_{in} gibt an, wie viel Prozent des Volumens der Segmentierung mit dem GS-Objekt überlappt. Der Volumenanteil des Tokens, das den Goldstandard überragt, wird durch O_{out} berechnet. Somit kann durch Anwendung von zwei Schwellwerten die Qualität der

Volumenübereinstimmung definiert werden. Formidentische Objekte erreichen mit $O_{in} = 100\%$ und $O_{out} = 0\%$ Idealwerte.

Da alle klinisch relevanten Anreicherungen vom Radiologen markiert wurden, bildet jede automatische Segmentierung ohne Überdeckung mit einer GS-Vorgabe die Menge der *zusätzlichen Funde*. Die Tabelle 4.3 wurde hierfür um die GS-Klasse 7 erweitert. Somit kann auch die Differenzierung unbedeutender Bildbereiche wie Artefakte von echten Anreicherungen trainiert werden. Die Anzahl an Objekten dieser Klasse wird durch die Qualität der Segmentierung bestimmt.

Für das anstehende Training werden drei Zielvariablen aus den GS-Klassen bestimmt und dem Merkmalsvektor angefügt: $Ziel_{Mal}$ ist die Menge der malignen Tumore (GS-Klasse 1), $Ziel_{Ben}$ umfasst alle benignen Tumore (GS-Klasse 2) und die Gruppe $Ziel_{Add}$ alle zusätzlichen Funde (GS-Klasse 7). Die Anreicherungen im Bereich der Brustwarzen (GS-Klasse 6) nehmen eine Sonderstellung ein und bleiben für das Training genauso unberücksichtigt wie die histologisch verdächtigen oder ungesicherten Läsionen (GS-Klasse 3 bis 5). Durch deren Existenz wurde jedoch sichergestellt, dass keine Läsionen fälschlicher Weise der Zielklasse $Ziel_{Add}$ zugeordnet wurden.

Zum Trainieren des Klassifikators (s. Kap. 4.5) werden nur die Instanzen derjenigen Token einbezogen, die entweder eine durch die Kriterien O_{in} und O_{out} angemessene Überdeckung mit den GS-Vorgaben der Klassen 1 und 2 besitzen ($Ziel_{Mal}$ und $Ziel_{Ben}$) oder gar keine Überdeckung ($O_{in} = 0\%$ und $O_{out} = 100\%$) mit einer GS-Klasse haben ($Ziel_{Add}$). Die Lerndaten umfassen also maximal so viele Instanzen wie es histologisch gesicherte Läsionen und durch die automatische Segmentierung erzeugte zusätzliche Funde gibt.

4.5.3 Neuronale Netze

Künstliche Neuronale Netze (NN) eignen sich als überwachte Klassifikatoren (s. Kap. 2.4.4) zum Trainieren komplexer Sachverhalte. Als Analogie und Inspiration der NN diene das menschliche Gehirn. Die ersten Entwicklungen der künstlichen Neuronennetze gehen auf die Arbeiten von Warren McCulloch und Walter Pitts aus dem Jahr 1943 zurück. Ein neuronales Netz besteht aus Knoten (*Neuronen*), die als Schalterfunktion Informationen aufnehmen und weiterleiten. Dazu sind sie mit anderen Knoten durch Kanten verbunden. Das Gewicht der Kante drückt die Stärke der Verbindung aus. Die Topologie Neuronaler Netze zur Lösung konkreter Anwendungsprobleme durch überwachtes Training ist für gewöhnlich in drei Schichttypen organisiert: einem Inputlayer, der jedes Merkmal des Inputvektors durch einen eigenen Knoten entgegennimmt, einem Outputlayer, der die Zielvariablen abbildet und einem oder mehreren dazwischenliegenden Hiddenlayer, wodurch das Lernen auch nicht linearer Zusammenhänge ermöglicht wird. Bei Feedforward Netzen empfängt ein Knoten nur die Ausgaben der Knoten der übergeordneten Schicht. Das Wissen eines Neuronalen Netzes ist in seinen Gewichten gespeichert, als Lernen kann man das Anpassen der Gewichte bezeichnen.

Neuronale Netze mit Hiddenlayern werden durch die Backpropagation Lernregel trainiert. Diese durchläuft drei Phasen. Zuerst werden dem Netz durch die vorgegebenen Inputvektoren Reize gesetzt und der Output des Netzes berechnet. In der Vergleichsphase

wird ein Fehlervektor zwischen der Ausgabe des Netzes und den vorgegebenen Werten der Zielklasse der Merkmalsvektoren bestimmt. Ist der Fehler zu hoch, wird als letzte Phase der innovative Kern des Backpropagation Verfahrens durchlaufen. Sie breitet den Fehlervektor in entgegengesetzter Richtung von dem Output- zum Inputlayer aus und modifiziert nach und nach die Gewichte mit dem Ziel, den Fehler zu minimieren. Die drei Phasen werden solange durchlaufen, bis die Vergleichsphase einen akzeptablen Fehler hervorbringt.

Künstliche Neuronale Netze eignen sich zur Diskriminierung von Merkmalsvektoren höherer Dimension, auch wenn die Beziehungen zwischen den Merkmalen nicht linear sind oder keine begründeten Hypothesen über die Art der Zusammenhänge aufgestellt werden können. Das Training Neuronaler Netze erfordert eine gewisse Fertigkeit. Es gibt zwar Empfehlungen für die Wahl der Netztopologie, letztendlich erfolgt sie jedoch experimentell. Durch eine ungeeignete Anzahl an Neuronen kann der Wirkungszusammenhang zwischen Ein- und Ausgabe auswendig gelernt (*Overfitting*, vgl. Kap. 4.6.4) oder unzureichend generalisiert werden (*Underfitting*) [BEPW05, HTF01].

In Kapitel 4.6.5 werden weitere alternative Klassifikatoren vorgestellt. Die erläuterten Vorgehensweisen für das Training im folgenden Kapitel sind für alle in der vorliegenden Arbeit untersuchten Klassifikatoren gültig.

4.5.4 Training

Das Training von Klassifikatoren versucht Regelmäßigkeiten bzw. Muster in den Instanzen der Lerndaten zu erkennen und diese zu generalisieren. Danach kann der Klassifikator für unbekannte Merkmalsvektoren die Zielklasse unter den Voraussetzungen vorhersagen, dass die Lerndaten möglichst sowohl die gleiche Verteilung hatten als auch alle vorkommenden Fälle abgedeckt haben wie die unbekanntenen Instanzen.

Die Leistungsfähigkeit eines Klassifikators hängt wesentlich von der Qualität der zur Verfügung stehenden Lerndaten ab, die Zuordnung der Merkmale zu den Zielklassen muss korrekt sein (s. Kap. 4.5.1). Darüber hinaus müssen noch weitere verfahrenstechnische Aspekte des Trainings beachtet werden.

Die Menge der Merkmalsvektoren spannt einen Raum auf, dessen Dimension der Anzahl an Merkmalen entspricht. Das Bestreben der Klassifikatoren ist es, die im Raum angeordneten Zielklassen geeignet zu unterteilen (vgl. Bäume in Kap. 4.6.5). Viele Klassifikatoren bestimmen zur Optimierung der Unterteilung ein Fehlermaß, beruhend auf der Länge der Vektoren oder der Abstände der Merkmale. Die Qualität der Fehlerbestimmung erhöht sich daher, wenn Werte der einzelnen Merkmale normiert und somit in einem vergleichbaren Wertebereich liegen. Mit der *Standardisierung* findet hier ein übliches Verfahren vor der Trainingsphase Anwendung, das den Wert eines Merkmals M zum Mittelwert μ und zur Standardabweichung σ aller gegebenen Merkmalswerte normiert (vgl. Formel 4.15).

$$M' = \frac{M - \mu_M}{\sigma_M} \quad (4.32)$$

Zudem muss sichergestellt werden, dass die Anzahl der Instanzen der zu unterscheidenden Zielklassen in etwa gleich groß ist, man spricht auch von der *Samplegröße*. Andernfalls

wird die Fehlerbestimmung überwiegend durch die Merkmale des größeren Samples geprägt. Manche Klassifikatoren berücksichtigen dann nicht ausreichend die Instanzen des kleineren Samples bei der Fehlerbestimmung.

Alternativ zur Anpassung können unterschiedliche Samplegrößen durch stärkere Gewichtungen einzelner Instanzen ausgeglichen werden. Zugleich kann hierdurch Einfluss auf die „Bedeutung“ einzelner Instanzen genommen werden (vgl. Boosting in Kap. 4.6.5)

Die Klassifikation der drei Zielvariablen ($Ziel_{Mal}$, $Ziel_{Ben}$ und $Ziel_{Add}$, s. vorheriges Kapitel) wurde auf zwei Klassifikatoren verteilt. Jeder für sich trifft nur eine binäre Aussage. Für den ersten Klassifikator werden die Instanzen von $Ziel_{Mal}$ und $Ziel_{Ben}$ zu einer gemeinsamen Zielklasse $Ziel_{Lesion}$ zusammengefasst. Er kann nach dem Training somit zwischen Läsionen und zusätzlichen Funden unterscheiden. Während die vorhergesagten zusätzlichen Funde für die weitere Klassifizierung keine Rolle mehr spielen, differenziert der zweite Klassifikator die verbleibenden Läsionen in gut- und bösartig.

Sogenannte multinomiale Klassifikatoren – wie beispielsweise Neuronale Netze – lassen auch mehrere Zielklassen zu. Dies trifft aber nicht auf alle in Kapitel 4.6.5 untersuchten Klassifikatoren zu.

4.5.5 Export

Die Klassifikatoren können nach dem einmaligen Training auf den extrahierten Merkmalsvektoren der Bilddaten aller zukünftig untersuchten Patientinnen angewendet werden. Durch Ausgabe der gelernten Zielklasse geben sie eine eigene Einschätzung über die Präsenz von zusätzlichen Funden, benignen oder malignen Läsionen.

Da die Prototypsoftware über keine GUI verfügt, werden sämtliche Ergebnisse wieder im DICOM-Standard (s. Anhang A) exportiert. Nach Abschluss der Registrierung werden sowohl die korrigierten Bilddaten ($R_{SN_3} - R_{SN_7}$, s. Abb. 4.1) als auch die verwendeten Differenzbilder ($D_{SN_4} - D_{SN_7}$) der Segmentierung persistiert. Nach Beendigung der Klassifikation werden die vorhergesagten Dignitäten der Objekte als farbige DICOM-Bilder exportiert. Der Bildinhalt setzt sich aus den Grauwerten des Differenzbildes D_{SN_7} mit einer voxelweisen roten Einfärbung aller malignen und grün für die benignen Objekte zusammen. Alle vorhergesagten zusätzlichen Funde bleiben dem Export vorenthalten.

Die exportierten Dateien können abschließend durch einen beliebigen DICOM konformen Viewer betrachtet werden.

4.6 Evaluierung

Der Diskussionsteil befasst sich ausgiebig mit der Bewertung der durch die Bildverarbeitung erzielten Ergebnisse. Folgend werden die dazu verwendeten Techniken vorgestellt.

In der Literatur werden zur Qualitätsbeurteilung von Bildverarbeitungsalgorithmen – wie beispielsweise der Registrierung – Phantome eingesetzt oder eine Bewertung der Verbesserung durch Experten vorgenommen [STCS+03, RHBK02, DSRR99].

Phantome wurden in der vorliegenden Arbeit prinzipiell während der Softwareentwicklung im Rahmen von Unittests [Ham04] eingesetzt. Damit wurde die allgemeine Korrektheit der implementierten Verfahren verifiziert und insbesondere die Registrierung, Segmentierung und Merkmalsextraktion anhand von variierten Phantombildern getestet. Da diese Untersuchungen im Ergebnis lediglich die Korrektheit der Algorithmen bestätigen, wird im Diskussionsenteil nicht näher auf die Untersuchung mit Phantomen eingegangen.

Eine Beurteilung der Ergebnisse durch Experten stellt eine subjektive Einschätzung dar. Diese Möglichkeit wird auch aufgrund des hohen Zeitaufwands für die Evaluierung der vielen Teilergebnisse aller Komponenten nicht erwogen. Lediglich zur Segmentierung wird durch Betrachtung der manuellen GS-Vorgabe eine Bewertung durch einen Expertenradiologen herangezogen.

Eine Evaluierung der zahlreichen Ergebnisse der einzelnen Softwarekomponenten wird in der vorliegenden Arbeit durch mathematische Verfahren möglichst automatisch vollzogen. Für die Registrierung wird durch Abstandsmaße die Ähnlichkeit zweier Bilder beschrieben (s. Kap. 4.6.1). Die automatische Segmentierung wird durch einen Vergleich mit der bereits gegebenen Goldstandardvorgabe des Expertenradiologen bewertet (s. Kap. 4.6.2) und die Aussagekraft und die Beziehung der 24 extrahierten Merkmale untereinander durch verschiedene statistische Verfahren der Multivariaten Analyse untersucht (s. Kap. 4.6.3). Abschließend wird in Kapitel 4.6.4 das Validierungsverfahren zur Beurteilung der Vorhersageleistung von Klassifikatoren erläutert. Weitere Klassifikatoren, die eine mögliche Alternative zu Neuronalen Netzen darstellen, werden in Kapitel 4.6.5 aufgezeigt.

4.6.1 Abstandsmaße

Um die Ähnlichkeit zweier Bildvolumen zu bewerten, finden Distanzmaße Anwendung. Sie drücken deren Unterschiede in einem einzigen Qualitätswert aus. Folgend werden zwei Abstandsmaße zur Qualitätsbeurteilung eingeführt.

SSD Die Summe der quadratischen Differenzen (*SSD*) ist ein Standardverfahren der Statistik. Es werden die quadrierten Abstände (Differenzen) der Grauwerte aller korrespondierenden Punktpaare zweier Bilder \mathfrak{S}_1 und \mathfrak{S}_2 summiert:

$$SSD_{\mathfrak{S}_1, \mathfrak{S}_2} = \frac{1}{|\mathfrak{S}_1|} \sum_{i=1}^{|\mathfrak{S}_1|} (SI_{\mathfrak{S}_1, p_i} - SI_{\mathfrak{S}_2, q_i})^2, \text{ mit } p_i \in \mathfrak{S}_1, q_i \in \mathfrak{S}_2 \quad (4.33)$$

Das gleiche Abstandmaß wird auch beim Tracking innerhalb des lokalen Fensters $WT_{x,y,z}$ bestimmt, um den Restfehler zu minimieren (s. Formel B.3). Eine Summe von Null ergibt sich nur für identische Bilder. Je höher der Wert des Maßes, desto mehr oder größere Abstände liegen vor.

MI Als weiteres beliebtes Verfahren für die Messung statistischer Unabhängigkeit von Signalen kommt die Mutual Information (*MI*) zum Einsatz [Zhu07]. Sie wird häufig in multi modalen Registrierungsalgorithmen zur Bestimmung des Restfehlers verwendet (s.

Kap. 3.3.1). Der Berechnung der MI liegen die Entropien (H) sowohl über den Grauwertistogrammen der Einzelbilder $H(\mathfrak{S}_1)$ und $H(\mathfrak{S}_2)$ als auch über einem $2D$ -Histogramm beider Bilder $H(\mathfrak{S}_1, \mathfrak{S}_2)$ zugrunde. Die Entropien werden aus den Häufigkeiten h aller vorkommenden Grauwerte $SI_{\mathfrak{S},i}$ in den Einzelbildern berechnet bzw. aus der Matrix aller Grauwertpaarungen $(SI_{\mathfrak{S}_1,i}, SI_{\mathfrak{S}_2,j})$ im Falle des $2D$ -Histogramms:

$$\begin{aligned} MI_{\mathfrak{S}_1, \mathfrak{S}_2} &= H(\mathfrak{S}_1) + H(\mathfrak{S}_2) - H(\mathfrak{S}_1, \mathfrak{S}_2), \text{ mit} \\ H(\mathfrak{S}) &= - \sum_{i=1} h(SI_{\mathfrak{S},i}) \log_2 h(SI_{\mathfrak{S},i}) \\ H(\mathfrak{S}_1, \mathfrak{S}_2) &= - \sum_{i,j=1} h(SI_{\mathfrak{S}_1,i}, SI_{\mathfrak{S}_2,j}) \log_2 h(SI_{\mathfrak{S}_1,i}, SI_{\mathfrak{S}_2,j}) \end{aligned} \quad (4.34)$$

Veranschaulicht misst die MI die Informationen, die beide Bilder teilen. Wenn \mathfrak{S}_1 und \mathfrak{S}_2 unabhängig sind, also \mathfrak{S}_1 keine Information über \mathfrak{S}_2 gibt und umgekehrt, nimmt die $MI_{\mathfrak{S}_1, \mathfrak{S}_2}$ den Wert Null an. Im anderen Extrem identischer Bilder entspricht die $MI_{\mathfrak{S}_1, \mathfrak{S}_2}$ jeweils den Entropiewerten der beiden Einzelhistogramme und erreicht den Maximalwert. Die Mutual Information wird daher prozentual zur Entropie bestimmt:

$$MI\%_{\mathfrak{S}_1, \mathfrak{S}_2} = \frac{MI_{\mathfrak{S}_1, \mathfrak{S}_2}}{H(\mathfrak{S}_1)} \cdot 100 \quad (4.35)$$

Je höher der Wert von $MI\%_{\mathfrak{S}_1, \mathfrak{S}_2}$, desto mehr Gemeinsamkeiten weisen die beiden Bilder \mathfrak{S}_1 und \mathfrak{S}_2 auf. Einen Wert von 100 % erreichen nur identische Bilder.

4.6.2 Goldstandardvergleich

Ziel der Evaluierung der Segmentierung ist die Feststellung, wie viele der GS-Läsionen im Bild erkannt und wie formgenau sie bestimmt wurden. Hierfür wird ein Vergleich zu dem durch den Expertenradiologen auf Voxelebene markierten histologisch gesicherten Goldstandard vorgenommen. Ein ähnlicher Vergleich erfolgte bereits bei der Erzeugung der Lerndaten in Kapitel (s. Kap. 4.5.2). Dort wurden der Überdeckungs- und Überhangsbereich der Segmentierung zum GS durch die beiden Überdeckungskriterien O_{in} und O_{out} berechnet (s. Formel 4.30 u. 4.31). Während bei der Lerndatenerzeugung die beiden Werte zur Feststellung der räumlichen Übereinstimmung mit dem GS durch Anwendung von zwei Schwellwerten benutzt wurden, werden nun die Kriterien selbst zur Beschreibung der Formgenauigkeit verwendet.

Wie bereits erläutert (s. Kap. 4.5.2), wurde die GS-Festlegung auf den unregistrierten Bilddaten vorgenommen, während die automatische Segmentierung auf die registrierten Volumen angewendet wurde. Daher muss der GS vor dem Vergleich auf den Zeitpunkt der nativen Serie registriert werden.

4.6.3 Multivariate Analyse

Das Erkennen von Mustern mit Hilfe statistischer mathematischer Verfahren auf Datenbeständen wird als *Data Mining* bezeichnet. Das zentrale Werkzeug im Data Mining Prozess stellt die *Multivariate Analyse* dar. Sie betrachtet Abhängigkeitsstrukturen und damit das Zusammenwirken mehrerer Merkmale oder Instanzen. Unterschieden werden die Methoden einerseits in *strukturentdeckende* Verfahren, die die Beziehungen der einzelnen Variablen aufdecken (s. Cluster- und Hauptkomponentenanalyse) und andererseits in *strukturprüfende* Verfahren, die die Kausalbeziehungen der Variablen prüfen (s. Regressionsanalyse). Folgend werden zusammen mit dem Streudiagramm vier Analyseverfahren vorgestellt [BEPW05].

Vor der Analyse werden die Merkmale entsprechend der Vorbereitung der Trainingsphase (s. Formel 4.32) normiert.

Clusteranalyse Allgemein unterteilt die Clusteranalyse Objekte in Gruppen (Cluster), so dass die Objekte innerhalb derselben Cluster möglichst große Ähnlichkeiten aufweisen, während sie zu anderen möglichst verschieden sind. Als Ähnlichkeitsmaß dienen die Abstände der Objekte, die im aufgespannten Vektorraum als Punktwolke vorliegen. Gruppieren werden kann nach den Instanzen oder nach den Merkmalen.

In der Praxis werden häufig agglomerative (anhäufende, Bottom-Up) hierarchische Verfahren angewendet. Dabei werden schrittweise einzelne Objekte zu Clustern zusammengefasst, zuerst diejenigen mit minimalen Distanzen, dann nach und nach Cluster mit größeren Abständen. Das dabei entstehende Dendrogramm lässt sich als Baumstruktur visualisieren (s. Abb. 6.11 u. 6.12). Die Blätter des Dendrogramms zeigen die untersuchten Merkmale als einzelne Cluster, die Knoten vereinigen alle ihre untergeordneten Cluster. Die Länge einer Kante zwischen den Knoten gibt einen visuellen Hinweis über den Grad der Ähnlichkeit, je kürzer die Kante ist, desto größer ist die Ähnlichkeit.

Hauptkomponentenanalyse Mit der Hauptkomponentenanalyse wird der Fragestellung nachgegangen, ob sich eine Vielzahl an Merkmalen, die einen bestimmten Sachverhalt darstellen, auf wenige Faktoren (die Hauptkomponenten) reduzieren lassen. Während bei der Clusteranalyse der Merkmalsraum anhand von Abständen unterteilt wird, stellt die Varianz und damit die Streuung der Merkmale das zentrale Analysekriterium dar.

Die Punktwolke der untersuchten Merkmale (s. Tab. 4.2) bildet einen 24-dimensionalen Raum im kartesischen KOS. Es wird nun ein neues KOS derart in die Punktwolke platziert, dass die Varianz der Merkmale entlang der ersten Achse maximal wird. Durch Rotation wird die zweite Achse des KOS, die senkrecht auf der ersten steht, so ausgerichtet, dass die Varianz entlang ihrer Richtung am zweitgrößten wird usw. Dies wird fortgesetzt, bis alle 24 Achsen ausgerichtet sind. Die Gesamtvarianz der Merkmale ist die Summe der „Achsenvarianzen“.

Wenn bereits die ersten Achsen den größten Anteil der Gesamtvarianz abdecken, erscheinen die durch die Achsen dargestellten Faktoren ausreichend für den Gesamtinformationsgehalt aller Merkmale. Jeder Faktor kann aus einer Linearkombination der Merkmale zusammengesetzt werden, er ist häufig inhaltlich nicht mehr interpretierbar.

Als Ergebnis der Hauptkomponentenanalyse werden die Faktoren (s. Abb. 6.13 u. 6.14) und die Beziehung der Merkmale zu ihnen visualisiert, 2D- oder 3D-Graphen beschränken die Darstellung auf die ersten beiden oder drei Faktoren. Je größer die Entfernung eines Merkmals zu einer Achse ist, desto stärker ist seine Beziehung zu diesem Faktor. Merkmale an ähnlichen Positionen im Diagramm haben auch eine ähnliche Beziehung auf die dargestellten Faktoren.

Regressionsanalyse Die Regressionsanalyse bildet ein häufig eingesetztes und flexibles statistisches Verfahren der multivariaten Analyse. Während die Cluster- und Hauptkomponentenanalyse (s. oben) Strukturen der Merkmale aufdecken, kann die Regression durch Untersuchung von Kausalbeziehungen (Ursache – Wirkung) die Abhängigkeiten von Variablen überprüfen (strukturprüfendes Verfahren). Das Verfahren wird hier eingesetzt, um durch eine Merkmalsselektion eine Reduktion des Merkmalsraums zu erreichen. Die Leistung eines Klassifikators kann häufig durch eine Reduktion auf wenige Merkmale optimiert werden. Je höher die Dimension des Merkmalsvektors ist, desto mehr Instanzen werden auch benötigt, um den Merkmalsraum repräsentativ zu füllen. Ein Training mit nur den signifikanten Merkmalen führt durch niedrigere Raumdimensionen häufig zu verbesserten Vorhersagen.

Ziel der Regressionsanalyse ist es, die Beziehungen zwischen der abhängigen Zielklasse Y zu den unabhängigen Merkmalsvariablen (M) festzustellen. Es wird davon ausgegangen, dass die beobachteten Variablen mit einem zufälligen Fehler behaftet sind, so dass die Zusammenhänge geschätzt werden müssen. Da hier nur Klassifikatoren mit binären Aussagen eingesetzt werden, wird der Wert der Zielvariable mit 0 oder 1 kodiert. P ist die Wahrscheinlichkeit, mit der die abhängige Variable den Wert 1 annimmt $P(Y = 1)$. Zur Untersuchung der Einflüsse auf binäre Zielklassen findet im Folgenden die logistische Regression Anwendung, da für die lineare Regression besondere Voraussetzungen wie die Normalverteilung nicht gegeben sind. Jedem Merkmal M_x wird ein Koeffizient (b_x) zugewiesen. Die Koeffizienten der Modellfunktion werden mittels der Maximum-Likelihood-Methode geschätzt:

$$\ln \frac{P(Y = 1)}{P(Y = 0)} = \ln \frac{P(Y = 1)}{1 - P(Y = 1)} = b_0 + b_1 * M_1 + b_2 * M_2 + \dots + b_j * M_j \quad (4.36)$$

Durch den t -Test kann anhand der geschätzten Koeffizienten stochastisch untersucht werden [FKPT04], welchen Einfluss ein Merkmal auf die Zielklasse hat (s. Tab. 6.5). Je kleiner im Ergebnis die Überschreitungswahrscheinlichkeit des p -Wertes, desto signifikanter ist diese Beziehung.

Ausgehend von allen Merkmalen wird eine Teilmenge bestehend nur aus den Merkmalen mit den signifikanten Beziehungen gebildet. Die Teilmenge wird erneut analysiert, da ein zuvor signifikantes Merkmal in ihr wieder bedeutungslos werden kann. Die Selektion ist beendet, wenn die Analyse nur noch signifikante Merkmale in der Teilmenge hervorbringt.

Streudiagramm Durch ein Streudiagramm [Han02] können Abhängigkeiten statistischer Merkmale graphisch dargestellt werden (s. Abb. 6.16). Jedes Einzeldiagramm zeigt die Gegenüberstellung aller Wertepaare zweier Merkmale als Punktwolke im kartesischen Koordinatensystem. Das Diagramm erlaubt eine einfache optische Überprüfung, ob die Merkmalswertepaare durch eine gleichmäßige Streuung Unabhängigkeit oder unerwünschte Korrelationen zeigen. Die Merkmalswerte wurden wie auch bei Anwendung der Klassifikatoren standardisiert (s. Formel 4.32). Daher ist bei der Bewertung der Verteilung nur der Wertebereich -1 bis $+1$ der Diagrammachsen interessant.

4.6.4 Kreuzvalidierung

Die Leistungsfähigkeit von Klassifikatoren wird anhand der Vorhersagen auf unbekanntem Merkmalsvektoren bewertet. Dazu werden üblicherweise die gegebenen Lerndaten (s. Kap. 4.5.2) in eine Trainings- und Testmenge unterteilt. Erstere wird für das Training der Klassifikatoren benutzt, letztere dient nur zur Validierung der Vorhersage. Als Qualitätsmaß wird die Fehlerrate bestimmt, die sich aus der Anzahl der falsch vorhergesagten Klassen auf der Testmenge ergibt. Folgend wird das Prinzip von der Aufteilung der Trainings- und Testmengen zur Gewinnung der Fehlerrate erläutert.

Bei Gleichsetzung der Trainings- und Testmenge würde man Vorhersagen auf Instanzen treffen, die dem Trainingsprozess schon bekannt waren. Man erhielte nur eine Aussage darüber, wie gut der Klassifikator das Gelernte wiedergeben kann. Eine solche Fehlerrate gibt den *Resubstitutionsfehler* an, der einen sehr optimistischen Wert darstellt. Schließlich können Klassifikatoren ihre Vorgaben auswendig lernen und würden somit keinen Fehler produzieren. Aber wie gut wäre die Leistung auf unbekanntem Daten?

Soll die Fehlerrate für die Vorhersage auf unbekanntem Daten bestimmt werden, müssen die Instanzen der Testmenge unabhängig zur Trainingsmenge sein und dürfen keinesfalls zum Trainieren verwendet werden. Gleichzeitig müssen die enthaltenen Instanzen beider Mengen repräsentativ sein. Je mehr Trainingsdaten vorliegen, desto bessere Klassifikatoren können trainiert werden. Dagegen gibt eine umfangreichere Testmenge eine genauere Schätzung der Fehlerrate.

Auf Lerndaten, die wenig Instanzen umfassen, ist eine Kreuzvalidierung nach Leave-One-Out sinnvoll [HTF01]. Sie vermeidet überlappende Testmengen, indem nur eine Instanz aus den Lerndaten entnommen und für die Validierung zurückgelegt wird. Auf allen anderen Instanzen kann trainiert werden. Das Auslassen der jeweils nächsten Instanz und erneutes Training auf den veränderten Lerndaten wird so oft wiederholt wie es Instanzen in den Lerndaten gibt. Jede Instanz wird dadurch genau einmal der Testmenge zugeteilt. Der Vorteil dieses iterativen Verfahrens liegt in der optimalen Ausnutzung aller Lerndaten und der repräsentativen Testmenge, die in der Summe alle Instanzen enthält. Nachteilig ist der große Rechenaufwand, da während jeder Iteration ein neuer Klassifikator trainiert werden muss.

Das Iterieren über die einzelnen Instanzen wird folgend als *objektbasierte* Auswertung bezeichnet, also einer Aussage pro Läsion. In der objektbasierten Validierung werden nur die Objekte betrachtet, die durch die Software erfolgreich segmentiert wurden (s. Kap. 6.4.1

Kontingenztafel		Histologie	
		Klasse A	Klasse B
Vorhersage	Klasse A (positiv)	richtig-positiv (RP)	falsch-positiv (FP)
	Klasse B (negativ)	falsch-negativ (FN)	richtig-negativ (RN)

Tabelle 4.4: Kontingenztafel

u. 6.4.3). Sie wird nun um die Auflösung der Abhängigkeiten von den zu einer Patientin zugehörigen Instanzen erweitert. Das Iterieren erfolgt dazu über die Anzahl an Untersuchungen. Der Testmenge werden jeweils *alle* Instanzen der ausgelassenen Patientin zugeordnet. Da eine Patientin von mehreren Tumoren befallen sein kann, wird dadurch sichergestellt, dass beispielsweise zwei benachbarte Karzinome – mit wahrscheinlich abhängigen Morphologien und Kinetik – nicht in Trainings- und Testmenge aufgeteilt werden. Somit lassen sich als Alternative zur objektbasierten Auswertung nun *patientenbasierte* Auswertungen durchführen. Die Lerndaten enthalten jedoch aufgrund der Überdeckungskriterien nicht die Instanzen aller segmentierten Objekte, da Läsionen mit unzureichender Übereinstimmung mit dem GS dem Training vorenthalten werden. Um während der Auswertung eine Aussage für die Instanzen aller Segmentierungen leisten zu können, werden zusätzlich Testdaten durch Aushebeln der Überdeckungskriterien ($O_{in} \geq 0,0$ und $O_{out} \leq 100,00$) erzeugt. Anstatt den Fehler pro Patient nur auf den Instanzen der Lerndaten zu bestimmen, werden alle Instanzen dieser Patientin aus den Testdaten herangezogen. Dies ist möglich, da die Lerndaten eine Teilmenge der Testdaten darstellen. Die patientenbasierte Auswertung betrachtet die Statistik aus der Sicht des GS, unzureichende oder fehlende Segmentierungen werden als Fehler betrachtet und zusätzliche Funde werden als solche erkannt und bewertet (s. Kap. 5.4 u. 6.4.4).

Der Fehler pro Iteration wird auf allen Instanzen der Testmenge bestimmt. GS-Objekte, die aufgrund gescheiterter Segmentierung nicht in den Testdaten enthalten sind, werden bei der objektbasierten Auswertung dementsprechend nicht in der Statistik berücksichtigt.

Die Gesamtfehlerrate setzt sich aus der Summe der Einzelfehler aller Iterationen zusammen; sie kann anhand der Kontingenztafel eingesehen werden (s. Tab. 4.4). Stimmt eine Vorhersage mit der Zielklasse überein, wird sie als *Richtig* gewertet ($RP = \text{richtig-positiv}$ oder $RN = \text{richtig-negativ}$). Bei Diskrepanz wird sie dagegen als *Falsch* vermerkt ($FP = \text{falsch-positiv}$ oder $FN = \text{falsch-negativ}$). Die Begriffe *Positiv* (Klasse A) und *Negativ* (Klasse B) definieren sich je nach untersuchtem Sachverhalt, beispielsweise kann auf *Bösartigkeit* (=Positiv) und *Gutartigkeit* (=Negativ) getestet werden oder *Gesund* (=Positiv) und *Krank* (=Negativ).

Die Trefferrate (*Sensitivität*) und Richtignegativrate (*Spezifität*) werden anhand der Kontingenztafel berechnet. Zudem gibt die *Treffgenauigkeit* (engl.: *Accuracy*) die Rate der richtig klassifizierten Objekte an.

$$\text{Sensitivität} = \frac{RP}{RP + FN} \quad (4.37)$$

$$\text{Spezifität} = \frac{RN}{RN + FP} \quad (4.38)$$

$$\text{Treffgenauigkeit} = \frac{RP + RN}{RP + FP + RN + FN} \quad (4.39)$$

Nach Abschluss der Kreuzvalidierung, die lediglich ein Verfahren zur Bestimmung der Vorhersageleistung darstellt, kann im Übrigen ein Training über alle Instanzen erfolgen, möchte man den möglichst besten Klassifikator zur zukünftigen Anwendung auf unbekannte Instanzen erzeugen. Besonders bei Validierung durch Leave-One-Out kann angenommen werden, dass dieser Klassifikator eine sehr ähnliche Vorhersageleistung zu den vielen einzelnen Klassifikatoren zeigen wird, die während den Iterationen erzeugt wurden.

4.6.5 Alternative Klassifikatoren

Dieses Kapitel beschreibt weitere im R-Tool [GI] verfügbare Klassifikatoren, die eine Alternative zu den Neuronale Netzen darstellen [HTF01, BEPW05, Han02, FKPT04]. Die Normierung der Daten wurde, wie in Kapitel 4.5.4 erläutert, für alle Klassifikatoren gleichsam ausgeführt. Auch wurden immer zwei Klassifikatoren mit jeweils binärer Ausgabe (Läsion vs. zusätzlicher Fund und Maligne vs. Benigne) erzeugt.

Als konkurrierende Klassifikatoren zu den Neuronale Netzen werden mehrere statistische Verfahren betrachtet. Aus dem bekannten Verfahren der logistischen Regression (s. Kap. 4.6.3) kann direkt ein Klassifikator (*LogReg*) erzeugt werden, da mit nur zwei Zielklassen ein Spezialfall der Regressionsanalyse betrachtet wurde. Wenn ein Merkmalsvektor durch die Modellfunktion, bestehend aus den zuvor geschätzten Koeffizienten, vorhergesagt wird, kann der Ergebniswert mit Hilfe eines Schwellwertes wieder in eine der beiden Zielklassen zugeteilt werden ($Zielklasse_1 \leq 0,5 < Zielklasse_2$). Im Gegensatz zur linearen Regression liefert die logistische Regression bei der binären Betrachtung nur Ergebniswerte im Bereich 0,0 bis 1,0.

Ein weiteres klassisches statistisches Verfahren ist die Diskriminanzanalyse, die Fischer [Fis36] erstmals im Jahr 1936 beschrieben hat. Die Diskriminanzanalyse sucht Grenzen zwischen den Zielklassen anhand der Merkmalsausprägungen der gegebenen Instanzen. Die Merkmale werden dabei als Zufallsvariablen interpretiert.

Je nach erfüllter Voraussetzung erzielen die Verfahren optimale Diskriminanz. Dazu müssen bei der linearen Diskriminanzanalyse (*LDA*) die Merkmale sowohl normalverteilt sein als auch gleiche Varianzen besitzen. Sind die Varianzen unterschiedlich, erzielt die quadratische Diskriminanzanalyse (*QDA*) bei gleicher Normalverteilung noch optimale Ergebnisse.

Hastie et al. [HTF01] erweiterten die klassischen Ansätze der Diskriminanzanalyse. Sie zeigten die Äquivalenz der *LDA* mit der linearen Regression unter entsprechender Kodierung der Zielklassenzugehörigkeit. Der Ansatz der flexiblen Diskriminanzanalyse (*FDA*)

ersetzt die lineare Regression durch nicht lineare Verfahren, so dass eine Unterteilung des Merkmalsraums durch nichtlineare Trennflächen möglich wird.

Support Vector Machines (*SVM*) unterteilen den Merkmalsraum anhand mehrdimensionaler Hyperebenen. Um eine nicht lineare Trennung der Instanzen zu erreichen, wird der gesamte Merkmalsraum in eine weitaus höhere Dimension überführt. Dort kann eine lineare Trennung vorgenommen werden, die nach der Rücktransformation in den Ausgangsraum zu nicht linearen Hyperflächen führt. Da einerseits das Hochtransformieren rechenintensiv und andererseits das Rücktransformieren zu extrem komplexen Hyperflächen führen kann, werden die Trennflächen durch gut handhabbare Kernfunktionen beschrieben. Mit *LinSVM* werden lineare und mit *RadialSVM* nicht lineare radialbasis Funktionen untersucht. Der besondere Vorteil der SVM bei der Unterteilung der Zielklassen liegt darin, dass um die Klassengrenzen ein möglichst breiter Bereich frei von Instanzen bleibt.

Das Verfahren des *k-nearest Neighbor* (*knn*) hält alle Instanzen der Lerndaten als Modell vor. Zur Vorhersage einer unbekanntem Instanz werden ihre *k* nächstgelegenen Nachbarn im Merkmalsraum bestimmt und die mehrheitlich beteiligte Zielklasse vorhergesagt.

Andere Verfahren unterteilen den Merkmalsraum mit Hilfe von Baumstrukturen. Während die zuvor vorgestellten Verfahren alle Merkmale im Raum gleichzeitig betrachten, werden die Merkmale zur Erstellung von Klassifikationsbäumen sequenziell verwendet. Dadurch müssen die Merkmale für die Trainingsphase nicht mehr wie bisher normalisiert werden. Ausgehend von der Wurzel stellen die Knoten des Baumes binäre Entscheidungen dar, je nach Erfüllung wird der linke oder rechte Ast zum nächsten Knoten verfolgt. Die Blätter des Baumes enthalten die Zielklassen.

Ein Klassifikationsbaum (*Tree*) wird nach dem Top-Down Prinzip generiert. In jedem Schritt wird beispielsweise basierend auf der Entropie das jeweilige Merkmal bestimmt, welches die beste Unterteilung (minimale Unreinheit) der Instanzen leistet. Mit den verbleibenden Merkmalen kann der Baum weiter verzweigt oder auch „zurückgeschnitten“ werden, bis ein auf der Reinheit basierendes Abbruchkriterium erreicht wird.

Random Forest [Bre01] basiert auf den Klassifikationsbäumen, erzeugt von ihnen aber eine Vielzahl verschiedener Bäume (eben einen ganzen „Wald“). Das Wachstum der einzelnen Bäume enthält Zufallsvarianten, beispielsweise bei der Auswahl von Merkmalsteilmengen. Zur Vorhersage klassifizieren alle Bäume des Waldes die unbekanntem Instanz, die mehrheitlich beteiligte Zielklasse bildet das Resultat.

Boosting ist ein allgemeines Verfahren, um die Rate der korrekten Vorhersagen eines Lernalgorithmus zu erhöhen. Es produziert aus einer Familie von „schwachen“ Klassifikatoren eine „stärkere“ Kombination. Ein neues Klassifikationsmitglied wird auf modifizierten Lerndaten trainiert, deren Veränderung sich aus der Vorhersageleistung des zuvor erstellten Klassifikators ergibt. Das Boosting ermittelt alle falsch vorhergesagten Instanzen und gewichtet sie für den nächsten Trainingsprozess stärker. Die Software *R* bietet basierend auf Bäumen ein adaptives (*Boost.Ada*) und verallgemeinertes Boosting (*Boost.Gbm*) an.

Sowohl Random Forest als auch das Boosting sind Metaklassifizierer, sie enthalten eine eigene Merkmalsselektion (IS = Input Selection, vgl. auch Regressionsanalyse in Kap. 4.6.3). Diese beiden Klassifikatoren werden unter den Namen *Boost.Ada IS* und *Random Forest IS* evaluiert.

Kapitel 5

Ergebnisse

Die Prototypsoftware wurde auf den Bilddaten von 101 ausschließlich weiblichen Patientinnen ausgeführt. Ihr Durchschnittsalter lag bei 53,9 Jahren, die jüngste Patientin war 12, die älteste 79 Jahre, der Median lag bei 55 Jahren. Jede Patientin wurde jeweils nur durch eine Untersuchung in das Kollektiv aufgenommen. Alle Messungen erfolgten unter Verwendung derselben Hardware, demselben Protokoll und einheitlicher Handhabung der KM-Dosierung (s. Kap. 4.1). Die durchschnittliche Laufzeit der Software bis zum Export der Ergebnisse lag bei etwa 33 Minuten.

Die manuelle Festlegung des Goldstandards auf dem Kollektiv durch einen Expertenradiologen (s. Kap. 4.5.1) umfasste 87 maligne und 54 benigne histologisch gesicherte Befunde. Eine einzige Untersuchung enthielt überhaupt keine histologisch gesicherte Veränderung. Bei weiteren 35 Patientinnen wurden nur gutartige Läsionen festgestellt. 65 Frauen waren von wenigstens einem malignen Tumor befallen. Von ihnen hatten 6 sowohl maligne als auch benigne Befunde. Der vollständige vom Radiologen markierte GS, der neben den malignen und benignen sowohl verdächtige und unsichere Tumore als auch Brustwarzen (GS-Klasse 1 bis 6) umfasst, kann der Tabelle 5.1 entnommen werden. Tabelle 5.2 führt im Detail dazu die Histologien (vgl. Tab. 2.2) der GS-Klassen 1 und 2 auf.

Die Größe der malignen Tumore lag im Bereich von 20 bis 57.910 mm^3 (Median 1197 mm^3), die gutartigen Tumore waren im Mittel kleiner (12 bis 221.734 mm^3 , Median 207 mm^3). An dem mit Abstand größten Tumor mit einem Durchmesser von knapp 9 cm war die jüngste Patientin erkrankt. Es handelte sich als Sonderform des gutartigen Fibroadenoms um einen seltenen *Phylloides tumor*.

5.1 Registrierung

Die Qualität der Registrierung (s. Kap. 4.2) lässt sich durch verschiedene Parameter beeinflussen. Tabelle 5.3 listet die finalen Einstellungen auf, sie wurden experimentell durch Auswertung der Abstandsmaße (s. Kap. 4.6.1) ermittelt.

Ursprung	GS	Klasse	Häufigkeit (n = 565)
histologisch gesichert	maligner Tumor	1	87
	benigner Tumor	2	54
histologisch ungesichert	maligner verdächtiger Tumor	3	109
	benigner verdächtiger Tumor	4	162
	unsicherer Tumor	5	57
	Brustwarze	6	96

Tabelle 5.1: manuelle Goldstandardfestlegung des Kollektivs

Dignität	Tumortyp	Häufigkeit (n = 141)
maligne	Duktales Karzinom	61
	Lobuläres Karzinom	15
	Medulläres Karzinom	2
	Muzinöses Karzinom	1
	DCIS	4
	andere	4
benigne	Fibroadenom	19
	Papillom	3
	Mastopathie	9
	andere	23

Tabelle 5.2: Tumortypen der histologisch gesicherten GS-Klassen 1 und 2

Parameter	Wert	Bedeutung
$WF_{x,y,z}$	5x5x5	Feature Suchfenster
$TH_{Feature}$	0,02	Relative Featurequalität
$WS_{x,y,z}$	7x7x7	Feature Abstandsfenster
$WT_{x,y,z}$	7x7x7	Feature Trackingfenster
TH_{Trans}	$\begin{pmatrix} 0,005 \\ 0,005 \\ 0,005 \end{pmatrix}$	Vektor der Trackinggenauigkeit
TH_{Iter}	15	maximale Trackingiteration
$G_{Distance}$	128	Punktabstand initiales Vektorfeldgitter

Tabelle 5.3: Registrierungsparameter

Durch die Parametereinstellungen wurden auf der nativen Aufnahme im Mittel knapp 17.000 Features bestimmt. Die Bildschicht in Abb. 5.1 a stellt sie als rote und grüne Punkte dar. Diese verteilen sich in der ganzen Brust, überwiegend im kontrastreicheren Drüsen- gewebe. Dagegen wurden im homogenen Bereich der Luft keine und im Fettgewebe kaum Features gefunden. Manche Features hinterlassen den Eindruck, außerhalb der Brust zu liegen. Erst durch die angrenzenden Nachbarschichten kann im Dreidimensionalen erkannt werden, dass sie sinnvolle Features der Brust sind. Durch das Beibehalten nur der lokal stärksten Features wird eine gute Gleichverteilung innerhalb der Brust erreicht. Der Übergang von der Brust zur umgebenden Luft stellt zwar den kontrastreichsten Bereich im Bild dar, dennoch ist die Qualität der Features im Verhältnis zum Brustinneren niedriger, da die minimalen Eigenwerte an charakteristischen Ecken höhere Werte als an flächigen Über- gängen annehmen. Im Mittel konnten etwa 65 % der Features erfolgreich bis auf die fünfte Nachkontrastserie getrackt werden, ihr Anteil ist in Abb. 5.1 a grün eingefärbt. Abb. 5.1 b stellt die erfolgreich getrackten Positionen der Features auf der Serie SN_7 dar. Aufgrund von Verschiebungen in z-Richtung sind viele der auf der nativen Messung bestimmten Fea- tures in die Nachbarschichten verfolgt worden. Das Tracking eines Features bedurfte in den erfolgreichen Fällen meistens nur 5 bis 10 Iterationen.

Die Qualität der Bildverbesserung kann auch ohne Evaluierung bereits visuell deutlich wahrgenommen werden (s. Abb. 5.2 u. Abb. 6.1), besonders im Übergang von Hautbereich der Brust zur Luft. Während die Subtraktionen nicht deckungsgleicher Überlagerungen, die in negativen Werten resultieren, im Bild dunkel dargestellt sind, erscheinen positive Werte hell und bergen die Gefahr, im nachfolgenden Arbeitsschritt als KM-Anreicherungen segmentiert zu werden. In Abb. 5.2 a der registrierten gegenüber der unregistrierten (b) Differenzen ist deutlich die Reduktion des Bewegungsartefakts am Rand der Brust im mittleren Bildbereich zu beobachten. Als einfaches Bewertungsmaß sind in Tabelle 5.4 die durchschnittliche Signalintensität der registrierten und unregistrierten Differenzvolu- men der nativen zu allen Nachkontrastserien zusammengestellt. Es zeigt sich eine Minde- rung der Signalintensität im registrierten gegenüber dem unregistrierten Differenzvolumen. Verdeutlicht wird dies auch in den Abbildungen Abb. 5.2 c und d durch Einfärben aller Bildpunkte mit einer Signalintensität größer als 80 SI. Diese Einfärbung entspricht einer einfachen Schwellwertsegmentierung (vgl. Kap. 4.3), die bereits visualisiert, wie wichtig die Registrierung als Vorverarbeitung ist. Während die Segmentierung des Tumors (zu se- hen in der Brust der rechten Bildhälfte) auf den unregistrierten Bilddaten nicht von dem bewegungskompromittierten Parenchym abzutrennen ist, erscheint die gleiche Anreicherung in den registrierten Bilddaten isoliert (vgl. Kap. 6.2.4). Auch würden unregistriert die Bewegungsartefakte mit positiven Differenzwerten im Übergang der Brust zur Luft als potentielle Läsionen segmentiert. Im Diskussionsteil (s. Kap. 6.1) wird durch umfangreiche Evaluierungen gezeigt, dass die Minderung der Signalintensitäten durch eine Reduktion der Bewegungsartefakte bei Beibehaltung des KMs erreicht wird. Kapitel 6.1.4 zeigt den Registrierungsalgorithmus in Anwendung auf andere Bilddaten.

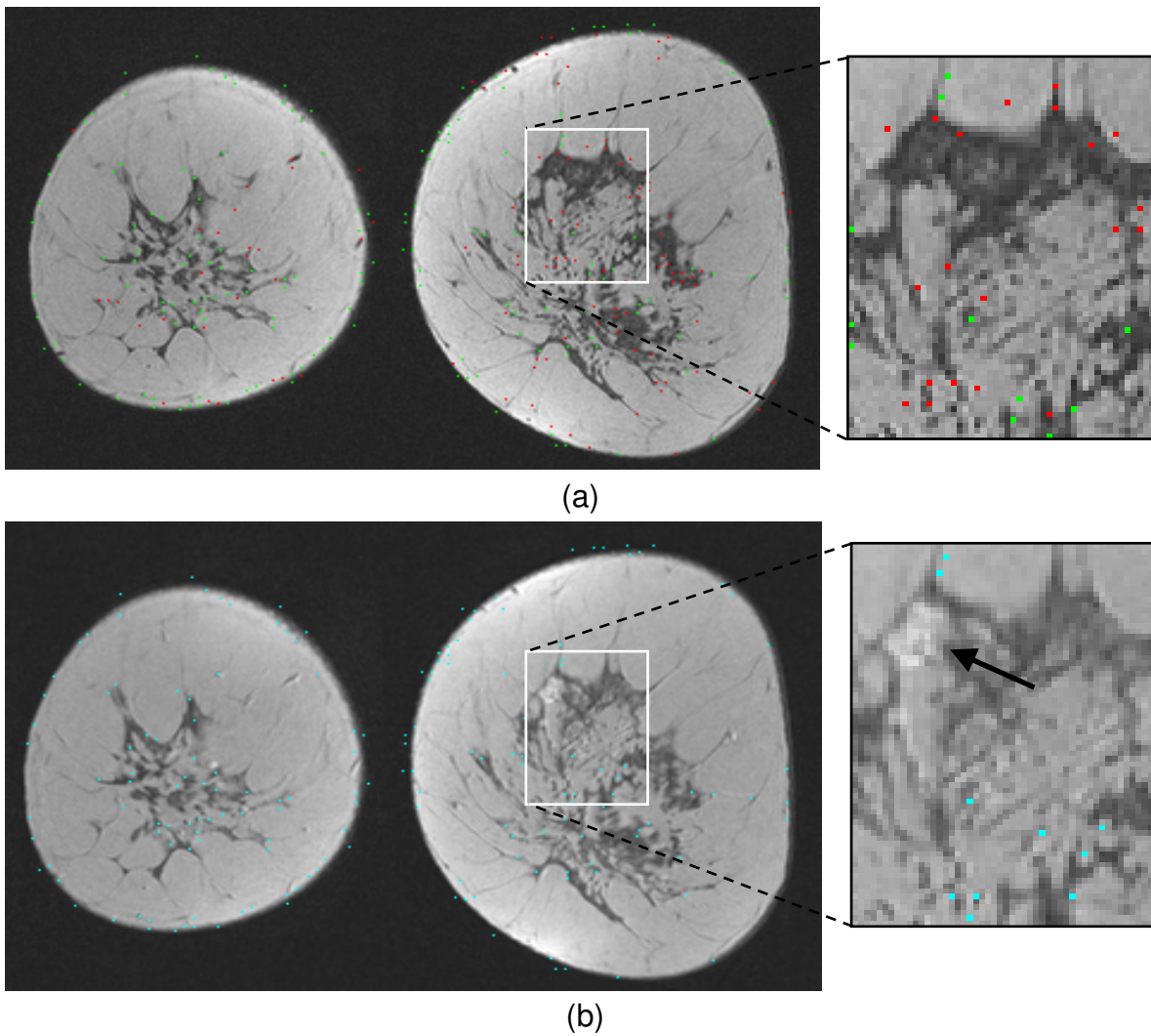


Abbildung 5.1: Visualisierung der (a) gesuchten und (b) getrackten Features mit Ausschnittsvergrößerung. Der Pfeil deutet die Lage des Tumors an.

unregistriert		registriert	
$SN_3 - SN_2$	4,27	$R_{SN_3} - SN_2$	3,08
$SN_4 - SN_2$	4,38	$R_{SN_4} - SN_2$	3,55
$SN_5 - SN_2$	5,46	$R_{SN_5} - SN_2$	4,42
$SN_6 - SN_2$	5,98	$R_{SN_6} - SN_2$	4,83
$SN_7 - SN_2$	7,30	$R_{SN_7} - SN_2$	5,86
$SN_8 - SN_2$	8,02	$R_{SN_8} - SN_2$	6,60

Tabelle 5.4: Vergleich der durchschnittlichen Signalintensität registrierter und unregistrierter Bildvolumen

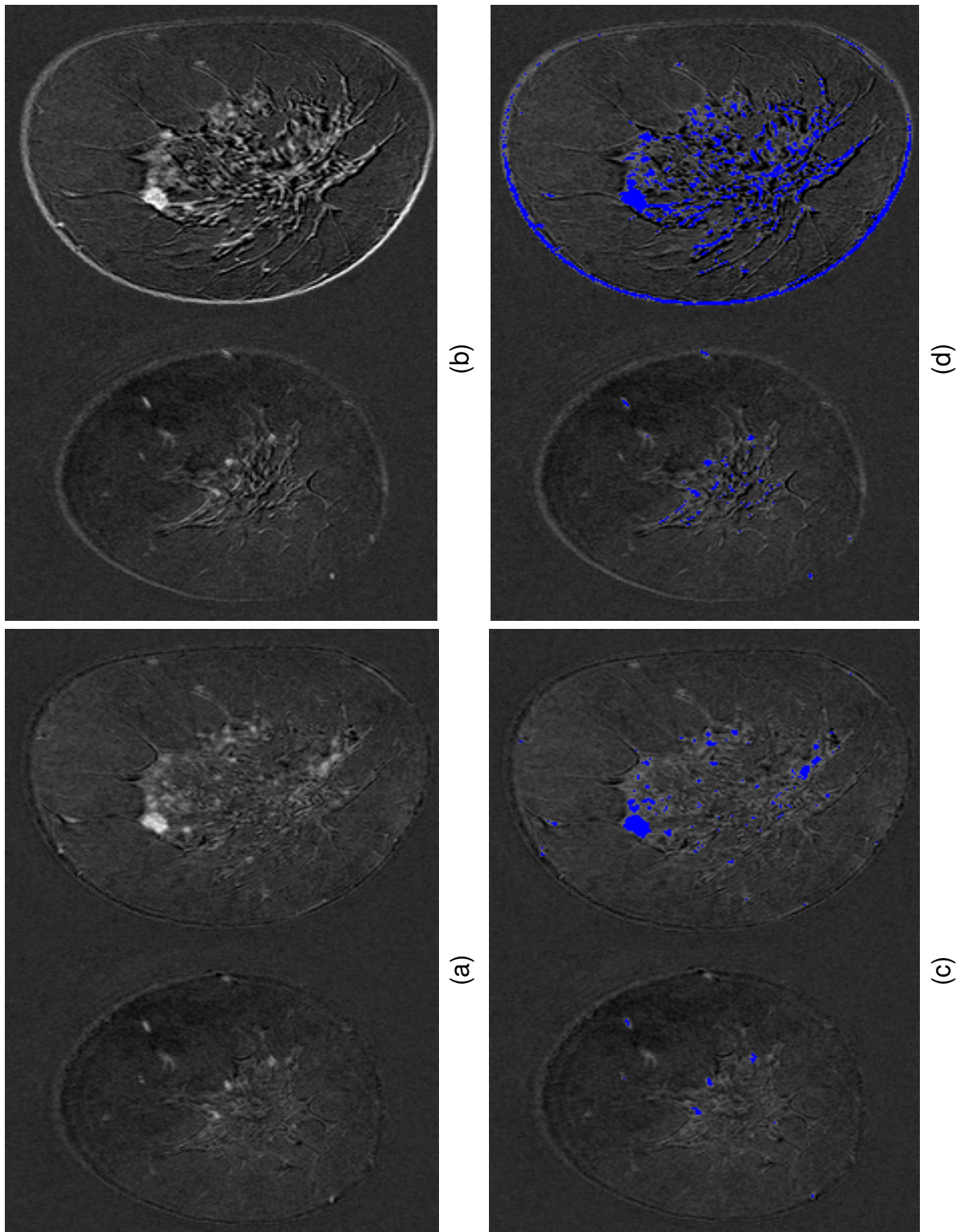


Abbildung 5.2: Unterschiede der (a, $R_{SN_7} - SN_2$) registrierten und (b, $SN_7 - SN_2$) unregistrierten Differenzbilder. Die Abbildungen c und d markieren alle Voxel mit SI größer als 80.

5.2 Segmentierung

Als mögliche Funktion f_{proj} (s. Formel 4.5 in Kap. 4.3.2) zum Erstellen des Projektionsvolumens wurde die Mittelwertbildung (f_{avr} , s. Abb. 5.3 a) und die Übernahme des maximalen Signalwerts (f_{max} , s. Abb. 5.3 c) untersucht.

$$\begin{aligned} V_{proj} &= f_{max}(D_{SN_4}, D_{SN_5}, D_{SN_6}, D_{SN_7}) \\ SI_{V_{proj},p} &= \max(SI_{D_{SN_4},p}, SI_{D_{SN_5},p}, SI_{D_{SN_6},p}, SI_{D_{SN_7},p}), \forall p \in V_{proj} \end{aligned} \quad (5.1)$$

$$\begin{aligned} V_{proj} &= f_{avr}(D_{SN_4}, D_{SN_5}, D_{SN_6}, D_{SN_7}) \\ SI_{V_{proj},p} &= \frac{SI_{D_{SN_4},p} + SI_{D_{SN_5},p} + SI_{D_{SN_6},p} + SI_{D_{SN_7},p}}{4}, \forall p \in V_{proj} \end{aligned} \quad (5.2)$$

Auf den erzeugten Volumen beider Projektionsfunktionen wurde die Segmentierung angewendet und dabei vornehmlich die Funktion für die individuelle Schwellwertberechnung $f_{Thres}(SI)$ aus dem Hotspot der Läsion und der übrigen Parameter variiert. Als erfolgreich wurde eine Segmentierung gewertet, wenn sie möglichst viele Läsionen des GS erkannte. Die Qualität der Übereinstimmung wurde anfänglich für ausgewählte Fälle visuell geprüft, später durch geeignete Definition der Überdeckungskriterien ($O_{in} > 10\%$ und $O_{out} < 50\%$, s. Kap. 4.5.2) automatisch über das gesamte Kollektiv bestimmt.

Auf der jeweils erfolgreichsten Segmentierung beider Projektionsvarianten wurde als Ausgangsbasis die Nachsegmentierung wiederum mit experimentell erarbeiteten Parametern untersucht. Als Bewertungsmaß wurde die Übereinstimmung mit dem GS durch $O_{in} > 15\%$ und $O_{out} < 50\%$ definiert. Die Anzahl der dadurch erkannten GS-Objekte und die resultierenden Schwellwertfunktionen $f_{Thres}(SI)$ sind in Tabelle 5.5 für beide Projektionsfunktionen aufgeführt. Als streng monoton fallende Projektionsfunktion war bereits die Festlegung linearer Beziehungen erfolgreich.

Durch die Maximumfunktion f_{max} wird sichergestellt, dass jede Anreicherung – unabhängig auf wie vielen Einzelserien erkennbar – im Projektionsvolumen vertreten ist. Dies hat Vorteile vor allem beim Erkennen von schwach anreichernden Läsionen. Die Segmentierung basierend auf der Maximumfunktion erkennt dagegen mehr gutartige Tumore. Nachteil der Maximumfunktion ist, dass Bewegungsartefakte und Bildrauschen stärker in das Projektionsvolumen eingehen. Dadurch ließen sich Läsionen teilweise nicht mehr vom angrenzenden Gewebe abgrenzen, drei Karzinome waren dadurch nicht mehr segmentierbar (s. Abb. 5.3 b und d, Markierung in a und c).

Die Mittelwertbildung unterdrückt dagegen das Vorkommen von Bewegungsartefakten und Bildrauschen, aber auch schwach anreicherndes Gewebe wie die vor allem nur auf den späten Bildfolgen erkennbaren gutartigen Läsionen. Die Intensitäten der im Projektionsvolumen abgebildeten Läsionen waren zwar niedriger als bei Verwendung der Maximumfunktion, dennoch blieben alle Karzinome segmentierbar. Um die erschwerte Abgrenzung auf dem Maximumvolumen zu leisten, musste der individuelle Schwellwert höher eingestellt

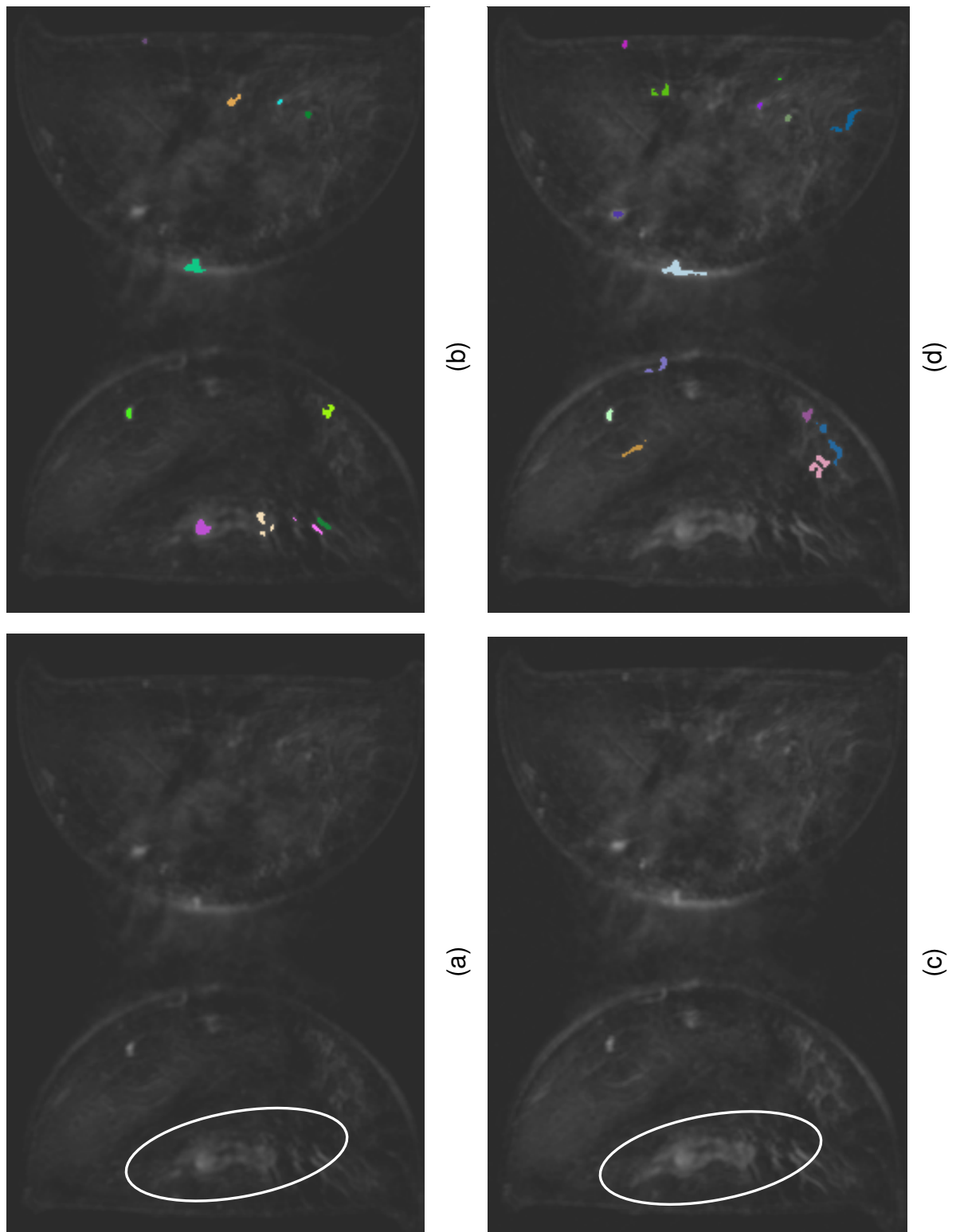


Abbildung 5.3: Projektionsvolumen durch (a und b) Mittelwertbildung und (c und d) Maximumfunktion mit entsprechender Segmentierung.

Projektion	$f_{Thres}(SI)$ V_{proj}	$f_{Thres}(SI)$ D_{SN_X}	Maligne (n = 87)	Benigne (n = 54)
f_{max}	$SI \cdot 0,60$	$SI \cdot 0,85$	84	49
f_{avr}	$SI \cdot 0,55$	$SI \cdot 0,85$	87	46

Tabelle 5.5: Vergleich der Schwellwertfunktion und erkannten GS der untersuchten Projektionsfunktionen

Parameter	Wert	Bedeutung
F_{Proj}	F_{avr}	Projektionsfunktion
TH_{Step}	3	Schrittweite der Iteration
TH_{Noise}	20	Schwellwert Bildrauschen
$f_{Thres}(SI)$	$0,55 \cdot SI$	individueller Schwellwert Projektionsvolumen
$f_{Thres}(SI)$	$0,85 \cdot SI$	individueller Schwellwert Differenzvolumen

Tabelle 5.6: Segmentierungsparameter

werden als beim Mittelwertvolumen. Die Schwellwertberechnung der Nachsegmentierung blieb dagegen identisch, da sie aus den Signalwerten der vorsegmentierten Läsion auf den Differenzvolumen berechnet wurde (s. Tab. 5.5). Das Kapitel 6.2.1 befasst sich ausgiebig mit der Anwendung der Schwellwerte.

Den Verlust von gutartigen Läsionen in Kauf nehmend, wurde folgend die Mittelwertfunktion F_{avr} verwendet. Die finalen Segmentierungsparameter sind in der Tabelle 5.6 zusammengefasst. Das Ergebnis der Segmentierung einer sowohl gut- als auch bösartigen Läsion ist in Abb. 5.4 und 5.5 dargestellt.

Das Projektionsvolumen einer Bildserie setzt sich aus etwa 10 Millionen Voxel zusammen. Im Mittel zeigen mit etwa 350.000 Voxel nur 3,5 % aller Bildpunkte eine Anreicherung von mehr als 20 SI (s. zyane und orange Einfärbung in Abb. 5.4 c und 5.5 c). Durch Befolgen der Segmentierungsheuristik auf dem Projektionsvolumen kann die Segmentierung aufgrund der Regeln und Überlagerungen teilweise eine Bewertung der Anreicherungen vornehmen und stärkere Ansammlungen hoher Signalintensitäten von den angrenzenden schwächeren unterscheiden (s. orange Einfärbung in Abb. 5.4 c und 5.5 c). Dadurch werden für die weitere Verarbeitung viele Bildbereiche ausgeschlossen, lediglich 6 % der angereicherten Voxel wurden als 3D-Tokenobjekte in die Intermediate Repräsentation überführt (s. Abb. 5.4 d und 5.5 d). Somit werden in deren abstrakten Darstellung in der Summe mit ca. 21.000 Voxel nur noch 0,21 % der ursprünglichen Bildinformation für die nachfolgende Bildverarbeitung verwaltet. Im Mittel werden pro Patient ca. 134 Token segmentiert. Die Menge des vorgegebenen GS (Klassen 1 bis 6, s. Tab 5.1) beläuft sich dagegen auf nur etwa 5,6 Objekte pro Patient. Die überwiegende Mehrzahl der automatisch segmentierten Token stellt demnach zusätzliche Funde dar. Sie zeichnen sich gegenüber den Läsionen durch eine

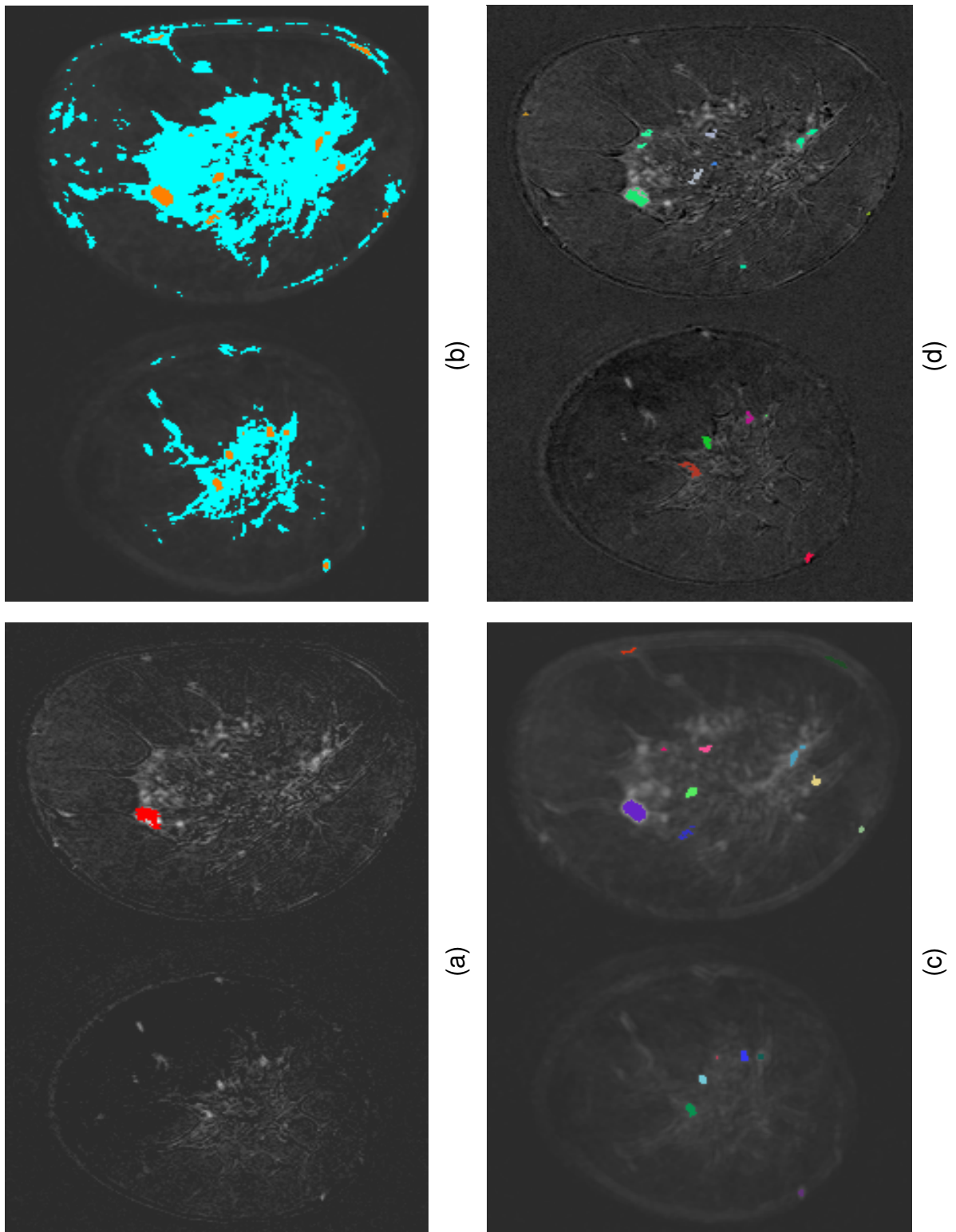


Abbildung 5.4: Segmentierung einer malignen Läsion: (a) registrierte Goldstandardvorgabe (b) Abgrenzung der Token (orange) gegenüber allen anreichernden Strukturen (zyan) (c) Segmentierung auf dem Projektionsvolumen (d) Nachsegmentierung einer einzelnen Serie

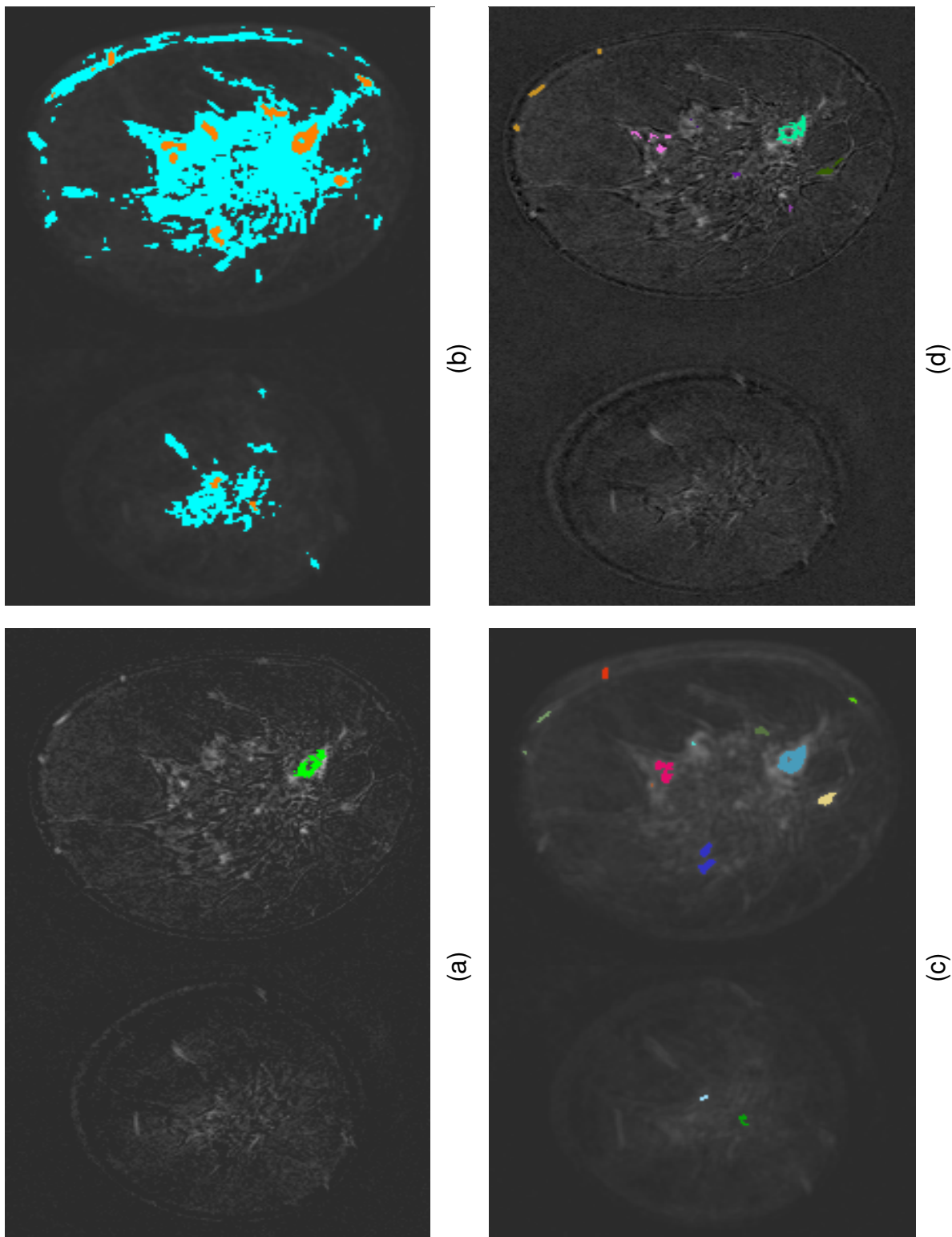


Abbildung 5.5: Segmentierung einer benignen Läsion: (a) registrierte Goldstandardvorgabe (b) Abgrenzung der Token (orange) gegenüber allen anreichernden Strukturen (zyan) (c) Segmentierung auf dem Projektionsvolumen (d) Nachsegmentierung einer einzelnen Serie

Parameter	Wert	Bedeutung
TH_{Wasser}	750	Schwellwert T_2 Segmentierung

Tabelle 5.7: Morphologieparameter

geringere Volumengröße aus. Dem Durchschnittsvolumen von 914 mm^3 aller GS-Objekten stehen nur 169 mm^3 der segmentierten zusätzlichen Funde gegenüber (s. a. Kap. 6.2.2). Nach visueller Prüfung wurden alle 87 histologisch gesicherten Karzinome durch die Segmentierung erkannt, dagegen konnten von den 54 benignen Tumoren 8 nicht segmentiert werden. Somit verfügten – neben der bekannten Patientin ohne histologisch gesicherten Tumor – drei weitere gesunde Patientinnen über keine segmentierten GS-Objekte der Klassen 1 und 2. In Kapitel 6.2.3 wird die Formgenauigkeit zu den GS-Vorgaben untersucht.

5.3 Merkmalsextraktion

Nach erfolgreicher Segmentierung der Token wurden 15537 Objekte morphologisch untersucht. Der einzig verbleibende Parameter zur Bestimmung des Wassergehalts auf der T_2 -gewichteten Aufnahme wurde auf 750 SI festgelegt (s. Tab. 5.7). Er eignete sich zum Segmentieren vorhandener Gefäße und Zysten.

Für eine zeitliche Optimierung erfolgten keine Berechnungen für klinisch unbedeutende Token, deren Volumen geringer als 10 mm^3 (= 13 Voxel) waren. Durch diese einfache Größenfilterung wurden 56 % der Objekte bereits vorklassifiziert und als zusätzlicher Fund von der finalen Klassifikation (s. nächstes Kapitel) ausgeschlossen. Diese Maßnahme entfernte keine Token, die einem histologisch gesichertem GS-Objekt entsprachen. Für die verbliebenen Objekte wurden jeweils die 24 Merkmale berechnet und bilden die Grundlage für die sich anschließende Klassifikation.

Eine genaue Betrachtung des Zusammenwirkens der Merkmale und ihre Aussagekraft im Vergleich zum BI-RADSTM-Standard wird in Kapitel 6.3.1 behandelt. Kapitel 6.3.2 macht deutlich, wie die mathematisch motivierten Merkmale von dem Vorgehen des Radiologen abweichen. Weitere Diskussionspunkte sind die Robustheit der Merkmale gegenüber der Bildqualität (s. Kap. 6.3.3) und die Ergebnisse der Multivariaten Analyse (s. Kap. 6.3.4 u. 6.3.5).

GS	Klasse	Häufigkeit
maligner Tumor	1	86
benigner Tumor	2	44
maligner verdächtiger Tumor	3	72
benigner verdächtiger Tumor	4	123
unsicherer Tumor	5	39
Brustwarzen	6	81
zusätzliche Funde ($> 10 \text{ mm}^3$)	7	8771

Tabelle 5.8: Häufigkeiten der GS-Klassen in den Lerndaten

5.4 Klassifikation

Beim Exportieren der Lerndaten wurden nur Objekte berücksichtigt, die die Überdeckungskriterien beim GS-Vergleich mit mindestens 15 % Überlappung und einem Überstandsanteil von höchstens 50 % erfüllten ($O_{in} > 15\%$ und $O_{out} < 50\%$, s. auch Kap. 6.2.3). Dadurch blieben zwar ein maligner und zwei benigne GS-Tumore der Lernmenge vorenthalten (s. Kap. 6.4.6), dafür wurde aber sichergestellt, dass nur Instanzen von Objekten mit hoher Ähnlichkeit zum GS trainiert wurden. Desweiteren wurden die Instanzen aller Objekte zu den Lerndaten hinzugefügt, die gar keine Überdeckung mit den GS der Klassen 1 bis 6 hatten. Sie beschreiben die Menge der segmentierten zusätzlichen Funde. Die Häufigkeit der verschiedenen GS-Klassen unter den Lerndaten kann Tabelle 5.8 entnommen werden. Nicht zu den Lerndaten hinzugefügt wurden Objekte, die eine Berührung mit den GS-Klassen 3 bis 6 hatten. Sie werden in Kapitel 6.4.3 gesondert diskutiert.

Die Daten enthielten Instanzen aller 101 Patientinnen, wenn nicht aus den GS-Klassen 1 oder 2 dann zumindest aus der GS-Klasse 7. Durch die gewählten Überdeckungskriterien werden keine der segmentierten zusätzlichen Funde fälschlicherweise in Übereinstimmung mit einem Goldstandardtumor gebracht, was die Qualität des Trainings verschlechtert hätte.

Durch Anwendung der Regressionsanalyse (s. Kap. 4.6.3) als eine Methode der multivariaten Analyse wurde vor dem Trainingsprozess eine Selektion der nur erfolgsversprechenden Merkmale durchgeführt (s. Kap. 6.4.2). Von den ursprünglich 24 Merkmalen fanden lediglich 16 Merkmale zur Erkennung der zusätzlichen Funde Anwendung (s. Tab. 6.6). Zur Diskriminierung maligner und benignen Tumore konnte gar eine Reduktion nur auf die 6 Merkmale M_{Form4} , $M_{Kontrast}$, $M_{TTPHotspot}$, $M_{WashOutToken}$, $M_{AUCHHotspot}$, M_{Alter} erfolgen (s. Tab. 6.8).

Als Netztyp wurden für beide Klassifikatoren feedforward backpropagation Netze mit einem Hidden Layer bestehend aus 6 Neuronen verwendet. Für das Training des ersten Netzes NN_{LA} , das Läsionen von zusätzlichen Funden differenziert, gingen mit der Zielklasse $Ziel_{Lesion}$ 130 Instanzen in Überdeckung mit den GS-Klassen 1 und 2 ein, während die zweite Zielklasse $Ziel_{Add}$ 8771 Instanzen der zusätzlichen Funde umfasste (s. Tab. 5.8, s.

Gesundheitsvorhersage der GS-Vorgabe durch NN		Histologie	
		Krank	Gesund
Vorhersage	Krank (positiv)	57	10
	Gesund (negativ)	8	26

Tabelle 5.9: Kontingenztafel der Vorhersage der Gesundheit durch GS-Objekte (Sensitivität = 88 %, Spezifität = 72 %, Treffgenauigkeit = 82 %)

Kap. 4.5.2). Um ein Überbewerten der zusätzlichen Funde zu verhindern, wurde die kleinere Zielklasse entsprechend stärker gewichtet (s. Kap. 4.5.4). Auch für das zweite Training des Klassifikators NN_{MB} mussten die 44 Instanzen der Zielklasse $Ziel_{Ben}$ gegenüber den 86 Instanzen der $Ziel_{Mal}$ stärker gewichtet werden.

Die Bewertung der Vorhersageleistung beider Klassifikatoren wurde mit Hilfe des Leave-One-Out Verfahrens patientenbasiert durchgeführt (s. Kap. 4.6.4). Die dazu notwendigen Testdaten für die Fehlervorhersage enthielt gegenüber den Lerndaten zusätzlich alle Instanzen mit Berührung zu den GS-Klasse 1 und GS-Klasse 2, auch wenn sie die Überdeckungskriterien (s.o.) nicht erfüllten. Anhand der Histologien wurde das Patientenkollektiv von 101 Patienten in krank und gesund unterteilt (vgl. Tab. 4.4):

- Als *Krank* (Histologie Klasse A) gelten 65 Patientinnen, die von mindestens einem histologisch gesichert malignen Tumor befallen sind.
- Als *Gesund* (Histologie Klasse B) gelten 36 Patientinnen, die keine oder nur histologisch gesicherte gutartige Tumoren aufweisen.

Für die Vorhersage pro Instanz wurden beiden Klassifikatoren NN_{LA} und NN_{MB} hintereinander ausgeführt. Folgende Definition bestimmt, ob die Vorhersage der Software pro Patient richtig war (vgl. Tab. 4.4):

- Eine *krankte Patientin* (Vorhersage Klasse A) wurde richtig vorhergesagt, wenn mindestens eines ihrer histologisch gesicherten Karzinome (GS-Klasse 1) durch die Software als solches erkannt wurde.
- Eine *gesunde Patientin* (Vorhersage Klasse B) wurde richtig vorhergesagt, wenn keiner ihrer histologisch benignen Tumore (GS-Klasse 2) fälschlicher Weise als maligne klassifiziert wurde.

Bei der Vorhersage der kranken Patientinnen wurden 57 der 65 Erkrankten als richtig positive erkannt (Sensitivität 88 %). Mit Gesund wurden 26 der 36 Patientinnen als richtig negativ vorhergesagt (Spezifität 72 %, s. Tab. 5.9).

Die Testdaten enthielten 8771 zusätzlichen Funde, von denen 92 % durch das Netz NN_{LA} erkannt und aussortiert wurden. Dadurch verblieben 742 Funde, die durch das nachfolgende Netz NN_{MB} in maligne (33 %) oder benigne (67 %) vorhergesagt wurden.

Gesundheitsvorhersage aller Objekte durch NN		Histologie	
		Krank	Gesund
Vorhersage	Krank (positiv)	64	28
	Gesund (negativ)	1	8

Tabelle 5.10: Kontingenztabelle der Vorhersage der Gesundheit durch alle Objekte (Sensitivität = 98 %, Spezifität = 29 %, Treffgenauigkeit = 71 %)

Damit verbleiben pro Patient zusätzlich zu den erkannten GS-Tumoren im Durchschnitt 7,3 falsch klassifizierte zusätzliche (*falsch-positive*) Funde.

Auch sortiert das erste Netz fälschlicherweise Läsionen als zusätzliche Funde (*falsch-negative*) aus. Unter den gut achttausend als zusätzliche Funde vorhergesagten Objekten befanden sich auch 2 Karzinome und 7 benigne Läsionen. Sie erreichten den zweiten Klassifikator nicht und gingen somit als falsche Vorhersagen in die Statistik ein (vgl. Tab. 4.4).

Die falsch vorhergesagten zusätzlichen Funde haben enorme Auswirkungen, wenn die obige Auswertung der Sensitivität und Spezifität unter Neudefinition der Vorhersage gesunder und kranker Patientinnen wiederholt wird (vgl. Tab. 4.4):

- Eine kranke Patientin (Vorhersage Klasse A) wurde richtig vorhergesagt, wenn mindestens ein Karzinom (egal ob tatsächlich Karzinom oder falsch klassifizierter zusätzlicher Fund) durch die Software erkannt wurde.
- Eine gesunde Patientin (Vorhersage Klasse B) wurde richtig vorhergesagt, wenn kein Karzinom durch die Software erkannt wurde.

Durch diese Bewertung wurden bis auf eine alle kranken Patientinnen als richtig positiv erkannt, aber lediglich 8 der 36 gesunden Patientinnen als richtig negativ (s. Tab. 5.10). Während nun eine sehr gute Sensitivität von 98 % erreicht wird, sinkt die Spezifität mit 29 % auf einen unakzeptablen Wert.

Die zusätzlichen Funde, die durch das erste Neuronale Netz nicht erkannt werden können, werden als Folge zwangsweise durch den zweiten Klassifikator entweder in gut- oder bösartig vorhergesagt. Auch wenn die Mehrheit dieser Objekte als gutartig klassifiziert wird und damit keine Auswirkung auf die obige Statistik hat, verbleiben dennoch durchschnittlich 2,4 „maligne“ zusätzliche Funde pro Patient.

Die Untersuchung des Einflusses einzelner Merkmale auf die Vorhersageleistung (s. Kap. 6.4.1), die Auswertung der histologisch ungesicherten GS-Objekte (s. Kap. 6.4.3) und die Vorhersageleistung alternativer Klassifikatoren (s. Kap. 6.4.4) sind Bestandteil des Diskussionskapitels.

Kapitel 6

Diskussion

Die Aufteilung des Diskussionskapitels richtet sich nach den implementierten Teilkomponenten der Software, jede wird in einer eigenen Sektion abgehandelt (s. Kap. 6.1 - 6.4). Zu Beginn jeder Sektion werden zuerst allgemeine Aussagen zur Begründung und Bedeutung der gewählten Algorithmen diskutiert und ein Vergleich mit der Literatur gezogen. Der nächste Sektionsabschnitt geht näher auf die Algorithmen der Komponente ein und hinterfragt kritisch die Stärken und Schwächen der Berechnungen. Besonders wichtig ist dabei die Bewertung der Qualität der erzielten Resultate. Sie werden entsprechend der Methoden des Kapitels 4.6 jeweils zum Ende jeder Sektion evaluiert. Dem Leser mit gemindertem Interesse an den mathematischen Verfahren sei das jeweils abschließende Kapitel mit den Erkenntnissen und Grenzen empfohlen, das auch die Ergebnisse der Evaluationen zusammenfasst. In den übergreifenden Diskussionskapiteln wird ein Vergleich zu den kommerziellen CAD-Systemen gezogen (s. Kap. 6.5) und die Vorteile und Nachteile des vorgestellten Prototyps in der praktischen Anwendung beleuchtet (s. Kap. 6.6). Als Ausblick fasst Kapitel 6.7 die nächsten Schritte für eine Vollendung als robuste Software zusammen.

6.1 Registrierung

Die Aufgabe der Registrierung liegt in der Eliminierung von Bewegungen in Bildfolgen. Ohne diesen Vorverarbeitungsschritt wäre die gesamte Prototypsoftware nur auf die Verwendung bewegungsfreier bzw. -armer Bilddaten eingeschränkt. Im klinischen Alltag entstehen häufig Untersuchungen mit mehr oder weniger starker Bewegung. Ohne eine Bildregistrierung wäre mit einer erhöhten Anzahl falsch positiver Befunde, mit dem Übersehen der eigentlichen Befunde aufgrund der Kaschierung durch die Bewegung und mit einer kompromittierten Darstellung der Morphologie und Kinetik der erkennbaren Anreicherungen zu rechnen.

Im stärkeren Maße, als dies dem Radiologen die Diagnostik erschwert, haben Bewegungen Einfluss auf alle nachfolgenden Bildverarbeitungsalgorithmen der Software. Es muss mit einer erhöhten Anzahl zusätzlicher Segmentierungen und dadurch mit erhöhten Laufzeiten gerechnet werden. Andere Systeme im wissenschaftlichen Umfeld, welche beispielsweise aus-

schließlich die Klassifikation ohne notwendige Vorverarbeitung untersuchen, sind letztlich als experimentelle Systeme zu betrachten, welche lediglich unter optimalen Bedingungen arbeiten können. Um dies anhand einer Teilkomponente stellvertretend zu verdeutlichen, werden in Kap. 6.2.4 die Ergebnisse der Segmentierung mit und ohne Anwendung der Registrierung verglichen.

Die Literatur (s. Kap. 3.3.1) wartet mit einer Vielzahl an Registrierungsverfahren auf. Speziell für die Brust-MRT wurden bereits Untersuchungen über die Eignung verschiedener Registrierungsmethoden angestellt. Denton et al. [DSRR99] ließen die Ergebnisse rigider, affiner und nicht rigider Registrierung anhand von Differenzbildern zur nativen Serie durch zwei Radiologen unabhängig voneinander bewerten. Während alle drei Registrierungsverfahren als Verbesserung der Bildqualität wahrgenommen wurden, konnte mit der nicht rigiden Registrierung die signifikant beste Qualitätssteigerung nachgewiesen werden. Rigid und affine Verfahren konnten die fließenden Bewegungen des weichen Brustgewebes nicht wiedergeben, ihr Unterschied war nicht signifikant.

Mit der Aufgabenstellung der Bewegungskorrektur T_1 -gewichteter Aufnahmen wurde ein *mono modales* 3D-Verfahren entwickelt (vgl. Kap. 3.3.1). Durch das Verfolgen lokaler Bildausschnitte basierend auf dem optischen Fluss wurde eine *nicht rigide* Lösung implementiert und damit die derzeit erfolgsversprechendste Registrierungsmethode für Brustbilddaten umgesetzt. Der Vorteil der Implementierung des modifizierten Kanade-Lucas-Tomasi Algorithmus (s. Anhang B) als Teilschritt liegt in seiner schnellen Laufzeit, das 2D-Originalverfahren wird für Echtzeitanwendungen eingesetzt.

In den vorliegenden Bilddaten konnte immer wieder eine Bewegung vermehrt in Richtung der z-Achse zum Lagerungstisch des MRT-Gerätes beobachtet werden (s. Abb. A.2). Demnach sinken die Mammae nach anfänglicher Anspannung der Patientin im Laufe der Untersuchung tiefer in die Brustspule ein. Gerade bei der vorliegenden koronaren Bildakquisition kann eine 2D-Registrierung in Anwendung auf die einzelnen Bildschichten diese Bewegungsrichtung gar nicht erkennen, da sie durch die Bildschichten verläuft. Ein 2D-Verfahren würde erst effektiv, wenn es in transversaler oder sagittaler Ebene eingesetzt würde. Generell können durch 2D-Verfahren die Bewegungen immer nur in einer Orientierung behandelt werden. Die Aufgabe der Bewegungskompensation lässt sich durch ein 3D-Verfahren sinnvoller lösen als die mehrfache Hintereinanderanwendung eines 2D-Algorithmus in den unterschiedlichen Orientierungen.

Zu Beginn des Forschungsprojekts war die Verwendung der T_2 -gewichteten Messung nicht angedacht. Mit der Wahl eines mono modalen Registrierungsansatzes kann die T_2 -gewichtete Serie SN_1 nun nicht mehr in den Registrierungsprozess einbezogen (s. Abb. 4.1) werden. Da sie zeitnah zur nativen Serie gemessen wird (s. Abb. 4.3), sind noch keine größeren Verschiebungen im Bild zu erwarten. Dennoch ist ihre Registrierung erstrebenswert, möchte man sicherstellen, dass beispielsweise die Bestimmung des Wassergehalts durch M_{T_2} bei den feinen Gefäßstrukturen nicht durch frühzeitige Verschiebungen im Bild kompromittiert wird. Es verbleibt daher noch weiterer Forschungs- und Entwicklungsaufwand für eine multi modale Registrierungsvariante.

Der Registrierungsalgorithmus setzt sich aus drei sequentiellen Teilalgorithmen zusammen (s. Abb. 4.4). Das Tracking erkennt die Bewegungen direkt aus dem Vergleich der

verschobenen Bildteilbereiche durch Minimieren eines Restfehlers (s. Anhang B). Die Eingabe des Trackingalgorithmus sind Bilddaten unbekannter und variationsstarker Inhalte und Bewegungen. Das mathematische Verfahren führt auf den Bildinhalten nicht zwingend zu eindeutigen Ergebnissen. Auch Bildinhalte unterschiedlichen Inhaltes können im ungünstigen Fall in Übereinstimmung gebracht werden. Denn ähnlich den Abstandsmaßen (s. Kap. 4.6.1) wird die Gleichheit zweier Bildteilausschnitte anhand eines einzelnen Qualitätswertes bestimmt.

Wurden die Translationen jedoch einmal (egal ob richtig oder falsch) ermittelt, stellen sie als Eingabe für die darauffolgende Interpolation und Grauwerttransformation bereits abstrahierte Informationen dar, die durch die zugrundeliegende Mathematik robust zu eindeutigen und nachvollziehbaren Ausgaben verarbeitet werden können. Die Qualität der gesamten Registrierung ist somit primär durch das Tracking geprägt. Das nächste Kapitel beschäftigt sich ausgiebig mit den technischen Aspekten der drei Sequenzen und gibt eine Abschätzung, inwieweit die Registrierung Bildinhalte verfälschen kann. Besonderes Augenmerk liegt auf der möglichen Veränderung der Läsionen, hierzu wird in Kapitel 6.1.3 zusätzlich noch die Volumenveränderung anreichender Strukturen untersucht.

Der Ablauf der drei sequentiellen Teilschritte ist darauf optimiert, mehrere Bildfolgen zu verarbeiten. So werden erkannte Bewegungen der vorherigen Bildfolge als Ausgangspunkt für das Tracking auf die nachfolgende verwendet. Im gleichen Zuge werden Featurepositionen, die nicht erfolgreich getrackt werden konnten, die entsprechende Translation aus dem Vektorfeld vorgegeben.

Der entwickelte Algorithmus ist nicht auf die Registrierung von Brust MRT-Bildern beschränkt. Der flexible Ansatz kann auch andere Bildinhalte oder Bilder anderer Modalitäten erfolgreich verarbeiten. Dazu ist lediglich eine Anpassung der Parameter (s. Tab. 5.3) notwendig. Untersuchungen auf weiterem Bildmaterial werden in Kapitel 6.1.4 beschrieben.

Die Bewertung der Bildverbesserung führt zu ambivalenten Aussagen. Beurteilungen durch den Radiologen resultieren zudem nur in einer subjektiven Aussage. Kapitel 6.1.2 zeigt zwei mathematische Verfahren zu Evaluation der gesteigerten Bildqualität.

6.1.1 Qualitätsabschätzung des Algorithmus

Die funktionalen Bilddaten unterliegen durch Bewegungen und KM-Anflutung gleich zweier Veränderungen. Während es die Bewegung zu eliminieren gilt, muss die Anreicherung unverändert erhalten bleiben.

Tracking Wie im vorherigen Kapitel bereits dargelegt, ist das Tracking als erster Teilschritt der drei sequentiell aufeinander abgestimmten Algorithmen grundlegend für die resultierende Qualität der Registrierung. Hier entscheidet sich anhand der erkannten Translationen, ob die später transformierten Bildinhalte unerwünschte Verfälschungen enthalten.

Nicht rigide Registrierungsverfahren stehen allgemein in der Kritik, die Kontrastmittel aufnehmenden Strukturen in ihrer Größe zu verändern [TSC⁺00]. Die Gefahr der Volumenveränderung ist dann gegeben, wenn der Algorithmus intensitätsbasiert arbeitet und somit voraussetzt, dass sich die Intensitätswerte der Bildpunkte zwischen Ausgangs- und Zielbild kaum verändern. Ist diese Bedingung beispielsweise durch den Anreicherungsprozess mit

KM nicht erfüllt, kann die Registrierung im ungünstigsten Fall diesen Prozess revidieren. Rigide und affine Registrierungen können ein solches Verhalten nicht hervorbringen, aber nicht rigide Verfahren erlauben es, die Abstandsverhältnisse von Strukturen zu verändern. Dadurch kann eine Anreicherung im Zielbild, die im Ausgangsbild noch nicht zu erkennen war, im ungünstigsten Fall für eine Erhöhung der Bildähnlichkeit in einen Punkt reduziert und durch Ausweiten des nicht anreichernden Nachbargewebes ersetzt werden.

Rücker et al. [RSH⁺99] stellen ein nicht rigides Registrierungsverfahren vor, deren Transformationsschritt Kontrollpunkte durch Free Form Deformation [SP86] zum Darstellen der Bewegung definiert. Das Kriterium für die Verschiebung der Kontrollpunkte basiert auf der normalisierten Mutual Information aus den Bildgrauwerten (s. Kap. 4.6.1). Rohlfing et al. [RMBJ03] konnten auf den Bilddaten von 17 Patientinnen zeigen, dass Rückers Registrierung das Volumen von Läsionen im Mittel um 26 % (Standardabweichung 22 %) reduzierte; im Extremfall lag ein Volumenverlust von 78 % vor.

Da auch die hier vorgestellte nicht rigide Registrierung direkt auf den Bildgrauwerten arbeitet, wird explizit die Volumenveränderung in den nachfolgenden Kapiteln evaluiert.

Bei näherer Betrachtung der Ergebnisse des Featuretrackings (s. Abb. 5.1) fällt auf, dass Features im Inneren und Randbereich von Läsionen kaum erfolgreich getrackt werden können. Auch im langsamer anreichernden Drüsenparenchym ist die Erfolgsquote deutlich niedriger. Demnach stellt das Anreicherungsverhalten eine derart große Veränderung dar, dass das Tracking in solchen Bildbereichen unter den strikt gewählten Parametern scheitert (s. Pfeil in Ausschnittsvergrößerung Abb. 5.1 b).

Das Tracking (s. Anhang B) basiert auf der Berechnung von Gradienten im Bild. Dazu werden ausgehend von einem zentralen Bildpunkt die Signaldifferenzen zu den benachbarten Voxeln einbezogen. Das Anreichern mit KM verändert die Signalwerte im Bild, so dass sich ein neues Gradientenbild ergibt, das ein erfolgreiches Tracking kaum mehr ermöglicht. Der Algorithmus geht davon aus, dass die Signalintensitäten zwischen Ausgangs- und Zielbild ähnlich sind. Diese Voraussetzung ist aufgrund der KM-Anflutung nicht mehr gegeben.

Somit scheitert für gewöhnlich das Tracking in Bereichen mit KM-Aufnahmen. Dies ist ein positiver Effekt, denn im anschließenden Interpolationsschritt kann die Bewegungsinformation nur noch aus den Translationen der Umgebung abgeleitet werden, in der durch keine oder sehr wenige KM-Vorkommnisse Features korrekt getrackt wurden. Die so gewonnene Bewegungsinformation kann zwar abschließend durch die Bildtransformation die Position und auch Form von Läsionen verändern, ein signifikanter Einfluss auf ihr Volumen ist jedoch nicht zu erwarten. Eine Evaluation der Volumenveränderung ist noch einmal Gegenstand des Kapitels 6.1.3.

Wenn das Tracking aufgrund der Anreicherung einen Bildbereich nicht mehr wiedererkennen kann, erhöht sich die Gefahr, dass stattdessen in der lokalen Umgebung ein alternativer Bildbereich gefunden wird, der dem ursprünglichen Ort zufällig ähnelt. Es können Korrespondenzen zu nicht anreicherndem Gewebe gefunden werden mit der Folge, dass bei der abschließenden Grauwerttransformation der Bereich der Anreicherung durch gesunde Gewebestrukturen ersetzt würde. Der eingangs erwähnte Effekt der Volumenreduktion von Läsionen käme dann zum Tragen. Im Zentrum einer Läsion ist ein solch fehlerhaftes

Tracking unwahrscheinlich, da die Anreicherung großflächig und die Distanz zum gesunden Gewebe zu groß ist; aber in deren Randbereich ist dies vorstellbar.

Jedoch besitzt der Algorithmus einige Qualitätsmerkmale, um fehlerhaftes Tracking zu unterbinden. Zum einen wird das Tracking durch die vorherige Selektion von Features nur an markanten Positionen gestartet. Diese befinden sich vor allem im kontrastreichen Übergang vom Drüsenparenchym zum Fettgewebe oder auch im Hautbereich. Ebenso können je nach Typ (s. Tab. 2.1) Übergänge vom Tumor zum gesunden Gewebe charakteristische Bildbereiche darstellen (s. Abb. 5.1). Für ein inkorrektes Tracking müsste aber ein ebenso charakteristischer Bildbereich in der Nähe der Features existieren. Des Weiteren kann die Qualität des iterativen Algorithmus durch die Parameter kontrolliert werden. Für die Brustbilddaten wurden nur maximal 15 Iterationen ($TH_{Iter} = 15$) zugelassen, um eine Genauigkeit von TH_{Trans} (s. Tab. 5.3) zu erreichen. Diese experimentell sehr strikt gewählten Beschränkungen geben nur wenig Möglichkeit für ein fehlerhaftes Tracking in entferntere Regionen.

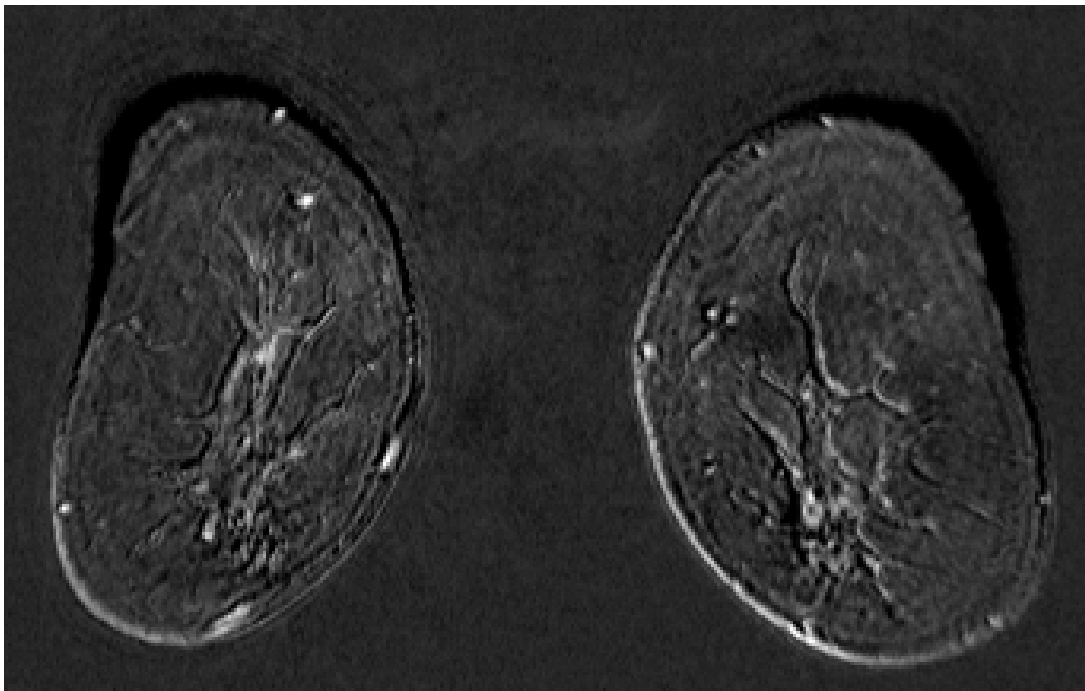
Das gewünschte Scheitern des Trackings in KM aufnehmenden Strukturen kommt jedoch nur dann zustande, wenn die Anflutung und damit die Bildveränderung ausreichend stark ist. Denn geringe Bildunterschiede werden vom Algorithmus durch Betrachten des Restfehlers kompensiert (s. Anhang B). Durch ihn soll ein erfolgreiches Tracking trotz Bildrauschens oder Anreicherung von gesunden Strukturen wie Fettgewebe ermöglicht werden (s. auch Kap. 6.1.3).

Großen Einfluss auf die Qualität der resultierenden Translationsvektoren hat vor allem die gewählte Fenstergröße für die Featuresuche und das Tracking. Die optimale Fenstergröße $WT_{x,y,z}$ zum Vergleich korrespondierender Bildausschnitte muss experimentell bestimmt werden. Mit Vergrößerung des Fensters $WT_{x,y,z}$ für die Gradientenberechnung erhöht sich auch die maximale erkennbare Distanz im Iterationsverfahren (s. Anhang B). Gleichzeitig wird der Einfluss des Bildrauschens geringer. Allerdings sinkt die Erfolgsquote des Trackings. Je größer das gewählte Fenster, desto mehr Veränderungen können im Fensterbereich liegen und das Erkennen von Übereinstimmung erschweren.

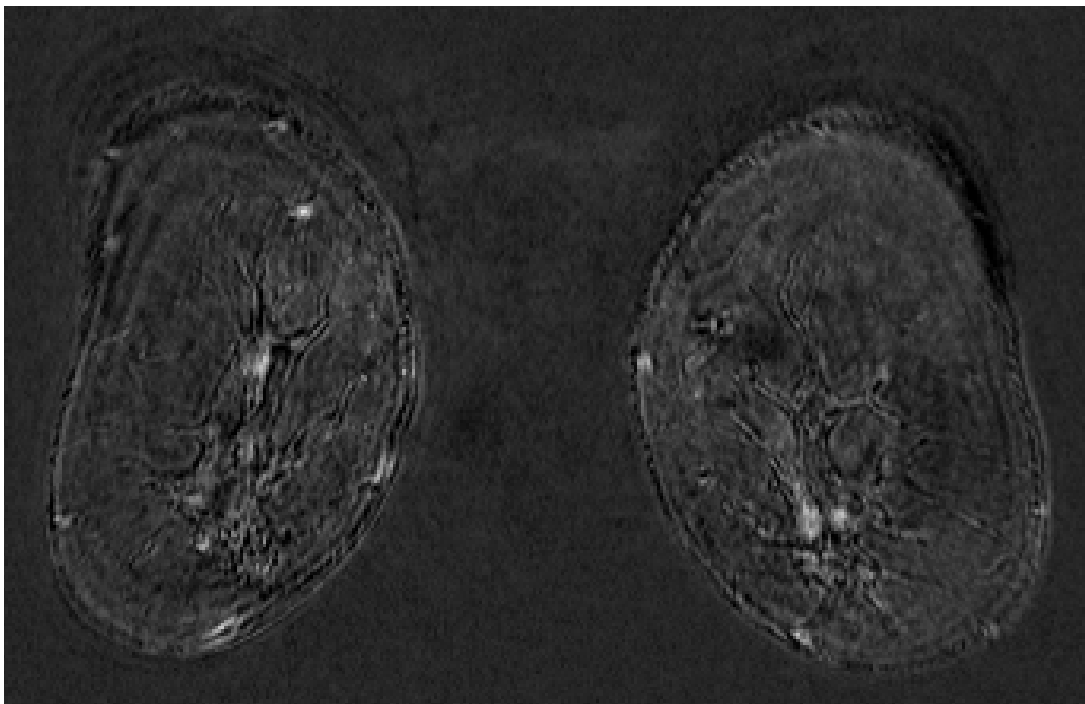
Eine Fenstergröße von 7x7x7 Voxel hat sich für die vorliegenden Brustbilddaten bewährt. Kleinere Fenster enthielten nicht genügend Bildinformation für ein erfolgreiches Tracking, die erkannte Translation verwies dadurch häufiger auf Bildbereiche ungleichen Inhalts.

Die als Kompromiss gewählte Fenstergröße war auch ausreichend beim Verfolgen größerer Bildbewegungen. In einem Beispiel (s. Abb. 6.1) konnten Verschiebungen von mehr als einem Zentimeter noch vollständig verfolgt werden (vgl. Kap. 6.1.4). Das unregistrierte Bild (a) zeigt im Übergang Haut zur Luft deutlich die schwarzen Säume im oberen Bildbereich, die bei der Bildsubtraktion durch die Auslösung nicht mehr überdeckender Strukturen aufgrund von Bewegungseinflüssen entstehen. Im registrierten Bild (b) sind diese nahezu vollständig kompensiert.

Vektoreninterpolation Da in Abhängigkeit der berechneten Featurepositionen Translationen nur für markante Bildbereiche vorliegen, ist die Aufgabe der Erzeugung des Vektorfeldes die Interpolation der Bewegungsinformation für jeden einzelnen Bildpunkt. Wie bereits erwähnt, ist die Interpolation ein sehr robustes Verfahren. B-Splines finden auf-



(a)



(b)

Abbildung 6.1: Gegenüberstellung eines (a) unregistrierten Subtraktionsbildes mit starkem Bewegungseinfluss mit einem (b) registrierten.

grund ihrer guten Biegeeigenschaften Anwendung. Vorteil des verwendeten Verfahrens ist die Lokalität der Berechnung auf nur $4 \times 4 \times 4$ Stützstellen und die gute Gesamtlaufzeit des Algorithmus. Anhand von Unittests wurde während der Implementierung beispielsweise getestet, ob Translationen derselben Länge und Richtung, die an einigen Stellen im Bildvolumen vorgegeben wurden, auf das gesamte Bild übertragen werden. Bei der Untersuchung einer Rotation eines synthetischen Würfels um seinen Schwerpunkt, die nur durch die Translation der acht Ecken erkannt wurde, konnte geprüft werden, dass die Bewegung auf die umliegenden Bildbereiche entsprechend übertragen wurde. Interessant ist das Verhalten der Interpolation, wenn Translationen widersprüchliche Informationen bereitstellen, beispielsweise zwei sich kreuzende Vektoren. Dies entspricht keiner physiologisch korrekt extrahierten Information aus den Brust MRT-Bilddaten, sondern kann nur durch fehlerhaftes Tracking verursacht werden.

Mit der gewählten Suchfenstergröße können beim Tracking der ca. 17.000 Features in wenigen Fällen ungleiche Bildausschnitte aufeinander abgebildet werden. Während die Translationsvektoren in der direkten Nachbarschaft die korrekte Verschiebung erfasst haben, zeigt der einzelne fehlerhafte Vektor in eine divergente Richtung oder besitzt eine abweichende Länge.

Die globale Auswirkung ist durch die darauffolgende iterative B-Spline Interpolation aller Translationen erst einmal gering (s. Anhang C). Beim Erzeugen des anfänglich groben Gitters überwiegt die Mehrheit der korrekt verfolgten Features. Erst bei feinerer Gitterstruktur wirkt sich der Fehler lokal auf die Bildqualität aus, es entstehen kreisförmige Registrierungsartefakte (s. Abb. 6.2). Sie sind kritisch, wenn sie im Bereich anreichernder Strukturen liegen und dadurch die Form oder Volumen der Läsion verändern (s. vorheriges Kapitel). Ein fehlerhaftes Tracking konnte jedoch nur in einem Fall im Bereich des mit Rückfaltungsartefakten (s. Kap. 4.2) durchgezogenen Thorax beobachtet werden.

Die meisten Gewebe des menschlichen Körpers können trotz des Einflusses von Bewegungen nicht komprimiert werden. Durch fehlerhaftes Tracking wird genau diese Invariante verletzt und Volumenverhältnisse werden verfälscht. Rohlfing et al. [RCM01, RMBJ03] stellen eine Optimierung durch zusätzliches Einfügen eines Regulierungsterms vor. Während der Vektorinterpolation wird anhand der Jacobi Determinante lokal entschieden, ob eine Kompression vorliegt. Die Stützstellen des Gitters werden immer derart gewichtet, dass das Bildvolumen erhalten bleibt. Durch diesen Regulierungsterm kann die Auswirkung falsch getrackter Features unterbunden werden.

Bildtransformation Die Aufgabe der Bildtransformation als letzter der sequentiellen Teilschritte liegt in der Übertragung der Grauwerte anhand des Vektorfelds an die neue Position. Solange die Bewegungsinformation des Vektorfelds kein Volumen eliminiert (s. auch Kap. 6.1.3), gehen durch die elastische Bildtransformation keine KM anreichernde Strukturen verloren. Eine Verfälschung der Signalwerte ist nicht zu befürchten.

Die registrierten Bilddaten zeigen im Vergleich zu den unregistrierten jedoch einen leicht geminderten Kontrast (Weichzeichnung). Dieser Effekt ist in der Literatur [Mei00] bekannt und auf die verwendete trilineare Interpolation zurückzuführen. Durch Anwendung aufwendigerer Interpolationsverfahren, die beispielsweise auf Splines basieren, kann die Bildschärfe erhalten werden.

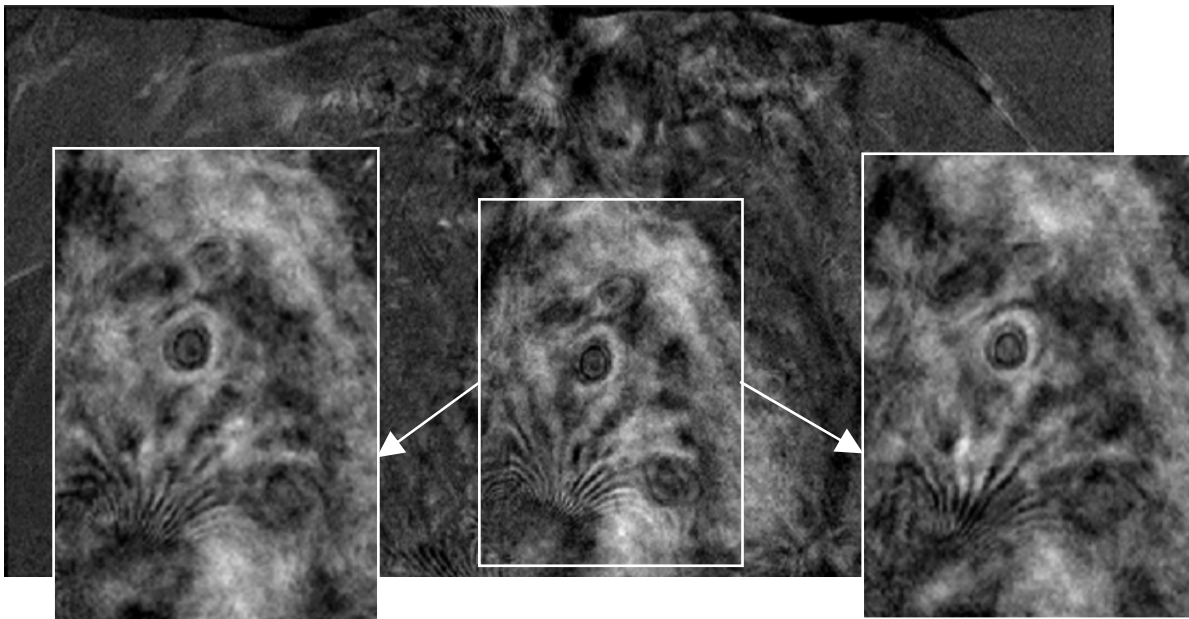


Abbildung 6.2: Kreisförmige Artefakte durch fehlerhaftes Featuretracking. Die Bildausschnitte zeigen die benachbarten Bildschichten.

6.1.2 Evaluation der Bildverbesserung

Eine Beurteilung der Qualitätsverbesserung durch die Registrierung erfolgt auf dem Vergleich der Distanzmaße (s. Kap. 4.6.1) und damit der Bildähnlichkeit der registrierten und unregistrierten Bilder jeweils zur nativen Aufnahme. Das Ergebnis ist nicht eindeutig interpretierbar, denn die funktionalen Bilddaten der Brust-MRT unterliegen zweier Veränderungen. Neben der Verschiebung der Bildinhalte durch die Bewegungen der Brust während der Bildakquisition führt auch das Anfluten durch die Applikation von KM zu unvorhersehbaren Veränderungen der Signalintensitäten (s. Kap. 6.1.3).

Unterschieden sich die Bilder nur anhand der Bewegung, wäre das erreichbare Optimum der Registrierung die Gleichheit der Bilder. Aufgrund der Anreicherung, die es bei der Registrierung unverändert zu erhalten gilt, bleibt das erreichbare Maß an Ähnlichkeit unbekannt.

Da das verwendete MRT-Protokoll zu Beginn zwei native Aufnahmen (SN_2 u. SN_3 , s. Abb. 4.3) akquiriert, ergibt sich die Möglichkeit, registrierte Bilder auch ohne KM-Anreicherung miteinander zu vergleichen. Leider liegen sie jedoch in verschiedenen Orientierungen vor und müssen daher vor der Registrierung transformiert werden (s. Anhang A.3). Auch für die nativen Serien bleibt die erreichbare Ähnlichkeit somit unbekannt, weil der Anteil an Veränderungen einerseits durch die KOS Transformation und andererseits durch die Bewegungen nicht unterschieden werden kann. Dennoch sollte hier die größte Ähnlichkeit erreicht werden.

Folgend werden Distanzwerte zur Bewertung der Veränderung der Bildähnlichkeit durch

Bild	Zeit (sek.)	SSD	MI%	Matrix	Orientierung
Unregistriert					
$SN_2 - SN_2$	0	0,00	100,00	high	koronar
$SN_3 - SN_2$	159	12,37	21,00	high	transversal
$SN_4 - SN_2$	255	11,11	23,99	low	koronar
$SN_5 - SN_2$	326	13,34	24,45	low	koronar
$SN_6 - SN_2$	397	14,50	24,14	low	koronar
$SN_7 - SN_2$	538	14,79	24,59	high	koronar
$SN_8 - SN_2$	697	19,14	20,49	high	transversal
Registriert					
$R_{SN_3} - SN_2$	159	8,15	24,08	high	transversal
$R_{SN_4} - SN_2$	255	8,76	27,14	low	koronar
$R_{SN_5} - SN_2$	326	10,43	26,66	low	koronar
$R_{SN_6} - SN_2$	397	11,26	26,39	low	koronar
$R_{SN_7} - SN_2$	538	11,74	27,27	high	koronar
$R_{SN_8} - SN_2$	697	11,16	22,32	high	transversal

Tabelle 6.1: Distanzmaße registrierter und unregistrierter Bildvolumen zur nativen Aufnahme

die Registrierung eingesetzt. Für die Statistik werden jeweils die Differenzbilder aller 101 Patientinnen der unregistrierten ($SN_4 - SN_8$) und der registrierten Nachkontrastaufnahmen ($R_{SN_4} - R_{SN_8}$) jeweils zur nativen Serie (SN_2) verglichen (s. Tab. 6.1). Die Tabelle zeigt zusätzlich den Zeitabstand zur nativen Serie in Sekunden, die Matrixgröße und die Orientierung der Bildserien. Die Abstandsmaße SSD und MI% wurde nur über den Bildbereich berechnet, den die koronaren und transversalen Bildserien im Patienten-KOS gemein haben (s. Anhang A.1).

Nach dem SSD Maß sind die Abstände der registrierten Bilddaten zur nativen Aufnahme deutlich geringer als die der unregistrierten. Die Registrierung macht die Bilder also ähnlicher. Die zweite native Serie SN_3 hat die größte Ähnlichkeit, die letzte Aufnahme eine geringere. Sowohl im registrierten als auch im unregistrierten Fall erhöht sich tendenziell der Abstand für die späteren Aufnahmen des Protokolls. Als Ursache kann neben den Bewegungen die KM-Anreicherung vermutet werden, da das kontinuierliche Anreichern der gutartigen und gesunden Gewebe größere Unterschiede mit fortschreitender Zeit verursacht.

Auch mit der MI% kann eine deutliche Steigerung der Bildähnlichkeit durch die Registrierung nachgewiesen werden. Für jede registrierte Serie wird ein höherer Wert – und damit größere Ähnlichkeit – gegenüber den unregistrierten erreicht. Jedoch zeigt ein Vergleich der Serien untereinander im Gegensatz zur SSD eine uneinheitliche Tendenz. Die wenigste Ähnlichkeit haben demnach die Serien SN_3 und SN_8 zur nativen Aufnahme. Bei erstgenannter Serie ist dies erstaunlich, müssten die beiden nativen Aufnahmen doch die

größte Ähnlichkeit aufweisen. Auch für SN_5 und SN_6 ist die Ähnlichkeit geringer als bei Serie SN_7 , obwohl letztere mehr KM und damit mehr Unterschiede aufweisen sollte.

Das unerwartete Verhalten der MI% ist in der Auswertung der Häufigkeit von Bildpunktpaaren im 2D-Histogramm begründet. Während beim SSD einzelne Punkte verglichen werden, wirkt sich bei MI% eine Weichzeichnung der Bilddaten genauso wie Rauschen wesentlich stärker auf die Berechnung der Entropie aus [Zhu07]. Eine Weichzeichnung ergibt sich auch bei KOS-Transformation der Serien SN_3 und SN_8 durch Anwendung der trilinearen Interpolation (s. vorheriges Kapitel). Aber auch die Serien SN_4 bis SN_6 enthalten eine wenn auch geringere Weichzeichnung, denn diese low-res Aufnahmen wurden durch das MRT-Gerät auf eine größere Bildmatrix interpoliert (s. Kap. 4.1.1). So ergibt sich für die Serie SN_7 die größte Ähnlichkeit; nur sie hat die gleiche Orientierung und Matrixgröße wie die native Aufnahme.

Mit der MI% kann wie mit der SSD die Erhöhung der Ähnlichkeit der registrierten gegenüber den unregistrierten Bilddaten nachgewiesen werden. Nach dem Distanzmaß SSD nimmt die Ähnlichkeit über die Zeit ab, die MI% ist dagegen für einen Vergleich der einzelnen Serien untereinander ungeeignet.

Die Bewertung der Bildähnlichkeit anhand des Abstandsmaßes stimmt nicht immer mit dem subjektiven Empfinden überein. Denn viele kleine, kaum wahrnehmbare Unterschiede im Bild können zu einer geringeren Ähnlichkeit führen als ein deutlich wahrnehmbarer.

Die Abstandsmaße können aufgrund der KM-Anreicherung und des Bildrauschens niemals den Idealwert für identische Bilder erreichen bzw. der erreichbare Idealwert ist nicht bekannt.

6.1.3 Einfluss auf Kontrastmittel anreichernde Strukturen

Beide Bewertungsmaße des vorherigen Kapitels zeigen, dass die Ähnlichkeit der Bilder durch Einsatz der Registrierung erhöht wird, ohne zu wissen, ob dies allein auf die Bewegungskompensation zurückzuführen ist oder auch durch die Eliminierung der KM-Anreicherung. Letzteres wäre sehr ungünstig und hätte gravierende Auswirkung auf alle nachfolgenden Bildverarbeitungsschritte. Im Kapitel 6.1.1 wurden Gründe genannt, warum eine Volumenveränderung von Läsionen durch die Registrierung nicht zu erwarten ist. Statt einer visuellen Überprüfung ergibt sich auch hier mit Hilfe der GS-Vorgaben die Möglichkeit einer automatisierten Evaluation.

Der GS wurde auf den unregistrierten Bilddaten der Serie SN_7 erhoben. Das bereits berechnete Vektorfeld der Registrierung (s. Kap. 4.2.1) kann zur Transformation auf den Zeitpunkt der nativen Serie verwendet werden. Somit kann ein Volumenvergleich zwischen den GS-Objekten im unregistrierten und registrierten Zustand vorgenommen und dadurch geprüft werden, ob das Vektorfeld volumenreduzierende Informationen im Bereich von Läsionen enthält.

Der Vergleich wurde für die Gruppe der histologisch gesicherten 87 malignen und 54 benignen Tumore separat durchgeführt. Für sie wurden jeweils die prozentualen Volumenänderungen Rel_{Vol} der registrierten (Vol_{reg}) zu den unregistrierten (Vol_{unreg}) Markierungen berechnet (s. Formel 6.1).

$$Rel_{Vol} = \frac{Vol_{reg}}{Vol_{unreg}} \cdot 100\% \quad (6.1)$$

Im Mittel konnte bei den Karzinomen ein Volumenanstieg von knapp 6 % mit einer Standardabweichung von 8,5 % festgestellt werden (s. Tab. 6.2). Der größte Volumenverlust lag bei 16,3 %, die maximale Zunahme bei 30,0 %. Ähnlich uneinheitlich waren die Ergebnisse der benignen Tumore, auch hier entstand ein Volumenanstieg im Mittel von 9,0 % bei einer Standardabweichung von 9,8 %. Der maximale Volumenverlust lag bei 9,3 %, während die maximale Zunahme einen Wert von 32,0 % erreichte. Nur eine deutliche Minderheit von 12 % der Karzinome und 10,2 % der benignen Tumore verloren überhaupt an Volumen. Die Standardabweichung verdeutlicht, dass größere Volumenänderungen nur Ausnahmen darstellen.

Die Registrierung tendiert nicht dazu, systematisch Anreicherungen auszulöschen, vielmehr ist im Mittel eine leichte Zunahme des Objektvolumens festzustellen. Eine mögliche Ursache dafür liegt in der Methodik der Evaluation, denn der GS stellt eine binäre Bildinformation dar. Sie kann bei der elastischen Transformation im Rahmen der Registrierung im Gegensatz zu den Grauwerten nicht interpoliert werden. Ein markierter Bildpunkt wird nach der Transformation dem nächstgelegenen Rasterpunkt zugewiesen. Auch kann der manuell selektierte Goldstandard (s. Kap. 4.5.1) auf die unregistrierten Differenzbilder nicht nur die Läsion, sondern aufgrund der noch enthaltenen Bewegung auch angrenzende Bewegungsartefakte einschließen (vgl. Kap. 6.2.3). Somit werden auch Bildbereiche, die keine Läsion darstellen, in die Statistik einbezogen.

Bei größeren Volumenveränderungen muss die Ursache jedoch in einem fehlerhaften Tracking gesucht werden. Eine erste visuelle Überprüfung der Läsionen mit extremen Volumenabweichungen zeigte aber keine wahrnehmbaren Anzeichen für eine Zerstörung von Bildinformation (vgl. Abb. 6.2). Auffällig ist jedoch, dass vor allem schwach anreichernde gutartige Befunde bzw. diffus wachsende Karzinome am empfänglichsten für Volumenveränderungen sind. Beide Typen zeichnen sich durch einen unscharfen Objektrand (s. Tab. 2.1) aus. Es bleibt zu vermuten, dass geringe Volumenabweichungen im Randbereich der Läsionen entstehen, die kaum wahrnehmbar sind und sich erst in der Gesamtbetrachtung bemerkbar machen. Mögliche Ursache können ungenau getrackte Features im Randbereich solcher Läsionen sein. Schwache Intensitätsänderungen stellen, wie in Kapitel 6.1.1 erläutert, in den Bildfolgen unter Umständen solche geringen Einflussgrößen dar, die wie das Bildrauschen vom Trackingalgorithmus als Restfehler aufgefasst werden. Auch die bei der Grauwerttransformation eingesetzte trilineare Interpolation kann durch geringe Weich-

Dignität	Rel_{Vol}	std Abw.	Min	Max
Maligne	+5,8 %	8,5 %	-16,3 %	30,0 %
Benigne	+9,0 %	9,8 %	-9,3 %	32,0 %

Tabelle 6.2: Relative Volumenänderung registrierter gegenüber unregistrierten GS-Läsionen

zeichnungen die Randschärfe beeinflussen und solche Volumenveränderungen verstärken.

Auch bei der Betrachtung beider Dignitäten in Tabelle 6.2 fällt auf, dass die Volumenbeeinflussung der schwach anreichernden gutartigen Tumore im Mittel höher als bei den stark anreichernden Karzinomen liegt.

Abschließend bleibt festzustellen, dass die Registrierung mit 5,8 % und 9 % mittlerem relativem Zuwachs einen überwiegend geringen Einfluss auf das Volumen anreichernder maligner und benigner Strukturen hat. In der Literatur konnten bei alternativen Verfahren Verkleinerungen im Extremfall von bis zu 78 % festgestellt werden. Das Mittel lag dort bei knapp 26 % [RMBJ03]. Durch Betrachtung von Gradienten und des Trackens nur markanter Bildinhalte ist die Erkennung vorhandener Bewegung wesentlich zuverlässiger als beispielsweise in der Veröffentlichung von Rückert et al. [RSH+99], die auf der Mutual Information basiert.

In der vorgelegten Arbeit bleiben alle histologisch gesicherten Strukturen durch den Registrierungsprozess erhalten. Da die Bilddaten durch Bewegung und KM-Aufnahme gleich zweier Veränderungen unterliegen, kann keine Bewertung vorgenommen werden, wie gut sich die transformierten Bilddaten dem tatsächlich vorliegendem Sachverhalt nähern, da dieser nicht bekannt ist. Die Evaluationen in der vorliegenden Arbeit zeigen jedoch, dass die Bildähnlichkeit durch die Registrierung erhöht werden konnte (s. vorheriges Kapitel). Dabei werden auch die Positionen der Läsionen nach Kompensation der Bewegung korrigiert und dadurch zwangsweise die Form der Läsion gewollt verändert. Der Vergleich zum registrierten GS lässt vermuten, dass Ungenauigkeiten des Trackings dabei einen zusätzlichen unerwünschten Einfluss auf die Läsion im Randbereich nehmen. Inwieweit die Ungenauigkeiten die Diagnose des Systems beeinflussen, beispielsweise durch abweichende Werte der Morphologieberechnung, kann nicht untersucht werden. Es ist aber anzunehmen, dass diese sehr gering sein dürften. Denn es gibt keinerlei Hinweise, dass Fehler der Registrierung beispielsweise einer gutartigen runden Läsion zu bösartigen Spikulierungen verhelfen. Gegenüber kleineren Veränderungen muss die Morphologie alleine schon wegen des Bildrauschens robust sein (s. Kap. 6.3.3).

6.1.4 Anwendung auf fremde Bildinhalte

Der Registrierungsalgorithmus ist nicht auf die Verarbeitung von Bilddaten mit Brüsten beschränkt. Als flexible Komponente können auch andere Inhalte korrigiert werden. Voraussetzung ist, dass es sich um mono modale 3D-Bilddaten mit – aber auch ohne – Anwendung von KM handelt.

Der vorhandene Algorithmus wurde nur durch Parameteranpassung auf Angiographie Bilddaten des Ober- und Unterschenkels angewendet. Während auch diese Bilddaten durch die MRT-Modalität erzeugt werden, diente als weiterer Test eine Thorax Aufnahme durch die *Computertomographie* (CT).

Angiographie Bei den Angiographiebilddaten wird KM wie bei der Brust-MRT mit dem Ziel eingesetzt, *Stenosen* (Verengungen) von Gefäßen sichtbar zu machen. Dazu wird eine native Messung und eine weitere mit KM-Applikation gemessen. Wiederum durch Bildsubtraktionstechnik lassen sich die angereicherten Gefäße darstellen. Im Vergleich zu

den Tumoren der Brust haben in der Angiographie Bewegungsartefakte auf die Beurteilung der feinen Gefäßstrukturen einen wesentlich stärkeren Einfluss.

Allgemein enthalten die Angiographiebilddaten weniger Kontrast als die Brustbilder, was das Bestimmen und Tracking der Features erschwert. Sowohl das Such- als auch das Trackingfenster mussten deswegen vergrößert werden ($WF_{x,y,z} = WT_{x,y,z} = 9 \times 9 \times 9$, vgl. Tabelle 5.3). Um die Qualität des Trackings weiter zu erhöhen, wurde durch Herabsetzen der Anzahl an erlaubten Iterationen ($TH_{Iter} = 5$) die Fehlerquote gesenkt.

Abb. 6.3 a und c zeigen die Subtraktionen unregistrierter Bilddaten. Im Übergangsbereich zur Luft und an der Knochenoberfläche sind die Bewegungsartefakte aufgrund der Verschiebungen gut erkennbar. Nach der Registrierung (s. Abb. 6.3 b und d) sind die Signalintensitäten in den Differenzbildern stark gemindert. Die Gefäße können nun deutlich anhand der hellen linienförmigen Struktur erkannt werden. Vor allem Features im Bereich des Knochengewebes, das in den Brustbilddaten nicht enthalten ist, erwiesen sich als gut verfolgbar.

Ein fehlerhaftes Tracking kann bei den Angiographiebilddaten wesentlich folgenreicher sein als dies bei der Brust der Fall ist, weil fehlerhafte Vektoren nicht vorhandene Stenosen im registrierten Bild erzeugen können (vgl. Abb. 6.2).

Um die Qualität und den Nutzen der Registrierung in der Angiographie abschließend bewerten zu können, bedarf es weiterer Studien.

Computertomographie Zum Testen des Registrierungsalgorithmus auf einer weiteren Modalität werden zwei Thoraxaufnahmen durch die CT verwendet, die in In- und Expiration erfolgten (s. Abb. 6.4 a und b). Die CT Bilddaten bestehen aus normierten Signalwerten. Sie stellen die Dichte von Geweben durch die hervorgerufene Abschwächung der eingesetzten Röntgenstrahlung in der Einheit Hounsfield dar [Kal06]. Hohe Dichte haben beispielsweise Knochen und wird entsprechend der hohen Signalwerte hell abgebildet, Luft mit der geringsten Dichte erscheint dagegen dunkel im Bild. In Abb. a und c erscheint die Lunge als dunkler Bereich im Bild mit den helleren Bronchien und Gefäßen, die Wirbelsäule und die die Lunge umgebenden Rippen sind dagegen hell abgebildet. Unterhalb des Thorax ist die Liege erkennbar, auf der der Patient während der Untersuchung gelagert ist. Das FoV geht über den durch die CT messbaren Bereich hinaus. Es lässt sich als kreisförmige Begrenzung der Röhre am Bildrand erkennen. Die Bildecken enthalten keine Informationen mehr. Durch Ein- und Ausatmen entstehen durch Heben und Senken des Brustkorbs große Verschiebungen vor allem in der oberen Bildhälfte.

Auf den kontrastreichen CT Bilddaten lassen sich die Features bereits innerhalb kleiner Fenster bestimmen und auch tracken ($WF_{x,y,z} = WT_{x,y,z} = 5 \times 5 \times 5$ Voxel). Wiederum wurden die Bewegungsinformationen vor allem anhand der Verschiebung der Knochen gut erkannt.

Abb. 6.4 c zeigt die Differenz der beiden unregistrierten Ausgangsbilder (a) und (b). Das Anheben der Bauchdecke durch das Einatmen ist durch die dunklen Differenzwerte im oberen Bildbereich deutlich zu erkennen. Sowohl im äußeren als auch im inneren Bereich der Lunge zeigen sich Differenzen im helleren Grau. Durch die Registrierung (s. Abb. 6.4 d) von Bild (b) auf (a) sind die Knochen im unteren Bereich des Bildes und die Bauchdecke wieder

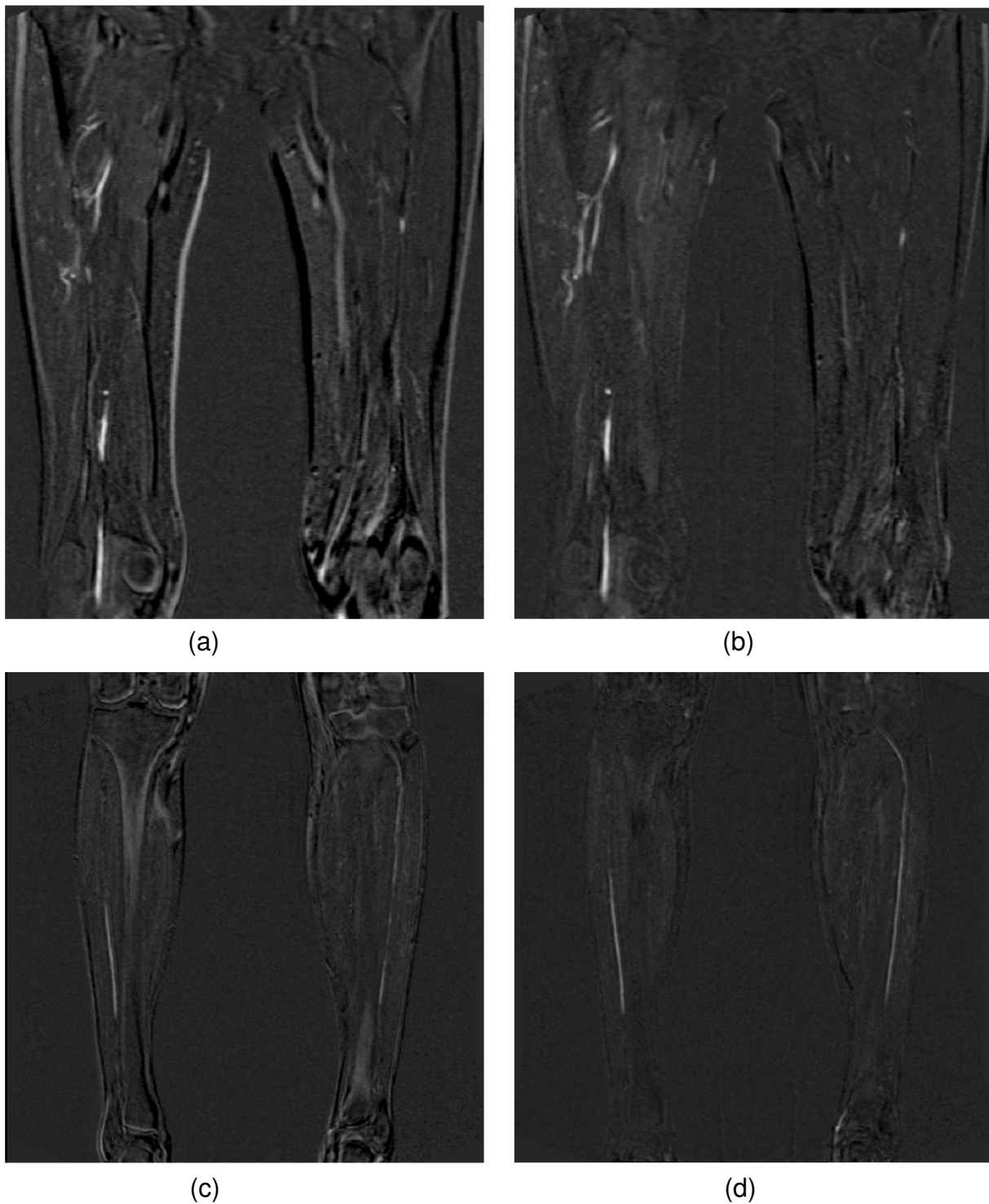


Abbildung 6.3: Gegenüberstellung (a und c) unregistrierter mit (b und d) registrierten Subtraktionen von Angiographiebildern.

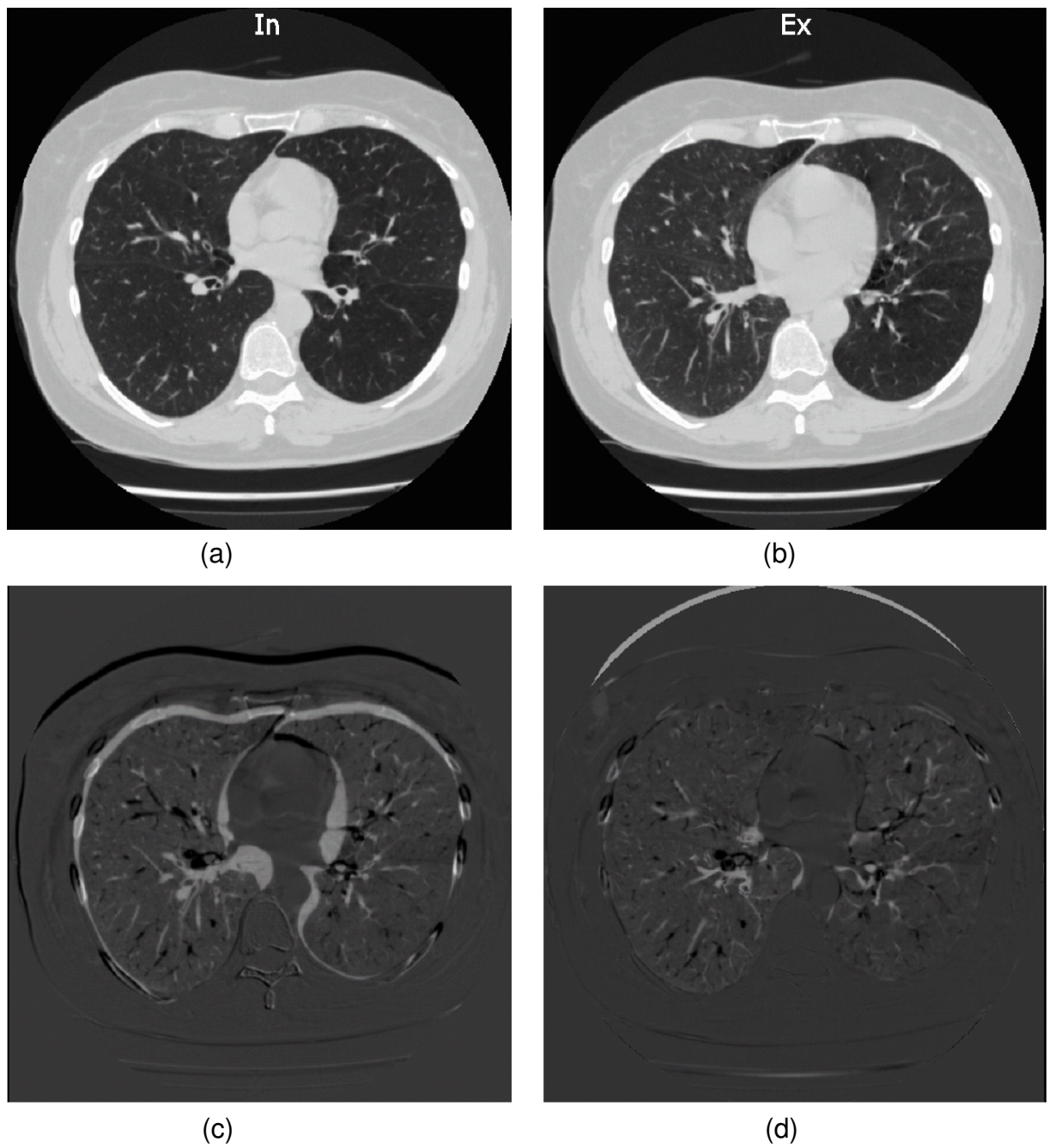


Abbildung 6.4: CT Bilddaten in (a) In- und (b) Expiration . Gegenüberstellung der Subtraktion (c) vor und (d) nach der Registrierung.

in Übereinklang gebracht. Die Bewegungsvielfalt innerhalb der Lunge konnte hingegen nicht mehr detailliert erkannt werden. Die Gefäße und Bronchien unterliegen einer derart starken Verformung, dass nur wenige Features erfolgreich getrackt werden konnten. Auch wenn die am Knochengewebe beteiligten Features im unteren Bildbereich fehlerfrei getrackt wurden, konnten die Rippen im oberen Bereich aufgrund der großen Distanz nicht mehr verfolgt werden.

Aber auch solch starken Verschiebungen, die jedoch für die Brustuntersuchung nicht zu erwarten sind (s. Abb. 6.1), können durch Erweitern des Registrierungsalgorithmus verarbeitet werden. Zur Steigerung der Reichweite behilft sich die Informatik anstelle einer Vergrößerung des Suchfensters mit einer hierarchischen Struktur verkleinerter Bildvolumen als Vorverarbeitungsschritt. Die oberen Ebenen, also die kleinsten Volumen der Bildpyramide, dienen zum Erkennen großer Translationen. Die exakten Verschiebungen werden darauf aufbauend auf den unteren und damit höher aufgelösten Bildebenen der Pyramide bestimmt [Bou02].

Interessant ist die Entstehung des hellen sichelförmigen Streifens direkt am oberen Bildrand (s. Abb. 6.4 d). Die Registrierung erkennt als globale Bewegungsinformation das Anheben der Bauchdecke und überträgt sie durch das anfänglich grobe Gitter des Vektorfelds auf das ganze Bild. Da die glatte Oberfläche der kreisförmigen CT-Röhre keine Erkennung von Features zulässt, wird bei der Verfeinerung des Gitters die globale Bewegung durch Mangel an Informationen in diesem Bereich nicht mehr korrigiert. Die kreisförmige Begrenzung im oberen Bild nimmt daher eine ellipsoide Form an.

Der vorgestellte Registrierungsalgorithmus erweist sich als eine flexible Komponente, die zur Qualitätssteigerung von Bilddaten bei verschiedenen Fragestellungen in der Medizin eingesetzt werden kann. Gegenwärtig hat der Algorithmus eine Limitierung bezüglich der maximal erkennbaren Distanz vorliegender Verschiebungen.

6.1.5 Erkenntnisse und Grenzen

Ungeachtet der nachfolgenden Softwarekomponenten zur Erkennung, Beschreibung und Bewertung der Läsionen, die den Prototyp erst zu einem Zweitmeinungssystem komplettieren, ist bereits das Ergebnis der Registrierung eine große Hilfestellung für den Radiologen. Sowohl die registrierten Bilddaten als auch die Differenzbilder werden vom System wieder im DICOM-Format (s. Anhang A) exportiert. Der Radiologe kann seinen gewohnten Diagnostikablauf nun direkt auf den registrierten Differenzbilddaten vornehmen, die um die Anzahl an falsch positiven Anreicherungen reduziert sind und die Morphologie und Kinetik im optimalen Fall ohne den Einfluss der Bewegung korrekt darstellen.

Das Tracking vorhandener Bewegungen ist der über die resultierende Qualität entscheidende Teilschritt der Registrierung (s. Kap. 6.1.1). Liegen fehlerhaft getrackte Vektoren vor, kann dies negative Auswirkung auf die Darstellung von Läsionen im registrierten Bild haben. Dies konnte nur in einer Messung im Thoraxbereich am Bildrand beobachtet werden (s. Abb. 6.2). Es wurden Lösungen vorgestellt, die fehlerhafte Vektoren im nachfolgenden Interpolationsschritt unterdrücken.

Die Registrierung stellt eine solide Komponente zum Einsatz auch auf anderen Bildmodalitäten dar (s. Kap. 6.1.4). Die Verbesserung der Bildqualität der Brust MRT-Bilddaten ist für den Beobachter deutlich wahrnehmbar und wurde auch statistisch durch Abstandsmaße verifiziert (s. Kap. 6.1.2). Verbleibende Abweichungen ergeben sich durch die Anreicherung mit KM. In Kapitel 6.1.3 konnte ergänzend zur statistischen Auswertung gezeigt werden, dass die anreichernden histologisch gesicherten Tumore durch die Registrierung erhalten bleiben. Dabei unterliegen die Tumore nur im geringen Maße dem in der Literatur bekannten Problem der Volumenreduktion, meist wurde gegenteilig eine leichte Volumenvergrößerung festgestellt. Somit liegt die Ursache für die Erhöhung der Bildähnlichkeit überwiegend in der Elimination von Bewegung bei gleichzeitiger Konservierung der Anreicherungen.

Derzeit existiert eine Limitierung bei der Erkennung größerer Verschiebungen, die von der gewählten Größe des Suchfensters beim Tracking abhängt. Auch wenn die Limitierung auf dem vorhandenen Kollektiv nicht zum Tragen kam, wurden auch hier durch einen pyramidalen Ansatz Verbesserungsmöglichkeiten aufgezeigt.

Insgesamt ist eine autonome Zweitmeinungssoftware ohne diesen erfolgreichen Vorverarbeitungsschritt nicht denkbar. Neben der Vereinfachung der nachfolgenden Bildverarbeitung werden die Morphologie und Kinetik erst korrekt abgebildet.

6.2 Segmentierung

Aufgabe der Segmentierung ist es, die relevanten von nicht relevanten Bildbereichen abzutrennen und erstere für die sich anschließende Bildverarbeitung bereit zu stellen. Von Interesse sind alle Strukturen der MRT-Brustbilddaten, die Anreicherungen zeigen. Die Segmentierung ist der entscheidende Schritt, der diese Prototypsoftware von anderen CAD-Systemen deutlich unterscheidet (s. Kap. 6.5). Denn sie erkennt automatisch jede isolierte Anreicherung als einzelnes Objekt und schafft somit erst die Voraussetzung für eine anschließende morphologische und kinetische Beurteilung pro Läsion. Die Literatur kennt für die Segmentierung der Läsionen als einzelne Objekte nur manuelle oder halbautomatische Verfahren. Ohne die Automatisierung der Segmentierungskomponente in der vorliegenden Arbeit müsste die Selektion der Läsionen durch Eingaben des Benutzers erfolgen.

Eine Segmentierung braucht Kriterien, anhand derer die Abgrenzung relevanter Bildbereiche vorgenommen werden kann. Da die Differenzbilder nach Elimination der Signale durch das Fettgewebe und aufgrund der Vorverarbeitung durch die Registrierung im Idealfall nur noch die KM anreichernden Strukturen enthalten, ist die Menge an nutzbaren Kriterien für eine Segmentierung sehr begrenzt. Ein Top-Down Ansatz scheidet zunächst einmal aus, da nur geringes Vorwissen über die Läsionen existiert. Sie können in ihrer Größe, Form, Textur oder Lokalisation erheblich variieren. Bekannte Top-Down Verfahren in der Literatur benötigen grobe manuelle Vorgaben [CGB06, ZBE⁺07, WSK⁺06]. Die Schwierigkeit beim alternativen Gruppieren der Voxel für ein Bottom-Up Vorgehen liegt in der Findung eines Homogenitätskriteriums. Beispielsweise variieren die Texturierung oder Randeigenschaften der Läsionen zu sehr, um als Kriterium geeignet zu sein. Zudem

stellen sie morphologische Eigenschaften dar, die zur späteren Klassifikation herangezogen werden. Als Grundlage für die Segmentierung könnten sie die Werte der entsprechenden Merkmalsberechnungen beeinflussen (s. Kap. 6.3.3).

Die Abgrenzung erfolgt hier durch die für KM-Konzentration stehenden Signalintensitäten (SI) der Differenzbilder und somit durch eine Schwellwertsegmentierung. Die Signalintensitäten sind jedoch nicht genormt. Aus ihnen alleine lässt sich nicht wie beispielsweise bei der CT (s. Kap. 6.1.4) direkt auf die Zugehörigkeit zu einem bestimmten Gewebe schließen. Das Anreicherungsverhalten jeder Läsion ist individuell und stellt sich aufgrund der Angiogenese in den unterschiedlichen Zeitpunkten der Dynamik durch einen sehr weiträumigen Wertebereich dar (s. Kap. 2.3.2). Daher wurde die Segmentierung um die Berechnung individueller Schwellwerte pro Läsion erweitert.

Allgemein wird der Segmentierungsschritt als eine globale Verarbeitung des gesamten Bildes verstanden. Jedes Region Growing leistet immer nur die Verarbeitung eines lokalen Bildbereiches und überführt ihn als zusammenhängendes Objekt in das Intermediate-Level. Um die Abarbeitung des Gesamtbildes zu erreichen, wurde der Segmentierungsprozess durch eine überschaubare Heuristik erweitert. Diese minimiert zudem die Anzahl zusätzlicher Funde (s. nächstes Kapitel), denn neben den Läsionen reichern auch Parenchym und Gefäße an. Diese Strukturen können aufgrund der Regeln teilweise erkannt und vom Segmentierungsergebnis ausgeschlossen werden.

In Abb. 5.4 und 5.5 sind die Ergebnisse der Segmentierung einer bös- und gutartigen Läsion derselben Patientin abgebildet. Abb. (a) zeigt die registrierte GS-Markierung des Radiologen in Überblendung mit dem registrierten Bildvolumen. In Kapitel 6.2.3 wird der Goldstandard benutzt, um die Qualität der automatischen Segmentierung zu beurteilen. Das Projektionsvolumen, das durch die Mittelwertbildung aller Serien weich gezeichnet wirkt, ist in Abb. (b) dargestellt. Alle segmentierten Token, von denen die Mehrheit zusätzliche Funde darstellen, sind durch Zufallsfarben hervorgehoben. Jede Farbe markiert eine einzelne Läsion. Kapitel 6.2.2 schlägt Verbesserungen zur Reduktion zusätzlicher Funde vor. In Abb. (c) ist noch einmal verdeutlicht, wie viel des angereicherten Volumens ($SI \geq 20$) bereits durch die Heuristik der Segmentierung ausgeschlossen wird (Zyan), während nur ein geringer Teil als Token (Orange) übernommen wird. Das nächste Kapitel diskutiert die Vorteile der Segmentierungsheuristik und des Einsatzes des Projektionsvolumens. Die Nachsegmentierung ist in Abb. (d) visualisiert, das Karzinom wurde auf der Serie SN_4 bestimmt, der gutartige Fund auf SN_7 . Die Form der nachsegmentierten Läsionen hat sich in Abb. (d) noch einmal deutlich gegenüber den Token im Projektionsvolumen (s. Abb. (c)) verändert. Die Gesamtzahl an Objekten der Einzelserie ist gegenüber dem Projektionsvolumen reduziert, da sie teilweise auf einer der anderen Bildserien nachsegmentiert wurden und somit nicht abgebildet sind.

Kapitel 6.2.4 arbeitet die Unerlässlichkeit der Registrierung als Vorverarbeitungsschritt heraus. Dazu wird die Segmentierung im Vergleich auch auf die Differenzbilder nicht registrierter Volumen angewendet.

6.2.1 Qualitätsabschätzung des Algorithmus

Die wesentlichen Bestandteile der Segmentierung sind die Heuristik, der Einsatz des Projektionsvolumens und die Nachsegmentierung.

Heuristik Durch die Heuristik (s. Kap. 4.3.1) und Überlagerungsregeln erhält A-priori-Wissen Einzug in den Segmentierungsalgorithmus. Der Regel (1) liegt zugrunde, dass Läsionen stärker anreichern als gesundes Gewebe (s. Kap. 2.3.2, s. Abb. 2.4). Somit wird die Reihenfolge der Bearbeitung festgelegt. Der Algorithmus berücksichtigt die signalstärksten Bildpunkte zuerst, weil sie die vermeintlich wichtigeren Bildinhalte darstellen.

Der Erfolg des Segmentierungsalgorithmus hängt im Wesentlichen von der Bestimmung der Schwellfunktion der Regel (2) ab. Nur auf der Grundlage von Schwellwerten erfolgt die Abgrenzung der relevanten Bildbereiche durch das Region Growing. Die endotumorale KM-Verteilung ist uneinheitlich und hängt von verschiedenen Gewebeeigenschaften wie der lokalen Gefäßdichte, dem Vorliegen lokal thrombosierter oder nekrotischer Areale und anderen Faktoren ab (s. Tab. 2.3). Das Verhältnis der SI des Hotspotvoxels zum Signalverhalten der gesamten Läsion kann daher sehr variabel sein, die Findung einer zufriedenstellenden monoton fallenden Funktion zur Schwellwertberechnung kann als anspruchsvoll angesehen werden.

Die untere Schranke von Regel (3) soll verhindern, dass durch die Bildsubtraktion ausgelöschte Bereiche aufgrund des Bildrauschens dennoch als Anreicherungen segmentiert werden. Damit wird das Abbruchkriterium formuliert, durch das die iterative Segmentierung gestoppt wird. In Kapitel 6.2.2 werden weitere Abbruchkriterien diskutiert.

Durch die Behandlung von Überlagerungen während der iterativen Segmentierung wird die Intension von Regel (1) weiter konsequent umgesetzt. Objekte, die durch einen höheren Schwellwert segmentiert wurden, erachtet das System wichtiger als angrenzende Areale, die durch Anwendung eines niedrigeren Schwellwerts entstehen. Durch diese Regel sollen Läsionen von gesunden Strukturen, die ebenfalls eine – wenn auch schwächere – Anreicherung zeigen, abgegrenzt werden.

Die Segmentierungsregeln wurde anfänglich für den Einsatz auf den 3D-Bildserien formuliert und die Segmentierung in der frühen Projektphase auf allen Serien einzeln nacheinander eingesetzt, die Verwendung des Projektionsvolumens wurde erst später eingeführt (s. u.). Bei der Segmentierung der Einzelserien zeigten sich jedoch Nachteile, da die Regeln idealisierte Annahmen voraussetzen.

So müssen die Bildpunkte mit den stärksten Anreicherungen nicht zwingend den Teil einer Läsion darstellen, sie können auch einem Gefäß zugehörig sein, verbliebene Bewegungsartefakte oder Bildrauschen darstellen. Besonders problematisch sind schwach anreichernde Tumore, die durch ein kräftiges Gefäß versorgt werden. Letzteres ist in der Regel gut kontrastiert und kann über stärkere SI verfügen als der Tumor selbst. Je nach Verlauf durch das diskrete Voxelraster kann das Gefäß aufgrund der limitierten Ortsauflösung in einem Bildpunkt in Mischung mit angrenzenden Strukturen dargestellt werden (*partial Volumeneffekt*). Dies beeinflusst die SI der „Gefäßvoxel“, wenn das Voxel nicht komplett im Inneren des Gefäßes liegt, sondern über den Rand hinausragt. Im ungünstigsten Fall wird der Hotspot nicht im Tumor, sondern im versorgenden Gefäß gefunden. Durch den

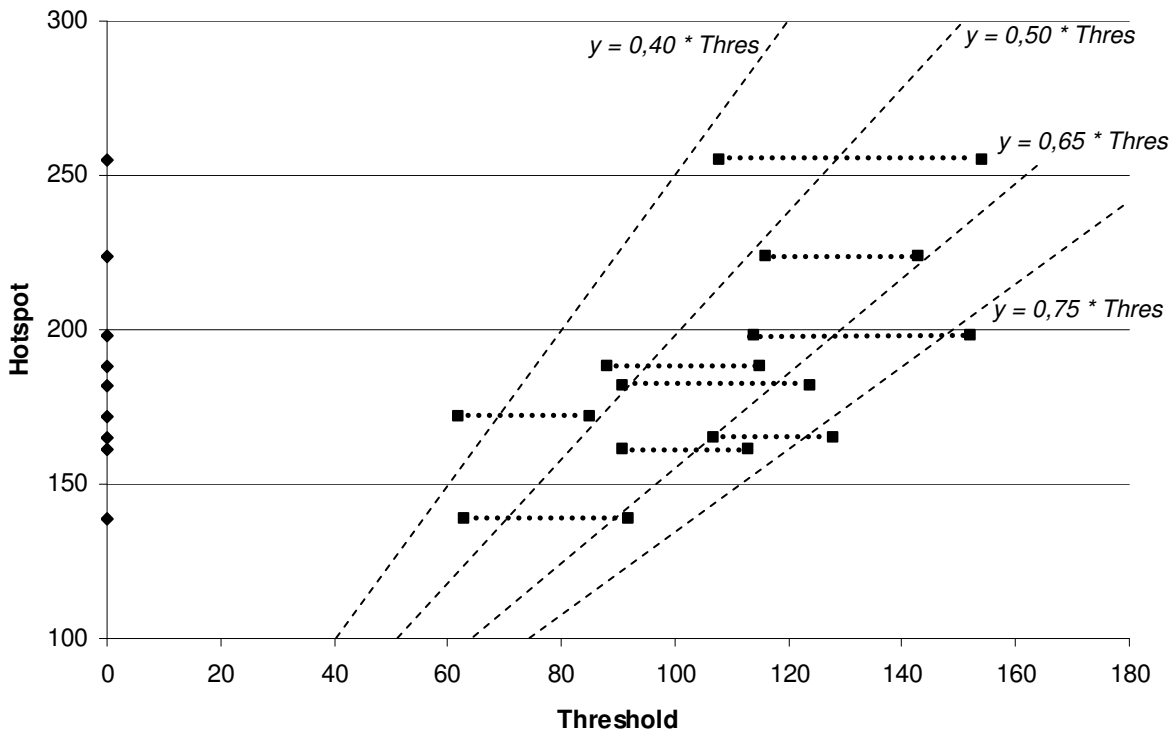


Abbildung 6.5: Signalintensitäten von Läsionenhotspots mit dem dazugehörigen Wertebereich optimaler Segmentierungsschwellwerte auf Bildeinzelserien.

partialen Volumeneffekt im Gefäßverlauf erreicht das Region Growing aufgrund der „Signalunterbrechungen“ aber nicht den Tumor. Wird der Tumor im späteren Verlauf durch einen niedrigeren Schwellwert segmentiert, der niedrig genug ist, um die Unterbrechung zu „überbrücken“, wird beim Erreichen des zuvor erkannten Gefäßsegments eine Überlagerung festgestellt. In der Konsequenz wird der gesamte Tumor vom Segmentierungsergebnis verworfen.

Das Erstellen einer Funktion für Regel (2) gestaltete sich bei der Segmentierung von Einzelserien schwierig. Eine Auswahl an Läsionen wurde in einem ersten experimentellen Schritt durch ein Region Growing mit unterschiedlichen manuell vorgegebenen Schwellwerten und Saatzellen segmentiert, um nach visueller Prüfung den optimalen Schwellwert festzustellen (s. Abb. 6.5). Dieser war nicht immer eindeutig, meist führte ein ganzer Bereich an Schwellwerten zu ähnlich guten Ergebnissen (gepunktete Intervalle). Gleichzeitig wurde der Signalwert des Hotspotsvoxels der Läsion zu jedem Wertebereich vermerkt (Markierung an der Hotspot Achse). Somit konnte der Bezug Hotspot SI zu dem Wertebereich des idealen Segmentierungsschwellwertes stichprobenartig für einige Läsionen aufgestellt werden; es wurde mit $y = Thres \cdot SI$ eine lineare Schwellwertfunktion aus den Daten abgeschätzt. Nach dieser Vorarbeit wurde mit dieser Funktion die Segmentierung mit verschiedenen Schwellwerten im Bereich $Thres \in \{0, 40 \dots 0, 75\}$ (s. gestrichelte Linie im Diagramm) nun auf alle Differenzserien (D_{SN_4} bis D_{SN_7}) des gesamten Kollektivs getestet

und nach erneuter visueller Analyse der Schwellwert $Thres = 0,65$ für das Gesamtkollektiv als geeignet erachtet.

Durch diese Schwellwertberechnung konnte zwar die Mehrheit der Läsionen zufriedenstellend segmentiert werden, in vielen Fällen waren jedoch deutliche Unter- oder Übersegmentierungen zu beobachten. In Abb. 6.5 lässt sich dies erkennen. Verläuft nämlich die gestichelte Linie einer Schwellwertfunktion links am Wertebereich vorbei, ist eine Übersegmentierung zu erwarten. Liegt sie dagegen über dem Bereich, führt das zu einer Untersegmentierung. Nur wenn die Funktion innerhalb des Bereiches abbildet, entspricht das Ergebnis der experimentell durchgeführten Segmentierung.

Die Beziehung zwischen SI des Hotspotvoxels und optimalem Schwellwert war nicht immer eindeutig. Mehrere Läsionen mit ähnlichen Hotspotsignalwerten bedurften sehr unterschiedlicher Schwellwerte für das Region Growing. Das Diagramm enthält Beispiele im Hotspotbereich von 160 SI.

Häufig wurde eine Läsion gleich in mehreren Einzelserien erkannt, wobei sie sich in ihrer Form sehr unterschiedlich darstellte. Meist war das Ergebnis nur in einer Serie zufriedenstellend.

Als Resümee der Segmentierung von Einzelserien muss festgestellt werden, dass der Informationsgehalt eines einzelnen Hotspotvoxels nicht ausreichend für die Schwellwertberechnung ist.

Projektionsvolumen Das Auswerten aller Einzelserien ist zuerst einmal sinnvoll, denn der zeitliche Verlauf des Anreicherungsprozess ist sehr variabel. Es ist nicht absehbar, auf welcher Serie die Läsionen deutliche Anreicherungen zeigen. Aufgrund des Anreicherungsprozesses kann sich dieselbe Läsion auf mehreren der aufeinanderfolgenden Zeitserien mit einer jeweils unterschiedlichen Morphologie darstellen. In der Folge könnte ein und dieselbe Läsion auf einer Serie als gutartig und auf einer anderen als bösartig klassifiziert werden. Hier entsteht ein Entscheidungsproblem um die Wahl der geeigneten Serie. Letztendlich möchte man nur eine Dignität pro Läsion vorhersagen. Zudem müsste zur Feststellung, ob zwei Token, die zu unterschiedlichen Zeitpunkten segmentiert wurden, überhaupt dieselbe Läsion darstellen, ein Positions- und Formvergleich der Objekte durchgeführt werden.

Durch den Einsatz des Projektionsvolumens entstehen viele Vorteile, durch die sich die zuvor genannten Schwierigkeiten bei der Segmentierung von Einzelserien verbessern lassen. Durch die Projektion als Erweiterung zum 4D-Vorgehen wird ein Zwischenschritt eingeführt, der sowohl den Positionsvergleich erübrigt als auch die Auswahl der geeigneten Einzelserie leistet. Somit kann vereinfacht eine Gesamtaussage für die Patientin berechnet werden.

Wie im Ergebnisteil bereits dargelegt, ist zur Erzeugung des Projektionsvolumens die Verwendung der Mittelwertfunktion vorteilhafter als die Maximumfunktion. Die Weichzeichnung durch die Mittelwertbildung (vgl. Mittelwertfilterung Kap. 2.4.2) reduziert Rauschen. Die Abbildung der Signalstärken von Bewegungsartefakten wird gegenüber den KM aufnehmenden Strukturen geschwächt. Als eine Folge werden im Projektionsvolumen weniger zusätzliche Funde segmentiert als auf jeder Einzelserie und somit die Anzahl an zusätzlichen Funden reduziert.

Die Projektion hat auch positiven Einfluss auf das zuvor beschriebene Problem signal-

reicher Gefäßversorgungen von Tumoren. Trotz Registrierung werden durch verbliebene geringfügige Bildverschiebungen die Abbildung der feinen Gefäße gegenüber den großvolumigen Tumoren geschwächt. Das Risiko, dass das versorgende Gefäß signalreicher erscheint und der Tumor aufgrund der Überlagerungsregel nicht segmentiert wird, ist wesentlich geringer.

Ganz entscheidend ist die Beteiligung des Projektionsvolumens an der Berechnung des Schwellwertes. Bei der Bearbeitung von Einzelerien konnte beobachtet werden, dass der Signalwert nur des einzelnen Hotspotvoxels nicht ausreichend für eine Ableitung des optimalen Segmentierungsschwellwertes durch die lineare Funktion ist. Durch die Projektion werden nun Informationen von vier Zeitserien einbezogen. Dadurch wird das Verhältnis zwischen Hotspot und optimalem Schwellwert durch die Mittelwertbildung deutlich stabiler.

Für die Vorsegmentierung auf dem Projektionsvolumen ist es besonders wichtig, möglichst alle Anreicherungen als Ausgangsbasis für die Nachsegmentierung zu erkennen. Es wird dabei nicht der Anspruch erhoben, den Randbereich der Objekte genau bestimmt zu haben. Die Morphologie der Läsion wird aufgrund der Projektion ohnehin nur angenähert dargestellt; hier ist die lineare Funktion zur groben Vorverarbeitung erst einmal ausreichend. Durch die Weichzeichnung ändert sich auch der Skalierungswert der Funktion f_{Thres} nun auf 0,55 gegenüber 0,65 der Einzelerien. Um möglichst viele Läsionen zu detektieren wurde der Skalierungswert defensiv gewählt (s. Tab. 5.6). Dadurch wird die Anreicherung im Projektionsvolumen tendenziell untersegmentiert und die Entstehung von unerwünschten Überlagerungen gemindert.

Da die Projektion die Morphologie der Läsionen verändert, muss zur genauen Formbestimmung eine Nachsegmentierung auf einer Einzelerie erfolgen. Die auf dem Projektionsvolumen erzeugten Token dienen als Schablonen, um die durchschnittliche SI pro Läsion für jede Einzelerie zu bestimmen. Jede Läsion wird derjenigen Einzelerie zugewiesen, auf der die maximale durchschnittliche SI erreicht wird (s. Abb. 4.1, $T_{SN_4-SN_7}$). Beginnend bei den signalstärksten Objekten, wird auf der entsprechenden Einzelerie durch den gleichen Algorithmus nachsegmentiert, der bereits auf dem Projektionsvolumen Anwendung fand.

Entscheidend ist, dass für die Nachsegmentierung nun der individuelle Schwellwert mit $f_{Thres}(SI) = 0,85 \cdot SI$ (s. Tab. 5.6) nicht aus dem Signalwert eines einzelnen Voxels, sondern aus dem erzielten durchschnittlichen Signalwert der Schablone berechnet wird. Er bezieht sich nun gleich auf eine Menge zur Läsion gehöriger Bildpunkte. Trotz Verwendung nur einer einfachen linearen Funktion kommt dem Ablauf zu Gute, dass zuerst mit der Segmentierung des Projektionsvolumens und anschließender Nachsegmentierung einer Einzelerie der individuelle Schwellwert *zweimal* pro Läsion berechnet wird.

Auch der zweite Skalierungswert von 0,85 wurde defensiv eingestellt. Dadurch wird – wie bereits beim Projektionsvolumen praktiziert – der Gefahr entgegen getreten, dass die Heuristik benachbarte Tumore aufgrund von Überlagerungen von der Ergebnisliste entfernt. Auch kann eine Übersegmentierung durch Anlagern benachbarter nicht anreichernder Strukturen die Form der Läsion stärker als eine Untersegmentierung verändern.

Der entwickelte Segmentierungsalgorithmus ist sehr sensitiv bei der Erkennung von Karzinomen. Jedoch konnten nicht alle gutartigen Läsionen segmentiert werden. Dies ist

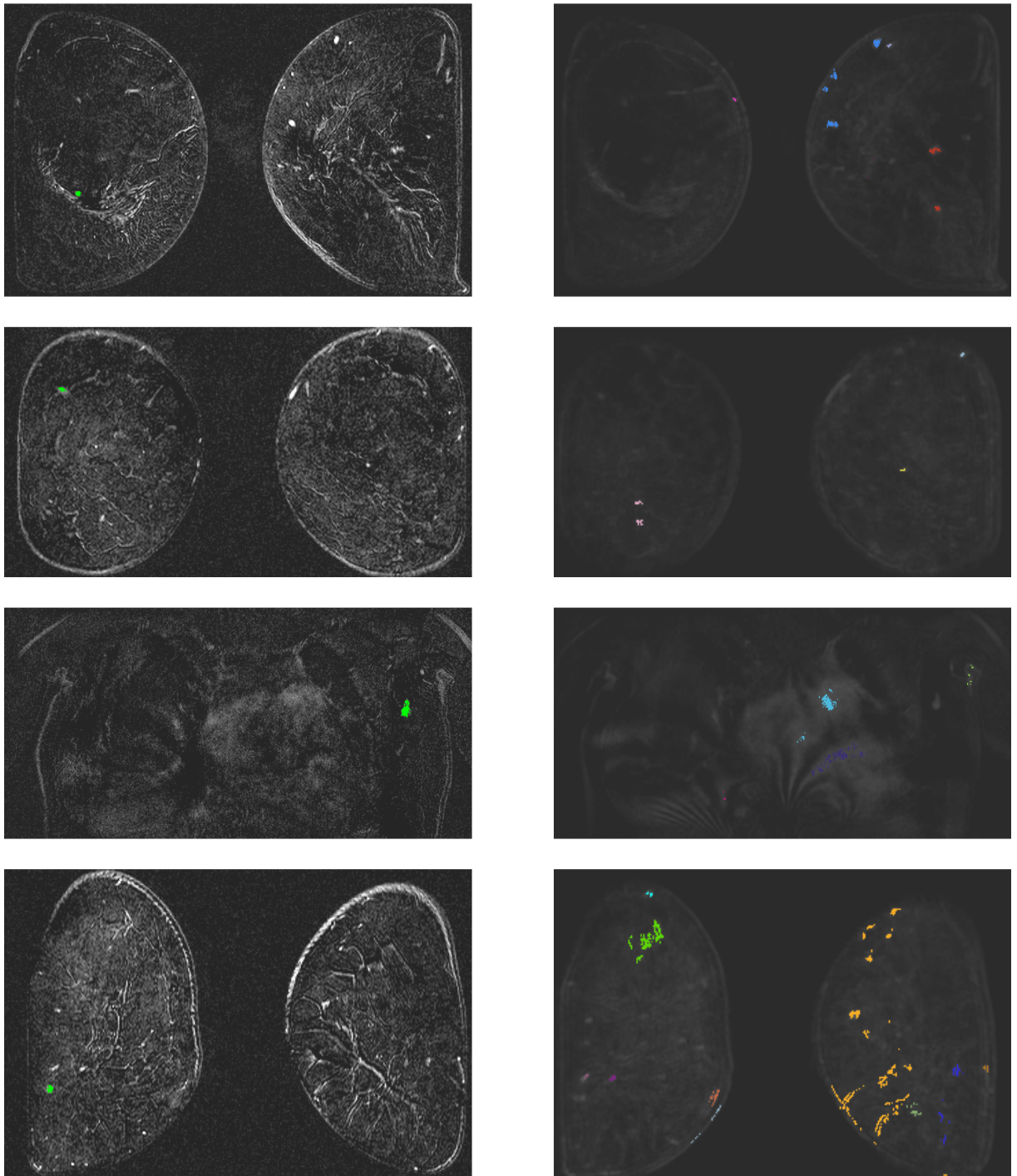


Abbildung 6.6: Vergleich verschiedener benigner GS-Vorgaben (grüne Markierung der Bilder der linken Spalte) und deren Vermissten in der Segmentierung des Projektionsvolumens (rechte Spalte).

nachvollziehbar, da die Segmentierung die Abgrenzung anhand der Signalanreicherung der Läsion vornimmt. Gutartige Funde reichern häufig nur auf den letzten Messungen und dort zudem nur schwach an. Die Mittelwertbildung des Projektionsvolumens kann deren Abbildung nicht mehr gewährleisten (s. Abb. 6.6). Das Versäumen der schwachanreichernden gutartigen Strukturen ist klinisch nicht weiter von Bedeutung. Die Segmentierung nimmt hier lediglich eine frühzeitige Klassifikation vor. Für den abschließenden Befund ist die Ausgabe vorhandener Karzinome maßgeblich.

Ob die Erkennungsrate der Segmentierung schwachanreichernder benigner Läsionen durch das Einbeziehen der späten transversalen Serie SN_8 erhöht werden kann, beschreibt eine noch ausstehende Aufgabe.

Nicht jede Erkrankung lässt sich durch ein Schwellwertverfahren geeignet segmentieren. Die KM-Verteilung innerhalb der Läsion und das Anreicherungsverhalten angrenzender Strukturen ist sehr variantenreich. Besonders mit Zunahme der Läsionengröße wird eine Segmentierung durch Schwellwerte unpräziser. Auch können großflächige Mastopathien oder diffuse Karzinome nicht immer als ein zusammenhängendes Objekt erkannt werden. In der Folge entstehen entweder Übersegmentierungen oder Untersegmentierungen. Letzteres erkennt einen einzelnen Erkrankungsherd gleich durch mehrere Objekte.

Das angenommene Verhältnis zwischen Hotspot und Schwellwert für die Nachsegmentierung hat sich für das gegebene homogene Kollektiv in der vorliegenden Arbeit bewährt. Inwieweit verschiedene Einflussfaktoren (s. Tab. 2.3) dies verändern können, ist Teil der Diskussion in Kapitel 6.3.3.

Abb. 6.7 zeigt den Fall eines 12-jährigen Mädchens mit einem großgewachsenen gutartigen Phylloides-tumor, der in den ventralen Bildschichten nicht mehr vollständig segmentiert wurde. Hier zeigt sich ein fallender Verlauf in den SI durch die Bildschichten, der im ventralen Bildbereich den Segmentierungsschwellwert unterschreitet. Je größer ein Tumor ist, desto mehr Inhomogenitäten können auftreten, die eine lokale Segmentierung durch Schwellwerte erschwert.

Auch können Gefäße nicht immer vollständig vom Objekt abgetrennt werden. Hier besteht die Gefahr der Verfälschung der sich anschließenden morphologischen Berechnungen (s. auch Kap. 6.3.3).

Durch Einsatz des Projektionsvolumens als Vorverarbeitung ergibt sich ein grobes, aber automatisiertes Zwischenergebnis. Für die Nachsegmentierung wurde der bestehende Algorithmus erneut unter minimaler Anpassung der Parameter angewendet. In der Literatur (s. Kap. 3.3.2) sind Top-Down Verfahren bekannt, die nach grober manueller Vorsegmentierung den Randbereich der Läsion automatisch erkennen [CGB06, ZBE⁺07, WSK⁺06]. Hier können sich Synergien ergeben, wenn das automatische Zwischenergebnis des Projektionsvolumens mit einem automatischen Top-Down Verfahren kombiniert wird. Als alternatives Nachsegmentierungsverfahren, das nicht direkt auf Schwellwerten beruht, bietet sich beispielsweise das Pyramidal Linking [LOPR97] an. Es nimmt als Bottom-Up Verfahren durch eine hierarchische Aufbereitung der Signalintensitäten in einer Bildpyramide (s. Kap. 6.1.4) für lokale Bereiche eine Binarisierung der Bildinformation aufgrund von Selbstähnlichkeit vor. Durch die lokale Sichtweise könnten dann alle Bereiche auch großer Läsionen wie in Abb. 6.7 genauer bestimmt werden.

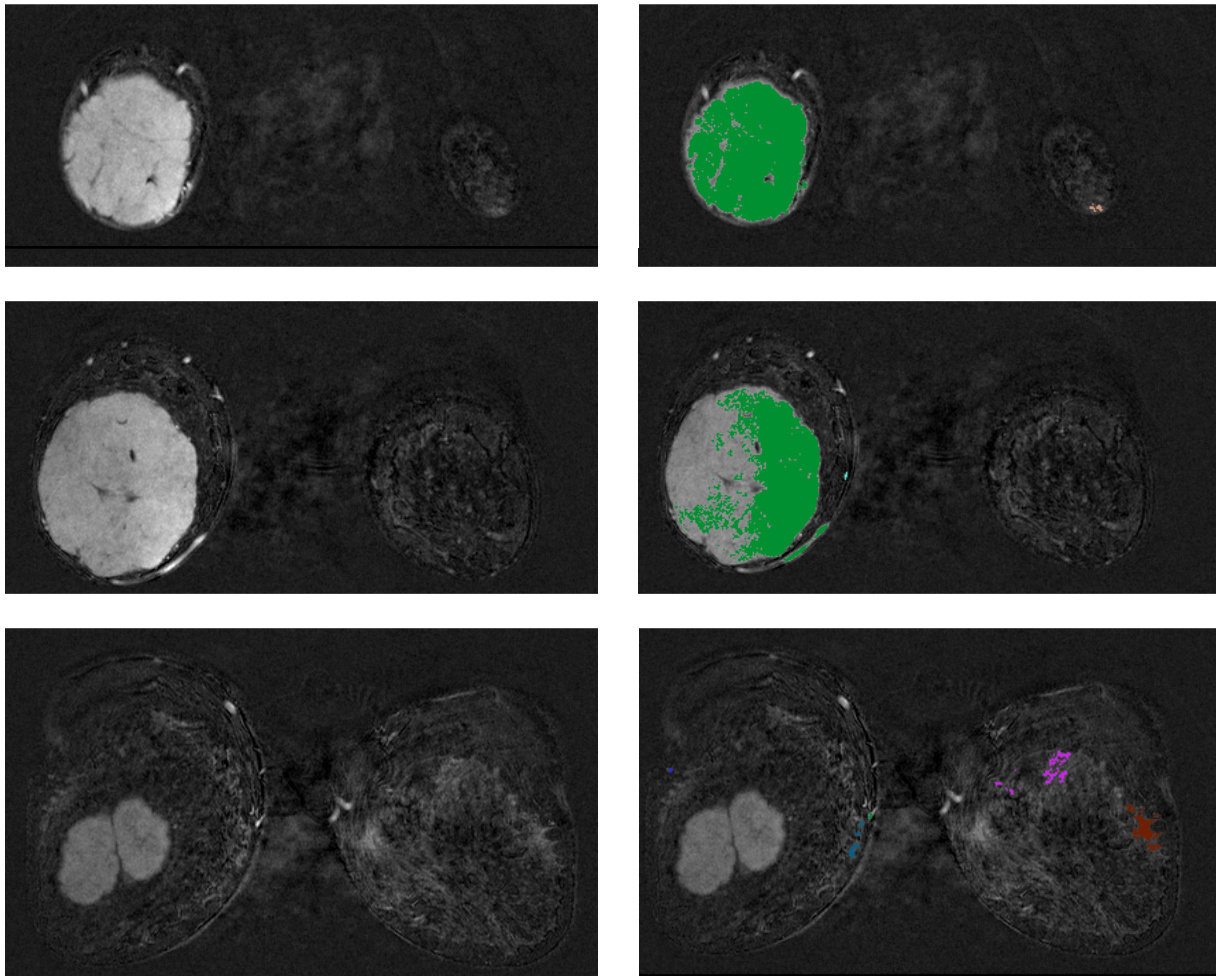


Abbildung 6.7: Verschiedene Bildschichten des Differenzvolumens derselben Patientin (linke Spalte) mit automatischer Segmentierung durch die Software (rechte Spalte).

6.2.2 Zusätzlich segmentierte Objekte

Die Zahl der als Tumor klassifizierten falsch positiven zusätzlichen Funde ist mit durchschnittlich 7,3 Objekten sehr hoch. Die Ursache dafür beginnt bereits mit der Segmentierung. Sie überführt mit ca. 134 erzeugten Token bei nur knapp 6 GS-Läsionen pro Patientin deutlich zu viel Bildinformation als relevant in das Intermediate- Level. Die Klassifikation erkennt die zusätzlichen Funde zwar mit 92 % sehr sensitiv (s. Kap. 5.4), aufgrund der großen Ausgangsmenge verbleiben letztendlich dennoch eine Vielzahl zusätzlicher Funde in der finalen Ausgabe der Software. Die zusätzlichen Funde der Segmentierung stellen sich häufig als kleine Objekte innerhalb des Bildrauschens dar, resultieren trotz Registrierung als längliche Objekte aus verbleibenden Bewegungsartefakten (vgl. auch Kap. 6.2.4) im Hautbereich oder beschreiben als größere Objekte Rückfaltungsartefakte besonders im Thoraxbereich (s. Kap. 6.4.5).

Zu deren Reduktion müssen verbesserte Regeln oder Kriterien in den Segmentierungsprozess einbezogen werden. Beispielsweise kann das Auswerten der zeitlichen Auflösung der Bilddaten sinnvoll sein. Gegenwärtig wird durch die Projektion der Informationsgehalt der Kinetik wenig genutzt. Die Kinetik wird erst während der Merkmalsextraktion für die anschließende Klassifikation umfassend charakterisiert (s. Kap. 4.4.3). Neben der Kinetik lassen sich noch weitere Informationen wie beispielsweise Kriterien aus der T_2 -gewichteten Messung heranziehen. Durch die Software bisher noch nicht berücksichtigt werden Informationen über die Lage der Läsionen in der Brust. Dadurch könnten zusätzliche Funde im Thorax- oder Mamillenbereich aufgrund der räumlichen Information frühzeitig erkannt werden. Als Bezugssystem wäre hierfür eine automatische Bestimmung von Positionsmarkern beispielsweise an den Brustwarzen und Brustbein Voraussetzung.

Um solche Kriterien in den Segmentierungsprozess einbeziehen zu können, müssen sie als Informationen pro Bildpunkt aufbereitet werden. Beispielsweise kann die Kinetik durch Bewertung der initialen und postinitialen Phase parametrisiert werden (s. Kap. 3.3.3). Die Position eines Bildpunktes kann als normiertes Abstandmaß im Raum zwischen Brustbein und Brustwarze formuliert werden.

Liegen mehrere Kriterien unterschiedlichen Grundcharakters pro Bildpunkt vor, ist die Bildung eines Merkmalsvektors und der Einsatz unüberwachter Low-Level Klassifikatoren für die Segmentierung sinnvoll (s. Kap. 2.4.4). Beschreiben die gesammelten Kriterien die anreichernden Bildbereiche eindeutig unterscheidbar zu den nicht relevanten Bereichen, werden die dazugehörigen Bildpunkte durch die Low-Level-Klassifikation derselben Klasse zugewiesen. Solche Segmentierungen, bei denen der Merkmalsvektor allerdings nur aus kinetischen Informationen besteht, sind in der Literatur hinreichend bekannt (s. Kap. 3.3.3). Sie dienen jedoch nur der Farbkodierung von Bildbereichen. Als zusätzlicher Schritt muss, wie beispielsweise in der vorliegenden Arbeit durch das Region Growing und Segmentierungsheuristik praktiziert, eine Segmentierung zur Gewinnung von 3D-Token erfolgen. Die Ergebnisklasse der unüberwachten Klassifikatoren beschränkt dabei nur die Menge an zu bearbeitenden Bildvoxeln. Die so erzeugten Token können das Segmentierungsergebnis des Projektionsvolumens ersetzen, abschließend kann unverändert der bestehende Nachsegmentierungsschritt auf den Einzelerien ausgeführt werden.

Der Nutzen aus der Kombination mehrerer Kriterien liegt in der Reduktion der Anzahl zusätzlicher Funde bereits im Segmentierungsschritt. Dadurch werden vermehrt Informationen, die als Merkmale in die High-Level-Klassifikation eingehen, in die Low-Level-Klassifikation verlagert. Es bleibt jedoch offen, ob die Anwendung dieser Kriterien in der Segmentierungs- oder Klassifikationskomponente sich als erfolgreicher erweist. Im High-Level-können Merkmale auf Grundlage der Token durch Betrachtung von Bildpunktmengen berechnet werden, während im Low-Level nur die Kriterien einzelner und damit informationsärmerer Bildpunkte verarbeitet werden können.

Eine weitere Möglichkeit zur Reduktion von zusätzlichen Funden ist die Verbesserung des Abbruchkriteriums. Derzeit stoppt die Segmentierung, wenn der berechnete individuelle Schwellwert f_{Thres} den Signalwert von SI_{Noise} aus Regel (3) unterschreitet, der fest auf einen Wert von 20 SI eingestellt wurde. Durch hohe Schwellwerte zu Beginn des iterativen Verfahrens wurden bereits zumeist die entscheidenden Läsionen segmentiert. Die niedrigen

Schwellwerte erkennen überwiegend Parenchym, Artefakten und Bildrauschen und resultieren vermehrt in zusätzlichen Funden. Durch das aktuell verwendete Abbruchkriterium wird auf jeden Fall sichergestellt, dass keine Läsion übersehen wird, da der Algorithmus stoisch die gesamte Brust abarbeitet. Die meisten Objekte resultierend aus Segmentierungen mit niedrigem Schwellwert lösen Überlagerungen aus; sie sind somit ohnehin kein Teil des Segmentierungsergebnisses (cyan zu orange Verhältnis in Abb. 5.4 u. 5.5). Sind Objekte mit schwacher Anreicherung jedoch gut abgegrenzt, wird keine Überlagerung ausgelöst; häufig erhält somit ein zusätzlicher Fund Einzug ins Intermediate-Level.

Eine Verbesserung wäre die Abschätzung des Wertes von SI_{Noise} aus dem Histogramm. Wie bei der Berechnung des Merkmals M_{MaxSI} kann dazu die Berechnung von Quantilen eingesetzt werden. Je stärker die Gesamtanreicherung im Bild ist, desto höher ist der berechnete Wert des Abbruchkriteriums und desto weniger schwach anreichernde Strukturen würden segmentiert.

Ein weiteres Abbruchkriterium lässt sich aus der Anzahl der bereits segmentierten Objekte herleiten. Der Algorithmus könnte gestoppt werden, wenn beispielsweise bereits 50 Token erzeugt wurden.

Entscheidend für die Wahl zusätzlicher Abbruchkriterien ist die Sicherstellung, dass dennoch alle relevanten Bildinformationen in Token überführt werden. Eine bereits erfolgreich angewendete Maßnahme zur Reduktion der Tokenanzahl um 56 % ist die Größenfilterung, die alle Objekte mit einem Durchmesser kleiner als 2,7 mm ($\approx 10 \text{ mm}^3$ bei Kugelform) in der Nachverarbeitung der Segmentierung verwirft. Tumore geringeren Durchmessers werden aufgrund noch ausbleibender Angiogenese durch die KM gestützte MRT kaum im Bild dargestellt (s. Kap. 2.1.3).

6.2.3 Vergleich zum Goldstandard

Folgend soll näher auf die Qualität der automatisch segmentierten Läsionen eingegangen werden. In der Literatur erfolgt eine Beurteilung der Segmentierungsergebnisse häufig nach visueller Bewertung von einem oder besser von mehreren Expertenradiologen. In der vorliegenden Arbeit liegt bereits eine manuelle GS-Segmentierung durch einen Expertenradiologen vor. Sie wurde zur Erzeugung der Lerndaten genutzt und wird darüber hinaus nun für eine automatische Evaluation der Segmentierung des Prototyps herangezogen.

Bewertet werden soll die Deckungsgleichheit der automatischen Segmentierung mit der GS-Vorgabe. Dazu werden die beiden Überdeckungskriterien O_{in} und O_{out} (s. Kap. 4.5.2) verwendet. Die Qualität wird folgend nach dem prozentualen Mindestanteil der Überlappung ($O_{in} = 30 \%$, 50% , 70% und 90%) gestaffelt, der Wert für die Übertagung wurde mit $O_{out} = 50 \%$ dagegen konstant gehalten. Tabelle 6.3 gibt an, wieviel Prozent der 141 GS-Läsionen den jeweiligen Überlappungswert erreicht haben.

Diese Ergebnisse der automatischen Segmentierung scheinen auf den ersten Blick eine schlechte Überdeckung mit dem GS zu bescheinigen, beispielsweise erreichen nur 12 % aller Läsionen eine Überlappung von mindestens 70 %. Jedoch müssen folgende Faktoren bedacht werden:

O_{in}	30 %	50 %	70 %	90 %
Läsionen	89 %	43 %	12 %	2 %

Tabelle 6.3: Relative Anzahl an automatisch segmentierten Läsionen, die einen der gestaffelten Überlappungswerte nach O_{in} mit dem GS erreichen

Der GS wurde auf den unregistrierten Differenzbildern der Serie SN_7 erhoben, während die automatische Segmentierung auf registrierten Subtraktionen arbeitet. Wie beim Erzeugen der Lerndaten (s. Kap. 4.5.2) muss der GS für einen Vergleich zuerst registriert werden. Die manuelle Segmentierung deckt die Anreicherung nicht immer exakt ab, denn für den Radiologen war auf den unregistrierten Bilddaten der Läsionenrand nicht immer deutlich erkennbar. Die Form der Läsion auf der Serie SN_7 kann sich zudem von den der Serien SN_4 , SN_5 oder SN_6 der Nachsegmentierung unterscheiden, GS und automatische Segmentierung werden häufig auf den Bildinformationen unterschiedlicher Zeitpunkte bestimmt. In Abb. 5.4 u. 5.5 ist ersichtlich, wie unvollständig der registrierte GS (a) die Form der Anreicherung auf der Serie der automatischen Segmentierung (d) abdeckt. Verstärkt werden die Ungenauigkeiten auch durch das Verfahren der GS-Registrierung, der aufgrund mangelnder Interpolationsmöglichkeiten die Binärinformation der manuellen Markierung nur der nächstgelegenen diskreten Rasterposition im Bild zugewiesen werden kann.

Auch hängen allgemein die Werte der Überdeckungskriterien stark von der Größe der Vergleichsobjekte ab, weil bei kleinen Objekten sich bereits durch minimale Verschiebungen der relative Anteil deutlich mindert.

Letztendlich wird der Vergleich mit einem GS durchgeführt, der durch nur einen Radiologen festgelegt wurde und daher als subjektiv anzusehen ist.

Die beiden segmentierten Läsionen in Abb. 5.4 u. 5.5 erreichen, obwohl sie sich sehr genau der sichtbaren Kontur der Anreicherung im Bild anpassen, nur eine Überlappung von 50 % bzw. 30 % im gutartigen Fall mit dem GS. Eine Überlappung von 90 % erreichten nur große Tumore wie der gutartige Phylloides tumor der 12 jährigen Patientin, obwohl gerade diese Segmentierung Mängel aufweist (s. Abb. 6.7)

Eine Überlappung von $O_{in} \geq 30\%$ stellt – nach visueller Überprüfung – bereits einen akzeptablen Wert dar, zur Erstellung der Lerndaten wurde sogar ein Schwellwerte von $O_{in} > 15\%$ und $O_{out} < 50\%$ gewählt (s. Kap. 5.4). Dadurch konnten mit 86 Objekten 99 % der bösartigen und mit 44 Objekten 81 % der gutartigen Läsionen in Übereinstimmung mit den GS-Vorgaben gebracht werden. Lediglich ein Karzinom erfüllte aufgrund einer Übersegmentierung den O_{out} Wert von 50 % nicht. Dieser Fall wird in Kapitel 6.4.6 ausführlicher vorgestellt.

Abschließend muss festgestellt werden, dass auf dem Boden der registrierten GS-Vorgabe zwar über ein Positionsvergleich eine vorhandene oder fehlende Segmentierung festgestellt werden kann. Aber für eine aufschlussreiche Aussage über die Formgenauigkeit ist sie nicht suffizient. Für die Erstellung der Lerndaten ist das Überdeckungskriterium jedoch anwendbar.

6.2.4 Einfluss der Registrierung auf die Segmentierung

Der Segmentierungsalgorithmus wurde versuchsweise auch auf den unregistrierten Bilddaten ausgeführt. Dazu wurden Differenzbilder aus den Serien TR_{SN_3} , SN_4 , \dots , SN_7 zur nativen Aufnahme SN_2 erzeugt. Wie im vorherigen Kapitel kann das Segmentierungsergebnis mit den GS verglichen werden, um den Einfluss der Registrierung zu bewerten.

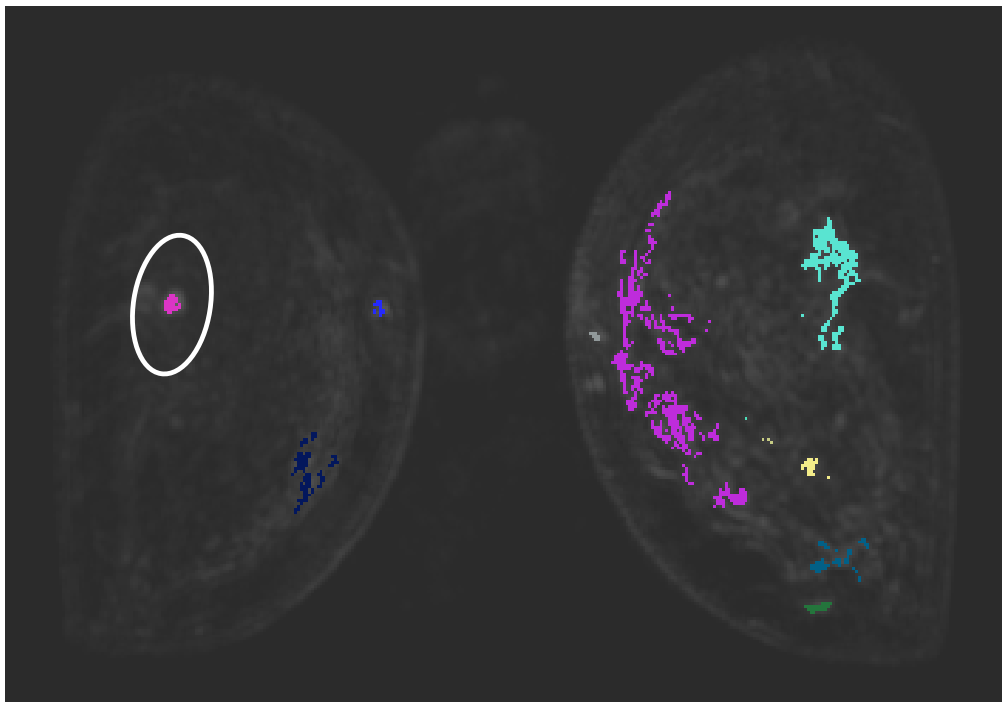
Im Ergebnis zeigten sich teilweise gravierende Unterschiede zu der Segmentierung auf registrierten Bilddaten. Es lassen sich in Abhängigkeit zur Stärke der im Bild enthaltenen Bewegung wiederkehrende Effekte beobachten. Der Grad an enthaltener Bewegung wurde auf den unregistrierten Bildsubtraktionen am Übergangsbereich Haut zur Luft bestimmt und in *gering* oder *stark* eingestuft.

Die Bilddaten von 69 Patientinnen enthielten nur geringe Bewegungseinflüsse. Die Segmentierung registrierter und unregistrierter Bilddaten führte dann auch zu ähnlichen Ergebnissen im Vergleich der Übereinstimmung mit den GS-Vorgaben. Nicht nur die Position und Form der segmentierten Läsionen waren sich ähnlich, auch die Gesamtanzahl an segmentierten Objekten glich sich. Die Verteilung der segmentierten zusätzlichen Funde wies jedoch deutliche Unterschiede auf (s. Abb. 6.8 a und b). Bereiche, die auf registrierten Bilddaten noch durch einzelne Objekte dargestellt wurden, werden auf den unregistrierten aufgrund der enthaltenen Bewegungsartefakte nun zu einem großflächigen vereint. Dadurch reduziert sich einerseits die Gesamtzahl an Objekten auf den unregistrierten Bildern, andererseits erhöht sie sich wieder durch die Segmentierung kleinerer Objekte in Bildbereichen, die nach der Registrierung ohne Bewegungseinfluss sind.

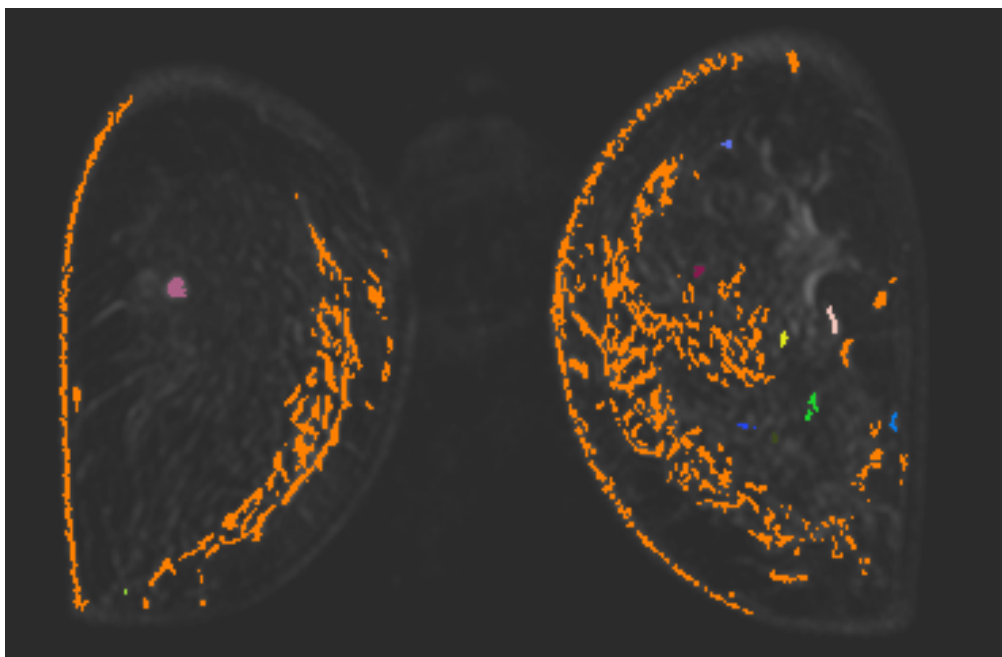
Das Auslassen der Registrierung hat dagegen auf den Bilddaten der verbleibenden 32 Patientinnen mit starken Bewegungseinflüssen gravierende Auswirkungen auf die Segmentierungsqualität. Die reichlich enthaltenen Bewegungsartefakte machen ein Abgrenzen der Läsionen kaum mehr möglich, die Anzahl an Gesamtobjekten verringert oder erhöht sich von Fall zu Fall. Als eine Folge entstehen starke Übersegmentierungen, in Abb. 6.9 a „verschmilzt“ der gefundene Tumor im registrierten Bild (weiße Markierung) zu einem einzigen Token im unregistrierten Bild, das die gesamte Brust vereinnahmt. In anderen Fällen (s. Abb. 6.9 d) kann der Tumor gar nicht segmentiert werden, hier verursachen benachbarte Bildartefakte höhere SI als die Erkrankung. Die Regeln der Segmentierung (s. Kap. 4.3.1) gehen jedoch davon aus, dass nur stärker anreichernde Bereiche Tumore beschreiben. Die schlechte Abgrenzung aufgrund der Bewegungsartefakte im Zusammenspiel mit den Regeln führen dazu, dass die eigentlichen Tumore von der Segmentierung kaum noch erkannt werden.

Das A-priori-Wissen der Regeln wurde unter der Voraussetzung aufgestellt, dass die Bilddaten keine Artefakte hoher SI enthalten. Dementsprechend verarbeitet die Segmentierung Bilddaten mit starken Bewegungseinflüssen nur mangelhaft. Besonders die verursachten Übersegmentierungen würden bei der anschließenden Morphologieberechnung zu völlig kompromittierten Werten führen.

Die Überdeckung der segmentierten Objekte mit den GS-Vorgaben (Überdeckungskriterien $O_{in} > 15\%$, $O_{out} < 50\%$) sind in Tabelle 6.4 aufgeführt, sie verschlechtert sich auf den unregistrierten Bilddaten erheblich (vgl. vorheriges Kapitel). Es können nur noch 56



(a)



(b)

Abbildung 6.8: Gegenüberstellung der Segmentierung (a) registrierter und (b) unregistrierter Projektionsvolumen. Die Markierungen zeigen die jeweilige Läsion.

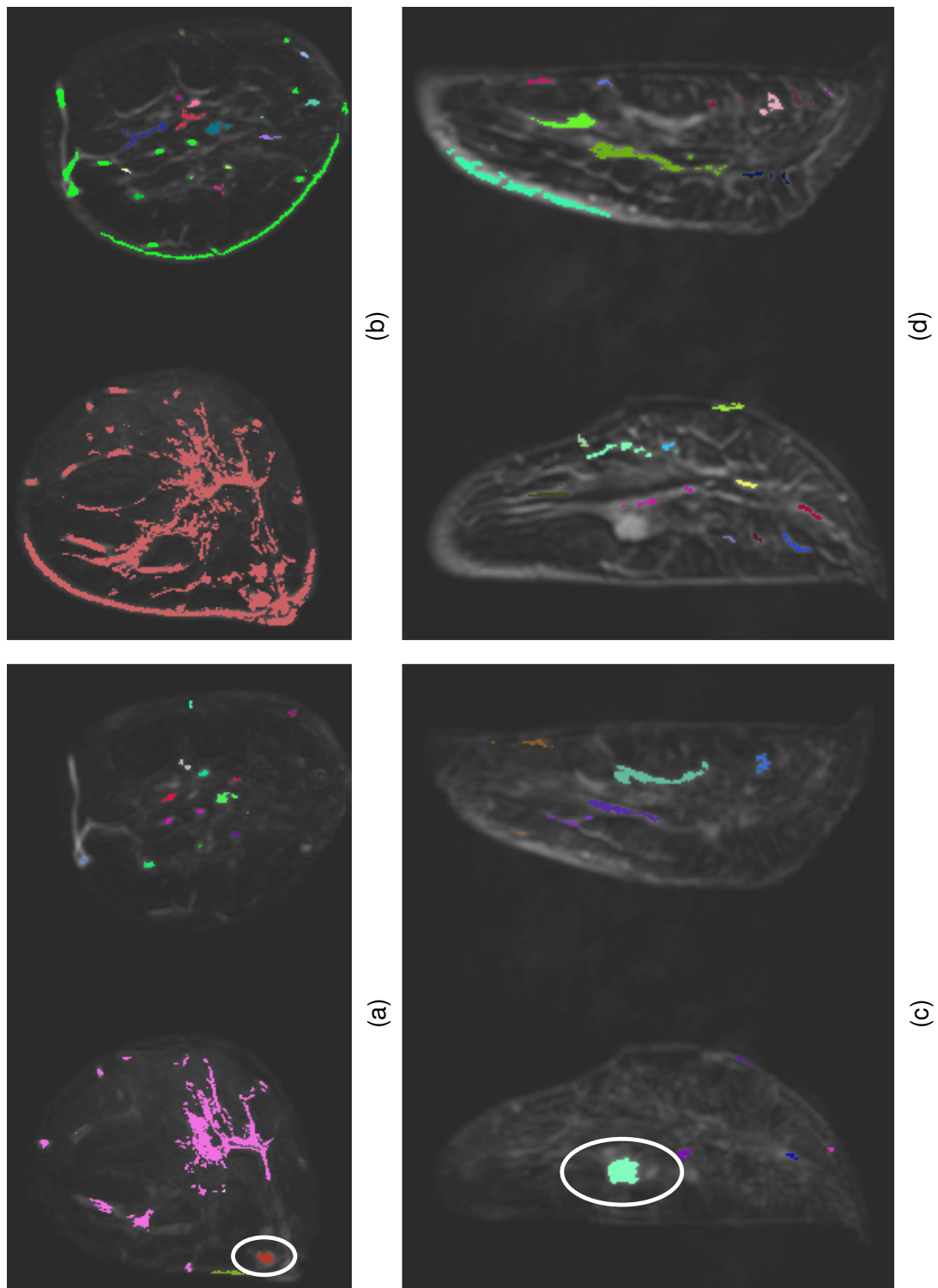


Abbildung 6.9: Gegenüberstellung der Segmentierung (a und c) registrierter und (b und d) unregistrierter Projektionsvolumen. Die Markierungen zeigen die jeweilige Läsion.

	registriert	unregistriert
maligne Läsionen	99 %	64 %
benigne Läsionen	81 %	31 %

Tabelle 6.4: Relativer Anteil an GS-Überdeckungen ($O_{in} > 15\%$, $O_{out} < 50\%$) maligner und benigner Läsionen auf registrierten im Vergleich zu unregistrierten Bilddaten

maligne und 17 benigne Objekte in Übereinstimmung gebracht werden. Viele der schwach anreichernden gutartigen Läsionen werden derart übersegmentiert, dass mehr als 50 % ihres Volumens außerhalb des GS-Objektes liegen.

6.2.5 Erkenntnisse und Grenzen

Die vorgestellte Segmentierung stellt ein sensitives Verfahren zur Erkennung von Karzinomen dar. Das Projektionsvolumen als Zwischenschritt vereinfacht die Berechnung von Aussagen pro Patient gegenüber der Segmentierung einzelner Serien. Auch verbessert es die Berechnung der individuellen Schwellwerte für die Nachsegmentierung einzelner Läsionen (s. Kap. 6.2.1). Ohne die Registrierung als Vorverarbeitung wäre eine Schwellwertsegmentierung nicht erfolgreich (s. Kap. 6.2.4). Die Segmentierung zeigt Schwächen in der Erkennung spätanreichernder gutartiger Strukturen, deren klinische Relevanz jedoch gering ist. Besonders die Segmentierung großer Läsionen scheint durch ein Schwellwertverfahren ungeeignet, meist werden nur Teilbereiche der Erkrankung erkannt. Hierzu wurden Verbesserungen vorgestellt.

Ungünstig ist letztlich die hohe Anzahl an zusätzlichen Funden, Verbesserungen zu deren Reduktion bereits im Segmentierungsschritt wurden in Kapitel 6.2.2 aufgezeigt. In der aktuellen Anwendung wird die Strategie verfolgt, jede noch so unscheinbare Struktur zuerst einmal zu segmentieren und erst später im Klassifikationsschritt anhand der gesammelten Merkmale zu entscheiden, ob das Objekt für die Diagnose relevant ist.

Eine abschließende Bewertung der Qualität der Segmentierung durch die GS-Vorgabe zeigte sich aufgrund verfahrenstechnischer Limitierungen als wenig praktikabel (s. Kap. 6.2.3). Denn der GS wurde auf der unregistrierten Bildserie SN_7 erhoben, während die automatische Segmentierung auf den registrierten Bildserien SN_4 bis SN_7 erfolgt. Zwar ist offensichtlich, dass sich die Segmentierung gut der Form der Anreicherungen im Bild anpasst. Dies müsste ergänzend durch die Bewertung mehrerer Radiologen jedoch noch statistisch evaluiert werden. Um nur die Anzahl der mit den GS-Vorgaben übereinstimmenden automatischen Segmentierungen zu bestimmen, ist die Anwendung der Überdeckungskriterien genügend.

6.3 Merkmalsextraktion

Die Merkmalsextraktion ahmt den Arbeitsschritt der Befundung des Radiologen nach, die gesammelten Informationen werden abschließend zur Vorhersage der Dignität verwendet.

Während die Algorithmen der Registrierung und der Segmentierung Parameter benötigen, um sie auf die Bearbeitung der variantenreichen Inhalte des Low-Level zu adaptieren, sind für die morphologischen und kinetischen Merkmalsextraktion auf den Token des Intermediate-Levels keine Einstellungsparameter mehr nötig. Denn die relevante Bildinformation des Low-Levels liegt im Intermediate-Level bereits abstrahiert vor (vgl. Kap. 2.4.1). Lediglich die Interpretation der T_2 -gewichteten Messung durch einen Schwellwert (s. Kap. 4.4.4) führt wegen Einbeziehung einer eigentlichen Segmentierungsaufgabe zu einem verbleibenden Parameter (s. Tab. 5.7).

Die Segmentierung erzeugt für jede verdächtige Läsion ein eigenes Token. Erst dadurch wird die Voraussetzung geschaffen, neben den kinetischen auch morphologische Kriterien zu berechnen.

Die Erhebung morphologischer Merkmale in der Literatur (s. Kap. 3.3.3) lässt sich in zwei Vorgehen unterscheiden. Häufig wurden durch einen Radiologen bewertete Merkmale, seltener algorithmische Berechnungen durch eine Software quantifiziert. Letzteres setzt entweder eine automatische oder eine zeitaufwändige manuelle Segmentierung der Läsionen voraus (vgl. Kap. 6.2.2).

Der Informationsgehalt mathematischer Merkmale und die durch einen Radiologen festgelegten visuellen Beschreibungen weichen voneinander stark ab (s. Kap. 6.3.2). Die Quantifizierung radiologischer Vorgaben umfasst meistens nur wenige Merkmale, diese sind dafür durch die visuelle Interpretation des Menschen bereits sehr informationsreich. Automatische Verfahren stellen dagegen mehr Merkmale zur Verfügung, von denen dann häufig die signifikantesten selektiert werden (s. Kap. 6.4.2). Folgend werden nur Publikationen diskutiert, die auf mathematischen Merkmalsberechnungen basieren.

Der vorgestellte Merkmalsumfang des Prototyps zur Läsionenbeschreibung liegt mit einer Anzahl von 24 im Vergleich zu bekannten Publikationen quantitativ höher. Mit Kinetik, Größen-, Form-, Textur- und Randschärfenbeschreibungen ist er auch wesentlich facettenreicher gegenüber den in der Literatur zumeist untersuchten Kinetiken und Randeigenschaften [HHS⁺86, HWP⁺89, GGB98, Kai93, Fis00, BMK⁺01, GDM⁺02]. Mit dem Einsatz von Algorithmen wie der Skelettierung wird hier auch ein neuer Ansatz zur Merkmalsbeschreibung in der Brust-MRT genutzt. Auf die Implementierung aufwendiger Einzelmerkmale wurde bewusst verzichtet, da auch die Laufzeit der Prototypsoftware für eine Diagnoseausgabe in einem akzeptablen Bereich liegen sollte (s. Kap. 6.3.2).

Im Gegensatz zu den morphologischen Merkmalen ist die Kinetik schon länger Gegenstand von Publikationen. Knopp et al. [KHR⁺02] quantifizierten verschiedene schwellwertbasierte Verfahren, manuelle Verfahren und ein einfaches pharmakokinetisches Modell. Sie konnten damit zeigen, dass zeitabhängige Schwellwertverfahren gegenüber reinen Schwellwertverfahren die Charakterisierung von Läsionen wesentlich verbessern. Letzteres Verfahren bezieht die Berechnung der Steigung auf einen zuvor definierten festen Zeitpunkt, während die zeitabhängigen Verfahren (s. Formel 2.1 u. 2.2) zur Maximalwertbestimmung

ein Zeitintervall zulassen. Bessere Ergebnisse zeigte das als manuelles Verfahren bezeichnete Vorgehen mit Betrachtung des TTP, auf dem auch die Berechnung in der vorliegenden Arbeit beruht. Das pharmakokinetische Modell erreichte eine zum manuellen Verfahren vergleichbare Leistung. Weiter zeigt Knopp auf, dass eine Normierung auf das Kontrastmittelangebot im Gefäßsystem (thorakalen Aorta) eine weitere Verbesserung darstellt. Das damit modifizierte manuelle Verfahren erreichte die höchste Treffgenauigkeit.

Pharmakokinetische Modelle (s. Kap. 3.3.3) wurden in der vorliegenden Arbeit nicht berücksichtigt, da sie für den Einsatz in eine automatische Software wenig praktikabel sind. Zum einen benötigen sie als Eingabefunktion die vorliegende Kontrastmittelkonzentration. Diese muss durch eine Messregion innerhalb eines Gefäßes möglichst nahe am Auswertungsort bestimmt werden. Die häufig dazu verwendete Aorta liegt jedoch außerhalb des FoV der koronaren Bilddaten; sie wäre jedoch am einfachsten zu segmentieren gewesen. Innerhalb der Brust befinden sich nur feinere Gefäße, die schwer segmentierbar sind und die KM-Konzentration aufgrund von Partialvolumeneffekten nicht zuverlässig zu bestimmen wäre. Zum anderen erfordern die pharmakokinetische Modelle eine höhere zeitliche Auflösung, die in üblichen Protokollen für die Brust-MRT nicht gegeben ist. Die Bilddaten müssen zur Bestimmung der Morphologie auch über eine ausreichend räumliche Auflösung verfügen [CGC02]. Pharmakokinetische Modelle basierend auf der Signalanreicherung nehmen zudem fehlerhaft an, dass die Anreicherung linear proportional zur KM-Konzentration ist [ABBM05].

Im nächsten Kapitel werden die 24 erhobenen Merkmale auf Vollständigkeit und Beschreibungskraft im Vergleich zum BI-RADSTM-Standard diskutiert. Kapitel 6.3.2 arbeitet die Vorgehensweise des Prototyps im Unterschied zum Radiologen in der Diagnoseerstellung heraus, während sich Kapitel 6.3.3 mit der Robustheit der Merkmale gegenüber Veränderungen in Bildqualität und deren besonderer Einfluss auf die Parametrisierung der Kinetik beschäftigt. Abschließend wird durch Cluster- und Hauptkomponentenanalyse (s. Kap. 4.6.3) das Zusammenwirken der Merkmale untersucht (s. Kap. 6.3.4 und 6.3.5). Während damit zunächst einmal die unbekanntenen Strukturen der Merkmale untereinander aufgezeigt werden, ist die Analyse der Beiträge der Merkmale zur Differenzierung maligner und benigner Tumore Bestandteil des Kapitels 6.4. Abschließend erfolgt die Zusammenfassung der Erkenntnisse und Grenzen (s. Kap. 6.3.6).

6.3.1 Merkmalsvergleich mit BI-RADSTM

BI-RADSTM standardisiert zur Beschreibung der Läsionen Terminologien, die der visuellen Wahrnehmung des Menschen von Mustern entsprechen. Folgend soll untersucht werden, welche Aussagen die algorithmisch motivierten Merkmale der vorliegenden Arbeit haben und inwieweit sie dem deskriptiven Umfang BI-RADSTM für die Befundung genügen.

Größe Zur Gruppe der Größen beschreibenden Merkmale gehören das Volumen M_{Volume} und die beiden Durchmesser $M_{MinDiameter}$ und $M_{MaxDiameter}$. Der Maximaldurchmesser ist ein durch die minimale Boundingbox geschätzter Wert (s. Abb. 4.6), der nur für rundliche Formen einen korrekten Wert annimmt. Im ungünstigsten Fall eines länglichen Objektes, das diagonal durch die Box verläuft, wird der tatsächliche Durchmesser unterschätzt.

Der BI-RADSTM -Standard sieht keine Größen beschreibenden Aussagen vor. Daher werden diese 3 Merkmale auch nicht zur Diskriminierung maligner und benigner Tumoren eingesetzt, sondern lediglich zum Klassifizieren von zusätzlichen Funden.

Form Die Form einer Läsion wird durch vier Merkmale beschrieben. M_{Form1} differenziert durch das Verhältnis minimaler und maximaler Durchmesser längliche von kompakten Objekten. Gleiches Augenmerk hat M_{Form2} , das das Verhältnis der Oberfläche zum Volumen betrachtet. Durch die Kugelform kann ein Volumen durch eine minimale Oberfläche begrenzt werden, jede andere Form vergrößert die Oberfläche im unterschiedlichen Maße. Für M_{Form3} wird auf das Skelettierungsergebnis zurückgegriffen. Objekte mit identischen Skelettliniendistanzen (s. Abb. 4.5 a und d) nehmen den Höchstwert von 1,0 an, während ein niedrigerer Wert auf uneinheitliche Abstände und somit eine unregelmäßige Form schließen lässt. Mit der fraktalen Dimension in M_{Form4} wird eine Aussage über die Raumfüllung gegeben.

BI-RADSTM sieht für die Beschreibung der Form Terminologien wie *rund*, *oval*, *lobuliert* und *irregulär* vor (s. Tab. 2.7). Unter den 24 Merkmalen der Software befindet sich jedoch keines, das diese Formbeschreibungen direkt wiedergeben könnte. Stattdessen liegen einfache mathematische Beschreibungen vor, die diese Aussagen wie beispielsweise M_{Form2} indirekt bereitstellen, jedoch keine visuelle Bedeutung haben (s. auch Kap. 6.3.2).

Begrenzung Zur Charakterisierung der Begrenzung werden zwei Merkmale berechnet, die sich grundlegend unterscheiden. $M_{Begrenzung1}$ basiert ähnlich dem Merkmal M_{Form3} auf dem Skelettierungsergebnis. Diesmal wird die Größe der Skelettlinie bewertet, denn je komplexer die Objektfläche ist, desto mehr Ausläufer und somit Skelettliniendistanzen bilden sich entsprechend der Anzahl der topologisch zu erhaltenden Ecken. Eine idealisierte Kugel würde in einem einzigen Skelettpunkt resultieren (s. Abb. 4.5 a).

Mit $M_{Begrenzung2}$ wird die Signalintensität des Objektrandes interpretiert, indem die Anreicherung des Objektinneren zum -äußeren ins Verhältnis gestellt wird. Da der Rand des Tokenobjektes durch den individuellen Schwellwert der Segmentierung festgelegt wird, hat eine Änderung des Wertes besonders großen Einfluss auf dieses Merkmal.

Während das erste Begrenzungsmerkmal BI-RADSTM -Eigenschaften (s. Tab. 2.7) wie *glatt*, *irregulär* oder *spikuliert* wiedergibt, bewertet $M_{Begrenzung2}$ eine *scharfe* oder *unscharfe* Randeigenschaft. Auch hier stellen die berechneten Merkmale keine binäre Aussage dar, sondern bilden den Quotienten in einem bestimmten Wertebereich ab.

Textur Besonders für die Beschreibung der endotumoralen Anreicherung existiert im BI-RADSTM -Lexikon aufgrund der guten Mustererkennung des menschlichen Auges eine Vielzahl an Terminologien (s. Tab. 2.7). Die Software ist nicht in der Lage, derart feine Unterscheidungen wiederzugeben. Sie beschreibt die Unregelmäßigkeiten in der Anreicherung durch drei Merkmale und bewertet dabei die Signalverteilung und Hohlräume im Tokenobjekt, die durch nicht oder nur schwach anreichernde Anteile im Objektinneren entstehen. $M_{Kontrast}$ wird durch die Varianz zum mittleren Signalwert bestimmt und entspricht unserem visuellen Verständnis des Kontrasts. Er unterscheidet homogene (geringe Varianz) von heterogenen (hohe Varianz) Anreicherungen. Die durch $M_{Nekrose}$ und $M_{Ringenhancement}$ beschriebenen Hohlräume beziehen sich direkt auf die BI-RADSTM -Terminologien Rand-

ständig und Septierung. Ersteres Merkmal bewertet den größten eingeschlossenen Hohlraum, letzteres die Gesamtheit aller Hohlräume. Auch hier stellen alle drei Merkmale keine deskriptive Aussage dar, vielmehr haben sie durch Fließkommawerte für den menschlichen Betrachter wenig Bedeutung.

Während die bisher diskutierten Merkmale alle zur Beschreibung von Herdbefunden dienen, standardisiert BI-RADSTM weitere Terminologien für nicht raumfordernde Anreicherung. Die Software kann zwischen den beiden Arten von Funden nicht unterscheiden. Unabhängig davon, ob die Heuristik der Segmentierung einen nicht raumfordernden Fund nur als ein Objekt oder als eine Vielzahl an Token detektiert, verarbeitet die Software jedes Objekt als einzelnen Herdbefund weiter. Die interne Charakteristik der raumfordernden Enhancements kann nicht erkannt werden, denn sie wird durch kein Merkmal beschrieben. Es gibt keinen Verarbeitungsschritt, der sich mit der räumlichen Anordnung der Token untereinander befasst. Daher kann die Software auch keine BI-RADSTM definierten Symmetrien durch Vergleich der Funde beider Brüste feststellen.

Kinetik Durch BI-RADSTM wird das Verhalten der Kinetik durch Bewerten der initialen und postinitialen Phase beschrieben (s. Abb. 2.4 b). Um eine Unabhängigkeit zur Bildinformation zu erreichen (s. Kap. 6.3.3), charakterisiert der Prototyp die Kinetik durch fünf normierte Merkmale. Für eine Vergleichsmöglichkeit der eigenen Merkmale wird folgend die klassische Berechnung durch Fischer (s. Formel 2.1 und 2.2) herangezogen, aus der die BI-RADSTM -Terminologien abgeleitet werden können.

Mit M_{TTP} wird ein Merkmal in der Einheit Sekunden berechnet, das die Aussage der initialen Phase charakterisieren soll. Die Kinetik typischer Karzinome zeigt maximale Signalintensitäten in der initialen Phase. Die BI-RADSTM -Terminologien *langsam*, *mäßig* oder *stark* sind Synonyme für die Steigung des Signals in dieser Phase. Je größer die Steigung, desto mehr spricht der Verlauf der initialen Phase für ein Karzinom. Nach Fischer wird die Steigung ausgehend vom SI Wert in der nativen Messung zur maximalen SI nach der KM-Gabe innerhalb der ersten ersten bis dritten Minute bestimmt (s. Formel 2.1). Durch Betrachten der verstrichenen Zeit zur maximalen Signalintensität kann M_{TTP} eine vergleichbare Information bereitstellen, da typische Karzinome dieses Maximum früh erreichen, während sich gutartige Tumore infolge des kontinuierlichen Anstiegs durch ein spätes Maximum auszeichnen. Ein kleiner Zeitwert signalisiert also Bösartigkeit, hohe Werte dagegen Gutartigkeit. Der Merkmalswert von M_{TTP} korreliert antiproportional zum Wert der Steigung. Jedoch wird die Steigung nach Fischer nur innerhalb der ersten bis dritten Minute bestimmt, während das Zeitmerkmal den Verlauf der gesamten Kinetik involviert.

Auch die postinitiale Phase wird klassisch durch die Steigung charakterisiert. BI-RADSTM standardisiert hierfür die Terminologien *Persistent*, *Plateau* oder *Wash-Out*. Auch diese Bestimmung leitet sich nach Fischer aus den Signalintensitäten ab. Es wird der Verlauf des Maximums der initialen Phase zum letzten Messwert bewertet (s. Formel 2.2). Das eigene Merkmal $M_{WashOut}$ folgt letztendlich der Berechnung des Fischer Merkmals. Allerdings bezieht sich die Steigung von $M_{WashOut}$ nicht auf das Maximum innerhalb der initialen Phase, sondern auf das Maximum zum Zeitpunkt M_{TTP} .

Mit den beiden Merkmalen M_{TTP} und $M_{WashOut}$ wird die initiale und postinitiale Phase charakterisiert und genügt somit bereits dem geforderten BI-RADSTM -Umfang. Zusätzlich

werden noch zwei Flächenmerkmale bestimmt; M_{AUC} und M_{AUCTTP} charakterisieren die durch die Kinetik gebildete Fläche und sind bezüglich der Baseline und dem maximalen Signalwert normiert; der Merkmalswert liegt immer im Wertebereich $M_{AUC}, M_{AUCTTP} \in [0, 0 \dots 1, 0]$. Der Nutzen beider Merkmale wird im Kapitel 6.3.3 im Zusammenhang mit der angestrebten Signalintensitätsunabhängigkeit diskutiert.

Da die initiale Phase durch einen Zeit- und nicht durch einen Signalwert charakterisiert wird, erfolgt keine Beschreibung basierend auf dem Absolutwert der maximalen Anreicherung. Folgend wird mit M_{MaxSI} auch ein Merkmal zugelassen, das den maximalen Signalwert beschreibt, ihn aber durch das Bildhistogramm normiert. Liegen allgemein im Bild die gemessenen SI höher, erhöht sich auch der oberste Quartilwert des Histogramms. Mit diesem Normierungswert kann für jede anreichernde Struktur eine relative Aussage zum Gesamtanreicherungsverhalten im Bild erzeugt werden. Allerdings wird der Quartilwert auch durch die Menge und Größe der anreichernden Strukturen und zusätzlich durch vorhandene Bildartefakte beeinflusst, die Normierung ist somit nicht gänzlich unabhängig.

Alle Kinetikmerkmale werden auf zwei alternativen Zeit/Signalintensitätskurven berechnet. Da besonders größere Läsionen ein sehr uneinheitliches endotumorales Anreicherungsverhalten zeigen können, kann mit der Hotspot ROI durch eine lokale Betrachtung eine abweichende Aussage entstehen. Der Hotspot wird in der vorliegenden Arbeit im signalstärksten Bereich der Läsion auf der Serie der Nachsegmentierung bestimmt. Er umfasst mindestens 20 Voxel. BI-RADSTM empfiehlt, den Hotspot einer Läsion durch einen Bereich von mindestens 3 Bildpunkten zu untersuchen. Bei mehreren vorhandenen Hotspots soll derjenige gewählt werden, der die verdächtigste Kinetik aufweist. Die Software führt einen solchen Vergleich nicht durch, sondern wählt immer den Bereich der stärksten Anreicherung.

Bei erfolgreicher Registrierung ist gewährleistet, dass sich unter dem Hotspotbereich in den anderen Bildserien der gleiche Gewebeabschnitt verbirgt.

6.3.2 Mensch versus Maschine

Das vorherige Kapitel macht deutlich, dass ein Vergleich der mathematisch motivierten Berechnungen mit den deskriptiven BI-RADSTM-Merkmalen kaum möglich ist. Mit visueller Wahrnehmung des Menschen und den mathematischen Werkzeugen der Software prallen zwei Vorgehensweisen aufeinander, die kaum Gemeinsamkeiten besitzen.

Als einfaches Beispiel kann die Bewertung der Form des Objektes herangezogen werden, die durch die BI-RADSTM-Terminologien *rund*, *oval*, *lobuliert* und *irregulär* beschrieben wird. Möchte man eine vorhandene Rundheit eines Objektes direkt als deskriptive Aussage mathematisch formulieren, müsste dazu beispielsweise das Formmodell einer Kugel durch Setzen des Mittelpunktes und Anpassen des Radius auf das Token gefittet werden. Dies ist bereits eine sehr rechenintensive Operation, denn der Abstand der Oberfläche des Modells muss mit der des Tokens verglichen werden. Der Gewinn dieser Berechnung liegt in einer einzigen beschreibenden Aussage. Bei einer etwas weitsichtigen Implementierung lässt sich durch das gleiche Modell auch gleichzeitig ein ovales Objekt fitten, wenn anstatt von einem zwei orthogonal zueinander stehende Radien variiert werden können. Zur Berechnung der

Aussage *lobuliert* bedarf es wesentlich komplexere Algorithmen.

Sowohl die Entwicklungs- als auch Ausführungszeit von Algorithmen zur direkten Herleitung einzelner deskriptiver Aussagen ist unverhältnismäßig. Die Beispiele machen aber auch deutlich, dass solche Berechnungen bereits Entscheidungen treffen müssen. Um wie viel darf die Fläche eines Tokens von der gefitteten Kugel abweichen, um noch das Kriterium der Rundheit zu erfüllen? Wie groß muss das Verhältnis der beiden Radien sein, um eine Kugel von einem ovalen Objekt zu unterscheiden? Die Herleitung solcher Aussagen stellen bereits Aufgaben einer Klassifikation dar.

Der Prototyp verzichtet gänzlich auf die Berechnung von Merkmalen, die eine für den Radiologen so wichtige visuelle Aussage haben. Stattdessen beschreitet die Software den Weg, durch einfache Merkmalsberechnungen möglichst facettenreiche Informationen zusammen zu tragen. Sie erzeugen keine binäre Aussagen wie *Rund* und *Nicht Rund*, sondern resultieren in Merkmalswerten in einem teilweise offenen Wertebereich. Die Merkmale enthalten dabei form-, begrenzungs- oder texturübergreifende Informationen. Beispielsweise wird das Vorhandensein von Hohlräumen durch die Texturmerkmale $M_{Nekrosis}$ und $M_{RingEnhancement}$ aufbereitet, es beeinflusst aber ebenso das Formmerkmal M_{Form4} . Auch $M_{Begrenzung1}$ gibt neben der Aussage über die Randeigenschaft Rückschlüsse auf die Form. Erschwerend können völlig unterschiedliche Objekte zu ähnlichen Merkmalswerten führen, so dass die Herleitung einer deskriptiven Aussage aus einem Einzelmerkmal kaum möglich ist. Die Information über die Rundheit eines Tokens steckt sicherlich in den berechneten Merkmalen, sie ist aber auf mehrere verteilt und bleibt dem menschlichen Betrachter verborgen.

Der Radiologe erkennt aufgrund der guten Mustererkennung des menschlichen Sehapparates deskriptive Aussagen aus dem Bildmaterial und leitet beispielsweise durch Anwendung des Göttinger Summenscores oder aufgrund seiner Erfahrung eine Diagnose ab.

In der vorliegenden Arbeit werden die Aussagen der Merkmale einzig dazu verwendet, Dignitäten direkt vorherzusagen ohne einen „Umweg“ über beschreibende Merkmale zu gehen. Der Trainingsprozess setzt die für den Menschen kryptischen Merkmalswerte in Bezug mit den bekannten Dignitäten, die Ableitung der Diagnose erfolgt (für den Menschen ebenfalls kryptisch) direkt aus dem Merkmalsvektor. Zuvor wurde das System auf den mit Merkmalsvektoren gefüllten mehrdimensionalen Raum trainiert. Dieser Prozess der Wissensaneignung ist gleichzusetzen mit den erworbenen Erfahrungswerten eines Radiologen.

Der Merkmalsraum könnte auch genutzt werden, um BI-RADSTM-Terminologien vorherzusagen. Die Informationen deskriptiver Aussagen lassen sich vermutlich aus der Fülle mathematischer Merkmale herleiten. Dazu müssten die bekannten Lerndaten (s. Kap. 4.5.2) neben der Zielklasse der Dignitäten zusätzlich um Zielklassen für die einzelnen BI-RADSTM-Terminologien ergänzt werden. Beispielsweise müsste durch einen Beobachter für jede Instanz manuell festgelegt werden, ob das Objekt eine runde Form besitzt. Mit diesem Rundheit GS kann dann über alle Instanzen ein Klassifikator trainiert und das Vorliegen der Rundheit als binäre Information vorhergesagt werden. Letztendlich könnte durch Bereitstellen eines GS für jede gewünschte Terminologie ein zuständiger Klassifikator trainiert werden, durch den genau diese eine deskriptive Aussage vorhergesagt werden kann. Wie auch bei der Dignitätsvorhersage könnte abschließend evaluiert werden (s. Kap. 4.6.4), wie

gut der Informationsgehalt der Merkmalsvektoren ist, um beispielsweise die Begrifflichkeit *Rund* vorherzusagen. Durch eine Merkmalsselektion (s. Kap. 6.4.2) ließen sich diejenigen Merkmale ermitteln, die zu dieser Aussage den signifikantesten Beitrag leisten.

Für die Vorhersage einer Diagnose kommt der Computer gänzlich ohne die Betrachtung von BI-RADSTM-Terminologien aus. Die Software könnte die zuvor erläuterten Klassifikatoren einsetzen, um BI-RADSTM konforme Aussagen im Rahmen einer Reportausgabe pro Patient zu leisten. Damit kann das System seine berechneten Kriterien für den Radiologen verständlich aufbereiten und die Computer Diagnose kann durch einen Aussenstehenden nachvollzogen werden. Damit würde die Grundidee BI-RADSTM – die standardisierte Befundaufzeichnung – durch die Software unterstützt und die Akzeptanz des Computersystems erhöht.

In der Literatur sind eine Vielzahl möglicher Kriterien für die Diagnostik aufgeführt [Kai07]. Die Diskussion hat deutlich gemacht, dass ein direktes Umsetzen jedes Kriteriums in einen eigenen Merkmalswert wenig sinnvoll erscheint. Die Aufgabe der Informatik ist es, die Informationen und Grundideen solcher Merkmale zu abstrahieren und daraus möglichst einfache mathematische Merkmalsberechnungen mit dem Ziel abzuleiten, den Merkmalsraum für die Klassifikation mit den entscheidenden Informationen anzureichern.

6.3.3 Robustheit der Merkmalsberechnung

In Tabelle 2.3 sind verschiedene Faktoren aufgezählt, die besonders die *Signalstärke* in Bildbereichen anreichernder Strukturen und damit die Kinetik beeinflussen. Darüber hinaus haben die Hardwarefaktoren Einfluss auf die Bildqualität, vor allem bezüglich des *Bildrauschens* und der *Ortsauflösung* und damit letztendlich auf die Morphologieberechnung.

Das untersuchte Kollektiv wurde immer unter gleichen Voraussetzungen gemessen (s. Kap. 4.1), so dass von gleichwertiger Bildqualität der einzelnen Untersuchung ausgegangen werden kann.

Für den universellen Einsatz einer automatischen Software ist es essentiell, die Merkmale unabhängig von der Bildqualität [KJM⁺06] zu bestimmen. Andernfalls müssten für jede Kombination aus Scanner, Sequenz und KM-Dosierung neue Klassifikatoren antrainiert werden. Aber auch im vorliegenden Protokoll werden die Morphologien durch Kombination von high- und low-res Messungen auf qualitativ unterschiedlichem Bildmaterial bestimmt [VTK04]. Erst durch Berechnung robuster Merkmale erreicht die Software eine flexible Anwendbarkeit.

Folgend wird über den Tellerrand des zur Verfügung stehenden Kollektives geschaut, um die Robustheit der Merkmale gegenüber der Bildqualität zu diskutieren. Die Diskussion gibt eine Einschätzungen wieder; eine Validierung beispielsweise durch Phantombilder ist kein Bestandteil der vorliegenden Arbeit.

Registrierung Die Vorverarbeitung durch die Registrierung ist gegenüber dem Bildrauschen, der Ortsauflösung und Signaländerungen durch das KM als robust einzuschätzen. Das Bildrauschen wird durch den Restfehler berücksichtigt und kann zusätzlich durch eine Parameteranpassung wie der Wahl einer entsprechenden Fenstergröße (s. Tab. 5.3) noch stärker unterdrückt werden. Das Tracking arbeitet lokal und ist somit unabhängig zur

Ortsauflösung. Wiederum kann durch eine Fenstergrößenanpassung der lokale Informationsgehalt zusätzlich verändert werden. Da die Registrierung ohnehin nur auf Bildbereichen ohne oder mit nur sehr geringem Anreicherungsverhalten erfolgreich arbeitet, ist eine durch das KM hervorgerufene Signaländerung in diesen Bildbereichen irrelevant für das Registrierungsergebnis (s. Kap. 6.1.1).

Segmentierung Die Segmentierung ist nur teilweise als robust einzuschätzen. Bildrauschen wird durch die Bildung des Projektionsvolumen unterdrückt und kann sich erst bei der Nachsegmentierung der Einzelserien bemerkbar machen.

Eine Änderung der Ortsauflösung bedeutet eine abweichende Größe der Bildmatrix, um die gleichen Bildinhalte darzustellen. Als Bottom-Up Verfahren ist die Matrixgröße irrelevant für die Segmentierung.

Die Signalintensitäten werden direkt für das Homogenitätskriterium des Schwellwertverfahrens verwendet. Somit hat eine Signaländerung wesentlichen Einfluss auf das Segmentierungsergebnis (s. Kap. 6.2.1). Da die Segmentierung jedoch mit individuellen Schwellwerten arbeitet, die relativ aus den lokal vorherrschenden Signalwerten abgeleitet werden, ist sie invariant gegenüber linearen Skalierungen der Grauwerte. Jedoch verändert die Verschiebung der Baseline (s. auch Abschätzung der Kinetik weiter unten) das Ergebnis und kann zu Über- oder Untersegmentierung der Läsionen führen. Dies hätte Auswirkung auf die Merkmalsberechnungen wie beispielsweise die Form. Besonders wenn durch die Änderung keine Spikulierungen oder andere relevanten Form- und Randeigenschaften mehr segmentiert werden, ist mit abweichenden Merkmalswerten zu rechnen und in der Folge auch mit einer verminderten Diskriminanzfähigkeit. Allgemein wird eine Untersegmentierung der eigentlichen Objektform eher entsprechen als eine Übersegmentierung, weshalb die Berechnung der individuellen Schwellwerte defensiv eingestellt wurde.

Morphologie Die Abschätzung der Robustheit der einzelnen morphologischen Merkmale erfolgt anhand der mathematischen Formeln.

Viele morphologische Merkmale werden nur aus der Betrachtung binärer Information der Token bestimmt, greifen also nicht auf die Grauwerte zurück. Daher sind sie robust gegenüber Bildrauschen und Signalstärke, auch wenn die Segmentierung zu ähnlichen Ergebnissen führt. Eine mögliche Beeinflussung ergibt sich jedoch durch die Ortsauflösung. Durch ihre Veränderung wird das gleiche Objekt mit einer abweichenden Menge an Bildvoxeln repräsentiert. Beispielsweise beeinflusst dies den Skelettierungsalgorithmus, der die Grundlage einiger Merkmale bildet. Die minimalen Distanzen als Ergebnis der Skelettierung bleiben durch Beachten des anisotropen Voxelspacing robust. Aber eine veränderte Anzahl an Tokenvoxeln führt auch zu einer abweichenden Skelettlinie. Somit bleibt der minimale Durchmesser $M_{Mindiameter}$, M_{Form1} und M_{Form3} auf Basis der minimalen Distanzen robust. $M_{Begrenzung1}$ ist dagegen abhängig von der Größe der Skelettlinie, die ins Verhältnis zur Tokengesamtgröße gesetzt wird. Dieses Verhältnis ist jedoch bei Änderung der Ortsauflösung nicht linear, der Einfluss auf die Voxelmenge des Tokens ist größer als auf die Skelettlinie. M_{Volume} und $M_{MaxDiameter}$ sind sehr robuste Merkmale, denn sie verarbeiten die binäre Information des Tokens unter Berücksichtigung der Ortsauflösung. M_{Form4} bezieht sich ebenfalls nur auf die Binärinformation. Die fraktale Dimension führt auch bei Änderung der Ortsauflösung aufgrund des Limes in der Theorie zu ähnlichen Merk-

malswerten (s. Formel 4.13). In der praktischen Anwendung zeigte sich jedoch, dass die Objektgröße bzw. die Ortsauflösung der Bilddaten allgemein viel zu gering für den Einsatz des Boxcounting Algorithmus ist. Für die Bildung der unendlichen Reihen stehen nur die Anfangswerte zur Verfügung, deren Umfang für eine Konvergenzberechnung zu gering ist. Aus diesem Grund beeinflusst allein schon die Positionierung des Gitters auf die Segmentierung den Ergebniswert. Die in dieser Form berechnete fraktale Dimension ist daher nicht robust gegenüber der Ortsauflösung (vgl. Kap. 6.5.3).

Wie auch $M_{Bergrenzung1}$ unterliegt mit M_{Form2} ein zweites Merkmal der Änderung der Ortsauflösung. Dieses Merkmal schätzt die Oberfläche anhand des diskreten Voxelrasters ab und bildet einen Quotienten zum Volumen. Erhöht sich beispielsweise die Ortsauflösung, wird die Kontur der Läsion besser angenähert, das Verhältnis Oberfläche zu Volumen bleibt dadurch nicht konstant.

Neben den zuvor diskutierten Merkmalen basierend auf der binären Information der Tokenelemente beziehen zwei Merkmale auch deren Signalintensitäten in ihre Berechnungen ein. Mit $M_{Kontrast}$ wird die Varianz berechnet, die durch eine Veränderung der Signalintensitäten direkt beeinflusst wird (z.B. Bildrauschen, Weichzeichnung oder Skalierung der Signalwerte, s. Kinetik weiter unten). Auch $M_{Bergrenzung2}$, das die Signalintensität des Objektrandes zum umliegenden Gewebe vergleicht, ist dann nicht robust, wenn sich durch einen Einflussfaktor die Signalstärke in der Läsion gegenüber dem angrenzenden Gewebe nicht linear entwickelt.

Der Wasseranteil durch das Merkmale M_{T2} wird auf der T_2 -gewichteten Messung bestimmt. Da sie nativ gemessen wird, unterliegt sie nicht den Veränderungen durch das KM. Jedoch können Rauschen oder andere signalbeeinflussende Faktoren aufgrund der einfachen Schwellwertanwendung zu abweichenden Merkmalswerten führen. Dagegen ist das Alter M_{Alter} als DICOM-Information absolut unabhängig von der Bildqualität.

Bei zwei der 12 statischen Merkmale ist eine Beeinflussung durch die Ortsauflösung zu befürchten ($M_{Bergrenzung1}$, M_{Form2}). Ob sie in der Praxis zum Tragen kommt, ist davon abhängig, wie stark die Ortsauflösung der Messungen variiert. Zwei weitere Merkmale beziehen die Signalintensitäten ein ($M_{Bergrenzung2}$, $M_{Kontrast}$). Hier ist eine Abhängigkeit zu der Signalstärke und dem Bildrauschen gegeben. Die übrigen Merkmale arbeiten nur auf den binären Token und sollten somit zu vergleichbaren Merkmalswerten führen, auch wenn die Segmentierung zu ähnlichen Objekten führt.

Kinetik Das Grundlagenkapitel (s. Kap. 2.3.2) hat dargelegt, dass die durch das KM hervorgerufenen Signalintensitäten im Bild genutzt werden, um Rückschlüsse auf das Vorliegen einer Hypervaskulierung von Läsionen und somit auf den Malignitätsgrad zu ziehen. Die folgende Untersuchung soll zeigen, dass Charakterisierungen der Kinetik ohne geeignete Normierung der zugrunde liegenden Signalintensitäten Fehler begehen. Denn es existieren viele Faktoren bei der Bilderzeugung mit unterschiedlich starkem Einfluss auf die Signalintensitäten (s. Tab. 2.3). Somit sind die Kinetiken in den Bilddaten verschiedener Untersuchungen ohne Normierung nicht mehr zueinander vergleichbar. Ziel der vorliegenden Arbeit ist jedoch eine robuste Charakterisierung der Kinetik.

Um den Einfluss von Veränderungen in den Signalintensitäten zu untersuchen, werden folgend verschiedene Beispielkinetiken herangezogen. Sie basieren teilweise auf Ergebnissen

von Phantommessungen mit unterschiedlicher KM-Anwendung [MKJ⁺08, ABBM05]. Mit der Berechnung nach Fischer (s. Formel 2.1 u. 2.2) wird eine zeitabhängige Schwellwertinterpretation den eigenen Merkmalsberechnungen zum Vergleich gegenübergestellt.

Abb. 6.10 a zeigt eine Zeit/Signalintensitätskurve, die der Kinetik eines typischen Karzinoms entspricht (s. Abb. 2.4 a). Der Verlauf der Kurve ist für ein einfacheres Auswerten durch einen Polygonzug aus Geraden abstrahiert. Die initiale und postinitiale Phase wird nach Fischer durch 100 % bzw. -20 % Steigung charakterisiert, das eigene Merkmal $M_{WashOut}$ entspricht dem Wert Fischers. Die maximale Signalintensität wird nach 2 Minuten (M_{TTP}) erreicht, die normierten Flächenberechnungen unterhalb der Kurve bis zur Baseline sind gegeben. Diese Kinetik soll folgend als Referenz dienen.

In Abb. 6.10 c wird eine Veränderung an der Hardware simuliert, beispielsweise durch den Einsatz einer anderen Sequenz bei der Bildakquisition. Als Folge werden die Signalintensitäten im gemessenen Bild verdoppelt, sowohl Baseline als auch maximaler Signalwert verschieben sich linear. Gegenüber solchen Einflüssen sind alle untersuchten Kinetikmerkmale robust und führen verglichen zum Referenzfall zu identischen Werten. Im Rückblick auf die Morphologiemerkmale sei angefügt, dass eine derartige Signalveränderung erheblichen Einfluss auf das Merkmal $M_{Kontrast}$ hat.

Interessant ist die Auswirkung, wenn nur die Signalintensitäten der KM anreichernden Strukturen beeinflusst werden. Abb. 6.10 d simuliert ein Verhalten nach Erhöhung des Bolus. Durch die höhere KM-Dosierung bleibt der relative Verlauf der Kurve vergleichbar zu (a), die gebildete Fläche als Indikator der geflossenen KM Menge vergrößert sich jedoch entsprechend. Die Baseline aus der nativen Messung bleibt davon unberührt, jedoch erhöht sich die maximale Intensität.

Diese extreme Erhöhung beeinflusst deutlich die klassischen Berechnungen nach Fischer. Besonders die Steigung der initialen Phase wächst nun auf das Dreifache, denn es wird der nun erhöhte maximale Signalwert mit dem unveränderten nativen Signal des Gewebes verglichen. Damit wird die Abhängigkeit zur KM-Dosierung durch die Parametrisierung nach Fischer offensichtlich. Auch die Charakterisierung der postinitialen Phase und auch das eigene Merkmal $M_{WashOut}$ werden – wenn auch im wesentlich geringerem Maße – davon beeinflusst. Im Gegensatz zur initialen Phase werden hierfür nur Werte für die Berechnung einbezogen, die alle dem Einfluss der KM-Konzentration unterliegen. Trotz der größeren Kurvenfläche bleiben die Merkmale M_{AUC} und M_{AUCTTP} aufgrund der Normierung zum Maximalwert gegenüber der veränderten Dosierung robust; dies hat jedoch auch Nachteile (siehe Abb. 6.10 e und f). Eine solche Normierung des Maximalwertes ließe sich jedoch nicht auf die zeitabhängige Schwellwertinterpretation anwenden, da sie immer zu identischen Steigungswerten führen und damit dem Merkmal die Grundlage entziehen würde.

Ergänzend zur Untersuchung der Morphologie zu Beginn des Kapitels sei bemerkt, dass das Merkmal $M_{Bergrenzung2}$ gegenüber einer solchen Signalveränderung nicht robust ist.

Das Beispiel in Abb. 6.10 b zeigt die Kinetik eines weiteren Karzinoms, das sich gegenüber (a) in seiner Gewebeeigenschaft unterscheidet. Es zeigt sich durch ein frühzeitiges und intensives Anreichern besonders maligne. Die klassische Parametrisierung nach Fischer kann die Beispiele (a) und (b) nicht unterscheiden, weil für die Charakterisierung der initialen Phase der Maximalwert innerhalb der ersten bis dritten Minute entscheidend ist.

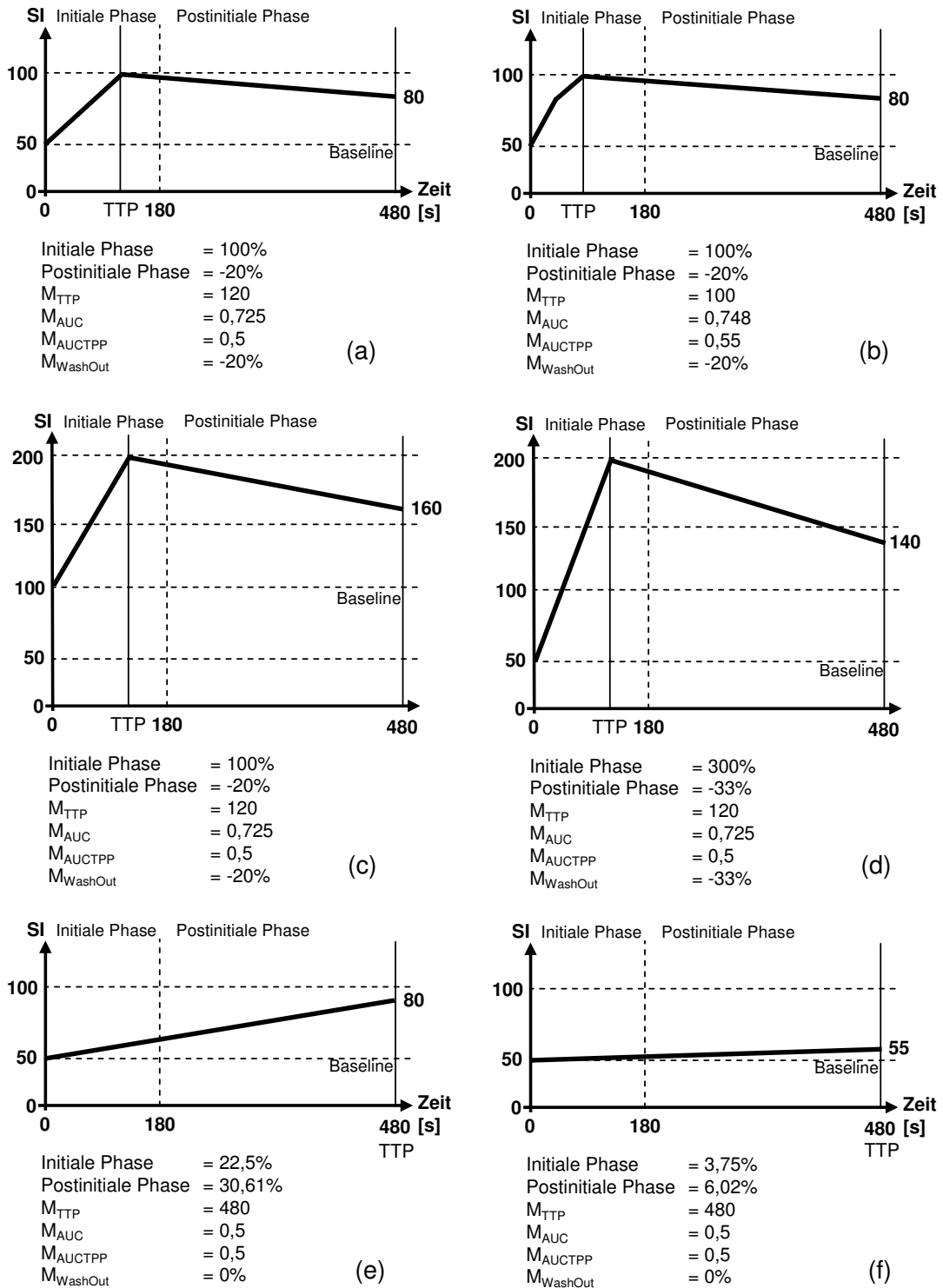


Abbildung 6.10: Schematische Zeit/Signalintensitätskurve zur Darstellung Kinetik beeinflussender Faktoren.

Der Zeitwert M_{TTP} und auch das Flächenmerkmal M_{AUCTPP} können dagegen beide das abweichende Verhalten charakterisieren.

Die Betrachtung der Kinetik einer typisch gutartigen Läsion (Abb. 6.10 e) zeigt für die Merkmale ein vielleicht unerwartetes Verhalten. Die klassische Variante beschreibt durch eine Steigung von gut 22 % und über 30 % eine für benigne Tumore typische schwache Anreicherung in der initialen und eine Persistenz in der postinitialen Phase. Das Zeitmerkmal M_{TTP} findet das Signalmaximum nun an der letzten Stützstelle bei 480 Sekunden und kann damit ebenso die Gutartigkeit durch eine lange Anreicherungsphase vom bösartigen Fall (a) unterscheiden. Das Merkmal $M_{Washout}$, das sich nicht wie bei Fischer auf die maximale Intensität innerhalb der ersten bis dritten Minute, sondern auf den gesamten Messzeitraum bezieht, kann daher keine positive Steigung annehmen. Der Merkmalswert liegt immer im Wertebereich Null oder kleiner, die Aussage des Merkmals bezieht sich nur auf die Charakterisierung des Wash-Outs und hat trotz ähnlicher Berechnung wenig Gemeinsamkeit zur Aussage durch Fischer. Die beiden Flächenmerkmale nehmen zudem identische Werte an, da der TTP auf der letzten Stützstelle liegt.

Zuletzt wird die Kinetik eines gesunden Gewebes wie beispielsweise Fett untersucht (Abb. 6.10 f). Die Steigungswerte der klassischen Fischerberechnung unterscheiden das geringere Anreicherungsverhalten vom vorherigen Fall durch einen noch geringeren Wert. Die eigenen Merkmale dagegen führen im Vergleich zu Fall (e) zu identischen Werten, da insbesondere die Flächenmerkmale zur maximalen Intensität normiert werden.

Die Untersuchung zeigt zweifellos, dass mit der Charakterisierung durch einfache Schwellwertverfahren Kinetiken, die beispielsweise auf unterschiedlicher KM-Konzentration basieren, nicht verglichen werden können. Die Merkmale M_{TTP} und auch die Flächenmerkmale sind gegenüber solchen Veränderungen robust und nehmen für typisch gut- und bösartige Verhalten auch differenzierbare Werte an. Durch die Flächennormierung geht jedoch die Aussage über die Stärke der Anreicherung verloren (vgl. Abb. 6.10 e und f). Sie ist besonders wichtig für die Unterscheidung der Bildartefakte und gesunden Geweben von den eigentlichen Tumoren. Aus diesem Grund wurde das Merkmal M_{MaxSI} initiiert. Da aber gerade solche Absolutwerte zu einer Abhängigkeit von der Bildqualität führen, wurde das Merkmal zur Anreicherung im Gesamtbild normiert. Hier besteht die Hoffnung, dass mit einer relativen Aussage zu den stärksten Anreicherungen im Bild intensitätsunabhängige Bewertungen erzeugt werden. Leider nehmen nicht nur die maximale Intensität, sondern auch die Anzahl und die Größe der anreichernden Strukturen und verbleibende Bewegungsartefakte Einfluss auf diese Normierung. Dieses Merkmal ist deswegen nicht gänzlich robust gegenüber Bildveränderungen und wird daher auch nicht zur Diskriminanz maligner und benigner Läsionen eingesetzt.

Solange die Bilddaten unter möglichst gleichen Voraussetzungen wie beispielsweise Einsatz desselben MRT-Geräts, dosierte KM-Anwendung und demselben Protokoll entstehen, sind die Aussagen der Kinetiken verschiedener Untersuchungen untereinander direkt vergleichbar. Die Parametrisierung kann mit einfachem Maß – sprich Berechnung der Steigung und Bewertung durch Schwellwerte – erfolgen.

Liegen jedoch Messungen verschiedener Institute vor, die durch Einsatz einer abweichenden Hardware (s. Tab. 2.3) erzeugt werden, kann die Diagnostik der Kinetik durch

solch einfache Schwellwertmethoden nicht mehr erfolgen bzw. bedarf einer individuellen Anpassung der Schwellwerte pro Institut.

Aber auch innerhalb desselben Instituts werden die Kinetik beeinflussende Faktoren wie beispielsweise die Dosierung des KMs nicht einheitlich gehandhabt. Häufig wird auf die Gewichtsangaben der weiblichen Patienten vertraut oder das Gewicht gar vom Untersucher geschätzt. Im klinischen Alltag herrscht keine Zeit für ein genaues Auswiegen der Patienten. Bewertungsverfahren, die Schwellwerte auf die absoluten Signalwerte der Kinetik anwenden, führen zwangsweise zu Ungenauigkeiten und begünstigen Fehlinterpretationen.

Bei exakter Handhabung müsste dieser Schwellwert für die Bewertung jeder einzelnen Läsion angepasst werden, da die Signalstärke schon alleine durch die Gewebeeigenschaft des Tumors nicht linear beeinflusst wird [ABBM05].

Mit den kinetischen Merkmalen der vorliegenden Arbeit muss aufgrund des Einsatzes von Zeitwerten und normierten Flächenberechnungen keine Anpassung bzgl. signalintensitätsbeeinflussender Faktoren vorgenommen werden. Dies ermöglicht einen flexiblen Einsatz des Systems. Lediglich beim Merkmal $M_{WashOut}$ konnte eine wenn auch geringere Abhängigkeit zur KM-Dosierung festgestellt werden. Vorteil der Software ist nicht nur die signalunabhängige Beurteilung pro Patient sondern sogar pro Läsion.

Da die Merkmale auf großen Voxelmengen der Token berechnet werden und auch der Hotspot mindestens 20 Voxel umfasst, sind die Merkmale robust gegenüber Bildrauschen und Ortsauflösung. Das Interpolationsverfahren, durch das die Signalwerte zwischen den Stützstellen berechnet werden, ist unzureichend für eine Extrapolation. Würde die Software also auf einem Protokoll angewendet, das einen Zeitraum von weniger als 513 Sekunden misst, könnten die fehlenden Messwerte nur ungenau vorhergesagt werden. Dies wäre jedoch notwendig, sollen Merkmale wie TTP vergleichbar bestimmt werden.

6.3.4 Clusteranalyse

Durch die Clusteranalyse wurden Dendrogramme (s. Kap. 4.6.3) sowohl mit allen Instanzen der Lerndaten (s. Abb. 6.11) als auch durch Auslassen der zusätzlichen Funde nur mit den Läsionen beschreibenden Instanzen (s. Abb. 6.12) erzeugt.

Auch wenn die klassischen Kinetikmerkmale nach Fischer nicht in das Training der Klassifikatoren der vorliegenden Arbeit eingehen, werden sie dennoch zur Ermöglichung eines Vergleichs zu den eigenen Kinetikmerkmalen folgend in der multivariaten Analyse berücksichtigt.

In beiden Diagrammen zeigen sich teilweise die gleichen Abhängigkeiten. So fällt auf, dass alle kinetischen Merkmale, die sowohl auf dem ganzen Objekt als auch auf dem Hotspot bestimmt wurden, ähnlich sind. Die Ähnlichkeit gilt auch für die beiden Merkmale $M_{Ringenhancement}$ und $M_{Nekrose}$, denn beide basieren auf der Berechnung der eingeschlossenen Hohlräume.

Die vier Merkmale $M_{Volumen}$, der minimale und maximale Durchmesser sowie M_{Form2} werden unter kurzen Distanzen gruppiert, da alle Merkmale sind zur Größe des Tokens abhängig sind.

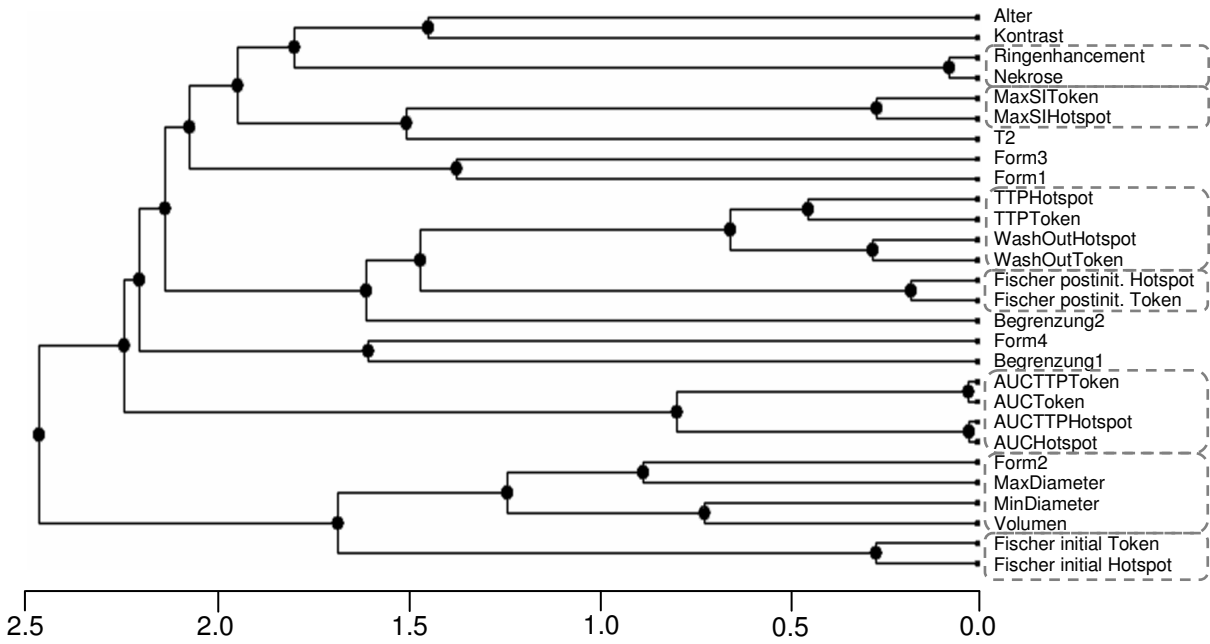


Abbildung 6.11: Dendrogramm über alle Instanzen.

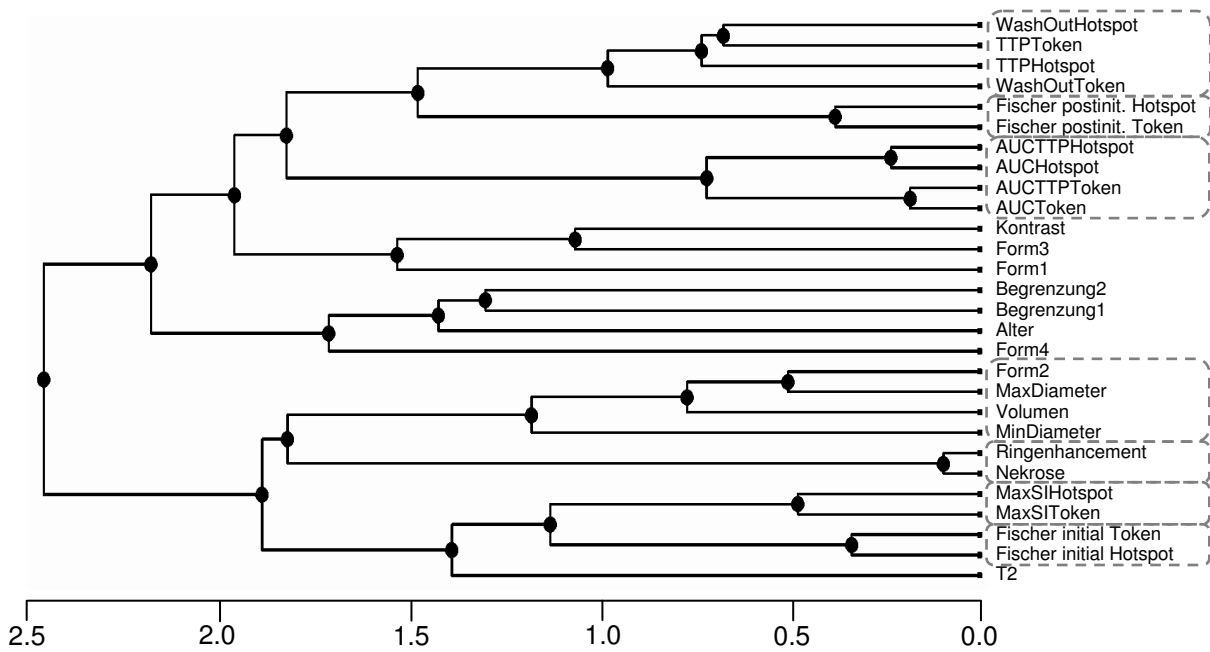


Abbildung 6.12: Dendrogramm über Läsionen beschreibende Instanzen.

Allgemein wenig Ähnlichkeit haben die Merkmale M_{Alter} , $M_{Kontrast}$, M_{T2} , alle Formen außer der bereits genannten M_{Form2} und die beiden Begrenzungsmerkmale. Sie nähren die Hoffnung, dass sie einen großen Beitrag zur Diskriminierung beisteuern (s. Kap. 6.4.2).

Die Ähnlichkeit der Kinetikmerkmale (s. Kap. 4.4.3) lässt sich am Besten am zweiten Diagramm (s. Abb. 6.12) nur der malignen und benignen Instanzen bewerten.

Nach BI-RADSTM wird die Kinetik klassisch durch initiale und postinitiale Phase charakterisiert, die bei Fischer auf genau zwei Merkmale übertragen werden. Diese werden im Dendrogramm erst im Wurzelknoten zu einem Cluster vereint, ihre Ähnlichkeit ist also minimal. Während das klassische Merkmal der initialen Phase dosisabhängig aus der maximalen Intensität der ersten bis dritten Minute hergeleitet ist (s. Kap. 6.3.1), wird die eigene Kinetik durch zeitabhängige oder normierte Merkmale interpretiert. Die beiden klassischen Grundverhalten der Kinetik werden nun auf vier Merkmale (M_{TTP} , M_{AUC} , M_{AUCTTP} , $M_{WashOut}$) abgebildet; durch das Dendrogramm können ihre Ähnlichkeiten untereinander und zu Fischers Berechnung verglichen werden.

Die Vermutung, dass das Merkmal M_{TTP} den Informationsgehalt der klassischen Charakterisierung der initialen Phase wiedergibt, hat sich in der Analyse nicht bestätigt. Diese beiden Merkmale zeigen wenig Ähnlichkeit, stattdessen ist die Affinität zu der klassischen Charakterisierung der postinitialen Phase wesentlich höher. Zusammen mit $M_{WashOuts}$ bilden M_{TTP} und das Fischer Merkmal sogar ein Cluster großer Ähnlichkeit. Denn das eigene Merkmal $M_{WashOut}$ wird nicht wie bei Fischer auf den Zeitraum der ersten bis dritten Minute bezogen, sondern auf den Zeitpunkt des Merkmals M_{TTP} . Somit nimmt das Merkmal $M_{WashOut}$ ein Plateau für einen späten TTP und somit typisch gutartiges Verhalten an und korreliert mit einer persistenten postinitialen Phase. Für einen frühen TTP beschreibt das Merkmal dagegen mit einem typisch bösartigen Verhalten ein Wash-Out Verhalten in der postinitialen Phase.

Auch die Merkmale der Flächenberechnungen der gesamten Kinetik (M_{AUC}) und des Abschnitts bis zum TTP (M_{AUCTPP}) ähneln sich. Beide Merkmale nehmen identische Werte an, wenn der TTP auf der letzten Stützstelle (also typischerweise im gutartigen Fall) ermittelt wurde.

Lediglich M_{MaxSI} zeigt eine erkennbare Entsprechung zur klassischen Charakterisierung der initialen Phase. Das Merkmal beinhaltet die maximale Signalintensität normiert zum Histogramm und nimmt für ein typisches malignes Verhalten einen höheren Wert an, genauso wie das klassische Merkmal dadurch eine größere Steigung erfährt.

Eine interessante, aber offene Frage ist, inwieweit sich das Ergebnis der Clusteranalyse speziell der kinetischen Merkmale verändern würde, wenn das untersuchte Kollektiv aus Untersuchungen unterschiedlicher Voraussetzungen (s. Tab. 2.3) bestehen würde.

6.3.5 Hauptkomponentenanalyse

In Abb. 6.13 und 6.14 sind jeweils die beiden ersten Faktoren der Hauptkomponentenanalyse (s. Kap. 4.6.3) dargestellt, die sich wie bei den beiden Dendrogrammen des vorherigen Kapitels wiederum durch die Wahl der Zielklassen unterscheiden.

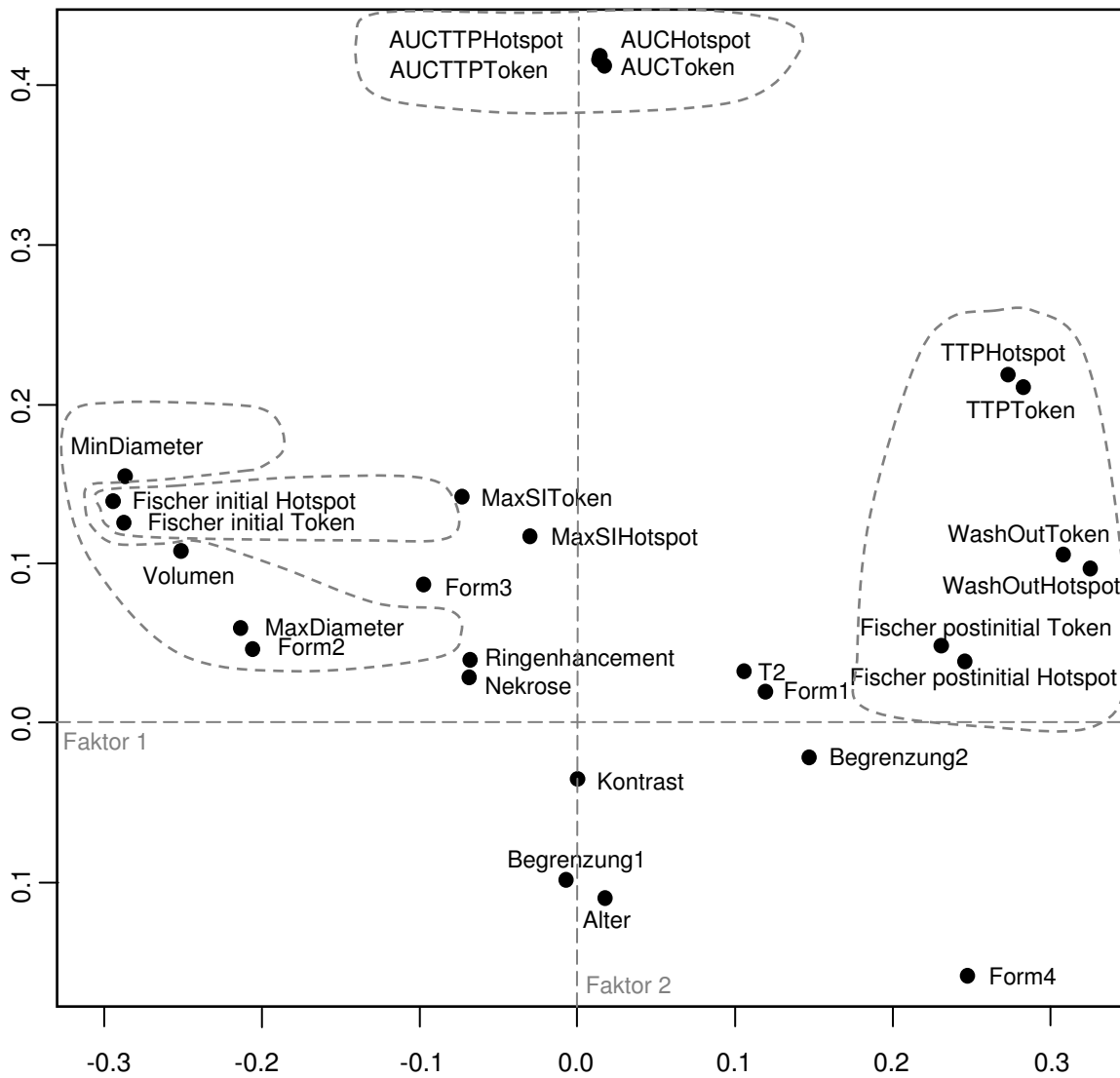


Abbildung 6.13: Hauptkomponentenanalyse über alle Instanzen.

Bei der Hauptkomponentenanalyse decken die beiden dargestellten ersten Faktoren nur etwa 36 % der Gesamtvarianz ab. Erst mit dem 13. Faktor wird ein akkumulativer Prozentanteil von über 90 % erreicht. Dies macht deutlich, dass der Gegenstand der vorliegenden Arbeit nicht auf wenige Faktoren reduziert werden kann. Der 2D-Darstellung der ersten beiden Faktoren kann also nicht entnommen werden, ob ein Merkmal – ohne erkennbare Beziehung im Diagramm – vielleicht auf einen der nicht dargestellten Faktoren große Auswirkungen hat. Es kann nur der Einfluss der Merkmale auf die beiden Faktoren bewertet werden, die sinngemäß nur ein Drittel des gesamten Sachverhaltes darstellen.

Besonders auffällig sind die Flächenmerkmale der Kinetik, die im Diagramm über alle Instanzen (s. Abb. 6.13) eine starke Beziehung zum ersten Faktor haben, während sie sich im zweiten Diagramm (s. Abb. 6.14) der Läsioneninstanzen deutlich verringern. In sie kann die Hoffnung gesetzt werden, einen Beitrag zur Differenzierung von Läsionen und zusätzlichen Funden zu leisten. Weiter zeigt die Auswertung nur der Läsioneninstanzen, dass fast sämtliche Kinetikmerkmale und die Größenmerkmale eine starke Beziehung auf den zweiten Faktor haben. Somit sind sie wichtig für die Darstellung des (groben) Sachverhaltes.

Dagegen leistet das Ringenhancement- und Nekrosemerkmal einen unbedeutenden Beitrag zum Sachverhalt. Beide Merkmale liegen nahezu im Ursprung des Diagramms.

Die Betrachtung des zweiten Diagramms nur der Läsioneninstanzen zeigt auf, dass viele Merkmale, die sich in der Clusteranalyse als wenig ähnlich erwiesen, nun vor allem im Randbereich liegen. Damit tragen sie viel zum Sachverhalt bei; sie haben eine große Beziehung entweder zu einem oder gar zu beiden Hauptkomponenten.

Merkmale, die in der Clusteranalyse (s. vorheriges Kapitel) jeweils Ähnlichkeit zu einer der beiden klassischen Kinetikcharakterisierungen zeigten, teilen sich auch im Hauptkomponentendiagramm in zwei Gruppierungen, getrennt durch den zweiten Faktor. Die Charakterisierung der initialen Phasen zeigt dabei einen größeren Einfluss auf den Sachverhalt als die postinitiale. Von der Merkmalsgruppe der postinitialen Phase zeigt vor allem das Merkmal M_{TTP} eine große Beziehung zu beiden Faktoren, während sie bei dem eigenen Wash-Out Merkmal sehr gering ist.

Die nach der Clusteranalyse vier Größen darstellenden Merkmale gruppieren sich ebenfalls und haben eine starke Beziehung teilweise zu beiden Faktoren. Demnach trägt die Größe viel zum Sachverhalt maligner und benigner Läsionen bei. Da durch den GS bekannt ist, dass maligne Läsionen im Mittel größer sind als benigne, werden diese Merkmale nicht zur Diskriminierung herangezogen. Auch sieht der BI-RADSTM-Standard keine Größeninformation zur Befundung vor.

Die gewonnenen Erkenntnisse lassen sich für die Optimierung von Klassifikatoren nutzen (s. Kap. 6.4.2).

6.3.6 Erkenntnisse und Grenzen

Durch die Merkmalsextraktion werden wie bei der Befundung Informationen über die segmentierten Objekte gesammelt. Die mathematischen Merkmale können den von BI-RADSTM geforderten Umfang nur zum Teil wiedergeben. Informationen über raumfordernendes Enhancement und Symmetrien werden von der Software gar nicht berechnet (s. Kap.

6.3.1). Die bereitgestellten Merkmale umfassen im Wesentlichen die Aussage der Kriterien, die im Göttinger Score (s. Kap. 2.3.5) Verwendung finden. Inwieweit die Merkmale die Unterscheidung von benignen und malignen Läsionen leisten können, wird in Kapitel 6.4.1 behandelt.

Es wurde in Kapitel 6.3.2 verdeutlicht, dass die Software die Diagnose ohne die für den Radiologen gewohnten visuellen Beschreibungen erzeugt. Es wurde die Möglichkeit aufgezeigt, durch Klassifikation die mathematischen Merkmale als BI-RADSTM-Terminologien für den Anwender verständlich aufzubereiten.

Bei der Untersuchung der Unabhängigkeit der Merkmalsberechnung wurde festgestellt, dass nicht alle Merkmale robust gegenüber den Bildveränderungen sind (s. Kap. 6.3.3). Die Skelettierung unterliegt der Ortsauflösung, andere morphologische Merkmale sind nicht auf die Signalintensitäten normiert ($M_{\text{Begrenzung1}}$, $M_{\text{Begrenzung2}}$, M_{Form2} , M_{Form4} , M_{Kontrast}). Die Charakterisierung der Kinetik pro Läsion ist hingegen weitestgehend unabhängig gegenüber absoluten Signalveränderungen. Dies ist besonders wichtig, da bekannte Faktoren (s. Tab. 2.3) den Vergleich der Charakterisierung verschiedener Untersuchungen nicht beeinflussen können. Da alle Bilddaten des verwendeten Kollektivs unter gleichen Voraussetzungen erhoben wurden, haben nicht robuste Merkmale keinen negativen Einfluss auf die Ergebnisse der vorliegenden Arbeit. Jedoch bedarf es einer Überarbeitung, wenn die Software mit den gleichen Klassifikatoren in unterschiedlichen Kliniken eingesetzt werden soll.

Durch die multivariate Analyse (s. Kap. 6.3.4 u. 6.3.5) können bereits vor der Klassifikation Urteile über die Verwendung einiger Merkmale getroffen werden. Größen beschreibende Merkmale sind kein Bestandteil des BI-RADSTM-Lexikons und werden in der folgenden Klassifikation auch nicht für die Diskriminierung der Karzinome herangezogen. Die Analyse hat aufgezeigt, dass sowohl die Merkmale M_{Volumen} , M_{Form2} als auch maximale und minimale Durchmesser einen auf die Größe bezogenen Sachverhalt darstellen. Da die malignen Tumore im Mittel größer sind als die gutartigen (s. Kap. 5), wäre durch größenabhängige Merkmale zudem eine Verzerrung des Sachverhaltes im Trainingsprozess zu befürchten. Alle vier genannten Merkmale finden daher nur zur Diskriminierung von zusätzlichen Funden Anwendung.

Auch macht die Analyse deutlich, dass sich unter den 24 Merkmalen keines ohne Verwendung von absoluten Signalintensitäten befindet, welches eine Alternative zur klassischen Charakterisierung der initialen Phase nach Fischer aufweist. Daher wird in Kapitel 6.4.1 gesondert der Frage nachgegangen, wie sich die Diskriminanzfähigkeit der klassischen Kinetikmerkmale gegenüber den eigenen Berechnungen verhält. Als einziges Merkmal hat M_{MaxSI} , das – wenn auch zum Histogramm normiert – in Abhängigkeit zu den absoluten Signalintensitäten berechnet wird, Ähnlichkeit mit der klassischen Charakterisierung. Trotz der Normierung wird auch dieses Merkmal nicht zur Diskriminierung der malignen Tumore herangezogen, um auf jeden Fall eine Signalunabhängigkeit der Klassifikatoren zu wahren.

6.4 Klassifikation

Als letzte Komponente wird von der Prototypsoftware durch die überwachte Klassifikation die eigentliche Diagnose auf Grundlage der extrahierten Merkmale geleistet. Von der Software wird für jedes Objekt eine eigene Einschätzung durch eine eindeutige Bewertung als zusätzlicher Fund oder Läsionen gegeben. Für jede erkannte Läsion erfolgt durch den zweiten Klassifikator eine eindeutige Vorhersage in gut- oder bösartig. Auch die Funktionalität dieser Komponente ist in den heute bekannten kommerziellen CAD-Systemen nur selten vorhanden (s. Kap. 6.5). Falls sie vorhanden ist, wird sie nur aufgrund manueller Vorgaben des Benutzers ermöglicht.

Es existieren eine Vielzahl an Klassifikationsverfahren im Data Mining Prozess (s. Kap. 4.6.5). Aufgrund von Vorstudien [VBK+03, VTK+05] fanden in der vorliegenden Arbeit Neuronale Netze Anwendung. Alternative Klassifikatoren dazu werden in Kapitel 6.4.4 untersucht und ein Vergleich ihrer Vorhersageleistung vorgenommen.

Wie ein Expertenradiologe das zugrunde liegende Kollektiv beurteilt hätte, wurde im Vergleich nicht untersucht. Es kann somit keine Aussage getroffen werden, ob die Vorhersage des Prototyps auf den vorhandenen Daten besser oder schlechter als ein Expertenradiologe abschneidet. Bezogen auf die Erkennung des Goldstandards liegt die Software mit 88 % Sensitivität und 72 % Spezifität im Bereich der von Radiologen erzielten Leistungen in der Literatur (vgl. Kap. 2.2.4).

Die Sensitivität der Erkennung von zusätzlichen Funden ist mit 92 % aufgrund ihres hohen Vorkommens letztendlich nicht ausreichend. Es verbleibt fast jeder zehnte zusätzliche Fund in der finalen Ausgabe und wird somit als gut- oder bösartige Läsion klassifiziert; dies macht eine Vorhersage auf Patientenbasis wenig praktikabel. Diese falsch positiven Funde können zwar vom Radiologen leicht von den tatsächlichen Läsionen unterschieden werden, aber die automatische Software würde fast jede Patientin als „krank“ vorhersagen.

Auf Kosten der Spezifität kann die Sensitivität des ersten Klassifikators zur Erkennung der zusätzlichen Funde erhöht werden, jedoch würden dann auch vermehrt Läsionen als zusätzlicher Fund aussortiert und unter Umständen maligne Läsionen verloren gehen. Dadurch würde sich die Sensitivität der Gesamtvorhersage vermindern. Beispiele für zusätzliche Funde sind in Kapitel 6.4.5 gegeben, sie liegen überwiegend im Thoraxbereich.

Der Informationsgehalt des Merkmalsvektors entspricht in etwa dem Umfang der verwendeten Kriterien des Göttinger Scores (s. Kap. 2.3.5). Auch die Leistung auf dem Goldstandard der vorliegenden Arbeit (s. Tab. 5.9) erreicht eine vergleichbare Treffgenauigkeit mit 82 % gegenüber der Publikation Fischers [FKG99] mit 85 %, wobei beim Prototypen die Sensitivität niedriger, aber die Spezifität höher liegt (88 %/72 % Sensitivität/Spezifität gegenüber 93 %/65 % bei Fischer, vgl. Kap. 6.4.4).

Auch ist auffällig, dass besonders die Vorhersage von kleinen Objekten weniger zufriedenstellend ist. Sie stellen sich wenig charakteristisch in der Morphologie dar. Hier ist eine Abgrenzung zu den zusätzlichen Funden besonders schwierig. BI-RADSTM beurteilt solch kleine Foki nur anhand der Kinetik, während der Prototyp auch die morphologischen Merkmale einbezieht. Als mögliche Verbesserung bietet sich das Trainieren von mehreren Klassifikatoren gestaffelt nach Objektgröße an. Derzeit ist hierfür jedoch das vorhande-

ne Kollektiv zu klein. Alternativ können die Größenmerkmale (M_{Volume} , $M_{MinDiameter}$, $M_{MaxDiameter}$ und M_{Form2}) mittrainiert werden, denn dies gibt dem Klassifikator die Möglichkeit zu erlernen, welche Merkmale erst in Abhängigkeit zur Größe signifikant werden. Derzeit werden die Größenmerkmale nicht trainiert, da sie kein Bestandteil des BI-RADSTM-Standards darstellen. Zudem wäre auch eine Verzerrung des zu lernenden Sachverhalts zu befürchten, da der maligne GS im Mittel größer als der benigne ist.

Die Vorhersageleistung des Prototyps kann nicht mit anderen Publikationen verglichen werden, weil sie dem vollautomatischen Prozess eines kompletten Diagnosesystems entstammt, während bekannte Veröffentlichungen unter manueller Vorarbeit Daten erheben. Einige Publikationen beschäftigen sich mit der Auswertung manuell erhobener Merkmale [VBK⁺03, SWBA]. Solche Datenmengen unterscheiden sich deutlich von denen durch automatisch berechnete Merkmale. Wie in Kapitel 6.3.2 ausführlich diskutiert, ist das Auge des Radiologen in der Lage, die Läsionen bei der Beschreibung bereits zu interpretieren. Er kann eine Läsion direkt als rundlich oder spikuliert bewerten, während ein Computersystem auf solche Aussagen erst einmal trainiert werden müsste. Der Radiologe fällt seine Entscheidungen häufig durch Betrachten einzelner Bildschichten und unterliegt in seiner Urteilsbildung weniger dem Bildrauschen, Bewegungsartefakten, Ortsauflösung oder anderen Einflussgrößen (s. Kap. 6.3.3). Die Leistung eines Trainings auf den manuell erhobenen Ausgangsdaten kann nicht mit den extrahierten mathematischen Merkmalen der vorliegende Arbeit verglichen werden.

Vomweg et. al [VBK⁺03] konnten auf manuell deskriptiven Merkmalen, erhoben auf 604 Läsionen, zeigen, dass NN befähigt sind, bessere Vorhersagen zu leisten als ein Expertenradiologe. Bei Szabo et al. [SWBA] zeigte ein Vergleich auf 104 Läsionen eine knappe Dominanz des Expertenradiologen gegenüber der NN.

In der Literatur werden besonders durch Untersuchungen kleiner Kollektive wie z.B. lediglich 14 Läsionen [TNB03] die besten Vorhersagen durch Klassifikatoren erzielt. Solch geringe Fallzahlen und somit wenig repräsentative Untersuchungen enden somit häufig in Überinterpretationen. Mit steigender Größe der quantifizierten Kollektive sinkt meist auch die Treffgenauigkeit. Eigene Erfahrungen und Auswertungen in der Literatur [DMK⁺99] machen deutlich, dass mindestens 200 Instanzen für eine aussagekräftige Statistik zur Vorhersage von Tumordignitäten benötigt werden.

Der Data Mining Prozess (s. Kap. 4.6.3) bietet viele Fehlerquellen und führt besonders in der Medizin nicht selten zu Fehlinterpretationen. Häufige Gründe sind neben den geringen Fallzahlen auch Ungenauigkeiten in den erhobenen Daten und Fehler in der statistischen Auswertung.

In der vorliegenden Untersuchung wird ein möglichst hohes Qualitätsmaß des Data Mining Prozess angestrebt. In Kapitel 4.5.1 wurde bereits auf die Erhebung des Goldstandards eingegangen. Alle Bilddaten wurden unter denselben Voraussetzungen erzeugt. Das System wurde nicht mit der Diagnose von Radiologen und damit einer subjektiven Einschätzung verglichen, sondern mit gesicherten Histologien. Die Vorhersage der trainierten Klassifikatoren wird durch eine Kreuzvalidierung nach der Leave-One-Out Methode evaluiert.

Die in der vorliegenden Arbeit auf einem Kollektiv von 101 Patientinnen mit 130 Läsionen

sioneninstanzen erhobenen statistischen Aussagen sind nur wenig repräsentativ. Denn die MRT-Untersuchung der Brust wird nur für bestimmte Indikationen eingesetzt (s. Kap. 2.2.4). Folglich lassen sich die statistischen Aussagen der vorliegenden Arbeit nicht verallgemeinern. Ob sich die Ergebnisse auf einem repräsentativen Patientenkollektiv wiederholen oder gar verbessern lassen, bleibt offen.

Die nachfolgenden Kapitel befassen sich durch Anwendung von Methoden der Multivariaten Analyse mit der Diskriminanzfähigkeit der Merkmale. Im nächsten Kapitel wird anhand semantischer Merkmalsteilmengen der Einfluss einzelner Merkmale auf die Differenzierung der Tumordignität untersucht. Die Selektion der signifikantesten Merkmale wird durch die Methode der Regressionsanalyse (s. Kap. 4.6.3) ermittelt. Durch dieses Wissen lassen sich leistungstärkere Klassifikatoren trainieren, wie sie im Ergebnisteil (s. Kap. 5.4) vorgestellt wurden.

Ebenfalls auf Objektbasis wird in Kapitel 6.4.3 die Vorhersage ungesicherter Histologien ausgewertet. Neuronale Netze sind in der medizinischen Anwendung sehr verbreitet. In Kapitel 6.4.4 werden alternative Klassifikationsverfahren miteinander verglichen. Die zusätzlichen Funde werden in Kapitel 6.4.5 näher vorgestellt. Nach Vorstellung eines einzelnen Problemfalls werden abschließend die Erkenntnisse und Grenzen (s. Kap. 6.4.7) aus den Untersuchungen zusammengefasst.

6.4.1 Merkmalskombinationen

Um den Einfluss verschiedener Merkmalskombinationen auf die Diskriminanzfähigkeit maligner und benigner Tumore zu untersuchen, wurde ein erneutes Training Neuronaler Netze mit Merkmalsteilmengen durchgeführt. Die Fehlerrate wurde wieder nach der Leave-One-Out Methode geschätzt (s. Kap. 4.6.4). Im Gegensatz zum Ergebnisteil (vgl. Kap. 5.4) wird sie nun auf Objektbasis ausgegeben. Die Evaluation wird nur auf den Lerndaten bestehend aus den 86 Instanzen der GS-Klassen 1 und den 44 Instanzen der GS-Klasse 2 ausgeführt. Die Diskriminanzierung der zusätzlichen Funde ist kein Bestandteil der Untersuchung. Um die Fehlerrate der verschiedenen Kombinationen miteinander vergleichen zu können, wurde immer die gleiche Netztopologie des Ergebnisteils verwendet.

Als Ausgangsbasis zum Vergleich der Merkmalskombinationen dient die Vorhersage eines Klassifikators, der über alle bis auf die größenabhängigen Merkmale ($M_{Volumen}$, $M_{MaxDiameter}$, $M_{MinDiameter}$, M_{Form2} , s. Kap. 6.3.6) trainiert wurde. Die Vorhersageleistung des auf den verbleibenden 20 Merkmalen trainierten Klassifikators NN_{Alle} resultiert in 85 % Sensitivität und 68 % Spezifität (s. Abb. 6.15).

Zur Beurteilung des Einflusses nur der morphologischen und kinetischen Merkmale wird folgend auf verschiedenen Merkmalsteilmengen trainiert.

Dem Klassifikator $NN_{Morphologie}$ dienen nur die morphologischen Merkmale M_{Form1} , M_{Form3} , M_{Form4} , $M_{Begrenzung1}$, $M_{Begrenzung2}$, $M_{Kontrast}$, $M_{Nekrose}$, $M_{Ringenhancement}$ und M_{T2} (s. Kap. 4.4.1 u. 4.4.2) als Eingabe. Der Klassifikator $NN_{Kinetik}$ kennt nur die zeit- und flächenbasierten kinetischen Merkmale (s. Kap. 4.4.3). Während das Merkmal M_{T2} der Morphologie zugeteilt wurde, konnte das letzte verbleibende Merkmal M_{Alter} keiner der

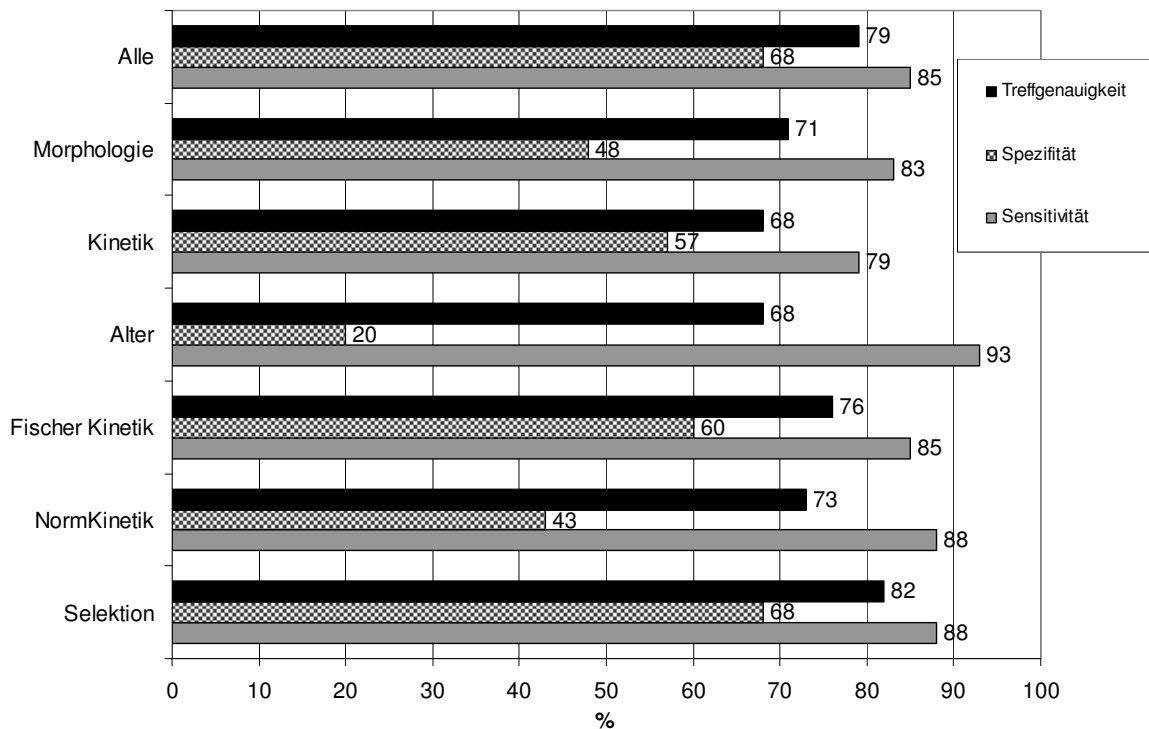


Abbildung 6.15: Vergleich der Vorhersageleistung auf Basis verschiedener Merkmalsteilmengen.

beiden Gruppen zugeordnet werden und stellt mit NN_{Alter} einen Klassifikator eines Einzelmerkmals dar (s. Abb. 6.15).

Das Alter stellt zwar ein sehr sensibles Merkmal (93 %) dar, ist dafür aber wenig spezifisch (20 %). Die Verschlechterung der Sensitivität der morphologischen und kinetischen Einzelklassifikatoren ist mit 79 % und 83 % gegenüber der 85 % über alle Merkmale (NN_{Alle}) eher als gering zu bewerten. Ihre Spezifität sinkt jedoch nun deutlich unter die 60 % Marke. Demnach wird eine hohe Spezifität erst durch gleichzeitiges Betrachten der Morphologie, Kinetik und des Alters erreicht. Viele der konkurrierenden Softwareprodukte beziehen jedoch nur kinetische Information in ihre (Low-Level) Vorhersagen ein (s. Kap. 3.2).

In Kapitel 6.3.4 konnte festgestellt werden, dass vor allem die Aussage der klassischen Charakterisierung der initialen Phase durch kein eigenes Kinetikmerkmal geleistet wird. Folgend soll der Unterschied der Vorhersageleistung der schwellwertabhängigen Berechnung nach Fischer (s. Formel 2.1 u. 2.2) zu den eigenen intensitätsunabhängigen Merkmalen untersucht werden. Die beiden Fischer-Merkmale der initialen und postinitialen Phase des Klassifikators $NN_{Fischer}$ erzielen mit 71 % nicht nur eine bessere Treffgenauigkeit als die eigenen normierten Kinetikmerkmale des Klassifikators $NN_{Kinetik}$ (68 %), sondern erreicht sogar die gleiche Sensitivität wie die Vorhersage über alle Merkmale (s. Abb. 6.15). Aber auch die Fischer-Merkmale überschreiten die 60 % Marke der Spezifität nicht.

Die Berechnung der Steigung durch Einbeziehen nicht normierter absoluter SI bei Fischer innerhalb der ersten bis dritten Minute der initialen Phase stellt demnach ein sehr wichtiges Diskriminanzkriterium aller untersuchten Merkmalskombinationen dar. Lediglich mit M_{MaxSI} (s. Kap. 4.4.3) befasst sich ein bisher außer Acht gelassenes Merkmal mit der Maximalanreicherung.

Folgend wird geprüft, ob sich der Informationsgehalt der Kinetikinterpretation nach Fischer durch M_{MaxSI} (als alternative Charakterisierung der initialen Phase) und den M_{TTP} (als alternative Charakterisierung der postinitialen Phase) durch den Klassifikator $NN_{NormKinetik}$ nachbilden lässt. Sie zeigten in der Hauptkomponenten- und Clusteranalyse (s. Kap. 6.3.4 u. 6.3.5) die größten Ähnlichkeiten.

Der Klassifikator $NN_{NormKinetik}$ (s. Abb. 6.15) ist sensitiver als $NN_{Kinetik}$ oder $NN_{Fischer}$, aber weniger spezifisch. Insgesamt ist seine Treffgenauigkeit höher als bei Fischer. Der Klassifikator stellt mit M_{MaxSI} somit eine gute Alternative dar. Das Merkmal unterscheidet sich in zwei wesentlichen Punkten von der klassischen Charakterisierung der initialen Phase. Neben der schon erwähnten Normierung zum Histogramm bezieht sich Fischers Merkmal auf das Zeitintervall der ersten bis dritten Minute, dagegen wird M_{MaxSI} auf dem gesamten Verlauf der Kinetik bestimmt. Gutartige Tumore reichern meist bis zum Ende der Untersuchung kontinuierlich an und erreichen dort höhere Signale als in der initialen Phase. Daher ist M_{MaxSI} gegenüber Fischers Merkmal nicht so spezifisch. Da M_{MaxSI} in der vorliegenden Arbeit ausschließlich zur Diskriminierung der zusätzlichen Funde herangezogen wird (s. Kap. 6.3.6), ist eine Beschränkung nur auf die initiale Phase jedoch nicht sinnvoll.

Wie in der Einleitung aufgeführt, kann von den hier geführten Statistiken auf keine Allgemeingültigkeit geschlossen werden. Bei dem vorliegenden Patientenkollektiv und berechneten Merkmalen konnte gezeigt werden, dass erst durch gleichzeitiges Betrachten der Morphologie und Kinetik die Spezifität bei der Differenzierung von Tumoren erhöht wird. Diese Aussage stimmt mit dem allgemeinen Konsens in der Fachliteratur überein [Fis00, Kuh00]. Ein Vergleich der signalabhängigen Berechnung nach Fischer zu den eignen signalunabhängigen Kinetikmerkmalen zeigt eine bessere Vorhersageleistung der Fischer-Merkmale. Durch Einbeziehen des normierten Merkmals M_{MaxSI} könnte jedoch eine vergleichbare Leistung erzielt werden.

Die Leistung von Klassifikatoren kann häufig durch die Wahl einer geeigneten Merkmalsteilmenge optimiert werden ($NN_{Selektion}$). Diese Selektion ist Gegenstand des nächsten Kapitels.

6.4.2 Merkmalsselektion

Durch Reduktion der Dimension des Merkmalsraums kann die Vorhersageleistung von Klassifikationen gesteigert werden. Mit der Regressionsanalyse (s. Kap. 4.6.3) steht dafür ein Standardverfahren der Mathematik zur Verfügung.

Zuerst erfolgt für den ersten Klassifikator NN_{LA} eine Untersuchung aller Merkmale bezogen auf die Zielklassen $Ziel_{Lesion}$ (GS-Klassen 1 und 2) und $Ziel_{Add}$ (GS-Klasse 7), die in der Modellfunktion durch die binären Zustände $Y = 0$ (Läsion) und $Y = 1$ (zusätzlicher

Fund) abgebildet wurden (s. Tab. 6.5).

Eine Reduktion auf wenige Merkmale gelingt bei den Zielklassen $Ziel_{Lesion}$ und $Ziel_{Add}$ nicht, es verbleiben immer noch 15 Merkmale (s. Tab. 6.6). Alle Größen-, Form- und Begrenzungsmerkmale gehören zur selektierten Teilmenge. Von den kinetischen Merkmalen leisten die Flächenmerkmale keinen Beitrag. Sie zeigten in der Hauptkomponentenanalyse (s. Kap. 6.3.5) noch einen starken Bezug auf den ersten Faktor. Die maximale Signalanreicherung M_{MaxSI} und M_{T2} hingegen zeigen eine hohe Signifikanz. Letzteres charakterisiert den Wassergehalt und leistet dadurch einen Beitrag beim Erkennen von Gefäßen. Die Maximalanreicherung durch M_{MaxSI} ist bei Läsionen für gewöhnlich höher als bei zusätzlichen Funden.

Die Analyse macht deutlich, dass keine herausragenden Merkmale für die Diskriminierung der zusätzlichen Funde vorliegen. Letztendlich wurden die Merkmale zur Unterscheidung maligner von benigner Läsionen entwickelt und für die Erkennung von zusätzlichen Funden einfach nur wiederverwendet.

Die Regressionsanalyse wurde mit den Zielklassen $Ziel_{Mal}$ ($Y = 0$) und $Ziel_{Ben}$ ($Y = 1$) wiederholt (s. Tab. 6.7). Wie in Kapitel 6.3.6 begründet, wurden für den zweiten Klassifikator NN_{MB} die Größen beschreibenden Merkmale ausgeschlossen und auch das Merkmal M_{MaxSI} nicht berücksichtigt. Anspruch der vorgestellten automatischen Software bleibt eine signalunabhängige Bewertung der Kinetik.

M_{T2} und $M_{AUCTTPHotspot}$, die zu Beginn noch eine Signifikanz aufwiesen, fielen nach weiteren Analysen der Teilmengen aus der Ergebnisliste heraus. Somit verblieben nur 6 Merkmale, darunter Form, Textur, kinetische Merkmale und Alter mit einer starken Signifikanz zu den Zielklassen (s. Tab. 6.8).

Die selektierten Merkmale waren bereits bei der Clusteranalyse auffällig (s. Abb. 6.12), denn sie zeigten im Dendrogramm wenig Ähnlichkeit untereinander und beschreiben somit unterschiedliche Facetten des Sachverhaltes. Bei der Hauptkomponentenanalyse (s. Abb. 6.13) zeigt M_{FormA} einen sehr starken und $M_{TTPHotspot}$ und $M_{AUCHotspot}$ einen etwas schwächeren Bezug auf beide Hauptachsen, während $M_{Kontrast}$, M_{Alter} und $M_{WashOutToken}$ noch auf eine der beiden Achsen einen starken Einfluss ausüben. Diese Merkmale haben also einen großen Einfluss auf den Gesamtsachverhalt.

Durch Streudiagramme (s. Kap. 4.6.3) wurde abschließend noch einmal überprüft, ob die Merkmale Abhängigkeiten untereinander aufweisen. Ein Diagramm wurde aus den selektierten Merkmalen der Regressionsanalyse mit den Zielklassen $Ziel_{Mal}$ und $Ziel_{Ben}$ des zweiten Klassifikators (s. Tab. 6.8) erzeugt. Es zeigt im Wertebereich von -1 bis $+1$ keinerlei Korrelationen unter den Merkmalswerten (s. Abb. 6.16). Im Sinne der Abhängigkeit wurden durch die Regressionsanalyse zweckmäßige Merkmale gewählt. Gleiches gilt auch für die Merkmalsselektion des ersten Klassifikators (s. Tab. 6.6), jedoch ist das Diagramm aufgrund der 16 verbleibenden Merkmale für eine Abbildung zu groß.

Zum Abschluss wurde ein weiteres NN mit den selektierten Merkmalen trainiert und zur Diskriminierung maligner und benigner Tumore auf den GS angewandt (vgl. vorheriges Kapitel). Durch Reduktion der Merkmale konnte die Vorhersageleistung von $NN_{Selektion}$ mit einer Sensitivität von 88 % bei Aufrechterhaltung der Spezifität von 68 % gegenüber NN_{Alle} (s. Tab. 6.15) noch einmal verbessert werden.

Läsion / zusätzlicher Fund	Koeffizient	Std. Abw.	t-value	p-value	code
$M_{Volumen}$	0.0105704	0.0018083	5.845	5.23e-09	***
$M_{MaxDiameter}$	0.0117054	0.0022441	5.216	1.87e-07	***
$M_{MinDiameter}$	-0.0665593	0.0034462	-19.314	< 2e-16	***
M_{Form1}	0.0043584	0.0021056	2.070	0.038484	*
$M_{Begrenzung1}$	-0.0060652	0.0017946	-3.380	0.000729	***
M_{Form3}	-0.0066864	0.0019364	-3.453	0.000557	***
M_{Form2}	0.0248686	0.0023292	10.677	<2e-16	***
$M_{Nekrose}$	0.0017499	0.0050509	0.346	0.729013	
M_{Form4}	-0.0157351	0.0026456	-5.948	2.82e-09	***
$M_{Ringenhancement}$	-0.0058579	0.0050229	-1.166	0.243550	
$M_{Begrenzung2}$	0.0051317	0.0012668	4.051	5.15e-05	***
$M_{Kontrast}$	-0.0043165	0.0011980	-3.603	0.000316	***
M_{T2}	-0.0027962	0.0012492	-2.238	0.025221	*
$M_{TTPToken}$	0.0060840	0.0023367	2.604	0.009238	**
$M_{TTPHotspot}$	0.0027609	0.0022640	1.219	0.222702	
$M_{WashOutHotspot}$	-0.0133802	0.0032528	-4.113	3.93e-05	***
$M_{WashOutToken}$	-0.0023383	0.0032753	-0.714	0.475303	
$M_{AUCToken}$	0.0271473	0.0583431	0.465	0.641725	
$M_{AUCHotspot}$	-0.0654760	0.0561282	-1.167	0.243426	
$M_{AUCTTPHotspot}$	0.0645634	0.0555713	1.162	0.245343	
$M_{AUCTTPToken}$	-0.0247497	0.0576623	-0.429	0.667775	
$M_{MaxSIToken}$	-0.0637284	0.0058027	-10.983	<2e-16	***
$M_{MaxSIHotspot}$	0.0342506	0.0058462	5.859	4.83e-09	***
M_{Alter}	0.0002795	0.0011667	0.240	0.810698	

Tabelle 6.5: Regressionsanalyse, Bestimmtheitsmaß R^2 : 0.1913, $R^2_{korrigiert}$: 0.1891, F-Test des gesamten Regressionsmodells: $F = 87.5$ mit p-Wert < 2.2e-16. Signifikanzkodierung: p-Wert: $0.0 < *** < 0.001 < ** < 0.01 < * < 0.05 < . < 0.1 < < 1.0$.

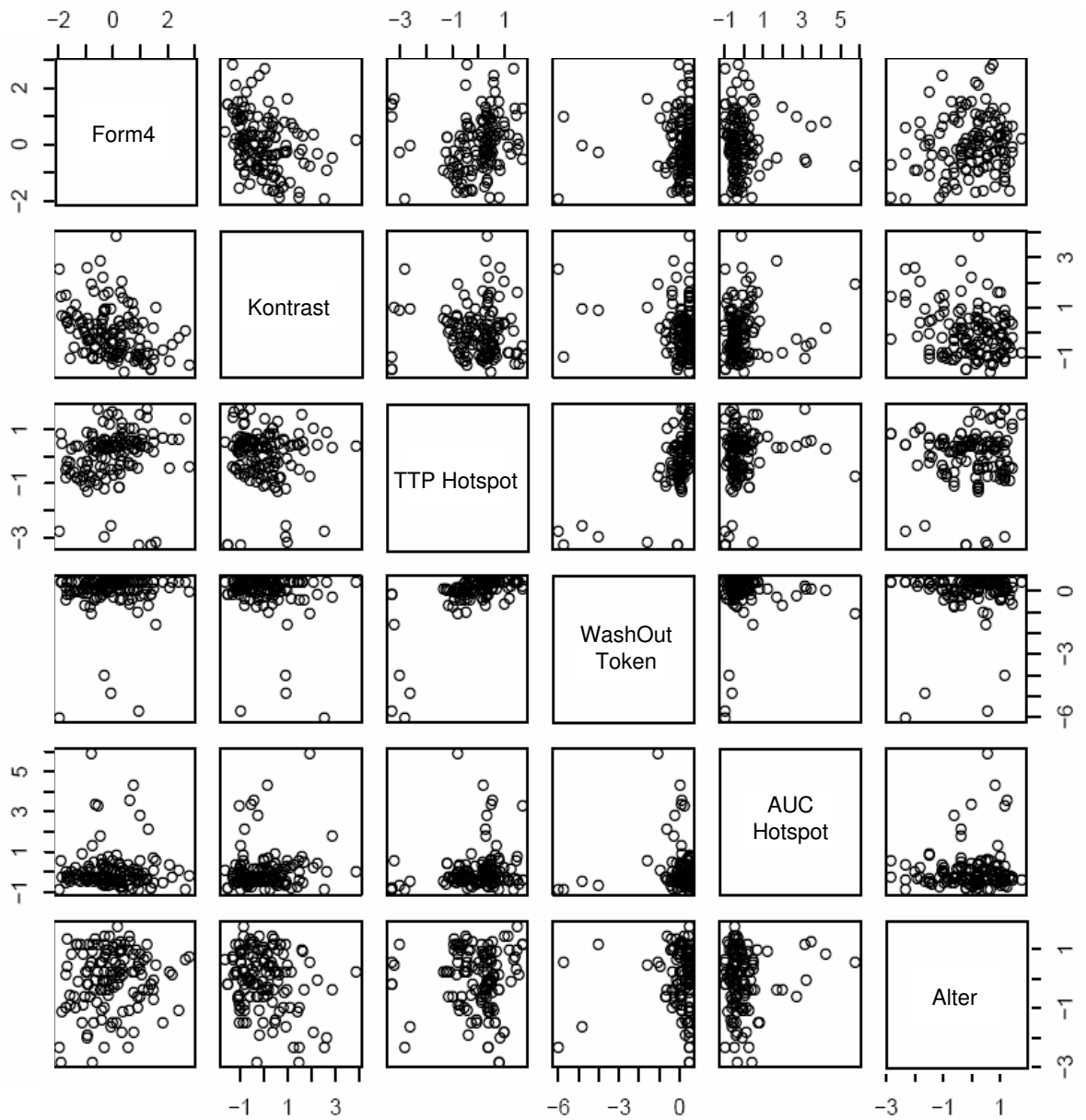


Abbildung 6.16: Streudiagramm der selektierten Merkmale zur Diskriminierung maligner und benigner Läsionen.

Läsion / zusätzlicher Fund	Koeffizient	Std. Abw.	t-value	p-value	code
$M_{Volumen}$	0.010596	0.001797	5.896	3.85e-09	***
$M_{MaxDiameter}$	0.011596	0.002242	5.171	2.38e-07	***
$M_{MinDiameter}$	-0.067127	0.003442	-19.502	< 2e-16	***
M_{Form1}	0.004474	0.002105	2.125	0.033605	*
$M_{Begrenzung1}$	-0.005980	0.001793	-3.335	0.000856	***
M_{Form3}	-0.007002	0.001934	-3.620	0.000296	***
M_{Form2}	0.024688	0.002327	10.608	< 2e-16	***
M_{Form4}	-0.016412	0.002635	-6.229	4.90e-10	***
$M_{Begrenzung2}$	0.005108	0.001259	4.057	5.01e-05	***
$M_{Kontrast}$	-0.004309	0.001196	-3.602	0.000318	***
M_{T2}	-0.002970	0.001242	-2.391	0.016823	*
$M_{TTPToken}$	0.006332	0.001511	4.192	2.79e-05	***
$M_{WashOutToken}$	-0.013615	0.001570	-8.670	< 2e-16	***
$M_{MaxSIToken}$	-0.063486	0.005713	-11.112	< 2e-16	***
$M_{MaxSIHotspot}$	0.034129	0.005759	5.926	3.22e-09	***

Tabelle 6.6: Regressionsanalyse, Bestimmtheitsmaß R^2 : 0.1893, $R^2_{korrigiert}$: 0.188, F-Test des gesamten Regressionsmodells: $F = 138.4$ mit p-Wert < 2.2e-16. Signifikanzkodierung: p-Wert: $0.0 < *** < 0.001 < ** < 0.01 < * < 0.05 < . < 0.1 < < 1.0$.

Die Reduktion auf 6 Merkmale aus ursprünglich 20 verbessert die Diskriminanzvorhersage maligner Tumore. Die hier selektierten Merkmale bilden somit auch den reduzierten Merkmalsvektor beider Klassifikatoren des Ergebnisteils (s. Kap. 5.4). Dort wurden sie jedoch für die Vorhersage auf Patientenbasis genutzt, während $NN_{Selektion}$ in diesem Kapitel für eine objektbasierte Aussage angewendet wurde.

Die Wahl der signifikantesten Merkmale aufgrund der Regressionsanalyse muss nicht zwingend zur optimalen Merkmalsauswahl und somit besten Diskriminierung führen. Eine Kombination weniger signifikanter Merkmale kann unter Umständen einen besseren Beitrag leisten. Sicherheit bringt daher nur das Ausprobieren aller Merkmalskombinationen. Dies ist häufig zu aufwendig. In [HTF01] wird ein Überblick verschiedener Verfahren und Strategien zur Merkmalsselektion gegeben. Auch durch ein Erweitern des Patientenkollektivs kann sich durch die hinzugekommenen Instanzen der Sachverhalt derart ändern, dass bei erneuter Analyse zuvor weniger signifikante Merkmale nun ausschlaggebend werden.

6.4.3 Vorhersage des ungesicherter Goldstandards

Folgend werden auf den histologisch ungesicherten GS-Vorgaben (s. Tab. 5.1) durch die bereits trainierten Klassifikatoren des Ergebnisteils Dignitätsvorhersagen geleistet. Dieser Teil des GS wurde nicht für das Training verwendet, sondern diente lediglich zur Feststellung der zusätzlichen Funde.

Maligne / Benign	Koeffizient	Std. Abw.	t-value	p-value	code
M_{Form1}	-0.008159	0.060594	-0.135	0.89314	
$M_{Begrenzung1}$	-0.002430	0.060289	-0.040	0.96792	
M_{Form3}	0.038101	0.045348	0.840	0.40266	
$M_{Nekrose}$	0.016865	0.303192	0.056	0.95574	
M_{Form4}	0.202491	0.065069	3.112	0.00238	**
$M_{Ringenhancement}$	-0.021340	0.298247	-0.072	0.94309	
$M_{Begrenzung2}$	-0.049304	0.039354	-1.253	0.21298	
$M_{Kontrast}$	0.114722	0.040899	2.805	0.00597	**
M_{T2}	0.066759	0.037834	1.765	0.08047	.
$M_{TTPToken}$	0.010992	0.062474	0.176	0.86067	
$M_{TTPHotspot}$	0.149515	0.056340	2.654	0.00916	**
$M_{WashOutHotspot}$	-0.203753	0.071312	-2.857	0.00513	**
$M_{WashOutToken}$	0.085746	0.086426	0.992	0.32335	
$M_{AUCToken}$	0.486683	0.307254	1.584	0.11612	
$M_{AUCHotspot}$	-0.605683	0.252284	-2.401	0.01807	*
$M_{AUCTTPHotspot}$	0.560747	0.241246	2.324	0.02198	*
$M_{AUCTTPToken}$	-0.446253	0.295164	-1.512	0.13348	
M_{Alter}	-0.081613	0.035512	-2.298	0.02348	*

Tabelle 6.7: Regressionsanalyse, Bestimmtheitsmaß R^2 : 0,5439, $R^2_{korrigiert}$: 0.4552, F-Test des gesamten Regressionsmodells: $F = 6.132$ mit p-Wert $8.44e-11$. Signifikanzkodierung: p-Wert: $0.0 < *** < 0.001 < ** < 0.01 < * < 0.05 < . < 0.1 < < 1.0$.

Maligne / Benign	Koeffizient	Std. Abw.	t-value	p-value	code
M_{Form4}	0.18078	0.03734	4.842	3.84e-06	***
$M_{Kontrast}$	0.12275	0.03745	3.277	0.00137	**
$M_{TTPHotspot}$	0.12070	0.04343	2.779	0.00632	**
$M_{WashOutToken}$	-0.12409	0.04466	-2.779	0.00633	**
$M_{AUCHotspot}$	-0.31677	0.14542	-2.178	0.03132	*
M_{Alter}	-0.08961	0.03366	-2.662	0.00882	**

Tabelle 6.8: Regressionsanalyse, Bestimmtheitsmaß R^2 : 0,4647, $R^2_{korrigiert}$: 0.4293, F-Test des gesamten Regressionsmodells: $F = 13.13$ mit p-Wert $1.649e-13$. Signifikanzkodierung: p-Wert: $0.0 < *** < 0.001 < ** < 0.01 < * < 0.05 < . < 0.1 < < 1.0$.

ungesichert GS	Klasse	Maligne	Benigne	zusätzlich. Fund	Σ
maligner verdächtiger Tumor	3	20	20	32	72
benigner verdächtiger Tumor	4	11	67	45	123
unsicherer Tumor	5	5	19	15	39
Brustwarzenanreicherung	6	31	17	33	81

Tabelle 6.9: Vorhersage auf dem ungesicherten Goldstandard

Die Objekte der automatischen Segmentierung wurde zum Erzeugen weiterer Testdaten auf Übereinstimmung mit den 424 ungesicherten GS-Objekten untersucht (vgl. Kap. 4.5.2), 315 Objekte erfüllten dabei die Überdeckungskriterien $O_{in} > 15\%$ und $O_{out} < 50\%$ (s. Tab. 5.8). Die Vorhersagen nach der objektbasierten Leave-One-Out Methode der beiden Klassifikatoren NN_{LA} und NN_{MB} (s. Kap. 5.4) auf diesen Testdaten sind in Tabelle 6.9 aufgeführt. Im Mittel werden mit ungefähr 40 % sehr viele Instanzen der betrachteten Dignitätsklassen als zusätzliche Funde vorhergesagt. Auffallend ist, dass die histologisch ungesicherten Anreicherungen im Vergleich zu den gesicherten im Durchschnitt über ein kleineres Volumen verfügen. Denn die Biopsie wird an den größeren Funden stellvertretend vorgenommen.

Mit 20 Objekten wurde etwa ein Drittel der verdächtigen Karzinome (GS-Klasse 3) als Karzinom vorhergesagt. Eine Mehrheit von 67 Objekten der benignen verdächtigen Objekte (GS-Klasse 4) wurde als gutartig erkannt. Für die unsicheren Läsionen (GS-Klasse 5), die mehrheitlich als benigne vorhergesagt wurden, liegt kein Anhaltspunkt für eine Bewertung vor. Die Einstufung der Vorhersage der Brustwarzenanreicherung (GS-Klasse 6) ist schwierig. Nach den BI-RADSTM-Kategorien stellen diese Bereiche keine Tumore dar. Sie müssten von der Software somit als zusätzliche Funde erkannt werden. Andererseits sind die Anreicherungen kompakt in der Form und zeigen als gesundes Gewebe ein eher kontinuierliches Anreicherungsverhalten. Eine Vorhersage als gutartige Läsion steht somit in keinem Widerspruch zu dem gelernten Sachverhalt der Neuronalen Netze. Mit ungefähr einem Drittel werden die Brustwarzenanreicherungen mit einer malignen Vorhersage definitiv falsch bewertet.

Die Aussage der gesamten Untersuchung ist gering, da die Zielklassen nicht auf gesicherten Histologien basieren.

6.4.4 Klassifikatorenvergleich

Folgend werden die Vorhersageleistungen verschiedener Klassifikatoren (s. Kap. 4.6.5) miteinander verglichen. Auch hier lassen sich die gewonnenen Aussagen nicht generalisieren, weil auf den wenig repräsentativen Lerndaten der vorliegenden Arbeit bestimmte Eigenschaften eines Klassifikators zu Vorteilen führen können, die auf repräsentativeren Lernlisten nicht entstünden.

Die Fehlerrate wird entsprechend dem Ergebnisteil wieder durch das Leave-One-Out

Verfahren (s. Kap. 4.6.4) auf Patientenbasis durch zwei hintereinander ausgeführte binäre Klassifikatoren bestimmt. Auch fand das Training aller untersuchten Klassifikatoren nur auf den selektierten Merkmalen (s. Kap. 6.4.2) statt. Da *Boost.Ada IS* und *Random Forest IS* als Metaklassifizierer über eine eigene Selektion verfügen, werden sie zusätzlich noch einmal auf allen Merkmalen trainiert. Beim *k*-nearest Neighbor Klassifikator wurde *k* experimentell auf 7 Nachbarn festgelegt. Abb. 6.17 zeigt in der Übersicht die Leistung der untersuchten Klassifikatoren durch Sensitivität, Spezifität und Treffgenauigkeit.

Das Neuronale Netz (NN) erreicht wie bereits im Ergebnisteil präsentiert (s. Tab. 5.9) eine Sensitivität/Spezifität/Treffgenauigkeit von 88 %/ 72 % /82 %. Die Diskriminanzanalyse in der flexiblen Variante (*FDA*) erreicht die gleiche Treffgenauigkeit mit einer etwas besseren Sensitivität auf Kosten der Spezifität. In der linearen Variante (*LDA*) liegt die Treffgenauigkeit nur ein Prozentpunkt niedriger. Die Vorhersage durch die lineare SVM (*LinSVM*) erreicht ebenfalls die Vorhersageleistung der NN. Sowohl *Tree* als auch *k*-nearest Neighbor haben eine geringere Spezifität bei der Vorhersage des Sachverhaltes. *Random Forest* weist gegenüber dem *Tree* eine bessere Sensitivität auf; das Boosting steigert dessen Treffgenauigkeit noch einmal. Dennoch bleiben die Leistungen der Entscheidungsbäume hinter denjenigen Verfahren zurück, die die Merkmale im Raum gleichzeitig betrachten.

Die autonome Merkmalsselektion der Metaklassifikatoren (*Random Forest IS* und *Boost.Ada IS*) brachte keine Verbesserung gegenüber der Auswahl durch die Regression (s. Kap. 6.4.2).

Unter allen getesteten Klassifikatoren ist besonders das Training Neuronaler Netze anspruchsvoll. Es gibt viele anzupassende Parameter (Anzahl der Hidden Layer, Anzahl der Neuronen, Aktivierungsfunktionen, ...). Letztendlich müssen die Einstellungen experimentell ermittelt werden. Jeder Trainingsprozess ist sehr rechenintensiv. Durch die Zufallsinitialisierung der Gewichte kann ein trainiertes Netz trotz gleicher Voraussetzungen kaum reproduziert werden. Auch kann aus dem Trainingsprozess eines NN nicht abschließend bestimmt werden, wie groß der Einfluss einzelner Merkmale auf die Zielklassen ist.

Die Diskriminanzanalyse bedarf dagegen keiner Parameteranpassungen. Bei der Betrachtung der Klassifikation als Regressionsproblem wird auf eine begründete mathematische Theorie zurückgegriffen. Das Verfahren kann auch zur Selektion von Merkmalen verwendet werden (s. Kap. 6.4.2), was besonders bei mehrdimensionalen Merkmalsräumen sinnvoll ist. Auch die SVM zeichnen sich durch ein einfacheres und reproduzierbares Training aus.

Die Neuronalen Netze erbringen auf dem vorhandenen Kollektiv keine erkennbar bessere Leistung als beispielsweise die SVM oder flexible Diskriminanzanalyse. Aufgrund der aufwendigen und auch fehleranfälligen Handhabung des Trainings Neuronaler Netze sind die beiden Alternativen zu bevorzugen.

Der Vergleich der verschiedenen Klassifikatoren macht auch deutlich, dass mit den vorhandenen Merkmalen eine bestimmte Fehlerrate nicht unterschritten werden kann. Der Grund dürfte daher nicht in dem gewählten Diskriminanzverfahren liegen, sondern im Informationsgehalt der zugrundeliegenden Instanzen der Lerndaten. Es liegt der Schluss nahe, dass sich erst durch Hinzufügen von Merkmalen mit neuem Informationsgehalt die Fehlerrate senken lässt.

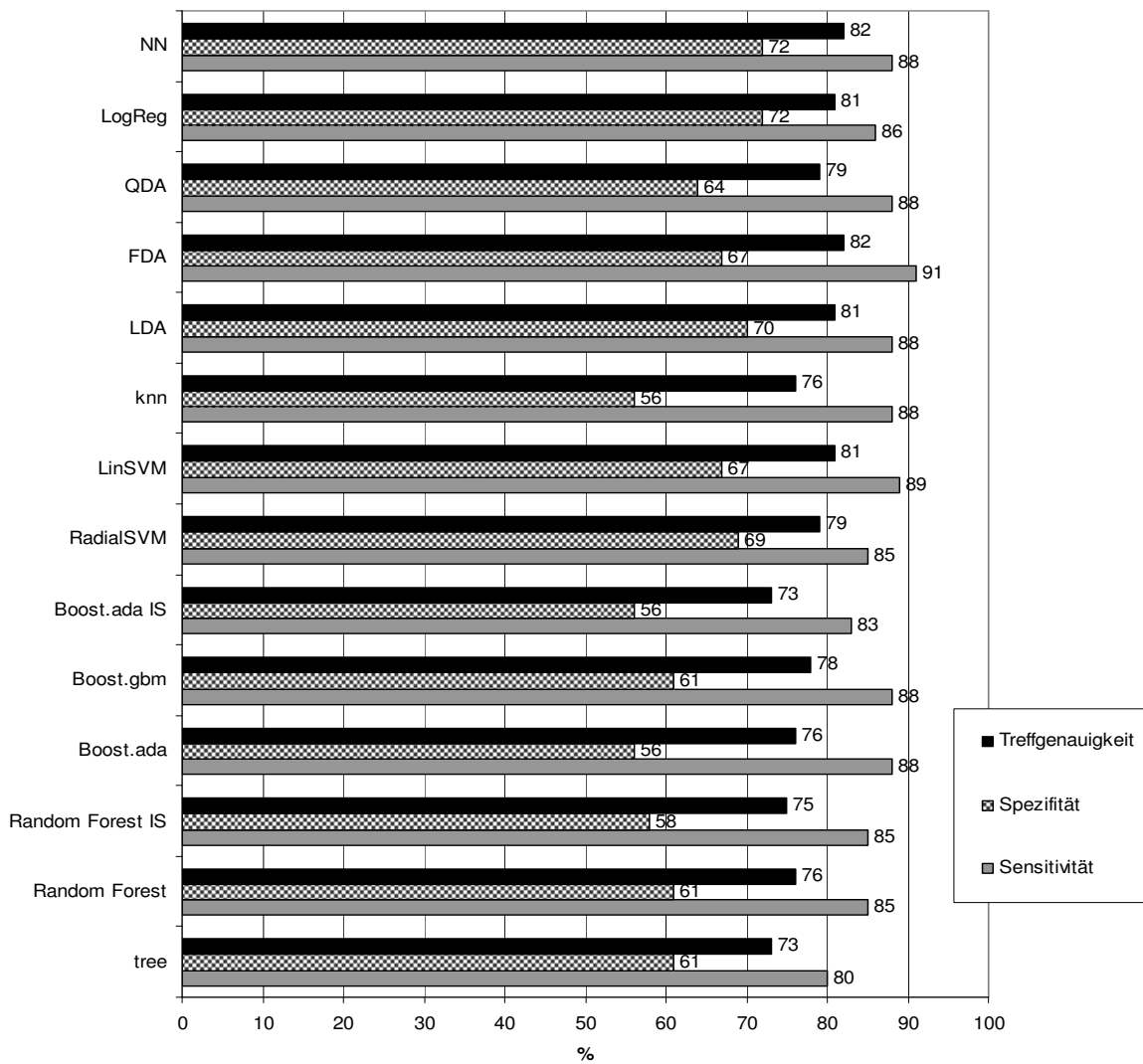


Abbildung 6.17: Vorhersageleistung verschiedener Klassifikatoren auf dem gesamten Kollektiv.

6.4.5 Beurteilung zusätzlicher Funde

Dem menschlichen Auge weit unterlegen ist die Software bei der Bewertung zusätzlicher Funde. Bildbereiche, die der Radiologe mit Leichtigkeit als unbedeutend interpretieren kann, arbeitet die Segmentierung aufgrund der Regeln stoisch ab, erzeugt daraus Token und klassifiziert sie im ungünstigen Fall als Läsion. Abb. 6.18 zeigt verschiedene Beispiele segmentierter zusätzlicher Funde. Die Mehrheit der zusätzlichen Funde liegt im Thorax- oder Brustwarzenbereich oder im Übergangsbereich der Brust zur Luft.

Die Entscheidung über zusätzliche Funde oder Läsionen beginnt mit der Segmentierung und endet durch die extrahierten Merkmale mit der anschließenden Klassifizierung. Die Merkmale wurden jedoch zur Diskriminierung der Tumordignitäten entwickelt. Für den Klassifikator NN_{LA} wurden sie einfach nur wiederverwendet. Auch ist die Menge an zusätzlichen Funden verglichen mit der Anzahl an Läsionen sehr groß, was das Trainieren der Neuronalen Netze erschwert.

Die Reduktion zusätzlicher Funde kann bereits bei der Segmentierung beginnen (s. Kap. 6.2.2). Wird die Anzahl der Objekte und damit der Instanzen verringert, ist das Trainieren der Klassifikatoren besser beherrschbar (s. Kap. 4.5.4). Des Weiteren müssen Merkmale entwickelt werden, die besser zur Differenzierung der zusätzlichen Funde geeignet sind. Beispielsweise können durch das Einbeziehen der Position des Objektes innerhalb der Brust Bereiche wie der Thorax oder die Brustwarzen als Artefakte ausgeschlossen werden. Dort sind keine Läsionen zu erwarten.

Die hohe Zahl an zusätzlichen Funden bereitet in zweierlei Hinsicht Schwierigkeiten. Zum einen verschlechtert sich die Lesbarkeit der Ausgabe des Prototyps, zum anderen ist für einen ungeübten Anwender ohne Hintergrundwissen über die implementierte Bildverarbeitung wenig verständlich, warum die Software solche einfachen Sachverhalte fehlerhaft deutet. Beides verschlechtert die Akzeptanz einer Diagnosesoftware.

Der Radiologe kann die von der Software produzierten zusätzlichen Funde sofort als solche deuten, seine Gesamtbeurteilung des Patienten anhand der computergenerierten Zweitmeinung muss dadurch nicht beeinträchtigt sein. Eine automatisch generierte Diagnose pro Patient durch die Software führt dagegen durch jeden verbleibenden zusätzlichen Fund zu einer Falschaussage (s. Tab. 5.10).

6.4.6 Problemfall

Abschließend wird auf einen Problemfall eingegangen, der zwangsweise immer eine falsche Vorhersage in den Statistiken verursacht. Eine Patientin erkrankte an einem bösartigem Tumor, in dessen angrenzendem Gewebe eine weitere gutartige Biopsie vorlag (s. Abb. 6.19 b). Zu einem Problem wurde der Fall deshalb, weil die Segmentierung diese beiden Objekte nicht zu trennen vermochte. Auch für das menschliche Auge ist hier kaum eine Grenze ersichtlich (s. Abb. 6.19 a). Das übersegmentierte Objekt überlagert somit beide GS-Vorgaben. Beim Export (s. Kap. 4.5.2) blieb dieses Objekt den Trainingsdaten vor-enthalten, da die Überdeckungskriterien (s. Formel 4.31) ein O_{Out} Wert von mehr als 50 % erkannte. Generell konnte der Prototyp zwar alle Karzinome im Patientenkollektiv seg-

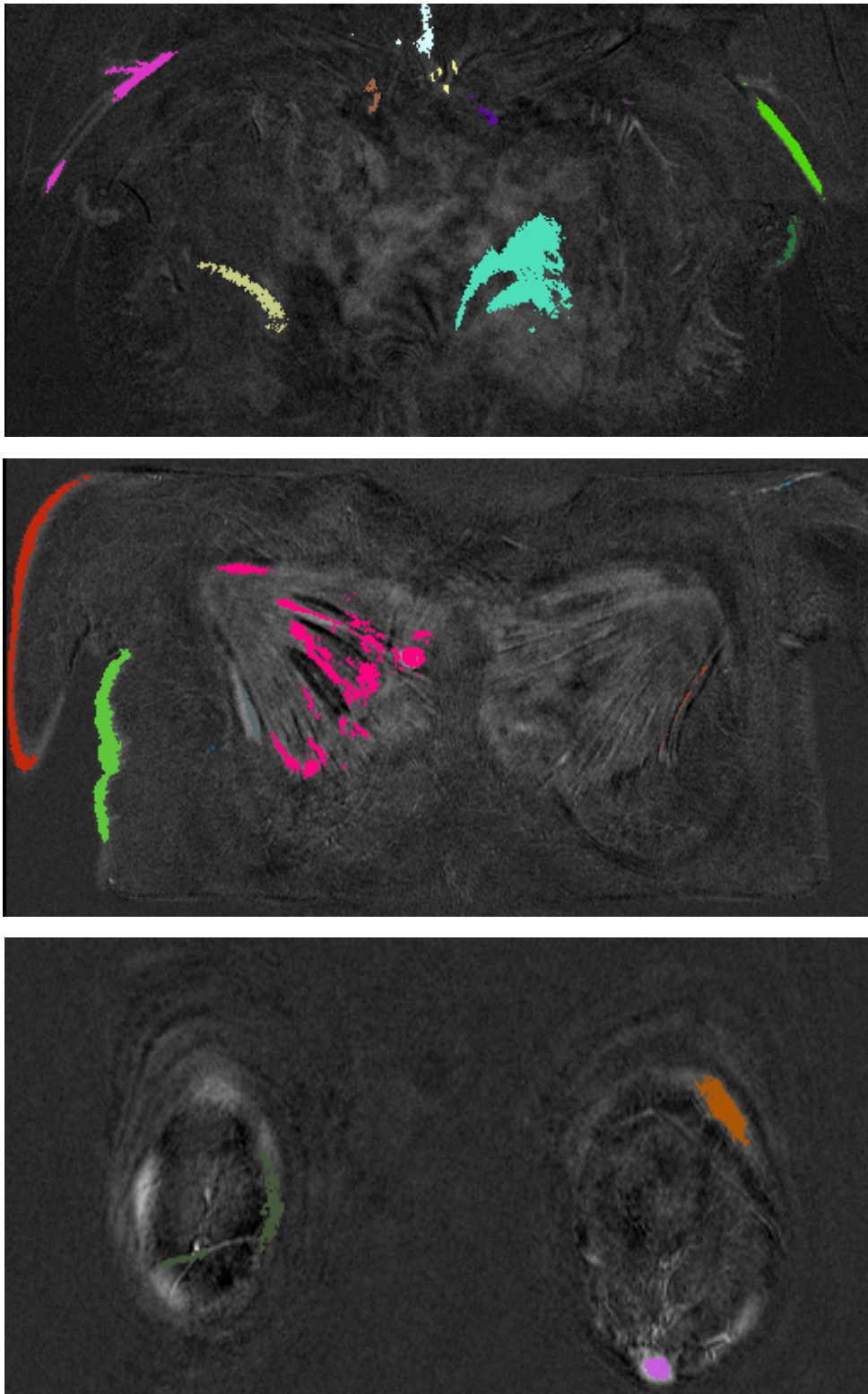
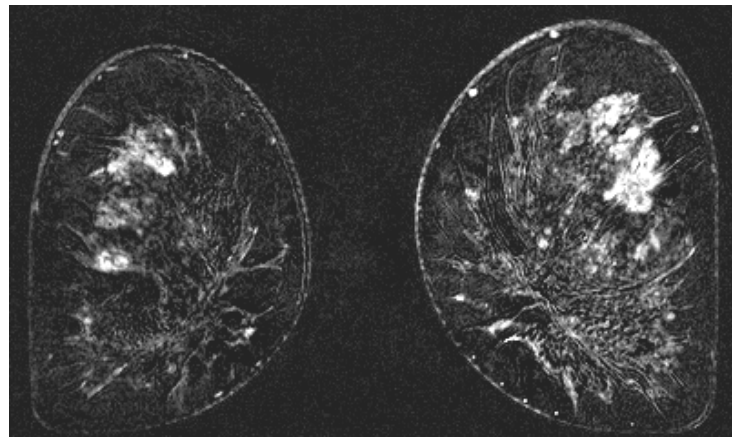
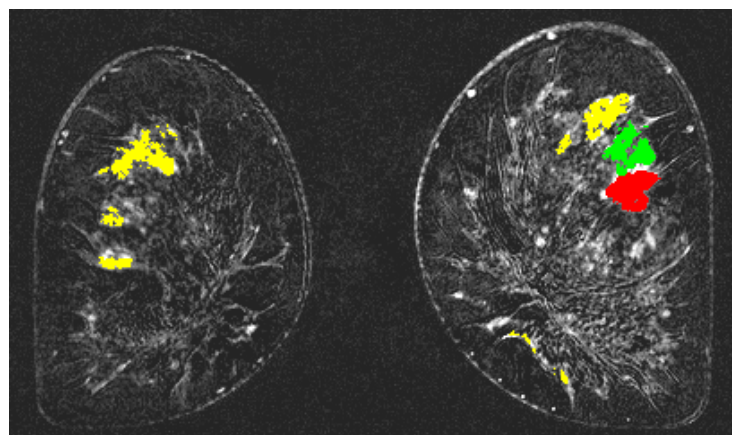


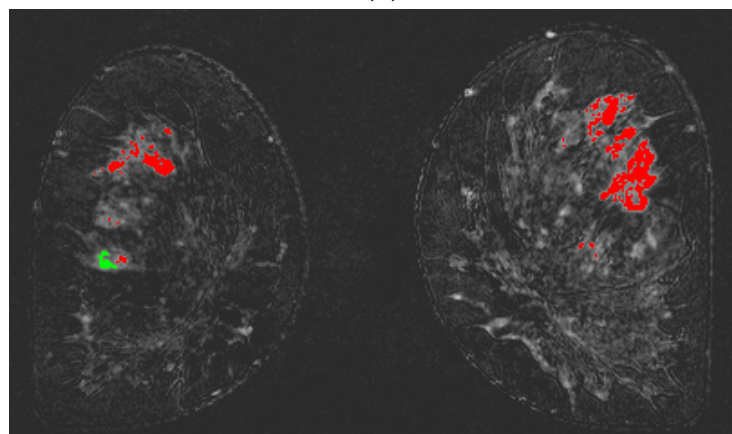
Abbildung 6.18: Bildbeispiele verschiedener segmentierter zusätzlicher Funde.



(a)



(b)



(c)

Abbildung 6.19: Problemfall mit falscher Vorhersage: (a) Tumor (oben rechts) in Subtraktionsserie, (b) GS mit direkt angrenzender benigner (grün) und maligner (rot) Histologie und weiterer benigner verdächtiger Tumore (gelb), (c) Vorhersage als Karzinom (rot) durch die Prototypsoftware.

mentieren, aber von den 87 malignen GS-Objekten erhielten aufgrund dieses Problemfalls nur 86 Einzug in die Lerndaten.

Die Testdaten für die Fehlerschätzung enthalten jedoch alle automatisch segmentierten Objekte. Das darin enthaltene übersegmentierte Problemobjekt wurde während der Kreuzvalidierung (s. Kap. 4.6.4) als Karzinom vorhergesagt (s. Abb. 6.19 c). Zwar vereint die Übersegmentierung sowohl gut- als auch bösartiges Verhalten, die Software erkannte im Problemfall aber den bösartigen Anteil. Die beiden angrenzenden GS-Vorgaben (s. Abb. 6.19 b) werden in den Testdaten nur durch das eine übersegmentierte Objekt repräsentiert. Im Rahmen der Fehlerabschätzung wird die vorhergesagte Dignität des Problemobjektes aber mit beiden GS-Vorgaben verglichen. Als Konsequenz wird die als Karzinom vorhergesagte Problemläsion beim Vergleich mit der gutartigen GS-Vorgabe als Fehler gewertet. Im umgekehrten Fall einer gutartigen Vorhersage für die Problemläsion wäre die maligne GS-Vorgabe als Fehler bewertet worden. Unabhängig davon, auf welche Dignität sich der Klassifikator für dieses Objekt festlegt, entsteht ein Vorhersagefehler.

6.4.7 Erkenntnisse und Grenzen

Die Prototypsoftware ist in der Lage, eindeutige Dignitätsvorhersagen pro Läsion zu geben. Die Vorhersageleistung des Prototyps auf dem GS als Ergebnis aller ausgeführten Teilkomponenten erreicht in etwa die in der Literatur publizierten Vorhersagen von Radiologen. Die hohe Anzahl verbleibender zusätzlicher Funde macht eine abschließende Aussage pro Patient nicht praktikabel. Die mangelnde Leistung der Vorhersage von zusätzlichen Funden liegt in der Wahl der Merkmale, deren Informationsgehalt auf die Differenzierung von gut- und bösartig abzielt und für die Diskriminierung von zusätzlichen Funden einfach nur wiederverwendet wurde.

Die Untersuchung in Kapitel 6.4.4 zeigt, dass neben den NN auch alternative Klassifikatoren vergleichbare Vorhersagen auf den vorliegenden Sachverhalt leisten. Mit SVM und flexibler Diskriminanzanalyse sind dies im Gegensatz zu NN einfach zu handhabende Verfahren. Auch macht der Vergleich aller Klassifikatoren deutlich, dass mit den extrahierten Merkmalen eine bestimmte Fehlerrate nicht unterschritten werden kann. Zur Verbesserung der Vorhersageleistung bedarf es der Erschließung zusätzlicher Informationen durch neue Merkmale.

Die Analyse des Einflusses verschiedener Merkmale auf die Diskriminanz macht deutlich, dass eine hohe Spezifität erst durch gemeinsames Betrachten von morphologischen und kinetischen Eigenschaften erreicht wird (s. Kap. 6.4.1). Viele CAD-Systeme betreiben jedoch nur kinetische Analysen. Die absolute SI der Kinetik (M_{MaxSI}) ist ein sehr sensibles Merkmal, das aber zugunsten einer signalunabhängigen Kinetikinterpretation nicht genutzt wird.

Die Reduktion der Dimension des Merkmalsraums durch Selektion nur der signifikantesten Merkmale verbessert die Vorhersageleistung (s. Kap. 6.4.2). Unter den sechs selektierten Merkmalen zur Differenzierung maligner und benigner Läsionen sind im Wesentlichen die Kriterien des Göttinger Scores enthalten (s. Tab. 2.4). M_{TTP} , $M_{WashOut}$ und M_{AUC} charakterisieren die Kinetik, $M_{Kontrast}$ gibt die KM-Verteilung innerhalb der Läsion wieder

Komponente	Breast <i>fTP</i>	DynaCAD TM	CADstream	ONCAD TM	Prototyp
Registrierung	A	A	A	-	A
Segmentierung	-	-	M	-	A
Kinetik	A	M	A	M	A
Morphologie	-	-	L	M	A
Klassifikation	L	-	L/M	H	H
Diagnose	M	M	M	M	A
Workflow	M	M	M	M	A

Tabelle 6.10: Komponentenvergleich kommerzieller Software mit dem Prototypen der vorliegenden Arbeit. Legende: (A) Automatisch, (M) Manuell durch den Benutzer, (L) Low-Level-Diagnose, (H) High-Level-Diagnose, (-) nicht vorhanden

und M_{Form4} die Form durch die fraktale Dimension. Während kein selektiertes Merkmal die Begrenzung beschreibt, wird stattdessen noch das Alter bewertet.

Auffällig ist, dass falsche Vorhersagen vor allem für Objekte geringen Volumens getroffen werden. Es kann vermutet werden, dass die Aussagen der Merkmale auf kleinen Objekten weniger diskriminant sind als die der großen, bzw. dass sich die Morphologien kleiner Objekte sehr ähneln. Hier zeigt sich die Limitierung durch die Ortsauflösung. Eine mögliche Verbesserung wäre das Training mehrerer Klassifikatoren gestaffelt nach Objektgröße oder aber auch die Hinzunahme des Größenmerkmals $M_{Volumen}$ in den Trainingsprozess. Denn auch BI-RADSTM empfiehlt, für Foki eine Beurteilung nur anhand der Kinetik vorzunehmen ganz ohne morphologische Betrachtung.

6.5 Vergleich zu kommerziellen Systeme

Ein Vergleich zu den kommerziellen CAD-Systemen (s. Kap. 3.2) ist nur auf einer sehr abstrakten Ebene möglich (s. Tab. 6.10). Einerseits haben die Bildverarbeitungskomponenten dieser Systeme mit dem Prototyp der vorliegenden Arbeit bis auf die Titulierung CAD funktional nicht viel gemein (s. Kap. 3.1). Andererseits liegen zu den wenigen Gemeinsamkeiten kaum wissenschaftlich fundierte Informationen über deren Funktionsweisen vor. Auch lassen sich die Ausgaben der Systeme nicht miteinander in einem Vergleich bewerten, da alle Konkurrenzprodukte manuelle Interaktionen und Vorgaben des Benutzers verlangen. Somit bliebe letztendlich nur ein Vergleich Radiologe als Systemanwender versus automatischer Prototyp. Folgend wird jede Teilkomponente des Prototyps mit den kommerziellen Systemen verglichen.

6.5.1 Registrierung

Einzig die Registrierung als Vorverarbeitung stellt eine Gemeinsamkeit zu einigen Konkurrenzsystemen dar. Der Prototyp wählt mit dem nicht rigiden Algorithmus (s. Kap. 4.2) den fortschrittlichsten Ansatz zur Registrierung der Brustbilder (s. Kap. 3.3.1).

Das Befundungstool CADstreamTM bietet nur eine 2D-Registrierung einzelner Bildschichten an, ein 3D ist aufgrund der hohen Laufzeit nicht marktreif. Die verfügbare Registrierung kann sowohl fettsupprimierte als auch ungesättigte Aufnahmen verarbeiten. Bedingt durch das 2D-Verfahren können Bewegungen in einer Raumebene nicht korrigiert werden. Auch *Breast fTP* und DynaCADTM verbessert die Bildqualität durch eine entsprechende Vorverarbeitung. Eine Bewertung der Qualität der verschiedenen Algorithmen ist, wie in Kapitel 6.1.2 erwähnt, objektiv kaum möglich. Auch die Feststellung, ob die konkurrierenden Registrierungen Volumen- und KM erhaltend sind wie der vorgestellte Algorithmus des Prototyps (s. Kap. 6.1.3), soll nicht Gegenstand der vorliegenden Arbeit sein, zumal ohnehin noch deutliche Unterschiede der Systeme an den folgenden Bildverarbeitungs-komponenten bestehen.

Die Software ONCADTM, die auf fettsupprimierten Aufnahmen arbeitet, verfügt über gar keine Registrierung. Daher muss bezweifelt werden, ob die vom System visualisierten Kinetikvorgaben überhaupt korrekt sind.

6.5.2 Segmentierung

Eine automatische Segmentierung (s. Kap. 4.3) der Läsionen als einzelne Objekte ist in keinem Konkurrenzprodukt vorhanden; hier ist der Prototyp allen kommerziellen Produkten überlegen. Die von einigen Produkten titulierte Segmentierung ist lediglich eine farbkodierte Darstellung kinetischer Informationen. Solche Verfahren sind nicht in der Lage, einzelne Läsionen zu erkennen und zu bewerten.

Lediglich CADstreamTM bietet auf der Basis der farbkodierten Darstellung eine – wenn auch manuelle – Segmentierung von Läsionen an. Hier muss der Benutzer den farblich hervorgehobenen Bereich markieren. Das System kann durch Bestimmen der Zusammenhangskomponenten aus der Farbdarstellung ein Token erzeugen. Allerdings werden auch Bereiche zusätzlicher Funde durch das Farboverlay zur Auswahl gestellt, die trotz Ausblendung des anreichernden Herzens in der Brust vorhanden sind. Hier muss der Benutzer die Unterscheidung tatsächlicher Läsionen von den zusätzlichen Funden leisten, indem er nur die relevanten Bildbereiche selektiert. Die erzeugten Token werden jedoch nicht genutzt, um weitere morphologische Merkmale zu berechnen oder eine Dignitätsvorhersage auf Objektebene zu treffen.

Um einen automatisierten Ablauf zu ermöglichen, segmentiert der Prototyp erst alle Anreicherungen, um sie anschließend auf Objektebene morphologisch und kinetisch zu bewerten und durch die Klassifikatoren selbständig zusätzliche Funde auszuschließen. CADstreamTM klassifiziert dagegen erst nur anhand kinetischer Informationen auf Low-Level-Ebene und kompensiert die nicht vorhandene Bewertung zusätzlicher Funde durch die Auswahl der eigentlichen Läsionen durch den Benutzer.

Weder *Breast fTP* noch DynaCADTM verlassen mit ihren Bildverarbeitungs-komponenten das Low-Level. Beide Tools bieten keine Segmentierung von Objekten und können somit keine morphologischen Aussagen über Läsionen machen, sondern lediglich einzelne Voxel mit Anreicherungen farblich hervorheben. Das Farboverlay umfasst auch Bildbereiche, die überhaupt keine Läsionen darstellen.

Die Software ONCADTM spart das Intermediate-Level komplett aus und kompensiert es durch die Leistung des Benutzers. Er muss die relevanten Bereiche ohne Unterstützung durch die Software selbst suchen. Der Anwender hat keine Möglichkeit seine Auswahl zu markieren, daher muss die folgende Morphologiebestimmung ebenfalls manuell erfolgen.

6.5.3 Merkmalsextraktion

Die Prototypsoftware erhebt auf der eigenen Segmentierung automatisch einen Merkmalsvektor aus 24 morphologischen und kinetischen Merkmalen für eine anschließende (High-Level) Klassifikation (s. Kap. 4.4).

Nur ONCADTM verwendet als einziges Konkurrenzprodukt ebenfalls einen Merkmalsvektor, der aufgrund fehlender Segmentierung jedoch fast ausschließlich aus deskriptiven Merkmalen des Benutzers besteht. Lediglich die fraktale Dimension [PL97], die wiederum erst durch manuelle Markierung im Bild von der Software berechnet werden kann, stellt das einzige (halb)automatisch generierte Merkmal dar. Die fraktale Dimension von Alan Penn unterscheidet sich wesentlich von der Berechnung in der vorliegenden Arbeit. Während der Prototyp eine reine Formbeschreibung des Objektes vornimmt, beschreibt Penn durch Einbeziehen der SI deren Texturierung. Sein Verfahren basiert auf einer 2D fraktalen Interpolationsfunktion, von der eine Mehrzahl generiert und mit der Varianz als Fehlermaß die beste Übereinstimmung geschätzt wird. Diese Berechnung muss als wenig robust gegenüber der Bildqualität eingeschätzt werden, da sich Abhängigkeiten sowohl zur Ortsauflösung als auch zu den Signalintensitäten ergeben.

Da ONCADTM keine Registrierung durchführt, muss die Korrektheit der angezeigten Kinetik angezweifelt werden. Ihre Bewertung durch den Benutzer wird ebenfalls in den Merkmalsvektor aufgenommen. Durch hohe Schichtdicken von 4 mm oder mehr lässt das verwendete Protokoll [PKT⁺01] vornehmlich nur – wie bei der fraktalen Dimension angewendet – 2D-Bildverarbeitung zu.

Allgemein erweckt die Software ONCADTM den Eindruck einer Forschungsplattform, die für den Einsatz im klinischen Alltag wenig tauglich erscheint.

Alle anderen Systeme betreiben keine Merkmalsextraktion, da ihnen entweder die Segmentierungskomponente (manuell oder automatisch) als Voraussetzung fehlt oder keine High-Level-Klassifikation angestrebt wird.

6.5.4 Klassifikation

Die Klassifikationsleistung der Systeme muss anhand der Ebene der Bildverarbeitung (s. Kap. 2.4.1) differenziert betrachtet werden.

DynaCADTM blendet die Kinetik als stufenlose Farbskala im Bild ein, die Software verfügt weder über eine High noch Low-Level-Klassifikation und leistet somit keinerlei Dignitätsvorhersage. *Breast fTP* nimmt dagegen eine Low-Level-Klassifikation in maligne, benigne oder Mehrdeutigkeit vor. Letzterer Software kann „Computer Aided Diagnosis“ nur auf der Low-Level-Ebene bescheinigt werden. Die Farbkodierung der Kinetik verteilt sich aber über den gesamten Bildbereich und schließt auch Bildartefakte, Gefäße und zusätzliche Funde ein.

Aus den farblichen Einblendungen der Low-Level-Klassifikation muss der Anwender zuerst die Läsionen als relevante Bildbereiche erkennen. Aufgrund der Beteiligung gleich mehrerer Farbkodierungen erhält er eine mehrdeutige Vorhersage, einen eindeutigen diagnostischen Schluss kann der Radiologe daraus nicht ziehen.

Auch CADstreamTM bietet lediglich eine Low-Level-Klassifikation mit drei resultierenden Farbkodierungen. Allerdings kann der Benutzer durch die manuelle Segmentierung Läsionen als Token definieren. Sie dienen jedoch lediglich zur Aufbereitung der Reportausgabe, in der für den gewählten Bildausschnitt die prozentuale Zusammensetzung der drei Farbkodierungen sowie der Abstand zur Haut und Mamille automatisch zusammengestellt werden. Eine Dignitätsvorhersage für das manuell bestimmte Objekt wird nicht geleistet, der Benutzer kann aber seine eigene Beurteilung unterstützt durch das BI-RADSTM-Lexikon im Report eintragen. Die Betrachtung der bloßen Kinetik ist wenig spezifisch (s. Kap. 6.4.1). Die Spezifität lässt sich erst durch Mitbetrachten auch morphologischer Eigenschaften erhöhen; diese Informationen werden aber von den genannten Produkten nicht erhoben.

Wie in Kapitel 2.3.2 erläutert, ist die Korrelation zwischen Dignität eines Tumors, seiner Gefäßdichte und den erzeugten Signalintensitäten durch das KM im Bild nicht proportional. Die Anreicherung jeder Läsion für sich stellt ein individuelles Verhalten dar. Die Berechnung der Farbkodierungen im Rahmen der Low-Level-Klassifikation erfolgt in den Konkurrenzprodukten jedoch global nach denselben Berechnungen über das gesamte Bild und ist somit weniger präzise als eine Kinetikauswertung pro Läsionenobjekt wie in der vorliegenden Arbeit.

Da die Kinetikberechnungen konkurrierender Systeme – wie in einer Livedemonstration während eines Senologiekongresses demonstriert – signalabhängig sind, müssen alle drei zuvor diskutierten Softwareprodukte eine Parametrisierung anbieten. Somit obliegt es den Einstellungen des Administrators, ob Bildpunkte als gut- oder bösartig beurteilt werden. Die Prototypsoftware verwendet Kinetikmerkmale, die wesentlich robuster gegenüber der Bildqualität sind (s. Kap. 6.3.3); hier sind keine Anpassungen nötig.

Lediglich ONCADTM (auf manuellen Merkmalen) und die Prototypsoftware führen eine echte High-Level-Klassifikation auf morphologischen und kinetischen Merkmalen aus. Im Gegensatz zum Prototyp verwendet ONCADTM keine modellbasierte Diskriminanz, sondern summiert ähnlich dem Göttinger Score die Merkmalswerte auf und vergleicht mit Scoringwerten bekannter Histologien. Hier werden die Möglichkeiten des mehrdimensionalen Merkmalsraums mit der Betrachtungsmöglichkeit der Abhängigkeiten einzelner Merkmale nicht – wie beispielsweise beim Trainieren der NN – zur Diskriminanz genutzt. ONCADTM klassifiziert anhand deskriptiver Merkmale, die der Anwender vorgibt (z.B. rund, oval, lobu-

liert), während die berechneten Merkmale des Prototypen zu kontinuierlichen numerischen Werten führen. Wie bereits in Kapitel 6.3.2 diskutiert, ist der Ursprung und somit Informationsgehalt der Merkmalsvektoren grundlegend verschieden, ein Vergleich der beiden High-Level-Klassifikationsverfahren somit wenig aussagekräftig.

Der Prototyp setzt die High-Level-Klassifikation neben der Dignitätsvorhersage auch zum Entfernen zusätzlicher Funde ein. Kein Konkurrenzsystem enthält einen Arbeitsschritt, der zusätzliche Funde innerhalb der Brust erkennen kann.

6.6 Nutzen als Zweitmeinungssystem

CAD-Systeme werden heute eingesetzt, um dem Arzt zusätzliche Informationen für die Diagnostik bereitzustellen. Das Erstellen der abschließenden Diagnose bleibt dem Mediziner vorbehalten. Ziel eines CAD-System ist es, die Qualität der Diagnose durch die bereitgestellten Informationen zu erhöhen.

Wie in Kapitel 3.4 dargelegt, unterstützt der vorgestellte Prototyp den Radiologen nicht nur durch zusätzliche Informationen, sondern ist auch in der Lage, eigenständige Diagnosen herzuleiten. Im vorherigen Kapitel wurde verdeutlicht, dass der Lösungsansatz des Prototyps mit der umfassenden Funktionalität über ein deutlich größeres Potential verfügt als bestehende Software kommerzieller Anbieter. Der Prototyp stellt ein Computer-Diagnose-System dar, dessen eigenen autonomen Einschätzungen als Zweitmeinungssystem eingesetzt werden können. Folgend wird diskutiert, wie groß der Nutzen der Systemausgaben des Prototyps für den Radiologen ist.

Mit der Vorverarbeitung durch die Registrierung leistet die Software eine notwendige Bildverbesserung. Ohne sie wären die nachfolgenden Bildverarbeitungsschritte kaum erfolgreich (s. Kap. 6.2.4). Aber auch für den Radiologen sind die registrierten Bilddaten ein zusätzlicher Gewinn. Zwar hat das menschliche Auge weit weniger Mühe als ein Computer, Bewegungsartefakte als solche zu erkennen. Aber die enthaltenen Bewegungen können zu verfälschten Ergebnissen in der morphologischen und kinetischen Analyse des Radiologen führen. Im schlimmsten Fall wird eine Anreicherung nicht mehr erkennbar abgebildet. Auch verlängert jede notwendige Unterscheidung zwischen tatsächlicher und zusätzlicher Anreicherung die Befundungszeit. Die Registrierung leistet also einen Beitrag zur Qualitätsverbesserung der Befundung und Zeitersparnis.

Der Vorteil der MRM gegenüber der Mammographie oder Sonographie ist die hohe Sensitivität, mit der Läsionen ab 3 mm Durchmesser im Bild erkannt werden können. Diese Erkennung leistet in der Software der Segmentierungsschritt. Die Software hat auf dem untersuchten Kollektiv gezeigt, dass sie in der Detektion von Karzinomen dem Radiologen in nichts nachsteht. Allerdings produziert sie nach der Klassifikation im Durchschnitt mehr als 7 zusätzliche Funde pro Patient. Hier ist es schwierig, den Nutzen der Software abzuwägen. Einerseits braucht das menschlichen Auge wegen seiner guten Mustererkennung kaum Unterstützung in der Lokalisation von Läsionen. Aufgrund des Zeitdrucks und der natürlichen Schwankung in der Leistungsfähigkeit von Menschen ist jedoch vorstellbar, dass vor allem kleine Befunde im klinischen Alltag übersehen werden können. Da mit den dyna-

mischen Aufnahmen mehr als 300 Bildschichten betrachtet werden müssen, ist die Gefahr des Versäumens bei kleinen, schwach anreichernden Strukturen, die sich nur auf wenigen Bildschichten zeigen, am größten. Hier kann die Einfärbung der klassifizierten Läsionen im exportierten DICOM-Standard hilfreich sein, das Augenmerk des Benutzers noch einmal auf bestimmte Bildbereiche zu lenken. Aber gerade dabei ist enorm wichtig, dass die Software wenig zusätzliche Funde produziert. Denn die Menge an hervorgehobenen falsch positiven Anreicherungen kann den Radiologen von den wichtigen Bildbereichen ablenken. Die überwiegende Mehrheit an zusätzlichen Funden entsteht bei der Software im Thoraxbereich. Somit hält sich die Überflutung mit irrelevanten Bildbereichen in Grenzen, der Anwender kann beim Scrollen die dorsalen Bildschichten einfach auslassen.

Wie in Kapitel 6.3.2 verdeutlicht, unterscheidet sich der Weg der Diagnosefindung zwischen Software und Radiologen grundlegend. Die Diagnostik der Mamma-MRT ist wenig standardisiert (s. Kap. 2.3.6). Das Potential einer Software liegt darin, durch eine standardisierte Ausgabe die Befundung zu vereinheitlichen. In der aktuellen Fassung beschränkt sich die Ausgabe der Software lediglich auf grüne und rote Einfärbungen als Bildoverlay. Um den Nutzen des Prototyps zu erhöhen, ist die Ausgabe eines strukturierten Reports notwendig. Die Software sollte dem Benutzer die relevanten Informationen zur Begründung der eigenen Diagnose durch die Terminologien des BI-RADSTM-Standards bereitstellen und somit diesen Standard konsequent unterstützen. In Kapitel 6.3.2 wurde diskutiert, dass dies durch Einsatz weiterer Klassifikatoren realisiert werden kann. Selbst bereits zur Verfügung stehende Informationen wie die Zeit/Signalintensitätskurve der Kinetik sind derzeit kein Bestandteil der Ausgabe.

Eine Einteilung in BI-RADSTM-Kategorien (s. Tab. 2.8) kann auf Basis des Ausgabe werts des NN geschehen, der im Wertebereich 0,0 bis 1,0 liegt. Das Minimum steht für sicher gutartig und das Maximum für sicher bösartig, die übrigen Kategorien können durch Setzen sinnvoller Schwellwerte innerhalb des Bereichs abgeleitet werden.

Gerade eine BI-RADSTM gerechte Aufbereitung würde dem Anwender eine enorme Zeit sparen. Das Sammeln der Kriterien ist mit ca. 5 Minuten pro Läsion der zeitaufwendigste Arbeitsschritt für den Radiologen in der Diagnostik. Hier steht noch eine wenig forschungs lastige Fleißaufgabe für die Reportgenerierung an, damit sich der Einsatz der vorgestellten Software in der Praxis bewährt.

Im klinischen Alltag bleibt trotz Standardisierung durch BI-RADSTM die Ableitung der Diagnose dem individuellen Vorgehen des Radiologen überlassen. Nach dem zeitaufwändigen Bestimmen der Kriterien folgt die Diagnose, deren Qualität in der Mamma-MRT in besonderem Maße durch die Erfahrung des Radiologen bestimmt wird. Hier liegt der große Vorteil der Prototypsoftware, die konkurrenzlos eindeutige Diagnosen ausgibt. Der Radiologe erhält die Möglichkeit, seine eigene Einschätzung mit der Zweitmeinung der Software zu vergleichen.

Nicht abgeschätzt werden kann der Einfluss der Software auf die eigentliche Entscheidungsfindung des Radiologen. Generell bleibt beim Einsatz von CAD-Systemen zu befürchten, dass als Folge falscher Vorhersagen durch die Software der Radiologe verunsichert wird und somit die Anzahl an eingeleiteten Biopsien steigen könnte. Studien in der Mammographie (s. Kap. 2.2.2) zeigen, dass sich die Anzahl der Biopsien ab den Einsatz

von CAD-Systemen nicht erhöht haben [NHP⁺95, DMK⁺99].

Fraglich bleibt, wie groß die Akzeptanz der Prototypsoftware ist. Die Algorithmen zur Diagnosefindung arbeiten fern des menschlichen Verständnisses. Auch begründet die Software ihre Entscheidung derzeit noch nicht, denn hierzu wäre wie erwähnt die Aufbereitung der Morphologien und der Kinetik durch die BI-RADSTM-Terminologien notwendig. Die Komplettierung der Software durch die Reportausgabe würde eine Akzeptanz erheblich erhöhen. Der Nutzen des Prototyps läge dann in der hohen Zeitersparnis für den Radiologen. Die Rechenzeit von 30 bis maximal 60 Minuten ist ausreichend schnell, um die Software in den radiologischen Alltag zu integrieren. Laufzeitoptimierung sind für die aus der Forschungsumgebung entstammenden Implementierung wie üblich möglich.

Interessant sind die Reaktionen von Radiologen bei der Vorstellung der Software, wobei sich ein Generationenkonflikt abzeichnet. Während meist von den Älteren immer wieder die Frage gestellt wird, ob man denn wirklich glaube, dass eine Software den Radiologen ersetzen könne, interessieren sich jüngere Kollegen für den Preis und die mögliche Zeitersparnis durch den Einsatz der Software.

6.7 Ausblick

Für eine weiträumige Untersuchung der Tauglichkeit der Software im klinischen Alltag muss eine multizentrische Studie angestrebt werden. Die verschiedenen Evaluationen der Diskussion haben gezeigt, dass vor dem Einsatz die Funktionalität in einigen Bereichen überarbeitet werden muss.

Wie bereits erwähnt, sind nicht alle für die Klassifikation erhobenen morphologischen und kinetischen Merkmale robust gegenüber der Bildqualität (s. Kap. 6.3.3). Erst wenn die Merkmale unabhängig zur Bildqualität extrahiert werden können, sind Vorhersagen durch Verwendung derselben Klassifikatoren möglich. Andernfalls müssten für jeden Einsatzort eigene Klassifikatoren antrainiert werden, wofür das Erheben eines histologisch gesicherten GS pro Institut von Nöten wäre. Das Hinzufügen neuer Merkmale verspricht zudem (s. Kap. 6.4.7), die zusätzlichen Funde zu reduzieren und die Vorhersageleistung der Dignitäten zu steigern.

Bisher wurde die Software auf Bilddaten identischer MRT-Protokolle eingesetzt. Für ein robustes Verhalten müssen für die erzeugten Bilddaten mögliche Variablen existierender MRT-Geräte und Protokolle berücksichtigt werden: Anzahl der Sequenzen, Kombinationen der Orientierung, Einsatz von fettsupprimierten Sequenzen, Zeitpunkt der KM-Gabe, dargestellter Bildausschnitt und weitere. Speziell auf den Einfluss der Bildqualität wurde ausführlich in Kapitel 6.3.3 eingegangen mit dem Ergebnis, dass nicht alle Merkmale unabhängig gegenüber solchen Faktoren sind.

Zur Bewältigung der vielen Akquisitions- bzw. Protokollvarianten müssen als weitere Vorverarbeitung möglichst viele Informationen durch Auswertung der DICOM-Tags (s. Kap. A) automatisiert erkannt werden. Es existiert bereits ein automatisches Angleichen, wenn einzelne Serien in unterschiedlichen Orientierungen vorliegen (s. Kap. A.3).

Die Verwendung von T_1 - und T_2 -gewichteten Sequenzen in Kombination mit und oh-

ne Fettsuppression stellt unter Umständen eine Aufgabe dar, die besser durch eine multimodale Registrierung gelöst werden kann. Ob die bestehende Registrierung auch fettsupprimierte Aufnahmen korrigieren kann, wurde noch nicht untersucht.

Unvorhersehbar ist das Verhalten der Segmentierung, wenn der Ausschnitt des Bildvolumens mehr als nur die Mammae darstellt. Sind weitere anreichernde Strukturen wie das Herz oder die Aorta dargestellt, würden sie zwangsweise als potentielle Läsionen segmentiert. Hier wäre wiederum ein Vorverarbeitungsalgorithmus hilfreich, der die Mammae erkennt und irrelevante Bildbereiche, die nicht zur Diagnostik beitragen, ausblenden kann.

Man kann an dieser Auflistung, die sicherlich keinen Anspruch auf Vollständigkeit stellt, erkennen, dass für die Kommerzialisierung der Prototypsoftware noch weiterer Entwicklungsaufwand betrieben werden muss. Auch wenn für viele der notwendigen Verbesserungen in der Literatur bereits Lösungen existieren, bedarf eine robuste Implementierung dennoch eines hohen zeitlichen Entwicklungsaufwands. Hieraus kann auch erahnt werden, warum kommerzielle Anbieter diesen Schritt bisher noch nicht gegangen sind, sondern vermehrt auf manuelle oder halbautomatische Verfahren setzen.

Kapitel 7

Zusammenfassung

Mit der vorliegenden Arbeit konnte erstmalig gezeigt werden, dass mit Hilfe komplexer Methoden der Bildverarbeitung eine Software konstruiert werden kann, welche in Anlehnung an die Vorgehensweise eines Radiologen in der Lage ist, automatisch Läsionen innerhalb der MRT-Bilddaten der weiblichen Brust zu erkennen und selbständig eine Diagnose für jede gefundene Läsion abzuleiten. Besonders hervorzuheben ist hierbei die vollautomatische Funktionsweise der Software, welche den kompletten diagnostischen Prozess ohne jegliche Benutzerinteraktion abwickeln kann.

Möglich wurde diese Software im Wesentlichen durch die konsequente Verwendung des Konzeptes der Low, Intermediate und High-Level-Funktionalität. Bis dato verfügbare Produkte unterstützen den Befunder meist nur auf Low-Level-Ebene, beispielsweise durch Bereitstellung von Farbkodierungen einzelner Bildpunkte für den Radiologen.

Während des ersten Schrittes der Vorverarbeitung korrigiert die Software während der Bildaufnahme entstandene Bewegungsartefakte durch eine dreidimensionale flexible Bildregistrierung. Dieser Vorverarbeitungsschritt ist aus Sicht des Computersystems absolut notwendig, weil hierdurch besonders die Anzahl der Bewegungsartefakte reduziert wird, welche eine Kontrastmittelaufnahme des Gewebes vortäuschen. Die dreidimensionale Bildregistrierung nützt jedoch nicht nur der Software in den darauffolgenden Arbeitsschritten, sondern ist auch eine visuelle Aufbereitung der MRT-Bilddaten. Besonders kleinste Läsionen könnten durch Verschiebungseffekte während der Bildakquisition in den darauffolgenden Subtraktionen kaschiert werden und somit dem menschlichen Betrachter während der Befundung entgehen. Somit würde ein menschlicher Benutzer des Systems bereits durch die alleinige Bereitstellung der jetzt registrierten Bildinformationen wesentlich profitieren.

Nach der Vorverarbeitung detektiert die Software automatisch verdächtige Anreicherungen in den Bilddaten als einzelne Objekte. Hierdurch ist es möglich, dem Benutzer während der späteren Befundung die auffälligen Bildareale durch eine farbliche Hervorhebung einzelner Läsionen kenntlich zu machen. Auch hier ergibt sich der Vorteil zunächst aus der maschinellen Bildverarbeitung, welche stoisch jede Bildserie analysiert, während der Anwender aufgrund der Flut der Bildschichten und der im klinischen Alltag herrschenden Zeitnot eine Läsion auch einmal „übersehen“ könnte. Ein Nachteil der computerisierten Objekterkennung ist jedoch die hohe Anzahl verdächtiger Objekte nach der eigentlichen

Bildsegmentierung, die zum großen Teil zusätzliche und somit unbedeutende Funde darstellen. Um diese von den eigentlichen Läsionen unterscheiden und um die Dignität der Läsionen bewerten zu können, müssen noch weitere Bildverarbeitungsschritte folgen. So berechnet die Software anschließend für jedes detektierte Objekt morphologische und kinetische Informationen. Aufgrund der Vorverarbeitung sind diese nicht durch Bewegungseinflüsse im MRT-Bild kompromittiert. Ein besonderer Anspruch der Software ist es, die kinetischen Informationen robust gegen Einflussfaktoren wie beispielsweise unterschiedliche Kontrastmitteldosierungen zu halten. Die gewonnenen Informationen über jedes einzelne Objekt werden in der Folge von einem trainierten Klassifikator (neuronales Netz) zur Differenzierung von zusätzlichen Funden und Läsionen verwendet. Somit kann dem Benutzer eine um die zusätzlichen Funde deutlich reduzierte Liste an möglichen Läsionen im Bild präsentiert werden. Da für diesen Schritt der Reduktion der Objekte bereits eine Klassifikation seitens des Computersystems notwendig ist, zeigt sich bereits hier eine deutliche Abgrenzung zu den aktuell verfügbaren Systemen.

Darüber hinaus verwendet das Computersystem einen zweiten trainierten Klassifikator, um die verbliebenen Objekte hinsichtlich möglicher Malignität zu bewerten. Es wird somit eine Verdachtsdiagnose ausgegeben, welche dann in der Folge für weitere Visualisierungsschritte für den Benutzer verwendet werden kann. Ebenso ist es möglich, die abgegebene Beurteilung jeder Läsion als Zweitmeinungssystem für den Benutzer einzusetzen. Die ausgegebenen Läsionen und berechneten morphologischen und kinetischen Eigenschaften können darüber hinaus für ein strukturiertes Reporting verwendet werden.

Basierend auf dem histologischen Goldstandard zeigt die Software bei der Vorhersage der Objekte auf den Bilddaten von 101 Patientinnen eine Sensitivität von 88 % bei einer Spezifität von 72 % und erreicht somit vergleichbare Erkennungsraten, wie sie auch von Radiologen in der Literatur aufgezeigt werden [Kai90, HKBL⁺92]. Durchschnittlich verbleiben pro Patient knapp 2,5 vermeintlich bösartige zusätzliche Funde, welche durch den menschlichen Benutzer zumeist einfach als verbliebene und falsch klassifizierte Bildartefakte erkannt werden können.

Zusammenfassend ist es – nach Kenntnis des Verfassers – zum ersten Mal gelungen, ein vollautomatisches Diagnosesystem für MRT-Bilddaten der weiblichen Brust zu erstellen. Eine erste Validierung auf 101 Bilddaten zeigt eine akzeptable Laufzeit des Systems. Vor der möglichen Entwicklung eines Produktes steht sicherlich noch der weitgreifende Aufwand der Robustmachung des Systems gegen unterschiedliche Einflussfaktoren, beispielsweise die Qualität der Bilddaten und unterschiedliche Bildakquisitionsverfahren.

Anhang A

Koordinatensystem nach DICOM

Der *DICOM* Standard (Digital Imaging and Communications in Medicine) wurde in Zusammenarbeit verschiedener Normierungsgremien in den Vereinigten Staaten, Europa, Japan, dem American College of Radiology (ACR) und der National Electrical Manufacturers Association (*NEMA*) entwickelt [Ame]. Das Ziel war eine Kommunikationsschnittstelle zwischen bildgebenden Geräten und beliebig anderen Modalitäten in der Medizin zu erarbeiten, um der wachsenden Bedeutung der digitalen Bildverarbeitung in der Medizin Rechnung zu tragen.

DICOM-Dateien enthalten sehr unterschiedliche Informationen. Es existieren u.a. Gruppen für Patientendaten, Aufbau der Studie und Serie oder der Bilddaten. Die kleinste Informationseinheit ist ein Attribut, das durch Identifikationszahlen für die schon erwähnte Gruppe und Elemente eindeutig definiert ist (Syntax: *0xGruppe*, *0xElement*, s. Tab. A.1). Alle spezifizierten Attribute des Standards sind in einem Datenwörterbuch erfasst. Jede DICOM-Datei besteht aus einer Vielzahl solcher Attribute.

A.1 Patientenkoordinatensystem

Durch die Lage des Patienten im MRT wird ein rechtshändiges kartesisches 3D-Koordinatensystem (*KOS*) definiert, dessen Ursprung fest durch den Mittelpunkt der Magnetronne vorgegeben ist. Die Lage des Patienten muss vor dem Start des Protokolls mitgeteilt werden, damit das MRT-Gerät das *Patienten-KOS* entsprechend ausrichten kann. Die positive x-Achse führt nach der Ausrichtung immer von der rechten Seite (*R*) zur linken Seite (*L*) des Patienten, während die positive y-Achse von Bauch- (*A*) zur Rückenseite (*P*) und die positive z-Achse von den Füßen (*I*) zum Kopf (*S*) des Untersuchten verläuft (s. Abb. A.1).

Die Unterschiede der *koronaren*, *sagittalen* und *transversalen* Schichtführungen sind in Abb. A.1 ebenfalls visualisiert. Sie beziehen sich immer auf das Patienten-KOS. Koronare Schichten verlaufen parallel zur *xz*-Ebene und unterteilen den Körper in eine Vorder- und Rückseite. Transversale Ebenen verlaufen parallel zur *xy*-Ebene, während die sagittale Ebene den Körper durch die *yz*-Ebene in eine rechte und linke Hälfte teilt.

Das *Patienten-KOS* lässt sich aus den Informationen mehrerer Attribute der einzelnen

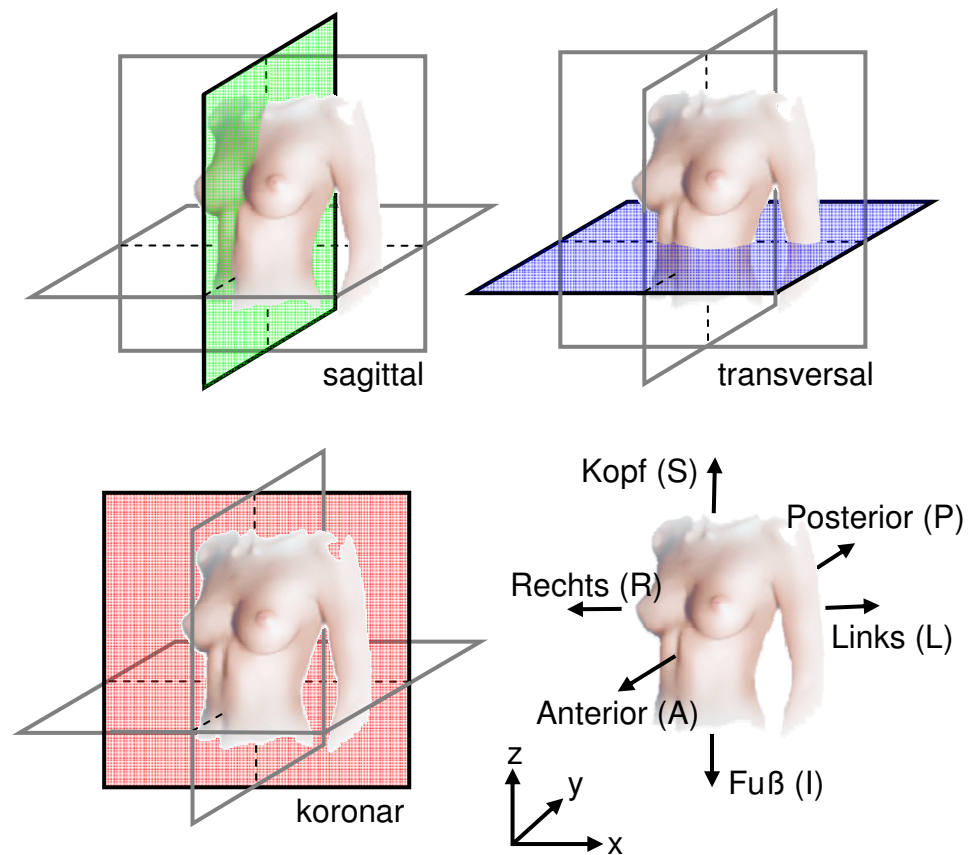


Abbildung A.1: Patienten Koordinatensystem nach DICOM mit Visualisierung der Orientierungen

Name	Attribut	Information
PixelData	(0x7FE0, 0x0010)	2D-Array mit Signalwerten
ImagePositionPatient	(0x0020, 0x0032)	3D-Vektor $IPP_{x,y,z}$ vom Ursprung des MRT zum Ankerpunkt der Bildebene
ImageOrientationPatient	(0x0020, 0x0037)	zwei rechtwinklig zueinanderstehende 3D-Vektoren $IOP1_{x,y,z}$ und $IOP2_{x,y,z}$, die die Bildebene aufspannen
PixelSpacing	(0x0028, 0x0030)	Ausdehnung $S_{x,y}$ der Pixel der Bildschicht
SliceThickness	(0x0018, 0x0050)	Dicke der Bildschicht (S_z)

Tabelle A.1: Liste verschiedener DICOM-Attribute, die das Patienten-KOS definieren.

DICOM Datei rekonstruieren (s. Tab. A.1). Maßgeblich sind der *IPP* und die beiden *IOP* Vektoren, die die Lage und Orientierung der einzelnen Bildschichten im Raum definieren (s. Abb. A.2, obere Bildhälfte). Die Einheit des Systems besteht in Millimeter, die räumliche Ausdehnung eines Bildpunktes (*Voxelspacing*) ergibt sich aus dem *PixelSpacing* und der *SliceThickness* ($S_{x,y,z}$).

Das Patienten KOS in Abbildung A.2 verdeutlicht, wie derselbe Raum der Brüste durch Sequenzen verschiedener Schichtorientierung (T_1 koronar, T_2fs koronar, T_1 transversal) mit abweichenden IOP und IPP Vektoren im DICOM-Standard definiert wird, wobei sich dennoch gleiche Koordinaten auf die gleichen Positionen des Patienten beziehen.

A.2 Modellkoordinatensystem

Die Bildvolumen werden aus den Signalmessungen der einzelnen DICOM-Dateien zusammengesetzt. Dazu erfolgt eine Sortierung der Dateien primär nach Seriennummern (s. Abb. 4.3) und sekundär nach Schichtpositionen (*IPP*, s. Abb. A.2 oben). Die Bilddaten werden abschließend aus der DICOM-Datei Schicht für Schicht in die systemeigene 3D-Bilddatenstruktur kopiert.

Die Datenstruktur enthält jedoch keine Informationen über die räumliche Orientierung der Bildvolumen, die Anordnung der Voxel wird nur durch ein *Modell-KOS* definiert. Es ist durch die Schichtführung geprägt, der Ursprung des Systems liegt in der ersten Bildschicht des Volumens, die *z*-Achse verläuft entlang der Bildschichten (s. Abb. A.2 unten). Entgegen dem Patienten-KOS ist die Einheit der Koordinaten *Voxel*. Die Größe der Volumen entspricht der Matrix des gesetzten FoV (s. Tab. 4.1).

Anhand der Modellkoordinaten sind Bildvolumen unterschiedlicher Sequenzen nicht miteinander vergleichbar, denn Voxel mit gleichen Modellkoordinaten beziehen sich auf unterschiedliche Positionen des Patienten. Der Bezug zum Patienten KOS kann erst durch Anwendung von affinen Transformationen hergestellt werden (s. nächstes Kapitel).

A.3 Koordinatentransformation

Neben den Bilddaten im Modell KOS kann auch deren räumlicher Bezug zum Patienten KOS aus den DICOM-Dateien extrahiert werden. Die Bildverarbeitung benutzt zur Speicherung solcher Daten 4x4 Translationsmatrizen (s. Formel A.1). Durch sie können affine Transformationen wie Translation, Rotation oder Skalierung auf die Bildkoordinaten angewendet werden [NH04]. Die Matrizen können direkt aus den Informationen der bekannten DICOM-Attribute (s. Tab. A.1) erzeugt werden. Die *Translation* (M_{Trans}) wird durch *ImagePositionPatient* Vektor zur ersten Bildschicht, die *Rotation* (M_{Rot}) durch die beiden Vektoren der *ImageOrientationPatient* und deren Kreuzprodukt und die *Skalierung* (M_{Skal}) durch das *Voxelspacing* definiert:

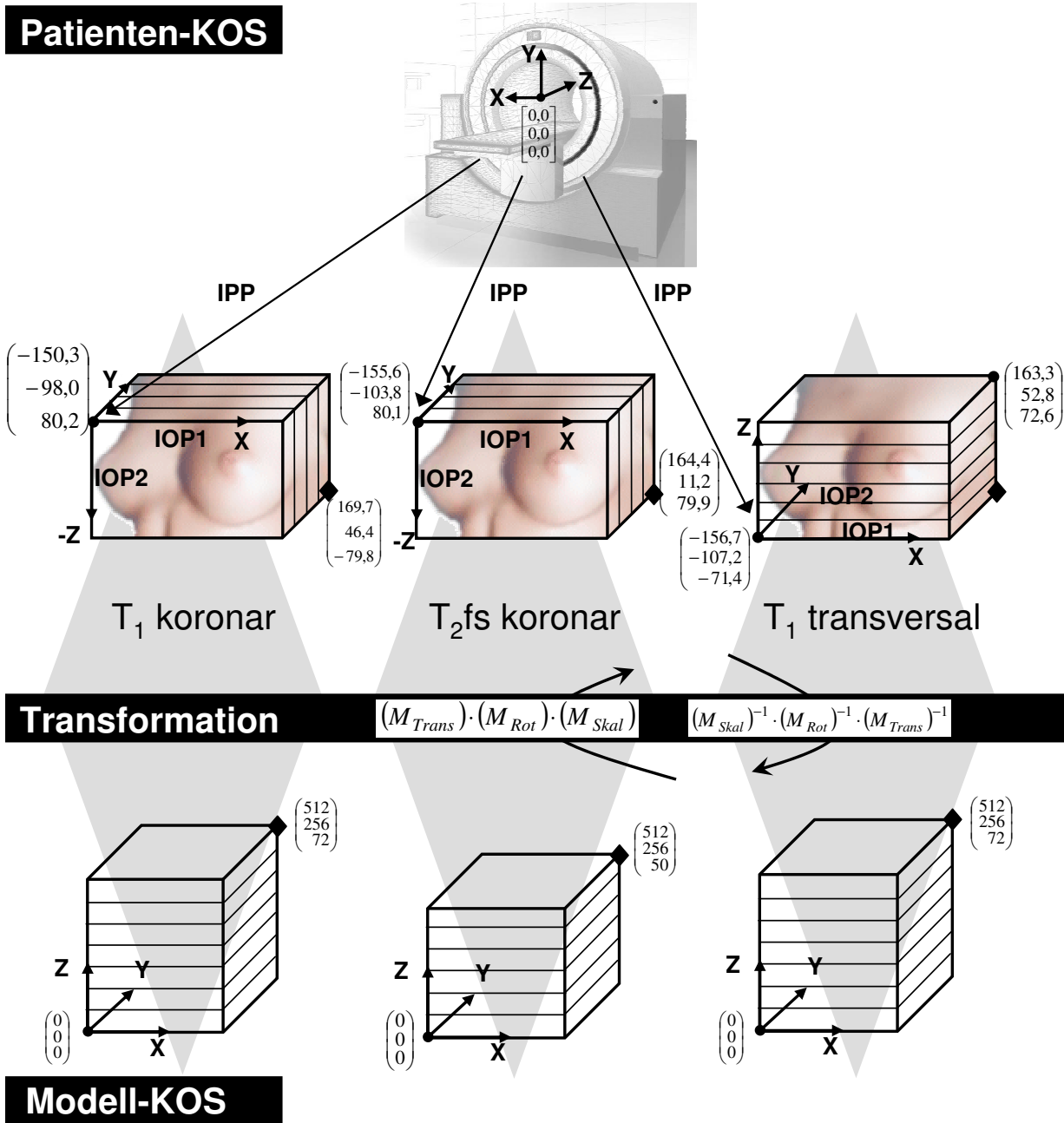


Abbildung A.2: Darstellung des Patienten-KOS mit dem Ursprung im MRT-Gerät (oben) im Vergleich zum Modell-KOS ohne räumlichen Bezug (unten) und der KOS-Transformation (Mitte) anhand dreier Sequenzen verschiedener Schichtführung

$$\begin{aligned}
M_{T_{Trans}} &= \begin{pmatrix} 1 & 0 & 0 & IPP_x \\ 0 & 1 & 0 & IPP_y \\ 0 & 0 & 1 & IPP_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
M_{Rot} &= \begin{pmatrix} IOP1_x & IOP2_x & IOP1_x \times IOP2_x & 0 \\ IOP1_y & IOP2_y & IOP1_y \times IOP2_y & 0 \\ IOP1_z & IOP2_z & IOP1_z \times IOP2_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
M_{Skal} &= \begin{pmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
\end{aligned} \tag{A.1}$$

Jede Koordinate des Modell-KOS eines Bildvolumens lässt sich durch Hintereinanderausführung der Skalierung, Rotation und Translation in das Patienten-KOS überführen (s. Abb. A.2, Mitte). Dort sind sie mit den transformierten Modellkoordinaten der Bildvolumen anderer Sequenzen vergleichbar. Durch die Inverse der Matrizen wird umgekehrt vom Patienten- ins Modell-KOS transformiert, wobei auch die Anwendungsreihenfolge der Matrizen umgekehrt werden muss. Durch die Kombination beider Transformationsrichtungen (s. Formel A.2) kann beispielsweise die Modellkoordinate eines koronaren Bildvolumens K durch die Matrix M_K in das Patienten-KOS und von dort weiter durch die inverse Matrix M_T^{-1} in die Modellkoordinate des transversalen Volumens T überführt werden:

$$\begin{aligned}
\begin{pmatrix} K_x \\ K_y \\ K_z \\ 1 \end{pmatrix} &= \underbrace{(M_{K,Skal})^{-1} \cdot (M_{K,Rot})^{-1} \cdot (M_{K,Trans})^{-1}}_{M_K^{-1}} \cdot \underbrace{(M_{T,Trans}) \cdot (M_{T,Rot}) \cdot (M_{T,Skal})}_{M_T} \cdot \begin{pmatrix} T_x \\ T_y \\ T_z \\ 1 \end{pmatrix} \\
\begin{pmatrix} K_x \\ K_y \\ K_z \\ 1 \end{pmatrix} &= (M_{T \rightarrow K}) \cdot \begin{pmatrix} T_x \\ T_y \\ T_z \\ 1 \end{pmatrix}, \text{ mit } M_{T \rightarrow K} = M_K^{-1} \cdot M_T
\end{aligned} \tag{A.2}$$

Anhang B

KLT Feature Tracking

Das Feature Tracking dient zur Erkennung und Verfolgung von Bewegungen in aufeinanderfolgenden Bildern. Mit Feature werden charakteristische Positionen im Bildvolumen bezeichnet, während *Tracking* ihr Wiederauffinden bzw. Verfolgen im zeitlich späteren Bild bedeutet.

Der Registrierungsalgorithmus in dieser Arbeit basiert auf dem Kanade-Lucas-Tomasi Feature Tracker (*KLT*) [Bur84, LK81]. Während die Veröffentlichung des KLT nur 2D-Bilddaten zu tracken vermag, ermöglicht der Ansatz dieser Arbeit ein 3D-Vorgehen.

Das Tracking findet in Bildserien Anwendung, die zeitlich nahe zusammen liegen und somit über ähnliche Bildinhalte verfügen:

$$SN_J(p) \approx SN_I(p) \tag{B.1}$$

$p_{x,y,z}$ stellt eine beliebige lokale Position der Bilder SN_I und SN_J dar, wobei SN_J zeitlich nach SN_I entstanden ist. Das Feature an der Position $p_{x,y,z}$ könnte durch die Signalintensität des entsprechenden Bildpunkts beschrieben werden; jedoch wäre dies nicht eindeutig in Bezug auf das Tracking. In der funktionalen MRT der Brust gibt es keine signifikanten Signalintensitäten, die sich eindeutig von allen anderen Werten unterscheiden würden. Zudem unterliegen die Signalwerte dem Bildrauschen. Der KLT bestimmt das Feature daher auf einer Menge von Bildpunkten innerhalb eines lokalen Fensters W , das auf die Position $p_{x,y,z}$ zentriert wird. Es wird angenommen, dass der Inhalt des Fensters in Bild SN_I – trotz einer möglichen Verschiebung um eine Distanz $d_{x,y,z}$ – dem im darauffolgenden Bild SN_J ähnelt. Als lokales Bildmodell ergibt sich:

$$SN_J(p) \approx SN_I(p - d) \tag{B.2}$$

Das Ergebnis des Trackings eines Features ist die unbekanntes Translation des 3D-Vektors $d_{x,y,z}$, der auch als der *optische Fluss* bezeichnet wird. Unterschiede der aufeinanderfolgenden Fenster, die sich nicht alleine durch die Translation erklären lassen (z.B. Bildrauschen), werden als der Restfehler ε aufgefasst. Bei perfekter Übereinstimmung wäre der Fehler Null. Das Registrierungsproblem besteht darin, den Translationsvektor d so zu wählen, dass der Restfehler möglichst gering wird:

$$\varepsilon = \int_W [SN_I(p-d) - SN_J(p)]^2 dp \quad (\text{B.3})$$

Zur Minimierung des Restfehlers in (B.3) wird ähnlich der Berechnung der Varianz die kleinste Quadrate Methode verwendet. Die Lösung der Gleichung ist jedoch zu aufwendig, effizienter kann fortgefahren werden, wenn $SN_I(p-d)$ linearisiert wird. Diese Vereinfachung setzt voraus, dass die gesuchte Translation $d_{x,y,z}$ kleiner als der Bereich des Fenster W ist. Zur Annäherung wurde die Taylor Reihe verwendet und anschließend auf den linearen Anteil verkürzt:

$$SN_I(p-d) \Rightarrow SN_I(p) - g \cdot d \quad (\text{B.4})$$

Bildlich dargestellt bezeichnet g einen 3D-Gradientenvektor, der – die Grauwerte des Bildvolumens als Funktionen aufgefasst – die Richtung und Stärke des steilsten Anstieges im Bild beschreibt. Die Gleichung (B.4) wird für die Differenzierung in (B.3) substituiert:

$$\varepsilon = \int_W [SN_I(p) - g \cdot d - SN_J(p)]^2 dp \quad (\text{B.5})$$

In (B.5) ist der Translationsvektor d nun unabhängig zum Bild SN_I . Die Gleichung kann über d differenziert werden, zur Bestimmung des lokalen Minimums wird die Funktion auf Null gesetzt. Durch anschließende geschickte Anwendung von Vektoranalysis kann auf die handhabbare Gleichung $G \cdot d = v$ verkürzt werden:

$$\begin{aligned} 0 &= \int_W (SN_I - SN_J - g \cdot d)g \, dA \\ \int_W (g \cdot d)g \, dA &= \int_W (SN_I - SN_J)g \, dA, \text{ mit } (g \cdot d)g = (gg^T)d \\ \underbrace{\left(\int_W gg^T \, dA \right)}_G d &= \underbrace{\int_W (SN_I - SN_J)g \, dA}_v \end{aligned} \quad (\text{B.6})$$

Das Ergebnis ist ein lineares System mit den drei Unbekannten x, y, z des gesuchten Translationsvektors $d_{x,y,z}$. G repräsentiert eine symmetrische 3x3 Matrix, die die Gradienteninformation innerhalb des Fensters W beinhaltet und v ist ein 3D-Vektor, der die Differenzen der Signalwerte der Bilder SN_I und SN_J in Bezug auf den Gradienten g beschreibt. Aufgrund der Linearisierung in (B.4) ist die Gleichung (B.6) nur annähernd bestimmt. Der genaue Translationsvektor $d_{x,y,z}$ kann dennoch durch Lösung von (B.6) durch das iterative Newton-Raphson-Verfahren [BS] berechnet werden. Durch dieses Standardverfahren der Mathematik lassen sich Näherungen von Nullstellen finden.

Da die Grauwerte der Bilddaten diskrete Funktionen darstellen, finden für alle aufgeführten Formeln statt der Integrale endliche Summen in der Umsetzung Anwendung.

Der KLT Algorithmus ist nur für Features an denjenigen Position im Bildvolumen erfolgreich, deren Fenster einen charakteristischen Bildausschnitt aufweisen. Features innerhalb homogener Bereiche mit ähnlichen Signalintensitäten oder flächigen Strukturen

können kaum Bewegungsinformationen wiedergeben. Vor dem Start des Trackings muss eine Beschränkung auf nur die erfolgversprechenden Features erfolgen, also Fensterbereiche mit einer charakteristischen Texturierung. Dazu kann dieselbe Gradientenmatrix G des Trackings verwendet werden. Ihre Eigenwerte sind hoch, wenn sich die Texturierung innerhalb des Fensters W deutlich vom Bildrauschen absetzt. Als Qualitätsmaß für ein Feature wird daher der minimale der drei Eigenwerte der Gradientenmatrix verwendet. Je höher sein Wert, desto erfolgversprechender ist das Tracking an dieser Position.

Anhang C

Vektorinterpolation

Das Featuretracking (s. Anhang B) führt zu Translationsaussagen einiger weniger Bildpunkte. Um die Verschiebungen für jeden Bildpunkt erheben zu können, bedarf es einer Interpolation der getrackten Translationsvektoren. Ein schneller Interpolationsalgorithmus basierend auf B-Splines-Approximation erfüllt den Anspruch, zufällig verteilte Daten interpolieren zu können [LWS97]. Ursprünglich für das Morphing von 2D-Gesichtsfotos konzipiert, wurde das Verfahren noch zu einem 3D-Ansatz erweitert.

Kern des Algorithmus ist ein 3D-Gitter Φ , das den gleichen Raum wie das Bildvolumen einnimmt. Es besteht aus Kontrollpunkten, die einen gleichmäßigen Abstand zueinander haben. Sie verhalten sich wie Gewichte, an die die gegebenen Vektoren geheftet werden können. Beim Gewichten wird der gegebene Translationsvektor $d_{x,y,z}$ nicht einem, sondern gleich 4x4x4 benachbarten Kontrollpunkten zugewiesen:

$$f_{Spline}(p_{x,y,z}) = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 B_l(u) B_m(v) B_n(w) \Phi_{i+l,j+m,k+n} \quad (C.1)$$

Abhängig vom Ursprung $p_{x,y,z}$ des Vektors $d_{x,y,z}$ wird der nächstgelegene Kontrollpunkt $\Phi_{i,j,k}$ gesucht, wobei $u, v, w \in [0, 0 \dots 1, 0)$ die relative Distanz zu diesem Punkt darstellt. Jeder Translationsvektor $d_{x,y,z}$, der dem Grid Φ zugewiesen wird, muss die approximierenden B-spline Funktion $f_{Spline}(p_{x,y,z}) = d_{x,y,z}$ erfüllen. Werden demselben Kontrollpunkt zu viele Translationsvektoren zugewiesen, gibt es keine eindeutige Lösung für die Verteilung auf die Kontrollpunkte $\Phi_{i+l,j+m,k+n}$. Eine zu geringe Anzahl führt dagegen zu einer unendlichen Menge an Ergebnissen. Daher werden die Kontrollpunkte mit der kleinste Quadrate Methode durch Minimierung von $\sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 \Phi_{l,m,n}^2$ berechnet. B_l , B_m und B_n sind Polynome dritter Ordnung für approximative Spline Interpolation:

$$\begin{aligned}
B_0(t) &= \frac{(1-t)^3}{6} \\
B_1(t) &= \frac{3t^3 - 6t^2 + 4}{6} \\
B_2(t) &= \frac{-3t^3 + 3t^2 + 3t + 1}{6} \\
B_3(t) &= \frac{t^3}{6}
\end{aligned} \tag{C.2}$$

Die Genauigkeit der Gewichtsfunktion f_{spline} hängt von der Dichte des Gitters Φ ab. Während eine grobe Struktur – mit einer geringen Anzahl an Kontrollpunkten aber großen Abständen – in einer weichen Verteilung resultiert, werden die vorgegebenen Translationsvektoren im allgemeinen nur ungenau approximiert. Je feiner das verwendete Gitter, desto mehr erhöht sich die Genauigkeit.

Das Interpolationsverfahren arbeitet iterativ. Es verwendet zu Beginn grobe Gitter und verfeinert sie mit jedem Schritt, bis für jeden Bildpunkt ein korrespondierender Kontrollpunkt existiert. In jeder Iteration werden die Kontrollpunkte mit den gegebenen Translationsvektoren gewichtet. Anschließend werden die Vektoren um den approximierten Anteil im Gitter verringert. Somit verändern sich die gegebenen Translationsvektoren und nähern sich mit jeder Iteration immer weiter einem Nullvektor.

Durch die Verwendung eines groben Gitters in den anfänglichen Iterationen nehmen Bildbereiche, für die keine Bewegungsinformationen getrackt werden konnten, die globale Bewegungsrichtung der nächstgelegenen Vektoren an. Bildregionen uneinheitlicher Bewegungsvektoren können dagegen erst durch die feinen Gitterstrukturen der letzten Iterationen genau approximiert werden.

Als Resultat stellt das feinste Gitter die Grundstruktur zum Interpolieren dar. Das Gitter entspricht einem *Vektorfeld*, das eine Translation für beliebige Koordinaten des Bildvolumens durch Formel C.1 erzeugt kann.

Stichwortverzeichnis

- A-priori-Wissen, [43](#), [44](#)
- Abstandmaß, [82](#)
- Accuracy, [87](#), *siehe* Treffgenauigkeit
- ACR, [36](#), [183](#)
- Adenokarzinom, [21](#)
- affine Transformation, [42](#), [50](#)
- American College of Radiology, [36](#), *siehe* ACR
- Angiographie, [116](#)
- anisotrop, [59](#)
- Anreicherung, [29](#)
- Artefakte
 - Bewegungs-, [31](#)
 - Rückfaltungs-, [61](#)
 - Registrierungs-, [111](#)
- AUC, [74](#)

- bösartig, *siehe* maligne
- Basalmembran, [19](#), [22](#), [23](#)
- Baseline, [74](#)
- Begrenzung, [34](#), [67](#), [70](#), [71](#), [139](#)
- benigne, [20](#), [21](#)
- Bewegungsartefakte, [31](#)
- BI-RADSTM, [36](#), [138](#)
- Bindegewebe, [19](#)
- Biopsie, [23](#)
- Blooming, [52](#)
- Boost.Ada, [89](#)
- Boost.Ada IS, [89](#), [167](#)
- Boost.Gbm, [89](#)
- Boosting, [81](#), [89](#)
- Bottom-Up, [43](#)
- Boxcounting, [70](#)
- BRCA 1, [24](#), [30](#)
- BRCA 2, [24](#), [30](#)
- Breast *fTP*, [47](#), [174](#)–[176](#)
- Brust, *siehe* Mamma
- Brustdrüse, [19](#)
- Brustwarze, [19](#)

- CAD, [45](#), [46](#)
- CADsciences, [47](#)
- CADstreamTM, [47](#), [174](#), [176](#)
- Clusteranalyse, [84](#), [149](#)
- Computer Aided Detection, [45](#)
- Computer Aided Diagnosis, [45](#)
- Computer Automatic Diagnosis, [54](#)
- Computer Diagnosis, [54](#)
- Computertomographie, [116](#), *siehe* CT, [117](#)
- ConfirmaTM, [47](#)
- CT, [116](#), [117](#)

- Data Mining, [84](#)
- Dendrogramm, [84](#)
- DICOM, [183](#)
- Differenzbilder, [31](#)
- digitale Bildverarbeitung, [39](#), [41](#)
- Dignität, [20](#)
- Dilatation, [43](#)
- Diskriminanzanalyse, [88](#), [167](#)
- Ductus lactiferi colligentes, [19](#)
- duktales Karzinom, [21](#), [34](#)
- duktales Karzinom in-situ, [22](#)
- DynaCADTM, [47](#), [174](#)–[176](#)
- dynamisches Merkmal, [67](#), *siehe* Kinetik

- Echogenität, [27](#)
- Ektomie, [25](#)
- elastische Transformation, [50](#)
- Enhancement, [31](#), *siehe* Anreicherung
- Entropie, [83](#)
- Epithel, [21](#)
- Erosion, [43](#)

- falsch-negativ, 87
falsch-positiv, 31, 87
Farbkodierung, 46
FDA, 88, *siehe* flexible Diskriminanzanalyse
Feature, 62
Fenster, 42
Fettgewebe, 19
Fettsättigung, 29
Fettsuppression, 29
Fettunterdrückung, 29
Fibroadenom, 21, 32, 34, 35
Fibrose, 34
Filter, 42
flexible Diskriminanzanalyse, 88, 167
fMRT, 29
Fokus, 36
Form, 34, 67, 69, 70, 139
Fourier-Transformation, 42
FoV, 29
fraktale Dimension, 48, 70
Frequenzraum, 42
fTP, 47, 174–176
- Göttinger Score, 35, 37
Glandula mammaria, 19
globale Transformationen, 42
Goldstandard, *siehe* GS
GS, 75
GUI, 47
gutartig, *siehe* benigne
- Hauptkomponentenanalyse, 84, 151
Herdbefund, 36
High-Level, 41, 43, 46, 48
High-Level-Diagnose, 46
high-res, 57
Histogramm, 42
Homogenitätskriterium, 43, 65
Hotspot, 33, 72
Hounsfield, 117
hyperintens, 29
hypointens, 29
- in-situ, 22, 25, 30
- Instanz, 78
Intermediate-Level, 41, 42
Interpolation, 42
 polynom, 73, *siehe* Polynominterpolation
 trilinear, 62, *siehe* trilineare Interpolation
 vektor, 63, *siehe* Vektorinterpolation
IOP, 185
IPP, 185
- k-nearest Neighbor, 89, 167
Kanade-Lucas-Tomasi, 63, 189
Karzinom, *siehe* Mammakarzinom
Kinetik, 29, 32, 36, 72, 140, 145
Klassifikation, 43, 45, 52, 75, 102, 156, 175
Klassifikator, 88
KLT, 189, *siehe* Kanade-Lucas-Tomasi
KM, 28
knn, 89, *siehe* k-nearest Neighbor
Kontingenztafel, 87
Kontrastmittel, 28, *siehe* KM
Koordinatensystem, 41, *siehe* KOS
koronar, 183
KOS, 41, 42, 183
- Läsionen, 21
LDA, 88, *siehe* lineare Diskriminanzanalyse
Leave-One-Out, 86
Lernen, 79
lineare Diskriminanzanalyse, 167
lineare Regression, 85, 88
LinSVM, 89, 167
Lipom, 21, 26
Lobi glandulae mammariae, 19
lobuläre Karzinom in-situ, 22
Lobuläres Karzinom, 21, 25
logistische Regression, 85, 88
LogReg, 88
Low-Level, 41, 42, 46, 47
Low-Level-Diagnose, 46
Low-Level-Klassifikation, 44, 46
Low-Level-Visualisierung, 46

- low-res, 57
Lymphknoten, 24, 30
- Magnetresonanzmammographie, 28
Magnetresonanztomographie, 28, *siehe* MRT
maligne, 20, 21
Mamma, 19
Mammakarzinom, 21
 duktal, 21
 in-situ, 22
 invasiv, 21
 lobulär, 21, *siehe* Lobuläres Karzinom
 medullär, *siehe* Medulläres Karzinom, 22
 muzinöse, *siehe* Muzinöses Karzinom, 22
 nicht invasiv, 21
Mammographie, 26
Maske, 42, *siehe* Fenster
Mastitis, 21
Mastopathie, 21, 24, 25, 35
Medulläres Karzinom, 22
Menstruation, 19
Merkmalsextraktion, 43, 45, 51, 67, 101, 137, 175
Merkmalsselektion, 85, 160
Merkmalsvektor, 43
Metastasen, 20, 24
 Lymphknoten, 25
Mevis, 47
MI, 82, 113
Mikroverkalkung, 26, 30
Milchdrüse, 19
Milchgang, 19, 21
Milchsäckchen, 19
MIP, 47
Modalitäten, 49
Modell-KOS, 185
mono modal, 50
Morphologie, 19, 28, 34, 36, 52, 144
morphologische Transformation, 43
morphologischer Operator, 43, 68, 71
Mortalität, 25
MPR, 47
MRM, 28
MRT, 18, 28
multi modal, 50
Multivariate Analyse, 84
Mutual Information, 82, *siehe* MI
Muzinöses Karzinom, 22
- Nekrose, 34, 71
NEMA, *siehe* National Electrical Manufacturers Association
Neoangiogenese, 22, 32, 34
neuronales Netz, 79, *siehe* NN
NN, 44, 79, 167
- ONCADTM, 48
ONCADTM, 174–176
optischer Fluss, 189
Ortsraum, 41
Overfitting, 80
- p-Wertes, 85
Palpation, 25
Papillom, 21
partial Volumeneffekt, 123
Patienten-KOS, 183
Peano-Kurve, 70
Perfusion, 73
peritumorales Ödem, 35
Permeabilität, 23, 32, 51
Phantom, 81
pharmakokinetisches Modell, 47, 51
Phylloidentumor, 91
Pixel, 41
Polynominterpolation, 73
Präkanzerosen, 21
Punktoperatoren, 42
- QDA, 88, *siehe* quadratische Diskriminanzanalyse
quadratische Diskriminanzanalyse, 88, 167
- Rückfaltungsartefakte, 61
RadialSVM, 89, 167
Random Forest, 89, 167
Random Forest IS, 89, 167

- Region Growing, 43, 65
 Region of Interest, 72, *siehe* ROI
 Registrierung, 45, 49, 60, 91, 105, 133, 143, 174
 Registrierungssartefakte, 111
 Regressionsanalyse, 85
 Relaxation, 28
 Resubstitutionsfehler, 86
 richtig-negativ, 87
 richtig-positiv, 87
 rigide Transformation, 50
 Ringenhancement, 34
 ROI, 72
 Rotation, 42, 185

 Saatzelle, 65
 sagittal, 183
 Samplegröße, 80
 Schallschatten, 27
 Schwellwertmethoden, 51
 Screening, 26, 27
 Segmentierung, 42, 45, 64
 Segmentierungsheuristik, 65, 123
 Segmentierungsregeln, 65
 Selbsttransformation, 42
 semantisches Netz, 44
 semimaligne, 20, 21
 Sensitivität, 87
 Sequenz, 29
 SI, 29
 Signalintensität, 29, *siehe* SI
 Skalierung, 42, 185
 SN, 57
 Sonographie, 27
 Spezifität, 87
 Spikulierung, 34
 Spin, 28
 Spule, 28
 SSD, 82
 Staging, 30
 Standardisierung, 80
 statisches Merkmal, 67, *siehe* Morphologie
 Stenosen, 116

 Support Vector Maschinen, *siehe* SVM
 SVM, 44, 53, 89, 167

 t-Test, 85
 T_1 , 28
 T_1 -gewichtet, 29
 T_2 , 28
 T_2fs , 57, *siehe* Fettsättigung
 T_2 -gewichtet, 29
 Testdaten, 87
 Textur, 34, 70, 139
 Token, 41
 Top-Down, 43, 44
 Topologie, 68, 79
 Tracking, 62, 107, 189
 Transformation
 affine, 42, *siehe* affine Transformation
 elastische, 50, *siehe* elastische Transformation
 Fourier-, 42
 Koordinaten-, 185, *siehe* affine Transformation
 Matrix, 185, *siehe* Transformationsmatrix
 morphologische, 43, *siehe* morphologische Transformation
 nicht rigide, 50, *siehe* elastische Transformation
 rigide, 50, *siehe* rigide Transformation
 Selbst-, 42, *siehe* Selbsttransformation
 Transformationsmatrix, 185
 Translation, 42, 185
 transversal, 183
 Tree, 89, 167
 Treffgenauigkeit, 87
 trilineare Interpolation, 62
 TTP, 73
 Tumor, 20
 benigne, 20
 maligne, 20
 semimaligne, 20

 Überdeckungskriterien, 78

- Überlagerungsregel, [66](#), *siehe* Segmentierungs-
heuristik
- Ultraschall, [27](#)
- Underfitting, [80](#)
- Unittest, [82](#)
- vascular epithelial growth factor, [22](#)
- VEGF, [22](#)
- Vektorfeld, [63](#), [194](#)
- Vektorinterpolation, [63](#), [109](#), [193](#)
- Vorverarbeitung, [42](#), [43](#)
- Voxel, [41](#)
- Voxelspacing, [68](#), [185](#)
- Wash-Out, [32](#), [36](#), [74](#)
- Wasserstoffatom, [28](#)
- Weichzeichnung, [42](#)
- Zeit/Signalintensitätskurve, [29](#), [33](#)
- zusätzlicher Fund, [79](#), [169](#)
- Zweitmeinung, [18](#), [177](#)
- Zyste, [21](#), [26](#), [35](#)

Abbildungsverzeichnis

2.1	Anatomie einer gesunden weiblichen Brust in der Sagittalansicht (von Patrick J. Lynch)	20
2.2	Unterscheidung invasiver und in-situ Varianten duktaler und lobulärer Karzinome	22
2.3	Häufigkeit des Karzinomauftritts in Abhängigkeit der Verteilung im Drüsengewebe.	24
2.4	(a) Typische Signalverläufe gesunder Gewebe gegenüber gut- und bösartigen Veränderungen. (b) Schematischer Verlauf der Kinetik in der initialen und postinitialen Phase nach BI-RADS TM	38
2.5	Verfahren der Bildverarbeitung in Zuordnung der Wissensrepräsentation	40
4.1	Ablauf der ausgeführten Bildverarbeitungsschritte der Prototypsoftware	56
4.2	Dynamische MRT-Untersuchung der Brust, der Bildausschnitt zeigt die verwendete Brustspule	58
4.3	Sequenzen und zeitlicher Ablauf des eingesetzten Protokolls	58
4.4	Teilschritte des Registrierungsalgorithmus	61
4.5	Beispiele von Skelettierungen im $2D$ mit den minimalen Randdistanzen	68
4.6	Voxeltoken mit minimaler Boundingbox	69
4.7	Interpretation der Kinetik durch Zeit- und Flächenmerkmale	73
4.8	Beispiele von GS-Vorgaben: (a) maligner Tumor , (b) maligner (rot) und verdächtiger (orange) Tumor, (c) benigner Tumor, jeweils auf unregistrierten Bilddaten	77
5.1	Visualisierung der (a) gesuchten und (b) getrackten Features mit Ausschnittsvergrößerung. Der Pfeil deutet die Lage des Tumors an.	94
5.2	Unterschiede der (a, $R_{SN_7} - SN_2$) registrierten und (b, $SN_7 - SN_2$) unregistrierten Differenzbilder. Die Abbildungen c und d markieren alle Voxel mit SI größer als 80.	95
5.3	Projektionsvolumen durch (a und b) Mittelwertbildung und (c und d) Maximumfunktion mit entsprechender Segmentierung.	97

5.4	Segmentierung einer malignen Läsion: (a) registrierte Goldstandardvorgabe (b) Abgrenzung der Token (orange) gegenüber allen anreichernden Strukturen (zyan) (c) Segmentierung auf dem Projektionsvolumen (d) Nachsegmentierung einer einzelnen Serie	99
5.5	Segmentierung einer benignen Läsion: (a) registrierte Goldstandardvorgabe (b) Abgrenzung der Token (orange) gegenüber allen anreichernden Strukturen (zyan) (c) Segmentierung auf dem Projektionsvolumen (d) Nachsegmentierung einer einzelnen Serie	100
6.1	Gegenüberstellung eines (a) unregistrierten Subtraktionsbildes mit starkem Bewegungseinfluss mit einem (b) registrierten.	110
6.2	Kreisförmige Artefakte durch fehlerhaftes Featuretracking. Die Bildausschnitte zeigen die benachbarten Bildschichten.	112
6.3	Gegenüberstellung (a und c) unregistrierter mit (b und d) registrierten Subtraktionen von Angiographiebildern.	118
6.4	CT Bilddaten in (a) In- und (b) Expiration . Gegenüberstellung der Subtraktion (c) vor und (d) nach der Registrierung.	119
6.5	Signalintensitäten von Läsionenhotsspots mit dem dazugehörigen Wertebereich optimaler Segmentierungsschwellwerte auf Bildeinzelserien.	124
6.6	Vergleich verschiedener benigner GS-Vorgaben (grüne Markierung der Bilder der linken Spalte) und deren Vermissten in der Segmentierung des Projektionsvolumens (rechte Spalte).	127
6.7	Verschiedene Bildschichten des Differenzvolumens derselben Patientin (linke Spalte) mit automatischer Segmentierung durch die Software (rechte Spalte).	129
6.8	Gegenüberstellung der Segmentierung (a) registrierter und (b) unregistrierter Projektionsvolumen. Die Markierungen zeigen die jeweilige Läsion.	134
6.9	Gegenüberstellung der Segmentierung (a und c) registrierter und (b und d) unregistrierter Projektionsvolumen. Die Markierungen zeigen die jeweilige Läsion.	135
6.10	Schematische Zeit/Signalintensitätskurve zur Darstellung Kinetik beeinflussender Faktoren.	147
6.11	Dendrogramm über alle Instanzen.	150
6.12	Dendrogramm über Läsionen beschreibende Instanzen.	150
6.13	Hauptkomponentenanalyse über alle Instanzen.	152
6.14	Hauptkomponentenanalyse über Läsionen beschreibende Instanzen.	153
6.15	Vergleich der Vorhersageleistung auf Basis verschiedener Merkmalsteilmengen.	159
6.16	Streudiagramm der selektierten Merkmale zur Diskriminierung maligner und benigner Läsionen.	163
6.17	Vorhersageleistung verschiedener Klassifikatoren auf dem gesamten Kollektiv.	168
6.18	Bildbeispiele verschiedener segmentierter zusätzlicher Funde.	170

6.19 Problemfall mit falscher Vorhersage: (a) Tumor (oben rechts) in Subtraktionsserie, (b) GS mit direkt angrenzender benigner (grün) und maligner (rot) Histologie und weiterer benigner verdächtiger Tumore (gelb), (c) Vorhersage als Karzinom (rot) durch die Prototypsoftware.	171
A.1 Patienten Koordinatensystem nach DICOM mit Visualisierung der Orientierungen	184
A.2 Darstellung des Patienten-KOS mit dem Ursprung im MRT-Gerät (oben) im Vergleich zum Modell-KOS ohne räumlichen Bezug (unten) und der KOS-Transformation (Mitte) anhand dreier Sequenzen verschiedener Schichtführung	186

Tabellenverzeichnis

2.1	Dignität von Tumoren und unterscheidbare Merkmale [TMV91]	21
2.2	Ursprungsgewebe und Häufigkeiten nach Biopsie verschiedener Tumorarten [HKS03, Fis00].	23
2.3	Einflussfaktoren auf die Bildqualität und der dargestellten Kinetik in der funktionalen MRT	33
2.4	Bewertung von morphologischen und kinetischen Kriterien durch den Göttinger Score (maximal 8 Punkte).	37
2.5	Durch die Scoringtable wird der Göttinger Summenscore (s. Tab. 2.4) abschließend bewertet.	37
2.6	Gegenüberstellung verschiedener Diagnoseverfahren bei der Tumordifferenzierung [FKG99].	37
2.7	Auszug definierter Terminologien des BI-RADS TM -Lexikon.	38
2.8	BI-RADS TM -Kategorien für die Brust-MRT.	39
4.1	Das Protokoll der Bilderhebung verwendet vier verschiedene Sequenzen unterschiedlicher Orientierung und Ortsauflösung.	59
4.2	Übersicht der berechneten kinetischen und morphologischen Merkmale . . .	76
4.3	Goldstandard Klassen	78
4.4	Kontingenztafel	87
5.1	manuelle Goldstandardfestlegung des Kollektivs	92
5.2	Tumortypen der histologisch gesicherten GS-Klassen 1 und 2	92
5.3	Registrierungsparameter	92
5.4	Vergleich der durchschnittlichen Signalintensität registrierter und unregistrierter Bildvolumen	94
5.5	Vergleich der Schwellwertfunktion und erkannten GS der untersuchten Projektionsfunktionen	98
5.6	Segmentierungsparameter	98
5.7	Morphologieparameter	101
5.8	Häufigkeiten der GS-Klassen in den Lerndaten	102
5.9	Kontingenztafel der Vorhersage der Gesundheit durch GS-Objekte (Sensitivität = 88 %, Spezifität = 72 %, Treffgenauigkeit = 82 %)	103

5.10	Kontingenztafel der Vorhersage der Gesundheit durch alle Objekte (Sensitivität = 98 %, Spezifität = 29 %, Treffgenauigkeit = 71 %)	104
6.1	Distanzmaße registrierter und unregistrierter Bildvolumen zur nativen Aufnahme	113
6.2	Relative Volumenänderung registrierter gegenüber unregistrierten GS-Läsionen	115
6.3	Relative Anzahl an automatisch segmentierten Läsionen, die einen der gestaffelten Überlappungswerte nach O_{in} mit dem GS erreichen	132
6.4	Relativer Anteil an GS-Überdeckungen ($O_{in} > 15$ %, $O_{out} < 50$ %) maligner und benigner Läsionen auf registrierten im Vergleich zu unregistrierten Bilddaten	136
6.5	Regressionsanalyse, Bestimmtheitsmaß R^2 : 0.1913, $R^2_{korrigiert}$: 0.1891, F-Test des gesamten Regressionsmodells: $F = 87.5$ mit p-Wert $< 2.2e-16$. Signifikanzkodierung: p-Wert: $0.0 < *** < 0.001 < ** < 0.01 < * < 0.05 < . < 0.1 < < 1.0$	162
6.6	Regressionsanalyse, Bestimmtheitsmaß R^2 : 0.1893, $R^2_{korrigiert}$: 0.188, F-Test des gesamten Regressionsmodells: $F = 138.4$ mit p-Wert $< 2.2e-16$. Signifikanzkodierung: p-Wert: $0.0 < *** < 0.001 < ** < 0.01 < * < 0.05 < . < 0.1 < < 1.0$	164
6.7	Regressionsanalyse, Bestimmtheitsmaß R^2 : 0,5439, $R^2_{korrigiert}$: 0.4552, F-Test des gesamten Regressionsmodells: $F = 6.132$ mit p-Wert $8.44e-11$. Signifikanzkodierung: p-Wert: $0.0 < *** < 0.001 < ** < 0.01 < * < 0.05 < . < 0.1 < < 1.0$	165
6.8	Regressionsanalyse, Bestimmtheitsmaß R^2 : 0,4647, $R^2_{korrigiert}$: 0.4293, F-Test des gesamten Regressionsmodells: $F = 13.13$ mit p-Wert $1.649e-13$. Signifikanzkodierung: p-Wert: $0.0 < *** < 0.001 < ** < 0.01 < * < 0.05 < . < 0.1 < < 1.0$	165
6.9	Vorhersage auf dem ungesicherten Goldstandard	166
6.10	Komponentenvergleich kommerzieller Software mit dem Prototypen der vorliegenden Arbeit. Legende: (A) Automatisch, (M) Manuell durch den Benutzer, (L) Low-Level-Diagnose, (H) High-Level-Diagnose, (-) nicht vorhanden	173
A.1	Liste verschiedener DICOM-Attribute, die das Patienten-KOS definieren. .	184

Literaturverzeichnis

- [ABBM05] ARMITAGE, P ; BEHRENBRUCH, C ; BRADY, M ; MOORE, N: Extracting and visualizing physiological parameters using dynamic contrast-enhanced magnetic resonance imaging of the breast. In: *Med Image Anal* 9 (2005), Nr. 4, S. 315–329 [32](#), [138](#), [146](#), [149](#)
- [ABN01] ABDOLMALEKI, P ; BUADU, LD ; NADERIMANSH, H: Feature extraction and classification of breast cancer on dynamic magnetic resonance imaging using artificial neural network. In: *Cancer Lett* 171 (2001), Nr. 2, S. 183–191 [52](#)
- [Ame] AMERICAN COLLEGE OF RADIOLOGY: *Digital Imaging and Communications in Medicine*. <http://medical.nema.org/>, [183](#)
- [Ame03] AMERICAN COLLEGE OF RADIOLOGY: *Breast Imaging Reporting and Data System*. <http://www.acr.org/>, 2003 [36](#)
- [AP03] ALTERSON, R ; PLEWES, DB: Bilateral symmetry analysis of breast MRI. In: *Phys. Med. Biol.* 48 (2003), Nr. 20, S. 3431–3443 [51](#)
- [Bak82] BAKER, LH: Breast Cancer Detection Demonstration Project: Five-Year Summary Report. In: *CA Cancer J Clin* 32 (1982), S. 194–225 [25](#)
- [BB82] BALLARD, DH ; BROWN, CM: *Computer Vision*. PrenticeHall, 1982 [41](#)
- [BEPW05] BACKHAUS, Klaus ; ERICHSON, Bernd ; PLINKE, Wulff ; WEIBER, Rolf: *Multivariate Analysemethoden*. Berlin : Springer, 2005 [80](#), [84](#), [88](#)
- [BFV02] BAUM, F ; FISCHER, U ; VOSSHENRICH, R: Classification of hypervascularized lesions in CE MR imaging of the breast. In: *Eur Radiol.* 12 (2002), Nr. 5, S. 1087–1092 [34](#)
- [BLGG91] BASSETT, LW ; LIU, TH ; GIULIANO, AE ; GOLD, RH: The prevalence of carcinoma in palpable vs impalpable, mammographically detected lesions. In: *AJR Am J Roentgenol* 157 (1991), Nr. 1, S. 21–24 [25](#)

- [BMK⁺01] BRIX, G ; MV, M H. ; KNOPP ; LUCHT, R ; DOLL, J ; JUNKERMANN, H ; HAWIGHORST, H ; HABERKORN, U: Comparison of pharmacokinetic MRI and [18F] fluorodeoxyglucose PET in the diagnosis of breast cancer: initial experience. In: *Eur Radiol* 11 (2001), Nr. 10, S. 2058–2070 52, 137
- [BMM97] BUADU, LD ; MURAKAMI, J ; MURAYAMA, S: Patterns of peripheral enhancement in breast masses: correlation of findings on contrast medium enhanced MRI with histologic features and tumor angiogenesis. In: *J Comput Assist Tomogr* 21 (1997), Nr. 3, S. 421–430 32
- [Bou02] BOUGUET, JY: *Pyramidal Implementation of the Lucas Kanade Feature Tracker: Description of the algorithm.* 2002 120
- [Bre01] BREIMAN, Leo: Random Forests. In: *Mach. Learn.* 45 (2001), Nr. 1, S. 5–32 89
- [BS] BRONSTEIN, I. N. ; SEMENDJAJEW, K. A.: *Taschenbuch der Mathematik.* Teubner Verlagsgesellschaft 190
- [BS91] BASSETT, LW ; SMITH, C K.: Breast sonography. In: *AJR Am J Roentgenol* 156 (1991), Nr. 3, S. 449–455 27
- [BSP⁺91] BRIX, G ; SEMMLER, W ; PORT, R ; SCHAD, LR ; LAYER, G ; LORENZ, WJ: Pharmacokinetic parameters in CNS Gd-DTPA enhanced MR imaging. In: *J Comput Assist Tomogr* 15 (1991), Nr. 4, S. 621–628 51
- [Bur84] BURT, PJ: The Pyramid as a Structure for Efficient Computation. In: ROSENFELD, A. (Hrsg.): *Multi-resolution Image Processing and Analysis.* Bd. 12, Springer Verlag, 1984, S. 6–35 189
- [CAD] CADSCIENCE: *Breast fTP.* – <http://www.cadsciences.com/>, 47
- [CGB⁺05] COTO, Ernesto ; GRIMM, Sören ; BRUCKNER, Stefan ; GRÖLLER, Meister E. ; KANITSAR, Armin ; RODRIGUEZ, Omaira: MammoExplorer: An Advanced CAD Application for Breast DCE-MRI. In: *Proceedings of Vision, Modelling, and Visualization 2005*, 2005, S. 91–98 51
- [CGB06] CHEN, W ; GIGER, M ; BICK, U: A Fuzzy C-Means (FCM)-Based Approach for Computerized Segmentation of Breast Lesions in Dynamic Contrast-Enhanced MR Images. In: *Academic Radiology* 13 (2006), Nr. 1, S. 63–72 51, 121, 128
- [CGC02] CALAMANTE, F ; GADIAN, DG ; CONNELLY, A: Quantification of perfusion using bolus tracking magnetic resonance imaging in stroke: assumptions, limitations, and potential implications for clinical use. In: *Stroke* 33 (2002), Nr. 4, S. 1146–1151 138

- [CL97] CREVIER, Daniel ; LEPAGE, Richard: Knowledge-based image understanding systems: a survey. In: *Comput. Vis. Image Underst.* 67 (1997), Nr. 2, S. 160–185 40
- [CMCD95] COSTELLO, P ; MCCANN, A ; CARNEY, DN ; DERVAN, PA: Prognostic significance of microvessel density in lymph node negative breast carcinoma. In: *Hum Pathol* 26 (1995), Nr. 11, S. 1181–1184 32
- [Con] CONFIRMA: *CADstream*. – <http://www.confirma.com/>, 47
- [CTC91] CIATTO, S ; TURCO, M R. ; CATARZI, S: Causes of breast cancer misdiagnosis at physical examination. In: *Neoplasma* 38 (1991), Nr. 5, S. 523–531 25
- [DGW97] DEGANI, H ; GUSIS, V ; WEINSTEIN, D: Mapping pathophysiological features of breast tumors by MRI at high spatial resolution. In: *Nat Med* 3 (1997), Nr. 7, S. 780–782 47
- [DMK⁺99] DOI, K ; MACMAHON, H ; KATSURAGAWA, S ; NISHIKAWA, RM ; Y, Y J.: Computer-aided diagnosis in radiology: potential and pitfalls. In: *Eur J Radiol* 31 (1999), Nr. 2, S. 97–109 45, 157, 179
- [DRG95] Empfehlungen des Ausschusses Magnetresonanztomographie für Qualitätsanforderungen der MRT der Mamma. Jahresbericht. In: *Deutsche Röntgengesellschaft* 43 (1995), S. 59–61 34
- [DSRR99] DENTON, E ; SONODA, L ; RUECKERT, D ; RANKIN, S: Comparison and Evaluation of Rigid, Affine, and Nonrigid Registration of Breast MR Images. In: *J Comp Ass Tom* 23 (1999), Nr. 5, S. 800–805 81, 106
- [EMPV95] ELSSEN, PA V. ; MAINTZ, JBA ; POL, E-JD ; VIERGEVER, MA: Automatic registration of CT and MR brain images using correlation of geometrical features. In: *IEEE Trans. Med. Images* 14 (1995), Nr. 2, S. 384–396 49
- [FASW92] FLICKINGER, FW ; ALLISON, JD ; SHERRY, R ; WRIGHT, JC: Differentiation of benign from malignant breast masses by dynamic time-intensity evaluation of contrast enhanced MRI. In: *Proceedings of the Society of Magnetic Resonance in Medicine*. Berlin, 1992 32
- [FDTC⁺99] FISHER, ER ; DIGNAM, J ; TAN-CHIU, E ; COSTANTINO, J ; FISHER, B ; PAIK, S ; WOLMARK, N: Pathologic findings from the National Surgical Adjuvant Breast Project (NSABP) eight-year update of Protocol B-17: intraductal carcinoma. In: *Cancer* 86 (1999), Nr. 3, S. 429–38 21
- [Fei91] FEIG, SA: Mammographic screening: an historical perspective. In: *Semin Roentgenol* 28 (1991), Nr. 3, S. 193–203 25

- [FGC94] FROUGE, C ; GUINEBRETIERE, JM ; CONTESSO, G: Correlation between contrast enhancement in dynamic magnetic resonance imaging of the breast and tumor angiogenesis. In: *Invest Radiol* 29 (1994), Nr. 12, S. 1043–1049 [32](#)
- [FH06] FISCHER, U ; HELBICH, T: *Mammographiefindung nach BI-RADS*. Georg Thieme Verlag, 2006 [36](#)
- [FHD02] FURMAN-HARAN, E ; DEGANI, H: Parametric analysis of breast MRI. In: *J Comput Assist Tomogr* 26 (2002), Nr. 3, S. 376–386 [47](#)
- [FHV⁺93] FISCHER, U ; HEYDEN, D von ; VOSSHENRICH, R ; VIEWEG, I ; GRABBE, E: Signalverhalten maligner und benigner Läsionen in der dynamischen 2D-MRT der Mamma. In: *Fortschr Röntgenstr* 158 (1993), S. 287–292 [30](#), [32](#)
- [Fis36] FISHER, RA: The use of multiple measurements in taxonomic problems. In: *Annals Eugen.* 7 (1936), S. 179–188 [88](#)
- [Fis98] FISCHER, U: Aktuelle Aspekte der dynamischen MR-Mammographie. In: *Habilitationsschrift Universität Göttingen* (1998) [32](#)
- [Fis00] FISCHER, U: *Lehratlas der MR-Mammographie*. Georg Thieme Verlag, 2000 [19](#), [22](#), [23](#), [29](#), [32](#), [33](#), [34](#), [35](#), [51](#), [137](#), [160](#), [205](#)
- [FKG99] FISCHER, U ; KOPKA, L ; GRABBE, E: Breast Carcinoma: Effect of Pre-operative Contrast-enhanced MR Imaging on the Therapeutic Approach. In: *Radiology* 213 (1999), Nr. 3, S. 881–888 [35](#), [37](#), [156](#), [205](#)
- [FKPT04] FAHRMEIR, L ; KÜNSTLER, R ; PIGEOT, I ; TUTZ, G: *Statistik*. Berlin : Springer, 2004 [85](#), [88](#)
- [FLS⁺94] FOX, SB ; LEEK, RD ; SMITH, K ; HOLLYER, J ; GREENALL, M ; HARRIS, AL: Tumor angiogenesis in node-negative breast carcinomas—relationship with epidermal growth factor receptor, estrogen receptor, and survival. In: *Breast Cancer Res Treat* 29 (1994), Nr. 1, S. 109–116 [32](#)
- [Fol92] FOLKMAN, J: The role of angiogenesis in tumor growth. In: *Semin cancer Biol* 3 (1992), Nr. 2, S. 65–71 [23](#), [34](#)
- [FPF⁺06] FORBES, F ; PEYRARD, N ; FRALEY, C ; GEORGIAN-SMITH, D ; GOLDHABER, DM ; RAFTERY, AE: Model-based region-of-interest selection in dynamic breast MRI. In: *J Comput Assist Tomogr* 30 (2006), Nr. 4, S. 675–687 [51](#)
- [Fri93] FRIEDRICH, M: The technique and results of mammography. In: *Radiologe* 33 (1993), Nr. 5, S. 243–259 [26](#)

- [FWB⁺96] FISCHER, U ; WESTERHOF, JP ; BRINCK, U ; A, A S. ; GRABBE, E: Das dukdale in-situ-Karzinom in der dynamischen MR-Mammographie bei 1,5T. In: *Fortschr Röntgenstr* 164 (1996), Nr. 4, S. 290–294 [30](#)
- [FWIH89] FOLKMAN, J ; WATSON, K ; INGBER, D ; HANAHAN, D: Induction of angiogenesis during the transition from hyperplasia to neoplasia. In: *Nat* 339 (1989), S. 58–61 [32](#)
- [FYM93] FIGUEROA, JA ; YEE, D ; MCGUIRE, WL: Prognostic indicators in early breast cancer. In: *Am J Med Sci* 305 (1993), Nr. 3, S. 176–182 [25](#)
- [GDM⁺02] GILHUIJS, KG ; DEURLOO, E ; MULLER, SH ; PETERSE, JL ; KOOL, LJ S.: Breast MR Imaging in Women at Increased Lifetime Risk of Breast Cancer: Clinical System for Computerized Assessment of Breast Lesions-Initial Results. In: *Radiology* 225 (2002), S. 907–916 [52](#), [137](#)
- [GFC⁺92] GOEDDE, TA ; FRYKBERG, ER ; CRUMP, JM ; LAY, SF ; TURETSKY, DB ; LINDEN, SS: The impact of mammography on breast biopsy. In: *Am Surg* 58 (1992), S. 661–666 [26](#)
- [GGB98] GILHUIJS, KG ; GIGER, ML ; BICK, U: Computerized analysis of breast lesions in three dimensions using dynamic magnetic-resonance imaging. In: *Medical Physics* 25 (1998), S. 1647–1654 [52](#), [137](#)
- [GGVMB⁺04] GARCÍA-GÓMEZ, JM ; VIDAL, C ; MARTÍ-BONMATÍ, L ; GALANT, J ; SANS, N ; ROBLES, M ; CASACUBERTA, F: Benign /malignant classifier of soft tissue tumors using MR imaging. In: *MAGMA* 16 (2004), Nr. 4, S. 194–201 [51](#)
- [GI] GENTLEMAN, R ; IHAKA, R: *The R Project for Statistical Computing.* – <http://www.r-project.org/>, [55](#), [88](#)
- [GLCF72] GIMBRONE, MA ; LEAPMAN, SB ; COTRAN, RS ; FOLKMAN, J: Tumor dormancy in vivo by prevention of neovascularisation. In: *J. Exp. Med* 136 (1972), S. 261–276 [23](#)
- [GML⁺96] GILLES, R ; MEUNIER, M ; LUCIDARME, O ; ZAFRANI, B ; GUINEBRETIERE, JM ; TARDIVON, AA ; GAL, M L. ; VANEL, D ; NEUENSCHWANDER, S ; ARRIAGADA, R: Clustered breast microcalcifications: evaluation by dynamic contrast-enhanced subtraction MRI. In: *J Comput Assist Tomogr* 20 (1996), Nr. 1, S. 9–14 [30](#)
- [GWB⁺94] GASPARINI, G ; WEIDNER, N ; BEVILACQUA, P ; MALUTA, S ; PALMA, P D. ; CAFFO, O ; BARBARESCHI, M ; BORACCHI, P ; MARUBINI, E ; POZZA, F: Tumor microvessel density, p53 expression, tumor size, and peritumoral lymphatic vessel invasion are relevant prognostic markers in node-negative breast carcinoma. In: *J Clin Oncol* 12 (1994), Nr. 3, S. 454–466 [32](#)

- [Ham04] HAMILL, P: *Unit Test Frameworks*. O'Reilly, 2004 82
- [Han00] HANDELS, H.: *Medizinische Bildverarbeitung*. Teubner Verlag, 2000 65, 70, 71
- [Han02] HANDL, A: *Multivariate Analysemethoden. Theorie und Praxis multivariater Verfahren unter besonderer Berücksichtigung von S-PLUS*. Berlin : Springer, 2002 86, 88
- [Hec94] HECKBERT, S: *Graphics Gem IV*. Academic Press, 1994 62
- [HES97] HULKA, CA ; EDMISTER, WB ; SMITH, BL: Dynamic echo-planar imaging of the breast: experience in diagnosing breast carcinoma and correlation with tumor angiogenesis. In: *Radiology* 205 (1997), Nr. 3, S. 837–842 32
- [HFH+93] HARMS, SE ; FLAMIG, DP ; HESLEY, KL ; MEICHES, MD ; JENSEN, RA ; EVANS, WP ; SAVINO, DA ; WELLS, RV: MR imaging of the breast with rotating delivery of excitation off resonance: clinical experience with pathologic correlation. In: *Radiology* 187 (1993), Nr. 2, S. 493–501 30
- [HHS+86] HEYWANG, SH ; HAHN, D ; SCHMIDT, H ; KRISCHKE, I ; EIERMANN, W ; BASSERMANN, R ; LISSNER, J: MR imaging of the breast using gadolinium-DTPA. In: *J Comput Assist Tomogr* 10 (1986), S. 199–204 29, 51, 137
- [HKBL+92] HEYWANG-KÖBRUNNER, SH ; BECK, R ; LOMMATZSCH, B ; KELLNER, W ; UNTOH, M ; NATHRATH, W: Contrast-enhanced MR Imaging of the breast: survey of 1200 patient examinations. In: *Radiology* 185 (1992), Nr. P, S. 246 30, 182
- [HKHP+94] HEYWANG-KÖBRUNNER, SH ; HAUSTEIN, J ; POHL, C ; BECK, R ; LOMMATZSCH, B ; UNTCH, M ; NATHRATH, WB: Contrast-enhanced MR imaging of the breast: comparison of two different doses of gadopentetate dimeglumine. In: *Radiology* 191 (1994), S. 639–646 32, 33
- [HKS03] HEYWANG-KÖBRUNNER, SH ; SCHREER, I: *Bildgebende Mammadiagnostik*. Georg Thieme Verlag, 2003 22, 23, 25, 26, 27, 205
- [HTF01] HASTIE, T ; TIBSHIRANI, R ; FRIEDMAN, J: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Berlin : Springer, 2001 80, 86, 88, 164
- [HWP+89] HEYWANG, SH ; WOLF, A ; PRUSS, E ; HILBERTZ, T ; EIERMANN, W ; PERMANETTER, W: MR imaging of the breast with Gd-DTPA: use and limitations. In: *Radiology* 171 (1989), Nr. 1, S. 95–103 30, 51, 137
- [JB98] JACKSON, VP ; BASSETT, LW: Breast sonography. In: *Breast Dis* 10 (1998), Nr. 3-4, S. 55–66 27

- [Jäh05] JÄHNE, B: *Digitale Bildverarbeitung*. Springer, 2005 [42](#), [43](#)
- [JTKG00] JOHNSON, KM ; TAO, JZ ; KENNAN, RP ; GORE, JC: Intravascular susceptibility agent effects on tissue transverse relaxation rates in vivo. In: *Magn Reson Med* 44 (2000), Nr. 6, S. 909–914 [32](#)
- [KAH⁺96] KUMAR, R ; ASMUTH, JC ; HANNA, K ; BERGEN, JR ; HULKA, C ; KOPANS, DB ; WEISSKOFF, R ; MOORE, R: Application of 3D registration for detecting lesions in magnetic resonance breast scans. In: *Proc. SPIE* Bd. 2710, 1996, S. 646–656 [50](#)
- [Kai90] KAISER, WA: Dynamic MR-Breast Imaging using a double breast coil: a step towards routine MR examination of the breast. In: *Front Eur Radiol* 7 (1990), S. 39–68 [30](#), [32](#), [182](#)
- [Kai93] KAISER, WA: MR-Mammographie. In: *Radiologe* 33 (1993), S. 292–299 [51](#), [137](#)
- [Kai07] KAISER, WA: *Signs in MR-Mammography*. Springer-Verlag New York, Inc., 2007 [v](#), [17](#), [51](#), [143](#)
- [Kal06] KALENDER, Willi A.: *Computertomographie, Grundlagen, Gerätetechnologie, Bildqualität, Anwendungen*. Publicis Corporate Publishing, 2006 [117](#)
- [KCHL99] KRISHNAN, S ; CHENEVERT, TL ; HELVIE, MA ; LONDY, FL: Linear motion correction in three dimensions applied to dynamic gadolinium enhanced breast imaging. In: *Medical Physics* 26 (1999), S. 707–714 [50](#)
- [KFB⁺98] KWAE, EA el ; FISHMAN, JE ; BIANCHI, MJ ; PATTANY, PM ; KABUKA, MR: Detection of suspected malignant patterns in three-dimensional magnetic resonance breast images. In: *J Digit Imaging* 11 (1998), Nr. 2, S. 83–92 [51](#)
- [KFHD02] KELCZ, F ; FURMAN-HARAN, E ; D, D G.: Clinical testing of high-spatial-resolution parametric contrast-enhanced MR imaging of the breast. In: *AJR Am J Roentgenol* 179 (2002), Nr. 6, S. 1485–1492 [47](#)
- [KGB⁺96] KERLIKOWSKE, K ; GRADY, D ; BARCLAY, J ; SICKLES, EA ; ERNSTER, V: Likelihood ratios for modern screening mammography. Risk of breast cancer based on age and mammographic interpretation. In: *JAMA* 276 (1996), Nr. 1, S. 39–43 [26](#)
- [KHR⁺02] KNOPP, MV ; HIMMELHAN, N ; RADELEFF, J ; JUNKERMANN, H ; HESS, T ; SINN, HP ; BRIX, G: Methodenvergleich zur Quantifizierung der Kontrastmittelanreicherung am Beispiel der dynamischen MR-Mammographie. In: *Radiologe* 42 (2002), S. 280–290 [137](#)

- [KJM⁺06] KUHL, CK ; JOST, P ; MORAKKABATI, N ; ZIVANOVIC, O ; SCHILD, HH ; GIESEKE, J: Contrast-enhanced MR Imaging of the Breast at 3.0 and 1.5 T in the Same Patients: Initial Experience. In: *Radiology* 239 (2006), S. 666–676 [34](#), [143](#)
- [KK89] KAISER, WA ; KESS, H: Prototyp-Doppelspule für die Mamma-MR-Messung. In: *Fortschr Röntgenstr* 1151 (1989), S. 103–105 [29](#)
- [KKM⁺99] KUHL, CK ; KLASCHIK, S ; MIELCAREK, P ; GIESEKE, J ; WARDELMANN, E ; SCHILD, HH: Do T2-weighted pulse sequences help with the differential diagnosis of enhancing lesions in dynamic breast MRI? In: *J Magn Reson Imaging* 9 (1999), Nr. 2, S. 187–196 [35](#)
- [KMK⁺99] KUHL, CK ; MIELCARECK, P ; KLASCHIK, S ; LEUTNER, C ; WARDELMANN, E ; GIESEKE, J ; SCHILD, HH: Dynamic breast MR imaging: are signal intensity time course data useful for differential diagnosis of enhancing lesions? In: *Radiology* 211 (1999), Nr. 1, S. 101–110 [32](#)
- [KSS95] KUHL, CK ; SEIBERT, C ; SOMMER, T: Fokale und diffuse Läsionen in der dynamischen MR-Mammographie gesunder Probandinnen. In: *Fortschr. Röntgenstr.* 163 (1995), S. 219–224 [32](#), [33](#)
- [KTMW⁺05] KIER, C ; TOTH, D ; MEYER-WIETHE, K ; SCHINDLER, A ; CANGUR, H ; SEIDEL, G ; AACH, T: Cerebral perfusion imaging with bolus harmonic imaging. In: *Proc SPIE* Bd. 5750, 2005, S. 437–446 [73](#)
- [Kuh00] KUHL, CK: MRI of breast tumors. In: *Eur Radiol* 10 (2000), Nr. 1, S. 46–58 [32](#), [34](#), [35](#), [66](#), [160](#)
- [KZ89] KAISER, WA ; ZEITLER, E: MR imaging of the breast: fast imaging sequences with and without Gd-DTPA. Preliminary observations. In: *Radiology* 170 (1989), S. 681–686 [30](#)
- [LC69] LYKKE, AW ; CUMMINGS, R: Inflammation in healing. 1. Time-course and mediation of exudation in wound healing in the rat. In: *Br J Exp Pathol* 50 (1969), Nr. 3, S. 309–318 [32](#)
- [LDC⁺98] LIU, PF ; DEBATIN, JF ; CADUFF, RF ; KACL, G ; GARZOLI, E ; KRESTIN, GP: Improved diagnostic accuracy in dynamic contrast enhanced MRI of the breast by combined quantitative and qualitative analysis. In: *Br J Radiol* 71 (1998), Nr. 845, S. 501–509 [30](#)
- [Lip06] LIPPERT, H: *Lehrbuch Anatomie*. Elsevier, 2006 [19](#), [22](#), [23](#), [24](#), [25](#)
- [LK81] LUCAS, BD ; KANADE, T: An Iterative Image Registration Technique with an Application to Stereo Vision. In: *IJCAI81*, 1981, S. 674–679 [189](#)

- [LKB00] LUCHT, R ; KNOPP, MV ; BRIX, G: Elastic matching of dynamic MR mammographic. In: *images. Magn. Reson. in Medicine* 43 (2000), S. 9–16
50
- [LML⁺02] LIBERMAN, L ; MORRIS, EA ; LEE, M Joo-Young ; KAPLAN, JB ; LATRENTA, LR ; MENELL, JH ; ABRAMSON, AF ; DASHNAW, SM ; DERSHAW, DJ Ballon D.: Breast lesions detected on MR imaging: features and positive predictive value. In: *AJR Am J Roentgenol* 179 (2002), Nr. 1, S. 171–178
34
- [LOPR97] LEHMANN, T ; OBERSCHELP, W ; PELIKAN, E ; REPGES, R: *Bildverarbeitung für die Medizin*. Springer Verlag, 1997 71, 128
- [LRC94] LEE, TC ; RANGASAMI, LK ; CHU, CN: Building skeleton models via 3-D medial surface/axis thinning algorithms. In: *CVGIP: Graph. Models Image Process.* 56 (1994), Nr. 6, S. 462–478. – ISSN 1049–9652 68
- [LSG⁺06] LINEY, GP ; SREENIVAS, M ; GIBBS, P ; GARCIA-ALVAREZ, R ; TURNBULL, LW: Breast lesion analysis of shape technique: Semiautomated vs. manual morphological description. In: *J Magn Reson Imaging* (2006) 52
- [LVG80] LOBREGT, S ; VERBEEK, W ; GROEN, FCA: Three-dimensional skeletonization: principle and algorithm. In: *IEEE Transactions on Pattern Analyses and Machine Intelligence* 2 (1980), Nr. 1, S. 75–77 68
- [LWS97] LEE, S ; WOLBERG, G ; SHIN, SY: Scattered data interpolation with multilevel B-splines. In: *Visualization and Computer Graphics, IEEE Transactions on* 3 (1997), Nr. 3, S. 228–244 193
- [Man87] MANDELBROT, BB.: *Die fraktale Geometrie der Natur*. Basel : Birkhäuser, 1987 70
- [MBB⁺95] MUSSURAKIS, S ; BUCKLEY, DL ; BOWSELEY, SJ ; CARLETON, PJ ; FOX, JN ; TURNBULL, LW ; HORSMAN, A: Dynamic contrast-enhanced magnetic resonance imaging of the breast combined with pharmacokinetic analysis of gadolinium-DTPA uptake in the diagnosis of local recurrence of early stage breast carcinoma. In: *Invest Radiol* 30 (1995), Nr. 11, S. 650–662 30
- [MBH97] MUSSURAKIS, S ; BUCKLEY, DL ; HORSMAN, A: Dynamic MR imaging of invasive breast cancer: correlation with tumour grade and other histological factors. In: *The British Journal of Radiology* 70 (1997), Nr. 833, S. 446–451
32
- [Mei00] MEIJERING, E: Spline Interpolation in Medical Imaging: Comparison with Other Convolution-Based Approaches. In: GABBOUJ, M. (Hrsg.) ; KUOSMANEN, P. (Hrsg.): *Signal Processing X: Theories and Applications - Proceedings of EUSIPCO 2000* Bd. 4, 2000, S. 1989–1996 111

- [Mev] MEVIS: *DynaCADTM*. – <http://www.mevis.de/mms/Produktinfo/DynaCAD.html>, 47
- [Mic] MICROSOFT®: *Visual Studio V6.0*. – <http://www.microsoft.com/>, 55
- [Mil87] MILLNER, R: *Ultraschalltechnik, Grundlagen und Anwendungen*. Physik Verlag, 1987 27
- [MKJ⁺08] MACIAK, A ; KRONFELD, A ; JANSEN, O ; VUCUREVIC, G ; STOETER, P: Validierung der Reproduzierbarkeit der Perfusionsmessung mittels eines neuartigen Kapillar-basiertem MR-Perfusionsphantoms. In: *43. Jahrestagung der deutschen Gesellschaft für Neuroradiologie* (2008) 146
- [Mor01] MORAVEC, Hans: *Mind Children. Der Wettlauf zwischen menschlicher und künstlicher Intelligenz*. Hamburg : Hoffmann und Campe, 2001 40
- [MV98] MAINTZ, J ; VIERGEVER, M: A survey of medical image registration. In: *Medical Image Analysis 2* (1998), Nr. 1, S. 1–36 49, 50
- [MVF⁺06] MAYER, D ; VOMWEG, TW ; FABER, H ; WEINHEIMER, O ; MATTIUZZI, M ; BUSCEMA, M ; DÜBER, C: Fully Automatic Breast Cancer Diagnosis in Contrast Enhanced MRI. In: SPRINGER (Hrsg.): *Int J CARS 2006, Suppl 1* Bd. 1, 2006, S. 325–343 v, 18
- [NAL⁺05] NATTKEMPER, TW ; ARNRICH, B ; LICHTHE, O ; TIMM, W ; DEGENHARD, A ; POINTON, L ; HAYES, C ; LEACH, MO: Evaluation of radiological features for breast tumour classification in clinical screening with machine learning methods. In: *Artif Intell Med 34* (2005), Nr. 2, S. 129–139 53
- [NH04] NISCHWITZ, A ; HABERÄCKER, P: *Masterkurs Computergrafik und Bildverarbeitung*. Vieweg, 2004 42, 185
- [NHP⁺95] NISHIKAWA, RM ; HALDEMANN, RC ; PAPAIOANNOU, J ; GIGER, ML ; LU, P ; SCHMIDT, RA ; WOLVERTON, DE ; BICK, U ; DOI, K: Initial experience with a prototype clinical intelligent mammography workstation for computer-aided diagnosis. In: *Proc. SPIE 2434* (1995), Nr. 65, S. 65–71 179
- [NSO⁺97] NUNES, LW ; SCHNALL, MD ; OREL, SG ; HOCHMAN, MG ; LANGLOTZ, CP ; REYNOLDS, CA ; TOROSIAN, MH: Breast MR imaging: interpretation model. In: *Radiology 202* (1997), Nr. 3, S. 833–841 34
- [Off] OFFIS: *DCMTK*. – <http://dicom.offis.de/>, 55
- [OKC⁺95] OBERMAIR, A ; KURZ, C ; CZERWENKA, K ; THOMA, M ; KAIDER, A ; WAGNER, T ; GITSCH, G ; SEVELDA, P: Microvessel density and vessel invasion in lymph-node-negative breast cancer: effect on recurrence-free survival. In: *Int J Cancer 62* (1995), Nr. 2, S. 126–131 32

- [ORS⁺97] OREL, SG ; REYNOLDS, C ; SCHNALL, MD ; SOLIN, LJ ; FRAKER, DL ; SULLIVAN, DC: Breast carcinoma: MR imaging before re-excisional biopsy. In: *Radiology* 205 (1997), Nr. 2, S. 429–436 30
- [OVT⁺91] OORD, JC van ; VLIET, AM van d. ; THYN, CJ ; MAK, B ; HOOGEBOOM, GJ: The value of ultrasound mammography in palpable breast masses. In: *Rofa* 155 (1991), Nr. 1, S. 243–259 27
- [Pea76] PEARLMAN, AW: Breast cancer–influence of growth rate on prognosis and treatment evaluation: a study based on mastectomy scar recurrences. In: *Cancer* 38 (1976), Nr. 4, S. 1826–1833 22
- [Pen] PENN, A: *ONCADTM*. – <http://www.oncad.com/>, 48
- [PKT⁺01] PENN, AI ; KUMAR, N ; THOMPSON, SF ; SCHNALL, MD ; WANG, F ; GATSONIS, C: Preliminary performance analysis of breast MRI CAD system. In: *Proc. SPIE* Bd. 4322, 2001, S. 1944–1953 48, 175
- [PL97] PENN, A ; LOEW, H: Estimating fractal dimension with fractal interpolation function models. In: *IEEE Trans Med Imaging* 16 (1997), Nr. 6, S. 930–937 48, 52, 175
- [PPF99] PARKIN, DM ; PISANI, P ; FERLAY, J: Estimates of the worldwide incidence of 25 major cancers in 1990. In: *Int J Cancer* 80 (1999), Nr. 6, S. 827–841 17
- [PTB⁺06] PENN, A ; THOMPSON, S ; BREM, R ; LEHMAN, C ; WEATHERALL, P ; SCHNALL, M ; NEWSTEAD, G ; CONANT, E ; ASCHER, S ; MORRIS, E ; PISANO, E: Morphologic blooming in breast MRI as a characterization of margin for discriminating benign from malignant lesions. In: *Acad Radiol* 13 (2006), Nr. 11, S. 1344–1354 52
- [RBC93] REINTGEN, D ; BERMAN, C ; COX, C: The anatomy of missed breast cancers. In: *Surg Oncol* 2 (1993), Nr. 1, S. 65–75 25
- [RCM01] ROHLFING, T ; CR MAURER, Jr.: Intensity-Based Non-rigid Registration Using Adaptive Multilevel Free-Form Deformation with an Incompressibility Constraint. In: *MICCAI '01: Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer-Verlag, 2001, S. 111–119 111
- [RHBK02] REICHENBACH, JR ; HOPFE, J ; BELLEMANN, ME ; KAISER, WA: Development and validation of an algorithm for registration of serial 3D MR breast data sets. In: *MAGMA* 14 (2002), Nr. 3, S. 249–257 81

- [RHW⁺98] ROSENBERG, RD ; HUNT, WC ; WILLIAMSON, MR ; GILLILAND, FD ; WIEST, PW ; KELSEY, CA ; KEY, CR ; LINVER, MN: Effects of age, breast density, ethnicity, and estrogen replacement therapy on screening mammographic sensitivity and cancer stage at diagnosis: review of 183,134 screening mammograms in Albuquerque, New Mexico. In: *Radiology* 209 (1998), Nr. 2, S. 511–518 26
- [RMBJ03] ROHLFING, T ; MAURER, C ; BLUEMKE, D ; JACOBS, M: Volume-preserving non-rigid registration of MR breast images using free-form deformation with an incompressibility constraint. In: *IEEE Trans. Med. Imag.* 22 (2003), Nr. 6, S. 730–741 108, 111, 116
- [Rob04] ROBERT KOCH INSTITUT: *Krebs in Deutschland 2003 - 2004*. <http://www.rki.de/>, 2004 17
- [RSH⁺99] RUECKERT, D ; SONODA, LI ; HAYES, C ; HILL, DL ; LEACH, MO ; HAWKES, DJ: Nonrigid registration using free-form deformations: application to breast MR images. In: *IEEE Trans Med Imaging* 18 (1999), Nr. 8, S. 712–721 108, 116
- [RSM⁺06] RUITER, NV ; STOTZKA, R ; MULLER, TO ; GEMMEKE, H ; REICHENBACH, JR ; KAISER, WA: Model-based registration of X-ray mammograms and MR images of the female breast. In: *Nuclear Science* 53 (2006), Nr. 1, S. 204–211 50
- [Sch05] SCHICK, F: Grundlagen der Magnetresonanztomographie (MRT). In: *Der Radiologe* 45 (2005), Nr. 1, S. 69–88 29
- [SFP88] SCHWARTZ, GF ; FEIG, SA ; PATCHEFESKY, AS: Significance and staging of nonpalpable carcinomas of the breast. In: *Surg Gynecol Obstet* 166 (1988), Nr. 1, S. 6–10 26
- [Siv05] SIVARAMAKRISHNA, R: 3D Breast Image Registration – A Review. In: *Technol Cancer Res Treat* 4 (2005), Nr. 1, S. 39–48 49
- [SKH⁺97] SITTEK, H ; KESSLER, M ; HEUCK, AF ; BREDL, T ; PERLET, C ; KÜNZER, I ; LEBEAU, A ; UNTCH, M ; REISER, M: Morphology and contrast enhancement of ductal carcinoma in situ in dynamic 1.0 T MR mammography. In: *Fortschr Röntgenstr* 167 (1997), Nr. 3, S. 247–51 30
- [SLQD⁺97] SINHA, S ; LUCAS-QUESADA, FA ; DEBRUHL, ND ; SAYRE, J ; FARRIA, D ; GORCZYCA, DP ; BASSETT, LW: Multifeature analysis of Gd-enhanced MR images of breast lesions. In: *J Magn Reson Imaging* 7 (1997), Nr. 6, S. 1016–1026 30
- [Soi98] SOILLE, P: *Morphologische Bildverarbeitung*. Springer Verlag, 1998 43

- [SOS97] SIEWERT, C ; OELLINGER, H ; SHERIF, HK: Is there a correlation in breast carcinomas between tumor size and number of tumor vessels detected by gadolinium-enhanced magnetic resonance mammography? In: *MAGMA* 5 (1997), Nr. 1, S. 29–31 32
- [SP86] SEDERBERG, TW ; PARRY, SR: Free-form deformation of solid geometric models. In: *SIGGRAPH Comput. Graph.* 20 (1986), Nr. 4, S. 151–160 108
- [SSB02] SHAHAR, KH ; SOLAIYAPPAN, M ; BLUEMKE, DA: Quantitative differentiation of breast lesions based on three-dimensional morphology from magnetic resonance imaging. In: *J Comput Assist Tomogr* 26 (2002), Nr. 6, S. 1047–1053 52
- [STCS+03] SCHNABEL, JA ; TANNER, C ; CASTELLANO-SMITH, AD ; DEGENHARD, A ; LEACH, MO ; HOSE, DR ; HILL, DLG ; HAWKES, DJ: Validation of nonrigid image registration using finite-element methods: application to breast MR images. In: *IEEE Trans Med Imaging* 22 (2003), Nr. 2, S. 238–247 81
- [STL] STLPORT: *STL*. – <http://www.stlport.org/>, 55
- [Sto04] STOER, J: *Numerische Mathematik 1*. Springer, 2004 73
- [SW07] STAUBER, Manfred ; WEYERSTAHL, Thomas: *Gynäkologie und Geburtshilfe*. Stuttgart : Georg Thieme Verlag, 2007 24
- [SWBA] SZABO, BK ; WIBERG, MK ; BONE, B ; ASPELIN, P: Application of artificial neural networks to the analysis of dynamic MR imaging features of the breast. In: *Eur Radiol* 14, Nr. 7 51, 53, 157
- [SWH+97] STOMPER, PC ; WINSTON, JS ; HERMAN, S ; KLIPPENSTEIN, DL ; MA, MA A. ; BLUMENSON, LE: Angiogenesis and dynamic MR imaging gadolinium enhancement of malignant and benign breast lesions. In: *Breast Cancer Res Treat* 45 (1997), Nr. 1, S. 39–46 32
- [SY88] SAMEJIMA, N ; YAMAZAKI, K: A study on the vascular proliferation in tissues around the tumor in breast cancer. In: *Jpn J Surg* 18 (1988), Nr. 3, S. 235–242 32
- [TBS95] TOFTS, PS ; BERKOWITZ, B ; SCHNALL, MD: Quantitative analysis of dynamic Gd-DTPA enhancement in breast tumors using a permeability model. In: *Magn Reson Med* 33 (1995), Nr. 4, S. 564–568 47
- [TBS+06] TEIFKE, A ; BEHR, O ; SCHMIDT, M ; VICTOR, A ; VOMWEG, TW ; THELEN, M ; LEHR, HA: Dynamic MR imaging of breast lesions: correlation with microvessel distribution pattern and histologic characteristics of prognosis. In: *Radiology* 239 (2006), Nr. 2, S. 351–360 32

- [TKT93] TOI, M ; KASHITANI, J ; TOMINAGA, T: Tumor angiogenesis is an independent prognostic indicator in primary breast carcinoma. In: *Int J Cancer* 55 (1993), Nr. 3, S. 371–374 [32](#)
- [TMV91] THEWS, G ; MUTSCHER, E ; VAUPEL, P: *Anatomie, Physiologie, Pathophysiologie des Menschen*. Wiss. Verl.-Ges., 1991 [19](#), [20](#), [21](#), [22](#), [205](#)
- [TNB03] TZACHEVA, AA ; NAJARIAN, K ; BROCKWAY, JP: Breast cancer detection in gadolinium-enhanced MR images by static region descriptors and neural networks. In: *J Magn Reson Imaging* 17 (2003), Nr. 3, S. 337–342 [52](#), [157](#)
- [TSC⁺00] TANNER, C ; SCHNABEL, JA ; CHUNG, D ; CLARKSON, MJ ; RUECKERT, D ; HILL, DLG ; HAWKES, DJ: *Volume and Shape Preservation of Enhancing Lesions when Applying Non-rigid Registration to a Time Series of Contrast Enhancing MR Breast Images*. 2000. – 327–337 S. [107](#)
- [VBK⁺03] VOMWEG, TW ; BUSCEMA, M ; KAUCZOR, HU ; TEIFKE, A ; INTRALIGI, M ; TERZI, S ; HEUSSEL, CP ; ACHENBACH, T ; RIEKER, O ; MAYER, D ; THELEN, M: Improved artificial neural networks in prediction of malignancy of lesions in contrast-enhanced MR-mammography. In: *Med Phys* 30 (2003), Nr. 9, S. 2350–2359 [52](#), [156](#), [157](#)
- [VFM07] VOMWEG, TW ; FABER, H ; MAYER, D: Verfahren und Computerprogrammprodukt zur Aufzeichnung biomedizinischer Bilder mit verminderter objektbewegungsbedingten Bildgebungsartefakten, Europäisches Patentamt, EP1844440. (2007) [61](#)
- [Vli01] VLIETSTRA, J: *Dictionary of Acronyms and Technical Abbreviations*. Springer, 2001 [45](#)
- [VTK04] VOMWEG, TW ; TEIFKE, A ; KUNZ, RP: Combination of low and high resolution sequences in two orientations for dynamic contrast-enhanced MRI of the breast: more than a compromise. In: *Eur Radiol* 14 (2004), Nr. 10, S. 1732–1742 [34](#), [57](#), [143](#)
- [VTK⁺05] VOMWEG, TW ; TEIFKE, A ; KAUCZOR, HU ; ACHENBACH, T ; RIEKER, O ; SCHREIBER, WG ; HEITMANN, KR ; BEIER, T ; THELEN, M: Selbstorganisierende neuronale Netze zur automatischen Detektion und Klassifikation von Kontrast(mittel)-verstärkten Läsionen in der dynamischen MR-Mammographie. In: *Rofo* 177 (2005), Nr. 5, S. 703–713 [156](#)
- [VTS02] VOMWEG, TW ; TEIFKE, A ; SCHREIBER, WG: Eine Kombination niedrig und hochauflösender dynamischer T1-gewichteter Sequenzen zur besseren Beurteilung der Morphologie Kontrastmittel aufnehmender Läsionen in der MRT der weiblichen Brust. In: *Fortschr. Röntgenstr.* 174 (2002), S. 1445–1449 [57](#)

- [Wei07] WEINHEIMER, O: *Über das Vermessen tubulärer Strukturen in der Computertomographie*, Universität Mainz, Diss., 2007 [68](#)
- [WFP92] WEIDNER, N ; FOLKMAN, J ; POZZA, F: Tumor angiogenesis: a new significant and independent prognostic indicator in early-stage breast carcinoma. In: *J Natl Cancer Inst* 84 (1992), Nr. 24, S. 1875–1887 [32](#)
- [WLM84] WEINMANN, HJ ; LANIADO, M ; MÜTZEL, W: Pharmacokinetics of GdDT-PA/dimeglumine after intravenous injection into healthy volunteers. In: *Physiol Chem Phys Med NMR* 16 (1984), Nr. 2, S. 167–172 [29](#)
- [WSK⁺06] WU, Q ; SALGANICOFFB, M ; KRISHNAN, A ; FUSSELL, DS ; M, MK: Interactive lesion segmentation on dynamic contrast enhanced breast MRI using a Markov model. In: *SPIE* Bd. 6144, 2006, S. 1487–1494 [51](#), [121](#), [128](#)
- [WSWF91] WEIDNER, N ; SEMPLE, JP ; WELCH, WR ; FOLKMAN, J: Tumor angiogenesis and metastasis–correlation in invasive breast carcinoma. In: *N Engl J Med* 324 (1991), Nr. 1, S. 1–8 [32](#)
- [ZBE⁺07] ZHENG, Y ; BALOCH, S ; ENGLANDER, S ; SCHNALL, MD ; SHEN, D: Segmentation and Classification of Breast Tumor Using Dynamic Contrast-Enhanced MR Images. In: *MICCAI (2)*, 2007, S. 393–401 [51](#), [121](#), [128](#)
- [Zhu07] ZHU, Yang-Ming: Mutual information-based registration of temporal and stereo retinal images using constrained optimization. In: *Comput. Methods Prog. Biomed.* 86 (2007), Nr. 3, S. 210–215 [82](#), [114](#)
- [ZJB⁺96] ZUO, CS ; JIANG, A ; BUFF, BL ; MAHON, TG ; WONG, TZ: Automatic motion correction for breast MR imaging. In: *Radiology* 198 (1996), Nr. 3, S. 903–906 [50](#)