



UNIVERSITÄT
KOBLENZ · LANDAU

Fachbereich 4: Informatik



Forschungsbereich

Life.Science Informatics (FIT.LIFE)

Entwicklung von Merkmalen zur Bestimmung räumlicher Ausbreitungsmuster in histopathologischen Gewebeschnitten des Mammakarzinoms

Diplomarbeit
zur Erlangung des Grades
DIPLOM-INFORMATIKER
im Studiengang Computervisualistik

vorgelegt von

Jan Bornemeier

Betreuer: Dipl.-Inform. S. Wirtz, Institut für Computervisualistik,
Fachbereich Informatik, Universität Koblenz-Landau

Betreuer: Dipl.-Inform. D. Zühlke, Life Science Informatics,
Fraunhofer-Institut für Angewandte Informationstechnik, Sankt
Augustin

Erstgutachter: Prof. Dr.-Ing. Dietrich Paulus, Institut für
Computervisualistik, Fachbereich Informatik, Universität
Koblenz-Landau

Zweitgutachter: Prof. Dr. Thomas Berlage, Life Science Informatics,
Fraunhofer-Institut für Angewandte Informationstechnik, Sankt
Augustin

Koblenz, im Dezember 2011

Kurzfassung

Die vorliegende Arbeit beschäftigt sich im Rahmen des Projektes ExPrimage mit der Entwicklung von Bildmerkmalen zur Bestimmung von Ausbreitungsmustern in digitalen histopathologischen Gewebeschnitten des Mammakarzinoms. Im Gegensatz zu anderen Anwendungen aus dem Gebiet der digitalen histopathologischen Bildanalyse wird in dieser Arbeit multimodales Bildmaterial der gesamten Gewebeebene für die Bildanalysen verwendet. Die entwickelten Analysen und Merkmale konzentrieren sich auf segmentierte Bildobjekte. Hierfür werden die vorverarbeitenden Schritte erweitert. Anhand von Anforderungen des Domänenexperten werden relevante Bildmerkmale recherchiert und implementiert.

Eingesetzt werden morphometrische Merkmale (Basis-Deskriptoren, Fourier-Deskriptoren und Moment Invariants) zur Bestimmung der Form und Größe des Tumors. Strukturelle Merkmale auf Basis von Graphen (Minimum-Spanning-Tree und Delaunay-Graph) werden zur Beschreibung der Verteilung des Tumors verwendet. Das RCC8-Histogramm wird zur Analyse räumlicher Zusammenhänge verschiedener Bildmodalitäten des Tumors auf Mehrfachrelationen, adaptive Schwellwerte und die Integration von Größeninformation erweitert. Zur Bestimmung von Koloalisationen in multimodalen Bildmaterial wird ein Ansatz auf Grundlage der Linearen Distanztransformation entwickelt.

Unter Verwendung von Clustervalidierungstechniken wird ein Konzept zur Validierung der Merkmale erstellt. Zudem werden die entwickelten Merkmale in ein überwachtes Lernverfahren integriert, um ihre Relevanz im Hinblick auf einen Zusammenhang mit dem Krankheitsverlauf der Patienten zu untersuchen.

Abstract

This diploma thesis is concerned with the development of image-descriptors for the detection of spreading-patterns in digital histopathological images of breast cancer. Contrary to other approaches in digital histopathological image-analysis this thesis deals with multimodal-images on the whole tissue-slide. The developed analyses are focused on segmented image-object-level. Preprocessing-steps are extended. Based on requirements given by the clinical expert relevant image-descriptors are researched and implemented.

Used are morphometric image-descriptors (Basic-Descriptors, Fourier-Descriptors and Moment Invariants) to characterize size and shape of the tumor. Descriptors based on graphs (minimum-spanning-tree and delaunay-graph) are used to describe the distribution of the tumor-tissue. The RCC8-Histogram is used to analyse relations between different image-modalities. In this work it is extended to multi-relations, adaptive thresholds and integration of size-information. Additionally a new approach to quantify colocalisations of different image-modalities is developed.

By using clustervalidation-techniques a concept to validate the image-descriptors is designed. Furthermore the image-descriptors are integrated in a supervised Learning-system to look for relevances related to the progress of the disease.

Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Die Vereinbarung der Arbeitsgruppe für Studien- und Abschlussarbeiten habe ich gelesen und anerkannt, insbesondere die Regelung des Nutzungsrechts.

Mit der Einstellung dieser Arbeit in die Bibliothek bin ich einverstanden. ja nein

Der Veröffentlichung dieser Arbeit im Internet stimme ich zu. ja nein

Koblenz, den 19. Dezember 2011

Danksagung

Mein Dank gilt in erster Linie meinen Betreuern Dietlind Zühlke (FIT) und Stefan Wirtz (Universität Koblenz-Landau), die mich jeder Zeit fachlich und menschlich unterstützt haben. Sie gaben mir Orientierung und Halt, wenn ich dachte beides zu verlieren. Dr. Gernoth Grunst (FIT) möchte ich für die konstruktiv kritischen Gespräche und die Motivation immer weiterzumachen danken.

Kerstin Röser (Universitätsklinikum Hamburg-Eppendorf) danke ich für die Begutachtung der Ergebnisse, die Erstellung der Beobachtungen bzgl. unterschiedlicher Ausbreitungsmuster und die konstruktiven Telefonkonferenzen.

Prof. Dr. Thomas Berlage (FIT) möchte ich für die Möglichkeit die Diplomarbeit am FIT durchführen zu können danken. Prof. Dr.-Ing. Dietrich Paulus (Universität Koblenz-Landau) danke ich für die fachliche Diskussion und die Möglichkeit meine Diplomarbeit extern durchführen zu dürfen.

Dem gesamten Team vom FIT und der LOCALITE GmbH danke ich für die freundliche Aufnahme und unermüdliche Unterstützung.

Meinen Freunden Ralf Lütje, Christian Buchner, Nils Hering und Matthias Häusler möchte ich für die menschliche Unterstützung und die Korrekturen danken.

Meinem Großvater Rudolf Schmidt danke ich für die finanzielle Unterstützung, die mir mein Studium erleichtert hat. Meiner Mutter Eva Bornemeier, meinem Vater Martin Bornemeier, meiner Schwester Antje Bornemeier (vielen Dank für die ausführlichen Korrekturen) und meinem Bruder Tim Bornemeier danke ich für den familiären Rückhalt.

Mein besonderer Dank gilt meiner Freundin Joanna Kosteczka, die mich stets bei Rückschlägen auffing und mich immer wieder aufbaute. Vielen Dank mein Engel.

Inhaltsverzeichnis

1	Einleitung und Motivation	21
2	Grundlagen des Maschinellen Lernens	27
2.1	Überwachtes Lernen - Heterogenous Learning Vector Quantization	28
2.2	Unüberwachtes Lernen - K-Means	30
2.3	Maschinelles Lernen in medizinischen Bildanalyzesystemen	33
2.3.1	Schwellwertverfahren von Otsu	34
2.3.2	Connected-Components	35
2.4	Methoden zur Evaluation und Validierung	35
2.4.1	Klassifikationsmaße	37
2.4.2	Hopkins-Index	37
2.4.3	Hauptkomponentenanalyse	39
2.4.4	Silhouetten-Wert	40
2.4.5	Normalized Variation of Information	41
3	Merkmale in der histopathologischen und medizinischen Bildanalyse	43
3.1	Morphometrische Merkmale	45
3.1.1	Fourier-Deskriptoren	46
3.1.2	Basis-Deskriptoren	48
3.1.3	Moment invariants	49
3.2	Strukturelle Merkmale	50
3.2.1	Minimum-Spanning-Tree	52
3.2.2	Delaunay-Graph	53
3.2.3	Graphmerkmale	54
3.3	Räumliche Relationen - RCC8-Histogramm	56
4	ExPrimage als Anwendungsszenario	61
4.1	Datenlage	62
4.2	Vorhandene Module und Abläufe	66
5	Merkmale und Bildanalysen zur Bestimmung von Ausbreitungsmustern auf Gewebeebene	71
5.1	Anforderungen	71

5.2	Erweiterung der vorverarbeitenden Analysen	73
5.2.1	Anwendung des Otsu-Verfahrens auf den Lab-Farbraum	78
5.2.2	Generierung der Bildobjekte	79
5.3	Morphometrische Analysen	80
5.3.1	Deskriptive morphometrische Statistiken	83
5.3.2	Clustering der morphometrischen Merkmale	84
5.4	Strukturelle Analysen	87
5.4.1	Knotenidentifikation	89
5.5	Relationale Analysen	92
5.5.1	Erweiterung des RCC8-Histogramms	93
5.5.2	Lineare Distanz-Quantifizierung	99
5.6	System zur Validierung und Evaluierung der Merkmale	102
5.6.1	Konzept	102
5.6.2	Erzeugung randomisierter synthetischer Daten	106
5.7	Implementation	108
5.7.1	Verwendete Technologien	108
5.7.2	Datenmodell	108
6	Experimente und Ergebnisse	113
6.1	Synthetische Daten - Clusteranalysen	113
6.1.1	Morphometrische Analyse - Form-Datenbank	114
6.1.2	Graphanalyse - DG	116
6.1.3	Graphanalyse - MST	118
6.1.4	Relationale Analyse - RCC8	120
6.1.5	Relationale Analyse - LDQ	121
6.2	Reale Daten - Clusteranalysen	124
6.2.1	Morphometrische Analysen - AE1AE3	124
6.2.2	Strukturelle Analysen - AE1AE3 - DG	127
6.2.3	Strukturelle Analysen - AE1AE3 - MST	127
6.2.4	Relationale Analysen - CD45	132
6.2.5	Relationale Analysen - ER	134
6.3	Reale Daten - Überwachtes Lernen	136
6.3.1	Morphometrische Analysen - AE1AE3	137
6.3.2	Strukturelle Analysen - AE1AE3 - DG	138
6.3.3	Strukturelle Analysen - AE1AE3 - MST	140
6.3.4	Relationale Analysen - CD45	141
6.3.5	Relationale Analysen - ER	142
6.3.6	Kombinationen mit klinischen Daten	143
7	Diskussion	147
8	Zusammenfassung und Ausblick	153
8.1	Zusammenfassung	153
8.2	Ausblick	157

<i>INHALTSVERZEICHNIS</i>	11
A Symbolverzeichnis	159
B Merkmalsauswahl für Testfälle	161
C Bilder	165
D PCA	175
E DVD-ROM	177
Literaturverzeichnis	178

Tabellenverzeichnis

2.1	Kontingenztafel	42
2.2	Beispielsergebnisse der normalized Variation of Information	42
5.1	Überblick zu verwendete Bildmodalitäten unterschiedliche Analyseschritte	80
5.2	Merkmalsauswahl zur Erfüllung der morphometrischen Anforderungen . .	81
5.3	Merkmalsauswahl zur Erfüllung der strukturellen Anforderungen	88
5.4	MST-Merkmalswerte von Tumorprototypen	91
5.5	DG-Merkmalswerte von Tumorprototypen	92
5.6	Merkmalsauswahl zur Erfüllung der relationalen Anforderungen	93
5.7	Erläuterungen zu verwendeten Schwellwerten zur Bestimmung der RCC8- Relationen.	96
6.1	Stichprobenanalyse Morphometrisch - AE1AE3	126
6.2	Stichprobenanalyse AE1AE3 - DG	129
6.3	Stichprobenanalyse AE1AE3 - MST	131
6.4	Stichprobenanalyse CD45	133
6.5	Stichprobenanalyse ER	136
6.6	Klassifikationsergebnisse morphometrisch AE1AE3	137
6.7	Klassifikationsergebnisse AE1AE3 - DG	139
6.8	Klassifikationsergebnisse AE1AE3 - MST	140
6.9	Klassifikationsergebnisse - CD45	142
6.10	Klassifikationsergebnisse ER	143
6.11	Klassifikationsergebnisse (Kombination mit klinischen Daten)	146
B.1	Testfälle DG	162
B.2	Testfälle MST	162
B.3	Testfälle RCC8 und LDQ	162
B.4	Testfälle Morphometrische Analysen	163

Abbildungsverzeichnis

1.1	Gewebeebene und Zellebene	22
2.1	Bilderkennungsablauf	33
2.2	N-Nachbarschaft	35
2.3	Berechnung des Hopkins-Index	39
2.4	Clustervalidierung mittels Silhouetten-Wert	41
3.1	Morphometrische Analysen auf Zellebene	45
3.2	Graphen auf Zellebene	52
3.3	Minimum-Spanning-Tree	53
3.4	Delaunay-Graph	54
3.5	Relationen des Region Connection Calculus (RCC8)	57
3.6	RCC8-Histogramm auf zellulärer Ebene	59
4.1	Entstehung der digitalen histologischen Schnitte im Projekt ExPrimage . .	63
4.2	Strukturelle Färbungen	64
4.3	Funktionelle Färbungen	65
4.4	Bildanalysen im Projekt ExPrimage	66
5.1	Tumorareale in SOM-Gewebecharakterisierung	74
5.2	Unterschiedliche Tumorstrukturen in AE1AE3	74
5.3	Beispiel unterschiedlicher Tumorrelationen	75
5.4	Tumorgewebedetektion in AE1AE3	77
5.5	Vorteil des Otsu3Lab	77
5.6	Detektion des Tumorgewebes in AE1AE3 mittels Otsu3Lab	78
5.7	Beispiel unterschiedlicher Bildobjekte	81
5.8	Anwendung der Fourier-Deskriptoren	83
5.9	Beispiel Moment Invariants	84
5.10	Schematischer Ablauf des Clusterings der morphometrischen Merkmale . .	86
5.11	Morphometrische Analysen (Ergebnisse ausgewählter Tumorprototypen) .	87
5.12	Knotenidentifikation für Graphberechnungen	89
5.13	Strukturelle Analysen (Ergebnisse ausgewählter Tumorprototypen (MST))	90
5.14	Strukturelle Analysen (Ergebnisse ausgewählter Tumorprototypen (DG)) .	92
5.15	Ablauf zur Bestimmung einer RCC8-Relation	97

5.16	Relationale Analysen (Ergebnisse ausgewählter Tumorprototypen (RCC8))	99
5.17	Ablauf der Linearen Distanz-Quantisierung	100
5.18	Periphere Regionen innerhalb zentraler Aushöhlungen bei LDQ	101
5.19	Relationale Analysen (Ergebnisse ausgewählter Tumorprototypen (LDQ))	101
5.20	Konzept zur Validierung der Merkmale	103
5.21	Randomisierte synthetische Daten	106
5.22	Klassendiagramm Package: FeatureComputation	109
6.1	Hopkins-Index und PCA-Plot (Form-Datenbank)	115
6.2	Interne und externe Clustervalidierung - Form-Datenbank	116
6.4	Interne und externe Clustervalidierung (Synthetisch DG)	117
6.3	Hopkins-Index und PCA-Plot (Synthetisch DG)	117
6.5	Beispielergebnisse synthetischer Graph-Analysen (DG)	118
6.6	Hopkins-Index und PCA-Plot (Synthetisch MST)	119
6.7	Interne und externe Clustervalidierung (Synthetisch MST)	119
6.8	Beispielergebnisse synthetischer Graph-Analysen (MST)	120
6.9	Hopkins-Index und PCA-Plot (Synthetisch RCC8)	121
6.10	Interne und externe Clustervalidierung (Synthetisch RCC8)	121
6.11	Beispielergebnisse synthetischer RCC8-Analysen	122
6.12	Hopkins-Index und PCA-Plot (Synthetisch LDQ)	122
6.13	Interne und externe Clustervalidierung (Synthetisch LDQ)	123
6.14	Hopkins-Index und PCA-Plot (Morphometrisch - AE1AE3)	125
6.15	Interne und externe Clustervalidierung (Morphometrisch)	126
6.16	Hopkins-Index und PCA-Plot (DG - AE1AE3)	128
6.17	Interne und externe Clustervalidierung (DG - AE1AE3)	128
6.18	Hopkins-Index und PCA-Plot (AE1AE3 - MST)	130
6.19	Interne und externe Clustervalidierung (DG - AE1AE3)	131
6.20	Hopkins-Index und PCA-Plot (Relationen - CD45)	132
6.21	Interne und externe Clustervalidierung (Relationen - AE1AE3)	133
6.22	Hopkins-Index und PCA-Plot (Relationen - ER)	134
6.23	Interne und externe Clustervalidierung (Relationen - ER)	135
6.24	Bestes Klassifikationsergebnis morphometrisch - AE1AE3	138
6.25	Klassifikationsergebnis AE1AE3 - DG	139
6.26	Bestes Klassifikationsergebnis AE1AE3 - MST	141
6.27	Bestes Klassifikationsergebnis CD45	142
6.28	Bestes Klassifikationsergebnis ER	143
6.29	Vergleich kID ohne und mit implementierten Merkmalen	145
C.1	Form-Datenbank	166
C.2	Visualisierung morphometrisches Clusterergebnis auf Form-Datenbank	167
C.3	Beispielbilder Morphometrisches Clustering (SOM)	168
C.4	Beispielbilder Morphometrisches Clustering (AE1AE3)	169
C.5	Beispielbilder Graphen (AE1AE3)	170
C.6	Beispielbilder Clustering der Graph-Subregionen	171

C.7 RCC8-Erweiterung synthetisches Beispiel 172

D.1 Dimensionsreduktion durch PCA 176

Algorithmusverzeichnis

1	K-Means	32
2	Connected-Components	35
3	Algorithmus von Prim R.	53
4	Bestimmung der RCC8-Relationen nach Boucheron [Bou08].	58
5	Erweiterung des Algorithmus von [Bou08] zur Bestimmung des RCC8-Histogramms	95

Kapitel 1

Einleitung und Motivation

Die Gefahr für eine Frau, im Laufe ihres Lebens an Brustkrebs zu erkranken, beträgt in den westlichen Ländern ca. 12% und ist unter den Krebserkrankungen die zweit häufigste Todesursache [Ame11]. Allein in Deutschland gibt es pro Jahr ca. 50.000 Neuerkrankungen, wobei die Mortalität bei ca. 18.000 Frauen liegt [RP05]. Trotz großer Erfolge in der Intervention des Mammakarzinoms in den letzten Jahren gibt es immer wieder, laut Diagnostik ähnliche Fälle, bei denen sich bei gleichen Therapien andere Krankheitsverläufe beobachten lassen.

Mit der Frage einer individuelleren Diagnostik und Verlaufsabschätzung beschäftigt sich der Forscherverbund ExPrimage, in dessen Rahmen diese Diplomarbeit am Fraunhofer Institut für Angewandte Informationstechnik (FIT) im Forschungsbereich Life Science Informatics durchgeführt wird. Dieser Forscherverbund versucht pathologische und mikrobiologische Hypothesen sowie objektive Beobachtungen in Bezug auf Brustkrebs mit Hilfe quantitativer Analysen zu stützen und neue Zusammenhänge zu erkennen. Zusammengesetzt aus Informationen verschiedener Ebenen soll ein *Multilayered-Model*¹ erstellt werden, das im Zusammenhang mit dem Patientenverlauf Aufschlüsse über eine individuelle Prognose gibt.

Das FIT beschäftigt sich in diesem Rahmen mit der Bildanalyse von digitalisierten histopathologischen² Gewebeschnitten des Mammakarzinoms auf Gewebeebene (siehe Abbildung 1.1), um eine individuellere quantitative Charakterisierung der Tumore zu erhalten. Auf Grundlage der Bildanalysen und klinischer Parameter der Patienten wird versucht, relevante Zusammenhänge mit dem Krankheitsverlauf zu finden. Eingesetzt werden hierzu Verfahren des Maschinellen Lernens (siehe Kapitel 2), welche die Bildanalysen mit klinischen Parametern der Patienten zusammenbringen. Elementarer Bestandteil der Verfahren sind quantitative Bildanalysen zur Merkmalsextraktion [Han09], um die Eigenschaften pathologisch relevanter Bildstrukturen und Zusammenhänge in kondensierter, für die maschinelle Weiterverarbeitung nutzbarer Form zu erfassen.

¹Ein Hybridmodell des Tumors, das unterschiedliche Aspekte erfasst

²Histopathologie: Mikroskopische Krankheitsdiagnostik durch Analysen von gefärbten Gewebeschnitten.

Die Domänenexperten³ sehen in den Ausbreitungscharakteristiken der Tumore auf Gewebeebene Informationen, welche prinzipiell relevante Hinweise zur Prognose von Krankheitsverläufen enthalten, und nennen diesbezüglich verschiedene Hypothesen. Es hat sich herausgestellt, dass grundlegende, bereits durchgeführte quantitative Bildanalysen, wie beispielsweise Gewebestatistiken, nicht ausreichen, um die gesuchten Eigenschaften des Tumors zu erfassen. Aufbauend auf den Hypothesen der Domänenexperten in Form von Anforderungen sollen adäquate Bildmerkmale höherer Semantik gefunden und eingesetzt werden. Eine Herausforderung liegt unter anderem darin die *Semantische Lücke*⁴ (engl. „Semantic Gap“)[LZLM07] zu beachten und zu verringern, um die natürlich sprachlich formulierten Anforderungen des Domänenexperten in adäquate Bildanalysen und Bildmerkmale zu übertragen.

Ziel der vorliegenden Arbeit ist es, den bestehenden Bildanalyseverlauf auf Gewebeebene zu erweitern, wozu neben der Weiterentwicklung der Vorverarbeitung vor allem das Recherchieren und Implementieren relevanter Bildmerkmale gehört, um die Ausbreitungs- und Verteilungsmuster des Tumors auf Gewebeebene zu erfassen und zu quantifizieren. Damit soll ein Beitrag zur Überprüfung der Hypothesen der Domänenexperten und zur umfassenderen Beschreibung der Tumore geleistet werden.

In [Nie03] wird, entsprechend den Einschränkungen durch die Semantische Lücke, das automatisierte systematische Auffinden von charakteristischen Bildmerkmalen als ungelöstes Problem beschrieben. Vielmehr basiert die Auswahl geeigneter Bildmerkmale in der Regel auf der Intuition des Anwenders:

„Usually the researcher has some notion about what features should be useful.“ [RB03]

Von daher kann die Auswahl nur auf Basis von Literaturrecherchen und Erfahrungen aus anderen Studien aufgebaut werden, welche mit den Hypothesen der Domänenexperten in Verbindung gebracht werden müssen.

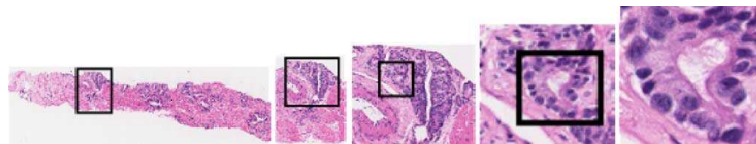


Abbildung 1.1: Unterschiedliche Auflösungsstufen von histologischen Schnitten (hier HE-Färbung). Links: Gewebeebene; Rechts: Zellebene (Quelle:[GBC⁺09]).

Die meisten Bildanalysesysteme in der digitalen Histopathologie arbeiten - wie der Domänenexperte - fast ausschließlich in ausgewählten Bereichen auf der zellulären Ebene (siehe Abbildung 1.1) und verwenden in der Praxis meist nur unimodales Bildmaterial (HE-Färbung (Hämatoxylin-Eosin-Färbung), siehe Abschnitt 4.1) [DY05, DHS⁺07, DAM⁺08, GBC⁺09, ACSGD10, Bou08]. Nur wenige Systeme verwenden Informationen und Bildmerkmale auf Basis des gesamten Gewebeschnittes.

³Pathologen und Mikrobiologen

⁴Anforderungen in Form von natürlicher Sprache welche in eine formale und reproduzierbare, weniger mächtigen Sprache (in diesem Fall Programmiersprache) übertragen werden.

In [DAB⁺04] werden quantitative Analysen auf der kompletten Gewebeebene angewendet, um gesundes Gewebe, Tumorstroma⁵ und Krebsgewebe auf histologischen Schnitten des Prostatakarzinoms zu unterscheiden. Unter Verwendung von unterschiedlichen Auflösungsstufen des Gewebeschnittes wird anhand von Texturmerkmalen (siehe Kapitel 3) durch Klassifikation ein automatisches Grading⁶ für den gesamten Gewebebereich durchgeführt. Bei den genannten Ansätzen wird ausschließlich eine Färbung⁷ berücksichtigt.

Dass der komplette Gewebeschnitt und unterschiedliche Färbungen pathologisch relevant sind, zeigt der Ansatz von Braumann et. al. [BEH⁺06], der versucht, verschiedene histologische Färbungen unter Verwendung von Registriertechniken in 3D zu rekonstruieren, um das Vordringen des Tumors in gesundes Gewebe darzustellen. In [GBC⁺09] wird in der Schlussbetrachtung die Wichtigkeit und der Mehrwert der Verwendung multimodaler Färbungen in Form von Informationsverdichtung betont. Im Rahmen des Projektes ExPrimage steht multimodales Bildmaterial (siehe Abschnitt 4.1) für jeden Patienten zur Verfügung, welches abhängig von der beinhalteten Information in den einzusetzenden Bildanalysen der vorliegenden Arbeit ebenfalls berücksichtigt werden muss.

Parallel zu der vorliegenden Diplomarbeit wird im ExPrimage-Kontext eine weitere Arbeit durchgeführt, welche die Textur innerhalb von Tumorregionen auf Gewebeebene [Kha11] untersucht.

Die vorliegende Arbeit beschränkt sich daher auf die Suche nach Merkmalen, die Ausbreitungsmuster auf segmentierten Bilddaten beschreiben. Auf eine grobe Gewebecharakterisierung durch die Self-Organizing-Map [KSH01] (SOM), die im Projekt ExPrimage unabhängig von der vorliegenden Arbeit erstellt wurde, kann zurückgegriffen werden. Zur Detailanalyse einzelner Gewebeschnitte müssen, ergänzend zur SOM, neue Verfahren gefunden werden. Im Rahmen der Erhebung verschiedener Merkmalsdaten ist es des Weiteren notwendig, ein geeignetes und möglichst flexibles Datenmodell zu finden, um auf sich ändernde Anforderungen reagieren zu können.

Merkmale, die bereits auf zellulärer Ebene erfolgreich eingesetzt werden, sollen für die Gewebeebene adaptiert und im Hinblick auf ihren diskriminierenden Charakter in dieser Auflösungsstufe untersucht werden. In Frage kommen dazu *morphometrische Merkmale* [RB03], welche versuchen, Größeneigenschaften und die Form einzelner Bildobjekte zu beschreiben [Bou08, DY05, GBC⁺09, NA08]. Des Weiteren *strukturelle Merkmale* [DY05] basierend auf Graphen [Bou08, DY05, GBC⁺09, ACSGD10, DHS⁺07, DAM⁺08] und ein Ansatz zur Bestimmung *räumlicher Relationen* [Bou08] auf Grundlage des *Region Connection Calculus 8* [RCC92] (siehe Kapitel 3).

Eine grundlegende Schwierigkeit bei den Analysen liegt in der fehlenden Ground-Truth bezüglich der zu untersuchenden Eigenschaften des Tumors auf Gewebeebene. Dieses Problem wird in [Bou08] als „*lack of ground truth*“ bezeichnet und lässt sich allgemein auf die Bildanalysen von histopathologischen Gewebeschnitten übertragen: „... *a true gold*

⁵Den Tumor versorgendes Gewebe

⁶Beurteilung der Malignität des Tumors

⁷Histologischer Schnitt mit markerspezifisch angefärbten Strukturen

standard for pathology specimens is not available in the same sense that it may be for other datasets ...“ [Bou08].

Aus diesem Grund müssen Wege für eine Validierung der entwickelten Merkmale gefunden werden. Hierzu werden die implementierten und entwickelten Merkmale zunächst anhand von synthetischen Daten auf ihre grundsätzliche Funktionsfähigkeit getestet. Im Verlauf des Projektes ExPrimage hat sich gezeigt, dass eine Einteilung der Tumore anhand spezifischer Eigenschaften durch Domänenexperten keine ausreichende Reliabilität erreicht, sondern erst ein maschineller Vorschlag ein valides *Post-Labeling*⁸ der Fälle durch den Domänenexperten ermöglicht. Getestet wird daher in einem ersten Schritt, inwieweit sich durch unüberwachtes Lernen bzw. einer Clusteranalyse (siehe Abschnitt 2.2), dem *K-Means*-Verfahren [XW05, JMF99], bereits diskriminierende, potenziell relevante Gruppen von Tumoren anhand der angewendeten Bildmerkmale darstellen.

Zur Validierung der Clusteranalysen muss zunächst ein geeignetes Konzept und geeignete Methoden gefunden werden. Da es sich bei der Clustervalidierung um ein mathematisch nicht korrekt gestelltes Problem handelt [Züh12], kann hier nur eine Kombination von Verfahren und Messungen Aufschluss über die Qualität geben. Zum Einsatz kommen in der vorliegenden Arbeit von daher verschiedene Clustervalidierungsverfahren (siehe Abschnitt 2.4), um die Objektivität der Ergebnisse zu gewährleisten. Die Ergebnisse werden zudem anhand von Stichproben mit den Beobachtungen des Domänenexperten (siehe Abschnitt 4.1) verglichen.

Des Weiteren werden die eingesetzten Merkmale daraufhin untersucht, ob sie Hinweise darüber geben, ob ein Patient überlebt, wieder erkrankt oder verstorben ist. Zur Validierung wird als Klassenzugehörigkeit (siehe Kapitel 2) der *Follow-up-Status*⁹ verwendet.

Dafür werden zunächst die entstandenen Cluster im Hinblick auf eine Korrelation mit dem *Follow-up-Status*, mit Hilfe der normalisierten *Variation Of Information* (siehe Abschnitt 2.4.5) untersucht. In einem zweiten Schritt werden die entwickelten Bildmerkmale mit den aussichtsreichsten Ergebnissen in ein überwachtes Lernverfahren (siehe Abschnitt 2.1), die *Heterogenous Vector Quantification* (kurz HLVQ) [Züh12] (siehe Abschnitt 2.1), integriert. Dieses Verfahren bewertet die Relevanz der eingesetzten Merkmale im Hinblick auf den Verlauf (*Follow-up-Status*) und bringt sie mit klinischen Patientendaten in Verbindung.

Anhand der *Klassifikationsrate*, der *Genauigkeit* und der *Trefferquote* (siehe Abschnitt 2.4) wird getestet, inwieweit die entwickelten Merkmale einen positiven Einfluss auf das Klassifikationsergebnis haben.

Die Gliederung der vorliegenden Arbeit ist so konzipiert, dass schrittweise in unterschiedlichen Detailstufen auf die behandelte Problematik hingeführt wird. Kapitel 2 beschreibt zunächst die Grundlagen des Maschinellen Lernens und gibt grundsätzliche Definitionen vor. In diesem Kontext werden die eingesetzten Lernverfahren und Validierungstechniken erläutert. Die vorliegende Arbeit wird an dieser Stelle in den Rahmen

⁸nachträgliche Klassifikation der Fälle anhand von computergestützt gefundener Prototypen

⁹Der *Follow-up-Status* gibt an, ob eine Frau fünf Jahre nach der chirurgischen Entfernung des Tumors gesund, wieder an Brustkrebs erkrankt oder verstorben ist.

maschineller Bildanalyssysteme eingeordnet. Kapitel 3 gibt einen Überblick zu aktuell in der Bildanalyse medizinischer und insbesondere histopathologischer Bilder eingesetzten Merkmalen. Im Verlauf dieser Ausführungen werden die verwendeten und implementierten Bildmerkmale beschrieben. Kapitel 4 beschreibt das Anwendungsszenario im Kontext des Projekts ExPrimage. Zunächst wird auf die vorhandene Datenlage und Prozessschritte eingegangen, um die Ausgangslage für die eingesetzten Analysen zu beschreiben. Kapitel 5 beschreibt die im Rahmen der vorliegenden Arbeit angewendeten Bildanalysen. Es werden zunächst Anforderungen definiert und daraufhin wird auf die entwickelten Bildanalysen zur Merkmalsextraktion eingegangen. Erweiterungen der Bildmerkmale und selbst entwickelte Ansätze werden vorgestellt. Das entwickelte Konzept zur Validierung der Ergebnisse wird dargestellt. Abschließend wird auf das entwickelte Datenmodell und verwendete Bibliotheken eingegangen. In Kapitel 6 werden die durchgeführten Experimente und Ergebnisse präsentiert. Zunächst werden die implementierten Abläufe und Merkmale an synthetischen und daraufhin an realen Datensätzen getestet. Diskutiert werden die Ergebnisse in Kapitel 7. Abschließend wird in Kapitel 8 eine Zusammenfassung und ein Ausblick gegeben.

Kapitel 2

Grundlagen des Maschinellen Lernens

Dieses Kapitel beschreibt die Grundlagen des Maschinellen Lernens und die in diesem Kontext in der vorliegenden Arbeit eingesetzten Methoden. Zunächst werden grundlegende Definitionen eingeführt. In Abschnitt 2.1 wird auf überwachte Lernverfahren und das in dieser Arbeit eingesetzte Verfahren *Heterogenous Learning Vector Quantization* eingegangen. In Abschnitt 2.2 wird das Vorgehen bei unüberwachten Lernverfahren sowie das verwendete *K-Means*-Verfahren erläutert. Es folgt ein Überblick zu grundlegenden Verarbeitungsschritten in maschinellen Bildanalyzesystemen. In diesem Kontext sollen die Analyseschritte der vorliegenden Arbeit eingeordnet und das verwendete *Schwellwertverfahren von Otsu* sowie die Bestimmung von *Connected-Components* vorgestellt werden. Eingesetzte Methoden zur Evaluation und Validierung der verwendeten Maschinellen Lernverfahren werden in Abschnitt 2.4 erörtert.

Maschinelles Lernen beschäftigt sich mit der automatischen algorithmischen Analyse von Daten mit dem Ziel, Regelmäßigkeiten (Muster) in diesen Daten zu erkennen und diese bestimmten Kategorien zuzuordnen [Bis06]. Als Muster bzw. nach Niemann [Nie03] komplexes Muster, wird im Kontext des Projektes ExPrimage die Gesamtsituation eines Patientenfalles gesehen. Ein komplexes Muster setzt sich aus verschiedenen einfachen Mustern zusammen, wozu im Projekt ExPrimage neben Bilddaten auch klinische Daten (siehe Abschnitt 4.1) zählen.

Die vorliegende Arbeit beschäftigt sich primär mit der Analyse von einzelnen Bilddaten, die nach Niemann [Nie03] als einfache Muster bezeichnet werden können. Im Rahmen der vorliegenden Arbeit wird ein einfaches Muster als allgemeine Bildfunktion f mit den Pixelkoordinaten x bzw. y als $f(x, y)$ bezeichnet. Durch f_{ζ}^{ϱ} wird eine differenziertere Beschreibung der Bildfunktion verwendet, wobei ϱ die Bildmodalität (siehe Abschnitt 4.1) und ζ den Bildfunktionstyp (z.B. Binärbild (*bin*), Grauwertbild (*gray*) oder RGB-Farbbild (*rgb*)) angibt.

Im Maschinellen Lernen werden Muster durch n Merkmalswerte m in einem n -dimensionalen Vektor

$$\mathbf{m} = \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} \in \mathbb{W}_M \quad (2.1)$$

beschrieben [Han09]. Die Dimension des Merkmalsvektors \mathbf{m} ist wesentlich geringer als die des Musters. Die Merkmale M_1, \dots, M_n mit den Wertebereichen W_{M_1}, \dots, W_{M_n} reduzieren die Information auf die zu untersuchenden Eigenschaften des Musters. Der Wertebereich der Merkmalsvektoren \mathbb{W}_M wird als

$$W_M = W_{M_1} \times \dots \times W_{M_n} \quad (2.2)$$

definiert [Han09]. Maschinelle Lernverfahren versuchen anhand einer vorgegebenen Menge von Mustern bzw. deren extrahierter Merkmalsvektoren einem Trainingsdatensatz k unterschiedlichen Kategorien bzw. Klassen $\Omega_1, \dots, \Omega_k$ zuzuordnen. Die Klassenzugehörigkeiten ω (auch Label genannt) des Trainingsdatensatzes bzw. einer Stichprobe $S = \{(\mathbf{m}, \omega) \in W_M \times \Omega\}$ bekannt [Han09]. Die Menge $\Omega = \{\Omega_1, \dots, \Omega_k\}$ wird Klassenmenge genannt. Die Stichprobe S wird verwendet, um die Parameter eines adaptiven Modells während der Trainingsphase (auch Lernphase genannt) anzupassen. In [Bis06] wird das Ergebnis der Ausführung eines maschinellen Lernverfahrens als Funktion $\delta(\mathbf{m})$ beschrieben, welche einem neuen Merkmalsvektor \mathbf{m} , der nicht zur Stichprobe gehört, automatisch einer Klasse Ω zuweist. Die Funktion $\delta(\mathbf{m})$ wird im Kontext von überwachten Lernverfahren (siehe Abschnitt 2.1) auch als Klassifikator bezeichnet [Nie03].

Die Merkmalsvektoren, welche nicht zur Stichprobe gehören, bilden den Testdatensatz. Die Funktion $\delta(\mathbf{m})$ wird während der Lernphase bestimmt. Die Fähigkeit neue Muster, die sich von bekannten Mustern (aus der Stichprobe) unterscheiden, korrekt einer Klasse zuzuordnen, wird Generalisierung genannt [Bis06]. Die Zuordnung neuer Muster anhand von $\delta(\mathbf{m})$ wird als Klassifikation bezeichnet [Nie03]. Da in der Praxis meistens nur eine kleine Untermenge aller möglichen Muster für das Lernen zur Verfügung steht und somit lediglich ein kleiner Teil aller möglichen Musterausprägungen erfasst werden kann, ist ein zentrales Ziel von maschinellen Lernverfahren die Generalisierung [Bis06]. Unterschieden werden können grundsätzlich zwei Ansätze des Maschinellen Lernens: überwachtes und unüberwachtes Lernen [Nie03]. In den folgenden Abschnitten wird im Hinblick auf die Anwendung in der vorliegenden Arbeit auf beide Ansätze und die eingesetzten Verfahren eingegangen.

2.1 Überwachtes Lernen - Heterogenous Learning Vector Quantization

Verfahren, die dem überwachten Lernen zugeschrieben werden, verwenden Stichproben S mit bereits klassifizierten Mustern, wie im vorigen Abschnitt beschrieben. Es ist also

Vorwissen bekannt anhand dessen Klassengrenzen mit Hilfe der Funktion $\delta(\mathbf{m})$ justiert werden können.

Das in der vorliegenden Arbeit angewendete Verfahren, die *Heterogenous Learning Vector Quantization* (kurz HLVQ) [Züh12], gehört zur Kategorie der überwachten Lernverfahren. Die Funktionsweise des Verfahrens soll im folgenden Abschnitt erläutert werden.

Heterogenous Learning Vector Quantization Die folgenden Ausführungen basieren, sofern nicht anders angegeben, auf [Züh12] und [Sch10].

In der Vektorquantifizierung wird die Klassifikation eines unbekanntes Musters durch eine *winner-takes-all* Entscheidung getroffen. Dem neuen unbekanntes Merkmalsvektor \mathbf{m} wird die Klasse Ω des, im Sinne eines definierten Abstandsmaßes $d(\mathbf{m}, \mathbf{w})$, nächsten Prototyps \mathbf{w} zugeordnet. Prototypen sind Vektoren im Merkmalsraum, welche die Charakteristika der Daten in ihrer Nachbarschaft widerspiegeln. Ihnen wird bei der Initialisierung eine Klassenzugehörigkeit $\omega_L = \delta(\mathbf{w}_L)$ zugewiesen.

Das Lernen erfolgt durch die iterative Anpassung der Prototypen an den Merkmalsraum der Stichprobe. Die Anzahl der Prototypen ist ein vorher zu definierender Parameter. Sie muss mindestens der Klassenanzahl k entsprechen. Die Lernphase des HLVQ-Verfahrens verläuft iterativ und minimiert die Kostenfunktion

$$E_{ob-GLVQ} = \sum_{n=1}^N \sum_{k=1}^K \Phi(\mu_k), \mu_k = \frac{D^+(\mathbf{w}_n, \mathbf{m}_k) - D^-(\mathbf{w}_n, \mathbf{m}_k)}{D^+(\mathbf{w}_n, \mathbf{m}_k) + D^-(\mathbf{w}_n, \mathbf{m}_k)} \quad (2.3)$$

über K Stichprobenvektoren und N Prototypen. Der Term μ_k wird relative Distanzdifferenz genannt. Er beschreibt sowohl den Abstand des betrachteten Stichprobenvektors zum nächsten Prototypen mit gleicher Klassenzugehörigkeit D^+ als auch den Abstand zum nächsten Prototypen mit nicht übereinstimmender Klassenzugehörigkeit D^- . Durch die Verwendung der Funktion Φ (z.B. Sigmoid-Funktion [SY96]) bekommt jedes μ_k , in Abhängigkeit des betrachteten Stichprobenvektors, ein individuelles Gewicht [SY96]. Die Distanzen D werden im HLVQ-Verfahren durch

$$D(\mathbf{w}_n, \mathbf{m}_k) = \sum_{j=1}^J \alpha_j d_j(\mathbf{w}_n, \mathbf{m}_k) \quad (2.4)$$

definiert. Das HLVQ-Verfahren unterteilt einen Merkmalsvektor in j unterschiedliche Merkmalsuntergruppen, welche mit den für diese Gruppen mathematisch adäquaten Abstandsmaßen d gemessen werden. Die Merkmalsuntergruppen werden entsprechend der zu verwendenden Metrik vor dem Lernprozess annotiert. Dabei gibt $\alpha \in \{0, 1\}$ die Gewichtung und Relevanz für die einzelnen Merkmalsuntergruppen an. Die Aktualisierung der Prototypen wird durch

$$w^+ \leftarrow w^+ + \Delta_{w^+}, \Delta_{w^+} = \epsilon_w \frac{\partial E_{vb-GLVQ}}{\partial w^+} \quad (2.5)$$

$$w^- \leftarrow w^- - \Delta_{w^-}, \Delta_{w^-} = \epsilon_w \frac{\partial E_{vb-GLVQ}}{\partial w^-} \quad (2.6)$$

bestimmt. Der inkorrekte Prototyp w^- stößt sich von der Stichprobe ab, während der korrekte Prototyp w^+ zur Stichprobe hingezogen wird. Der Lernprozess kann als Gradientenabstieg zu einem globalen Minimum verstanden werden. Die Lernrate ϵ_w gibt die Größe der Anpassungsschritte bzw. den Bruchteil der Länge des Differenzvektors an. Die Gewichtung α_j für jede Merkmalsuntergruppe wird während des Lernens ebenfalls mit Lernrate $\epsilon_\alpha \ll \epsilon_w$ angepasst. Die Lernraten werden im Laufe der Epochen verringert. Eine Epoche entspricht einem kompletten Durchlauf über alle Stichprobenelemente. Anhand eines Beispiels soll die Unterteilung in mehrere Merkmalsuntergruppen j deutlich werden: Der Beispielvektor

$$\mathbf{m}_k = \begin{pmatrix} \mu_{j=1} \\ \sigma_{j=1} \\ - \\ 15_{j=2} \\ - \\ 0.3_{j=3} \\ 0.7_{j=3} \\ - \\ \dots \end{pmatrix} \quad (2.7)$$

enthält drei unterschiedliche Merkmalsuntergruppen. Merkmalsuntergruppe $j = 1$ repräsentiert eine Normalverteilung, definiert durch Mittelwert μ und Standardabweichung σ . Das Merkmal $j = 2$ wird mit euklidischer Metrik gemessen. Merkmalsuntergruppe $j = 3$ repräsentiert eine Verteilung, welche durch Divergenzen gemessen werden.

Mit dem *HLVQ*-Verfahren ist es möglich, die relationalen klinischen Daten (siehe Abschnitt 4.1) der Patienten mit Merkmalen unterschiedlichster Metriken aus der Bildanalyse mathematisch korrekt zusammenzuführen und zur Klassifikation des komplexen Musters „Patientenfall“ im Bezug auf den Verlauf zu nutzen.

2.2 Unüberwachtes Lernen - K-Means

Das unüberwachte Lernen, auch *Clustering* oder *Clusteranalyse* genannt [JMF99, XW05], beschäftigt sich mit der Einteilung von Mustern f in Klassen, ohne dass eine Stichprobe S mit bereits klassifizierten Mustern vorliegt. Klasseneinteilungen in Form von Ω müssen in diesem Fall erst gefunden werden. Im Bereich der Clusteranalyse werden diese Einteilungen als Cluster C bezeichnet [XW05]. Clusteranalysen verfolgen das Ziel, eine endliche unklassifizierte Menge von Mustern in mehr oder weniger homogene, disjunkte Untergruppen (Cluster) zu strukturieren [XW05]. In der Literatur werden Cluster als intern

homogen und extern separiert beschrieben [XW05]. Zur Bestimmung von Clustern muss ein Distanzkriterium definiert werden. Die am weitesten verbreiteten Kriterien basieren auf dem quadratischen Fehler $d_e(\mathbf{m}_x, \mathbf{m}_y)^2$ unter Verwendung der euklidischen Distanz

$$d_e(\mathbf{m}_x, \mathbf{m}_y) = \left(\sum_{i=1}^n \|\mathbf{m}_x^i - \mathbf{m}_y^i\|^2 \right)^{1/2} \quad (2.8)$$

zwischen zwei Merkmalsvektoren \mathbf{m}_x und \mathbf{m}_y mit den Merkmalswerten i [XW05]. Clustering wird in Anwendungen eingesetzt, an denen eine klar klassifizierte Stichprobe nicht zur Verfügung steht, der Zusammenhang der Daten weitestgehend unbekannt ist oder die Datenlage zu umfangreich ist, um alle Aspekte menschlich objektiv einbeziehen zu können [Run10].

So werden in [GLB⁺10] beispielsweise auf Grundlage der Daten zur genetischen Expression durch eine Clusteranalyse 17 relevante Gruppierungen von Brustkrebstumoren identifiziert. Nach [JMF99] können Cluster-Algorithmen in hierarchische und partitionierende Algorithmen unterteilt werden. Als Partition wird das Ergebnis der Clusteranalyse, die Zuordnung der Muster zu Clustern, bezeichnet. Hierarchische Methoden erzeugen verschiedene vernetzte Serien von Partitionen während partitionierende Methoden nur eine Partition unter Angabe der Clusteranzahl K erzeugen. Das in der vorliegenden Arbeit eingesetzte Clusterverfahren *K-Means* gehört zu den partitionierenden Verfahren und soll im folgenden Abschnitt vorgestellt werden.

K-Means Die Motivation, das *K-Means*-Verfahren zu verwenden, resultiert aus seiner weiten Verbreitung und Anwendung in realen Applikationen sowie seiner geringen Zeit- und Speicherkomplexität [JMF99, XW05]. Die Cluster werden durch Prototypenvektoren ($\boldsymbol{\mu}$, siehe folgender Abschnitt) repräsentiert. Somit können zudem potenziell generalisierende typische Muster in der Nähe der Prototypenvektoren identifiziert werden, welche als Diskussionsgrundlage für die Domänenexperten herangezogen werden (siehe Abschnitt 5.6.1).

Die Clusteranzahl K muss initial vorgegeben werden. Im *K-Means*-Verfahren werden Clusterzentren bzw. Prototypen $\boldsymbol{\mu}$ für jedes Cluster C_k bestimmt wodurch die Cluster definiert werden. Ziel des *K-Means*-Verfahren ist das Auffinden dieser Clusterzentren und die Zuordnung von N Merkmalsvektors \mathbf{m} zu einem der Clusterzentren, sodass die Summe der quadratischen Distanzen zum nächsten Clusterzentrum minimal wird. Das Kriterium wird nach [XW05] definiert als

$$J(\mathbf{\Gamma}, \mathbf{M}) = \sum_{k=1}^K \sum_{n=1}^N \gamma_{kn} \cdot \|\boldsymbol{\mu}_k - \mathbf{m}_n\|^2 \quad (2.9)$$

wobei $\mathbf{\Gamma}$ die Partitionsmatrix mit

$$\gamma_{kn} = \begin{cases} 1 & \text{wenn } \mathbf{m}_n \in C_k \\ 0 & \text{sonst} \end{cases} \quad \text{mit } \sum_{k=1}^K \gamma_{kq} = 1 \quad \forall n \quad (2.10)$$

angibt. \mathbf{M} ist die Clusterzentrenmatrix mit $\mathbf{M} = [\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_K]$, wobei $\boldsymbol{\mu}_k$ das Clusterzentrum des k -ten Clusters ist und durch

$$\boldsymbol{\mu}_k = \frac{1}{N_k} \cdot \sum_{n=1}^N \gamma_{kn} \cdot \mathbf{m}_n \quad (2.11)$$

definiert wird. N_k ist die Anzahl von Merkmalsvektoren, die zu Cluster C_k gehören. Der Algorithmus zur automatischen Bestimmung der korrekten Partition im Sinne des definierten Kriteriums (2.9) wird nach [XW05] und [Run10] wie folgt notiert:

1. Initialisiere Clusterzentren $\boldsymbol{\mu}_k$ und bestimme die Clusterzentrenmatrix $\mathbf{M} = [\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_K]$
2. Ordne jeden Merkmalsvektor \mathbf{m}_n dem nächsten Cluster k mit $\boldsymbol{\mu}_k$ zu, sodass gilt $\|\boldsymbol{\mu}_k - \mathbf{m}_n\| = \min_{\{1, \dots, K\}} \|\boldsymbol{\mu}_k - \mathbf{m}_n\|$.
3. Bestimme entsprechend Schritt 2 die neue Partitionsmatrix $\boldsymbol{\Gamma}$
4. Berechne die neue Clusterzentrenmatrix \mathbf{M} anhand der neuen Partitionsmatrix $\boldsymbol{\Gamma}$ nach Formel (2.11)
5. Wiederhole Schritte 2 bis 4 bis keine Veränderung mehr in der Partitionsmatrix auftritt und somit $J(\boldsymbol{\Gamma}, \mathbf{M})$ minimiert wurde oder eine maximale Iterationsanzahl erreicht wurde.

Algorithmus 1: K-Means

Folgende Einschränkungen müssen bei der Verwendung des *K-Means*-Verfahren beachtet werden. Das *K-Means*-Verfahren generiert aufgrund der euklidischen Distanzen hypersphärische Cluster. Es werden also gleiche Varianzen im Merkmalsraum angenommen. Entscheidend für den Erfolg eines Clusterings ist die Anzahl der gewählten Cluster und die Initialisierung der Clusterzentren $\boldsymbol{\mu}_k$. Die Anzahl der Cluster kann durch Clustervalidierungsmethoden (siehe Abschnitt 2.4) geschätzt werden.

2.3 Maschinelles Lernen in medizinischen Bildanalyse-systemen

Bildanalyse-systeme verwenden Techniken des Maschinellen Lernens zur Bestimmung von Mustern in gegebenen Bilddaten. Dieser Abschnitt beschreibt die rudimentären Abläufe in Bildanalyse-systemen der Medizin und der digitalen Histopathologie anhand der Ausführungen in [Han09], [Nie03] und [GBC⁺09]. In diesem Überblick sollen die durchgeführten Analysen der vorliegenden Arbeit eingeordnet und angewendete Methoden erläutert werden. Angemerkt sei an dieser Stelle, dass der vorgestellte Ablauf nur ein Grundgerüst darstellt, an denen sich reale Anwendungen lediglich orientieren. Eine Übersicht zu den im Projekt ExPrimage angewendeten Analyseschritten wird in Kapitel 4 vorgestellt. Abbildung 2.1 zeigt eine Übersicht der grundlegenden Komponenten eines Bildanalyse-systems nach [Han09].

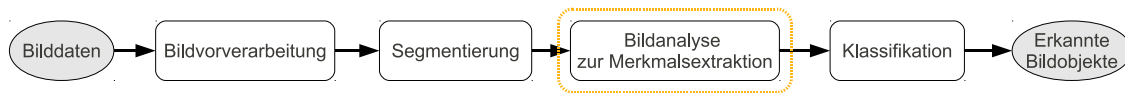


Abbildung 2.1: Grundlegende Komponenten eines medizinischen Bildanalyse-systems nach [Han09]. Die vorliegende Arbeit konzentriert sich auf die Erweiterung der Bildanalyse zur Merkmalsextraktion (orange markiert) und die dafür notwendigen Erweiterungen der Vorverarbeitung und Segmentierung.

Die aufgenommenen Bilddaten bzw. Bildmuster f werden zunächst einer *Bildvorverarbeitung* unterzogen. Das Bildmuster f kann durch verschiedene physikalische Einflüsse während der Aufnahme für die Analyse negativ beeinflusst werden. Durch die Bildvorverarbeitung werden die Bilddaten für die nachfolgenden Analyseschritte vorbereitet und optimiert. In der digitalen Histopathologie spielen vor allem Farb- und Beleuchtungsnormalisierung eine entscheidende Rolle [GBC⁺09].

Ziel der *Segmentierung* ist das Auffinden von Abgrenzungen unterschiedlicher Bildregionen bzw. Bildobjekte $R \subset f$ innerhalb eines Bildes. Sie ist zentraler Bestandteil in Bildanalyseverfahren der Histopathologie [DY05, GBC⁺09]. In der vorliegenden Arbeit ist die Segmentierung unterschiedlicher pathologischer Gewebestrukturen von Interesse. Sie ermöglicht die Analyse unterschiedlicher Gewebestrukturen auf morphologische Eigenschaften, ihre Verteilung und Relationen untereinander (siehe Abschnitt 5). In [OJJS96] wird das *Schwelwertverfahren von Otsu* [SRE⁺79] zur Segmentierung von Grauwertbildern digitaler Gewebeschnitte aufgeführt. Im Rahmen dieser Arbeit wird das *Schwelwertverfahren von Otsu* zur Segmentierung von Farbbildern erweitert 5.2.1. Daher soll das Verfahren in Abschnitt 2.3.1 erläutert werden. Regionen R repräsentieren in der vorliegenden Arbeit *Connected-Components* [Han09, BB05], auf dessen Generierung in Abschnitt 2.3.2 eingegangen wird.

Anschließend an die Segmentierung werden *Bildanalysen zur Merkmalsextraktion* zur Bestimmung markanter quantitative Kenngrößen (Merkmale) der Bilder f bzw. Bildobjekten R angewendet. Charakteristische Eigenschaften von Bildobjekten werden durch

die extrahierten Merkmale komprimiert repräsentiert. Diesem Kontext ordnet sich die vorliegende Arbeit primär zu. Auf in der Literatur verwendete Merkmale in der medizinischen und histopathologischen Bildanalyse wird in Kapitel 3 näher eingegangen.

Nachdem Merkmale zur Beschreibung der Bildmuster extrahiert wurden, werden diese als Merkmalsvektoren zur *Klassifikation* bzw. Erkennung der Bildmuster/Bildobjekte eingesetzt. In [Han09] beschränkt sich die Ausführung auf den Einsatz von überwachten Lernverfahren. Im Kontext der vorliegenden Arbeit wird unter diesem Schritt auch die Anwendung von unüberwachten Lernverfahren (siehe Abschnitt 2.2) verstanden.

2.3.1 Schwellwertverfahren von Otsu

Das Schwellwertverfahren von Otsu [SRE⁺79] bestimmt auf Grundlage des Histogrammes eines Bildes f einen globalen Schwellwert t , der die Pixelintensitäten des Bildes in zwei Klassen $\Omega_a(t)$ und $\Omega_b(t)$ aufteilt. Sei $p(i)$ die Wahrscheinlichkeit für das Auftreten einer Pixelintensität i (bestimmt aus dem normalisierten Histogramm) mit $0 < i < I$, wobei I die maximale Pixelintensität des betrachteten Bildes f darstellt. Die Wahrscheinlichkeit für das Auftreten von Pixelintensitäten der beiden Klassen in Abhängigkeit von t ist durch

$$P_{\Omega_a}(t) = \sum_{i=0}^t p(i) \qquad P_{\Omega_b}(t) = \sum_{t+1=0}^I p(i) \qquad (2.12)$$

gegeben. Die Varianzen zwischen den beiden Klassen in Abhängigkeit von t werden durch

$$\sigma_{\Omega_a}^2(t) = \sum_{i=0}^t (i - \bar{i}_{\Omega_a})^2 \cdot p(i) \qquad \sigma_{\Omega_b}^2(t) = \sum_{t+1=0}^I (i - \bar{i}_{\Omega_b})^2 \cdot p(i) \qquad (2.13)$$

definiert. Die Variablen \bar{i}_{Ω_a} bzw. \bar{i}_{Ω_b} repräsentieren die Mittelwerte der Pixelintensitäten in den einzelnen Klassen. Ziel des Verfahrens von Otsu ist es, die Varianz der Pixelintensitäten innerhalb der Klassen möglichst gering zu halten und zwischen den Klassen zu maximieren. Dies kann durch Maximieren des Kriteriums

$$Otsu(t) = \frac{P_{\Omega_a}(t) \cdot (\bar{i} - \bar{i}_{\Omega_a})^2 + P_{\Omega_b}(t) \cdot (\bar{i} - \bar{i}_{\Omega_b})^2}{P_{\Omega_a}(t) \cdot \sigma_{\Omega_a}^2(t) + P_{\Omega_b}(t) \cdot \sigma_{\Omega_b}^2(t)} \qquad (2.14)$$

zu $Otsu_{max}$ in der Iteration von t über den Wertebereich der Pixelintensitäten bestimmt werden. Dabei gibt \bar{i} den globalen Mittelwert der Pixelintensitäten an. Der Schwellwert t_{max} , welcher durch $Otsu_{max}$ bestimmt wurde, teilt die Pixelintensitäten von Bild f in zwei Klassen $i < t_{max} \rightarrow \Omega_a$ und $i \geq t_{max} + 1 \rightarrow \Omega_b$.

2.3.2 Connected-Components

Zusammenhangskomponenten bzw. *Connected-Components* werden in der vorliegenden Arbeit durch einen *Flood-Filling-Algorithmus* [BB05] berechnet. Definiert wird ein Connected-Component nach [Han09] als: „Sei ein N-Pfad definiert als eine Folge N-benachbarter Pixel. Eine Pixelmenge heißt N-zusammenhängend, wenn zu jedem Pixelpaar aus R ein N-Pfad existiert, der nur Pixel aus R enthält.“

Unterschieden wird zwischen einer 4- oder 8-Pixelnachbarschaft (siehe Abbildung 2.2). Ein Connected-Component R entspricht einer nach der gegebenen Definition zusammenhängenden Pixelmenge von Vordergrundpixeln eines Binärbildes. Ein Connected-Component R wird im Folgenden als Bildobjekt bezeichnet. Eingesetzt wird hierfür die entsprechende Funktionalität aus [Matc]. Die unterschiedlichen Bildobjekte R für ein Binärbild f_{bin} werden nach folgenden Schritten berechnet [Matc, BB05]:

1. Suche nach einem Pixel $p_{start} = 1 \in f_{bin}$, der noch keinem R zugeordnet wurde
2. Nehme p_{start} als Startpunkt und initialisiere ein neues Bildobjekt R_{new} , dem p_{start} zugeordnet wird
3. Ordne alle Pixel p , zu denen ein N-Pfad ausgehend von p_{start} existiert, dem neuen Bildobjekt R_{new} zu
4. Wiederhole Schritte 1 bis 3 solange Vordergrundpixel vorliegen, die noch keinem R zugeordnet wurden

Algorithmus 2: Connected-Components

Nach Ausführung des Algorithmus liegen separierte Bildobjekte R als unabhängige Pixelmengen vor, auf denen weiter operiert werden kann.

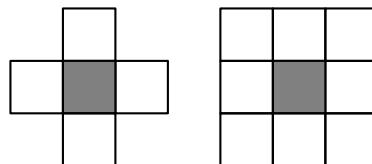


Abbildung 2.2: Unterschiedliche Pixelnachbarschaft (Grau = betrachteter Pixel): 4-Pixelnachbarschaft (links), 8-Pixelnachbarschaft (rechts).

2.4 Methoden zur Evaluation und Validierung

Nachdem einführend Grundlagen zu überwachten und unüberwachten Lernverfahren vorgestellt wurden, sollen im Folgenden die in der vorliegenden Arbeit angewendeten

Evaluations- und Validierungsmethoden für die eingesetzten maschinellen Lernverfahren *HLVQ* und *K-Means* erläutert werden. Dabei muss zwischen Methoden zur Bestimmung der *Klassifikationsgüte* des *HLVQ*-Verfahrens und der *Clustervalidierung* des *K-Means*-Verfahrens unterschieden werden.

Klassifikationsgüte Aufgrund der bekannten Klassenzugehörigkeit in überwachten Lernverfahren lassen sich klare Aussagen über die Leistung (Klassifikationsgüte) der Klassifikation treffen. Die angewendeten Evaluationsmaße zur Bestimmung der Performanz einer Klassifikation durch das *HLVQ*-Verfahren unter Verwendung der in dieser Arbeit entwickelten Merkmale werden in Abschnitt 2.4.1 erläutert.

Clustervalidierung Die Evaluation von Ergebnissen aus unüberwachten Lernverfahren sind aufgrund der fehlenden Klassenzugehörigkeit (siehe Abschnitt 2.2) wesentlich schwieriger. Einen ausführlichen Überblick zu dieser Thematik bieten die Arbeiten von Handl. et. al. [HKK05] sowie Runkler [Run10], an deren Ausführungen sich die folgenden Angaben, so weit nicht anders angegeben, orientieren.

Abschätzungen und statistische Tests, inwieweit Cluster in Datensätzen vorhanden sind, beschreiben die *Clustertendenz* [Fer00, HKK05, Run10], welche unabhängig von eingesetzten Clusteralgorithmen auf den Merkmalsdaten bestimmt werden kann. Um Ergebnisse verschiedener Clusteranalysen vergleichen und evaluieren zu können, braucht man möglichst objektive Kennzahlen, die eine Aussage über die Güte der Ergebnisse machen, da jeder Cluster-Algorithmus im Sinne seines Kriteriums die optimale Lösung findet [HKK05].

Inwieweit diese allerdings die zugrunde liegenden Daten widerspiegelt, kann durch die Cluster-Algorithmen nicht beantwortet werden. In [HKK05] wird dies als „*Non-significance of results in the absence of natural clusters*“ bezeichnet. Die Clustervalidierung überprüft das Vorhandensein von Gruppierungstendenzen und somit die Existenz von nicht-zufälligen Strukturen in Daten. Wie in [HKK05] bemängelt, geschieht eine objektive Verifizierung der Clusterergebnisse in der Literatur oft nicht, sondern gestaltet sich als manueller, langwieriger und subjektiver Prozess [HKK05].

Entscheidend für partitionierende Clusterverfahren ist die Schätzung der Anzahl zu verwendender Cluster K . Dieser Parameter muss, wie bei dem in dieser Arbeit verwendeten *K-Means*-Verfahren, vorher festgelegt werden (siehe Abschnitt 2.2).

Um diese Schwierigkeiten einzugrenzen, können Clustervalidierungstechniken eingesetzt werden. Dadurch kann zum einen die Güte der Cluster selbst bestimmt werden und zum anderen ein möglichst optimaler Parameter K geschätzt werden.

Man unterscheidet externe und interne Validierungsmethoden [HKK05]. Externe Methoden verwenden Wissen korrekter Klassenzugehörigkeiten ω der Merkmalsvektoren. Interne Methoden hingegen führen nur intrinsische Analysen auf den Ergebnispartitionen durch, um auf Grundlage methodenabhängiger Qualitätsmaße und Annahmen eine Aussage zu machen, wie gut die Cluster zu natürlichen Clustern (siehe Abschnitt 2.2) korrespondieren. Eine detaillierte Auflistung verschiedener Ansätze zur Clustervalidierung kann in [HKK05] gefunden werden.

Wie in der Einleitung (siehe Kapitel 1) bereits erwähnt, ist die Clustervalidierung ein im mathematischen Sinne nicht eindeutig gestelltes Problem. Eine eindeutige Lösung kann nicht bestimmt werden. Nur eine Kombination verschiedener Ansätze und Verfahren ist in der Lage, eine möglichst objektive Aussage über die Güte eines Clusterergebnisses zu machen [HKK05]. Zur Validierung der Ergebnisse der vorliegenden Arbeit muss eine dementsprechende geeignete Evaluierungsstrategie gefunden werden worauf in Abschnitt 5.6 eingegangen wird. Die in diesem Kontext angewendeten Methoden werden in den folgenden Abschnitten erläutert.

2.4.1 Klassifikationsmaße

Zur Evaluation von Klassifikationsergebnissen eines überwachten Lernverfahrens eignet sich die *Klassifikationsrate* (engl. *Ratio* oder *Accuracy*), die *Genauigkeit* (engl. *Precision*) und die *Trefferquote* (engl. *Recall*) [Run10, MKSW99]. Diese Evaluationsmaße verwenden die Anzahl der korrekt erkannten Muster TP (True-Positive), die Anzahl der korrekt als falsch erkannten Muster TN (True-Negative), die Anzahl der nicht korrekt erkannten Muster FP (False-Positive) und die Anzahl der nicht korrekt als falsch erkannten Muster FN (False-Negative) der Klassenmenge Ω [DY05]. Die Klassifikationsrate wird durch

$$\text{Klassifikationsrate} = \frac{TP}{n} \quad (2.15)$$

definiert, wobei n die Gesamtanzahl der verwendeten Muster angibt. Sie gibt den Anteil der korrekt klassifizierten Muster wieder und damit einen Überblick über die Gesamtleistung des Klassifikationssystems. Die Genauigkeit

$$\text{Genauigkeit} = \frac{TP}{TP + FP} \quad (2.16)$$

gibt Aufschluss darüber, wieviele Muster fälschlicherweise einer Klasse zugeordnet wurden. Ein hoher Wert repräsentiert eine hohe Genauigkeit. Die Trefferquote der Klassifikation wird durch

$$\text{Trefferquote} = \frac{TP}{TP + FN} \quad (2.17)$$

bestimmt. Sie gibt den Anteil der Falsch-Negativ klassifizierten Muster wieder. Ein hoher Wert repräsentiert einen geringen Falsch-Negativ-Anteil und damit eine hohe Trefferquote.

2.4.2 Hopkins-Index

Der *Hopkins-Index* [Run10], in der Literatur auch als *Hopkins-Statistik* bezeichnet [BD04], ist ein Verfahren zur Bestimmung der Clustertendenz eines gegebenen Merkmalsraumes

mit Y Merkmalsvektoren und basiert auf der Bestimmung der nächsten Nachbarn eines Merkmalsvektors \mathbf{m} . Die Definition und Implementation des Algorithmus orientiert sich an den Ausführungen in [Fer00], die einen verbesserte Approximation des Datenraumes vorstellen.

Zunächst wird die konvexe Hülle des Merkmalsraumes bestimmt. Innerhalb der konvexen Hülle des Merkmalsraumes werden gleichverteilt zufällig X Merkmalsvektoren $U = \{\mathbf{u}_1, \dots, \mathbf{u}_X\}$ generiert. Dabei ist $X \ll Y$ zu wählen. Danach werden X vorhandene Merkmalsvektoren $V = \{\mathbf{m}_1, \dots, \mathbf{m}_X\}$ gleichverteilt zufällig ausgewählt. Für die neu entstandenen Mengen U und V werden jeweils die Abstände $d_{u_1} \dots d_{u_X}$ bzw. $d_{v_1} \dots d_{v_X}$ zu den nächsten Nachbarn mittels der euklidischen Distanz (2.8) bestimmt. Der Hopkins-Index $h \in \{0, 1\}$ wird durch

$$h = \frac{\sum_{i=1}^X d_{u_i}}{\sum_{i=1}^X d_{u_i} + \sum_{i=1}^X d_{v_i}} \in \{0, 1\} \quad (2.18)$$

definiert. Die Nullhypothese des Hopkins-Index besteht in der Annahme, dass keine Clusterstrukturen im gegebenen Merkmalsraum enthalten sind. Die geringe Dichte in hochdimensionalen Merkmalsräumen führt bei der Anwendung von h zu instabilen Ergebnissen [Fer00]. Es wird daher in [HKK05] und [Fer00] vorgeschlagen, die Anwendung auf die ersten 2D-Hauptachsen-Projektionen (siehe Abschnitt 2.4.3) des Merkmalsraumes zu beschränken. Zur Verdeutlichung der Funktionsweise von h können drei typische Fälle unterschieden werden:

- $h \approx 0$: In diesem Fall sind die Abstände von U im Verhältnis zu V groß. Dies lässt auf eine regelmäßige Struktur schließen (siehe Abbildung 2.3 links oben).
- $h \approx 0.5$: Die Abstände von U im Verhältnis zu V sind ungefähr gleich. Dementsprechend deutet dieses Ergebnis auf zufällig verteilte Merkmalsvektoren hin (siehe Abbildung 2.3 rechts oben).
- $h \approx 1$: Die Abstände von U sind im Verhältnis zu den Abständen von V sehr groß. Dementsprechend können sich Clusterstrukturen im Merkmalsraum befinden (siehe Abbildung 2.3 unten links).

Da die Mengen U und V zufällig bestimmt werden, ist es sinnvoll, durch Mittelung der Ergebnisse h_μ verschiedener Durchläufe die Objektivität des Maßes zu erhöhen. Zur Beurteilung der Stabilität wird das Minimum h_{min} , das Maximum h_{max} jeder Iteration und die daraus resultierende Schwankung $h_{range} = h_{max} - h_{min}$ bestimmt [Fer00]. In Abbildung 2.3 (unten rechts) werden die Ergebnisse anhand von drei Beispieldatensätzen, welche die aufgezählten typischen Fälle repräsentieren, zusammengefasst.

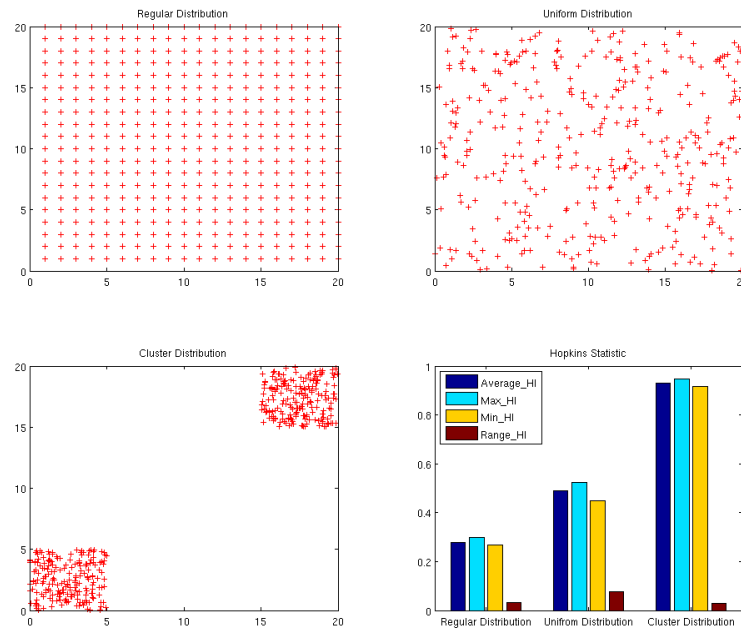


Abbildung 2.3: Beispiel *Hopkins-Index*: 2000 2D-Merkmalenvektoren regelmäßig verteilt (links oben, *Regular Distribution*), zufällig gleichverteilt (rechts, *Uniform Distribution*) und angeordnet in 2 Clustern (links unten, *Cluster Distribution*). Unten rechts wird der durchschnittliche *Hopkins-Index* jeder Verteilung (blau) sowie Minimum (gelb), Maximum (cyan) und die Abweichung (rot) angegeben.

2.4.3 Hauptkomponentenanalyse

Die Hauptkomponentenanalyse (kurz PCA (engl. Principle Component Analysis)) [Han09, Shl05] verfolgt das Ziel, durch Bestimmung von Hauptachsen bzw. Hauptkomponenten mit Hilfe einer kleineren Anzahl von Basisvektoren die Dimensionen des Merkmalsraumes zu reduzieren.

Die PCA versucht Redundanzen zu vermeiden und nur einen kleinen Teil der Informationen zu verlieren. Es wird angenommen, dass die Variablen statistisch unabhängig sind und die meiste Information in den höchsten Varianzen der Daten liegt (*Varianzkriterium*) [Han09]. Die Hauptkomponenten liegen nach der Hauptkomponentenanalyse sortiert nach dem *Varianzkriterium* vor. Die erste Komponente umfasst die größte Varianz der Merkmalsdaten, die zweite die nächst größere und so weiter.

Umfangreiche Merkmalsdaten können durch die PCA strukturiert, vereinfacht und durch Projektion auf die Hauptkomponenten mit den höchsten Varianzen visualisiert werden. Sie ermöglicht einen Einblick in hochdimensionale Datenräume [HKK05].

Nach einer Projektion der Merkmalsdaten auf Hauptkomponenten mit geringerer Dimension kann bestimmt werden, wie viel der Gesamtvarianz durch die Projektion abgedeckt wird. Wenn in den Projektionen der Merkmalsvektoren durch die PCA, beispielsweise in einem 3D-Datenraum, Clusterstrukturen zu erkennen sind, kann davon ausgegangen

werden, dass Cluster vorhanden sind. Der Umkehrschluss ist nicht gültig. Die PCA wird in der vorliegenden Arbeit als Teil des Validierungskonzepts (siehe Abschnitt 5.6.1) eingesetzt. Für die Erläuterung der einzelnen Berechnungsschritte der PCA sei an dieser Stelle auf Anhang D und [Shl05, Han09] verwiesen.

2.4.4 Silhouetten-Wert

Die *Silhouette-Width* (folgend Silhouetten-Wert genannt) nach [Rou87] berechnet für jeden Merkmalsvektor einen Wert $SV_{\mathbf{m}}$, der Aufschluss über die Separierung und Kompaktheit des Merkmalsvektors bezüglich seines zugeordneten Clusters C_k angibt. Aufbauend darauf kann sowohl für jedes Cluster C_k als auch für die gesamte Partition der durchschnittliche Silhouetten-Wert bestimmt werden, um so eine Aussage über die Güte der Clusteranalyse zu erhalten. Das Maß gehört zu den internen Clustervalidierungsmethoden [HKK05]. $SV_{\mathbf{m}}$ wird als

$$SV_{\mathbf{m}} = \frac{b_{\mathbf{m}} - a_{\mathbf{m}}}{\max\{a_{\mathbf{m}}, b_{\mathbf{m}}\}} \in \{-1, 1\} \quad (2.19)$$

definiert. Die Variable $a_{\mathbf{m}}$ bezeichnet die durchschnittliche Unähnlichkeit des Merkmalsvektors zu den anderen Merkmalsvektoren des gleichen Clusters. Zur Bestimmung von $b_{\mathbf{m}}$ wird die durchschnittliche Unähnlichkeit des Merkmalsvektors zu allen anderen Clustern, dem der betrachtete Merkmalsvektor nicht angehört, bestimmt. Das Minimum dieser Werte definiert $b_{\mathbf{m}}$. Als Unähnlichkeit wird im Rahmen der vorliegenden Arbeit die quadratische euklidische Distanz d_e^2 (siehe Formel (2.8)) zwischen zwei Merkmalsvektoren verwendet. Der Wertebereich von $SV_{\mathbf{m}}$ ist $\{-1, 1\}$.

Hohe Werte deuten auf gute Clusterstrukturen hin. Werte nahe 0 bedeuten, dass der Merkmalsvektor genauso gut einem anderen Cluster hätte zugeordnet werden können. Werte nahe -1 deuten darauf hin, dass der Merkmalsvektor einem falschen Cluster zugeordnet wurde. Der Wertebereich $0.75 < SV_{\mathbf{m}} \leq 1$ spricht für eine starke Strukturierung der Daten. Als mittlere Strukturierung wird der Wertebereich $0.5 < SV_{\mathbf{m}} \leq 0.75$ angegeben. Eine schwache Struktur liegt im Bereich $0.25 < SV_{\mathbf{m}} \leq 0.5$ vor. Bei Werten $0 < SV_{\mathbf{m}} \leq 0.25$ liegt keine Strukturierung vor. Abbildung 2.4 verdeutlicht die Zusammenhänge anhand eines Beispiels.

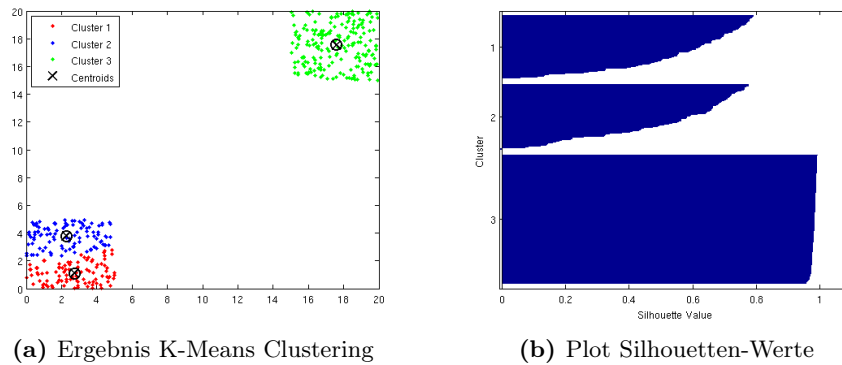


Abbildung 2.4: Clusterbewertung einer durch das *K-Means*-Verfahren durchgeführten Clusteranalyse (mit Clusteranzahl 3) mittels Silhouetten-Wert. Cluster 1 und 2 sind nicht korrekt und enthalten dementsprechend niedrigere Silhouetten-Werte als die Merkmalsvektoren in Cluster 3.

In (a) wird das Ergebnis eines 2D-Merkmalraumes mit offensichtlich zwei unterschiedlichen Clustern und die Ergebnispartition (farbig markiert), nach der Anwendung des *K-Means*-Verfahrens mit Clusteranzahl $K = 3$, dargestellt. In (b) werden die Silhouetten-Werte jedes Merkmalsvektors, aufgeteilt für jedes Cluster, angezeigt. Es ist deutlich zu erkennen, dass die Merkmalsvektoren aus Cluster 1 und 2 (welche korrekterweise ein Cluster bilden sollten) im Verhältnis zu den Merkmalsvektoren aus Cluster 3 wesentlich niedrigere Silhouetten-Werte haben.

2.4.5 Normalized Variation of Information

Die *Variation of Information* (kurz *VI*) bzw. die normalisierte *VI* nach [WXC09] (kurz VI_n) ist ein externes binäres Clustervalidierungsmaß [HKK05]. Insbesondere für die Anwendung auf Cluster, die durch das *K-Means*-Verfahren entstanden sind, berücksichtigt die VI_n die mathematischen Eigenschaften des *K-Means*-Verfahrens [WXC09]. Dieses Ergebnis erbrachte die Studie von Wu et. al. [WXC09], in der 16 unterschiedliche externe Validierungsmaße unter Verwendung des *K-Means*-Verfahrens untersucht wurden. Die VI_n wird in [WXC09] für schwierig zu unterscheidende Clusteranalysen vorgeschlagen, da sie sehr sensitiv gegenüber Clusterveränderungen ist. Die Berechnung der VI_n basiert auf der Kontingenztafel bzw. Kontingenzmatrix CM (siehe Tabelle 2.1).

		Partition P				
		P_1	P_2	\cdots	$P_{k'}$	Σ
Partition C	C_1	n_{11}	n_{12}	\cdots	$n_{1k'}$	$n_{1\cdot}$
	C_2	n_{11}	n_{12}	\cdots	$n_{1k'}$	$n_{1\cdot}$
	\cdot	\cdot	\cdot	\cdots	\cdot	\cdot
	C_k	n_{k1}	n_{k2}	\cdots	$n_{kk'}$	$n_{k\cdot}$
	Σ	$n_{\cdot 1}$	$n_{\cdot 2}$	\cdots	$n_{\cdot k'}$	n

Tabelle 2.1: Kontingenztabelle CM

Die Variable n_{ij} repräsentiert die Anzahl der Merkmalsvektoren, die Clusterpartition C_i zugeordnet wurden und Clusterpartition P_j (z.B. aus einer weiteren Analyse) angehören. Die VI kommt aus der Informationstheorie und ist eine Erweiterung der *Entropie* bzw. der *Mutual Information* [JPWPV03]. Die VI misst den Informationsverlust bzw. -gewinn von Clusterpartition C zur Clusterpartition P [Mar02]. Die in [WXC09] vorgeschlagene normalisierte Variante VI_n wird als

$$VI_n = 1 + 2 \frac{H(P) + H(C) - H(P, C)}{-(H(P) + H(C))} \in \{0, 1\} \quad (2.20)$$

definiert. Die marginalen Entropie¹ $H(C)$ bzw. $H(P)$ (vgl. [Sha48]) der zu vergleichenden Partitionen werden aus den Spalten bzw. Zeilen von CM berechnet. Die Verbundentropie² $H(C, P)$ (vgl. [Sha48]) wird ebenfalls aus CM berechnet. Niedrige Werte der VI_n deuten auf einen hohen Informationszusammenhang hin, während hohe Werte für wenig Informationszusammenhang stehen. Anhand der Kontingenztabelle(n) (siehe Tabellen 2.2), welche drei grundsätzlich zu unterscheidende Fälle der VI_n darstellen, wird die Interpretation dieses Maßes deutlich.

	P_1	P_2	P_3		P_1	P_2	P_3		P_1	P_2	P_3
C_1	15	15	15	C_1	27	0	27	C_1	45	0	0
C_2	15	15	15	C_2	0	27	0	C_2	0	45	0
C_3	15	15	15	C_3	27	0	27	C_3	0	0	45

(a) $VI_n = 1$ (b) $VI_n \approx 0.5$ (c) $VI_n = 0$

Tabelle 2.2: Beispiel von drei typisch zu unterscheidenden Ergebnissen der VI_n bei gegebenen Kontingenztabelle(n): a) kein Informationszusammenhang, b) kein eindeutiger Informationszusammenhang, c) vollkommene Informationsübereinstimmung.

¹ $H(\cdot) = -\sum_i p_i \cdot \log(p_i)$ mit $p_i = \frac{n_{i\cdot}}{n}$ bzw. $-\sum_j p_j \cdot \log(p_j)$ mit $p_j = \frac{n_{\cdot j}}{n}$
² $H(\cdot, \cdot) = -\sum_{ij} p_{ij} \cdot \log(p_{ij})$ mit $p_{ij} = \frac{n_{ij}}{n}$

Kapitel 3

Merkmale in der histopathologischen und medizinischen Bildanalyse

Dieses Kapitel gibt zunächst einen allgemeinen Überblick zu Bildmerkmalen (auch Deskriptoren genannt [NA08]), die in der digitalen histopathologischen und der medizinischen Bildanalyse eingesetzt werden. In diesem Rahmen werden verschiedene Merkmalskategorien definiert, welchen die in der vorliegenden Arbeit implementierten und eingesetzten Merkmale zugeordnet werden.

Eine umfangreiche Sammlung und einen Überblick zu Merkmalen im Anwendungsgebiet der medizinischen Bildanalyse bieten die Arbeiten von Rodenacker et. al. [RB03] (Zytologie und Histopathologie), Müller et. al. [MMBG04] (im Bereich des medizinischen Image-Retrievals) und die Ausführungen in [Han09]. Einen Überblick zu Merkmalen in der Bildanalyse von histopathologischen Bildern wird in Demir et. al. [DY05], Gurcan et. al. [GBC⁺09] und in der Einleitung von Altunbay et. al. [ACSGD10] gegeben. Anhand dieser Ausführung soll zunächst eine grundlegende Kategorisierung von Bildmerkmalen gefunden sowie deren Bedeutung erläutert werden.

In [RB03] werden Merkmale in morphometrische Merkmale, intensitätsbasierte Merkmale, texturbezogene Merkmale und strukturelle Merkmale eingeteilt. Im Bezug auf histopathologische Bildmerkmale wird in [ACSGD10] die gleiche Einteilung vorgenommen. In [DY05] werden zusätzlich fraktale Bildmerkmale genannt und strukturelle Merkmale als topologische Merkmale bezeichnet. In [GBC⁺09] wird zwischen *Object-Level Features* und *Spatially Related Features* unterschieden. *Object-Level Features* werden weiter unterteilt in „Form und Größe“, „Radiometrisch und Densitometrisch“, „Textur“ und „Chromatin Spezifisch“. Als *Spatially Related Features* werden Merkmale auf Grundlage von Graphstrukturen bezeichnet. In [DHS⁺07] und [DAM⁺08] werden diese Merkmale *Architectural Features* genannt. In [Bou08] werden räumliche Relationen zwischen verschiedenen Gewebetypen in Form von Histogrammen quantifiziert, welche in der Literatur noch keiner Kategorie zugeordnet sind. Die Arbeit von Bloch et. al. [Blo05] gibt einen zusammen-

fassenden Überblick zu „Fuzzy-Ansätzen“ von räumlichen Beziehungen/Relationen innerhalb segmentierter Bilddaten. Es werden Ansätze zur Beschreibung von Richtungen, Nachbarschaften und Überlagerungen beschrieben.

Handels [Han09] unterteilt Bildmerkmale anhand grundsätzlicher Analyseansätze in Texturanalysen, fraktale Analysen und morphologische Analysen. In [MMBG04] wird zwischen Farb-Level-, Grau-Level-, Form- und Texturmerkmalen unterschieden.

Anhand der vorigen Ausführungen wird ersichtlich, dass eine eindeutige Terminologie und Kategorisierung von Bildmerkmalen in der medizinischen und digitalen histopathologischen Bildanalyse in der Literatur nicht gegeben ist. Die aufgezählten Kategorien haben in ihrer Bedeutung teilweise Überschneidungen und teilweise nicht. Aufbauend auf [RB03] und [ACSGD10] wird in der vorliegenden Arbeit zwischen den folgenden Merkmalskategorien unterschieden, wobei die Kategorie „Relation“ hinzugefügt wurde:

- Intensität
- **Morphometrie**
- Fraktale
- **Struktur**
- Textur
- **Relation**

Intensitätsbasierte Merkmale beschreiben die Unterschiede in den Pixelintensitäten eines Bildes oder Bildausschnittes. Zu den wichtigsten intensitätsbasierten Merkmalen zählen der Mittelwert, die Standardabweichung, die Schiefe und der Exzess des vorher berechneten Histogramms eines Bildausschnittes. Fraktale Merkmale beschreiben die Selbst-Ähnlichkeit von untersuchten Bildbereichen. Nähere Ausführungen dazu können in [Han09] gefunden werden. In der histopathologischen Bildanalyse werden diese Merkmale sowohl für Grauwertbilder als auch für die einzelnen Farbkanäle von Farbbildern berechnet [DY05]. Texturmerkmalen werden in der digitalen Histopathologie häufig auf Bildausschnitte angewendet [GBC⁺09, DY05] und beschreiben strukturelle Veränderungen im Gewebe. Anwendung finden vor allem Haralick-Merkmale auf Grundlage der Kookkurrenz-Matrix [GBC⁺09, MMBG04, RB03]. Mit intensitätsbasierten und texturbasierten Merkmalen im Bezug auf die Analysen der Gewebeebene im Rahmen des Projekts ExPrimage beschäftigt sich die Arbeit [Kha11].

Für die vorliegende Arbeit sind Merkmale für Bildobjekte von Relevanz. Zu diesen Kategorien zählen die markierten: Morphometrie, Struktur und Relation. Ein Bildobjekt R wird hierbei als räumlich zusammenhängende Teilmenge eines Binärbildes f_{bin} , als Connected-Component, verstanden (siehe Abschnitt 2.3.2). Morphometrische Bildmerkmale beschreiben die Größe und Form eines Bildobjektes.

Unter strukturellen Merkmalen werden Merkmale verstanden, die auf Grundlage von Graphstrukturen berechnet werden und räumliche Verteilungen erfassen. Relationale Merkmale beschreiben Kolokalisationen und räumliche Relationen von Bildobjekten untereinander. Im Folgenden wird gesondert auf die im Fokus liegenden Kategorien und ihre Verwendung in der medizinischen und digitalen histopathologischen Bildanalyse eingegangen. Die eingesetzten und implementierten Merkmale sowie notwendigen Algorithmen werden in diesem Kontext vorgestellt.

3.1 Morphometrische Merkmale

Morphometrische Merkmale quantifizieren die Größe und Form von Bildobjekten. Entscheidend hierfür ist eine klare Segmentierung der Objektgrenzen. Sie werden in der histopathologischen Bildanalyse primär zur Beschreibung von Zellen und Zellkernen eingesetzt [ACSGD10]. Zu den am weitesten verbreiteten zählen die Fläche (Area), der Umfang (Perimeter), die Kompaktheit (Compactness), die Rundheit (Roundness), der maximale Durchmesser (Major axis), die Symmetrie (Symmetry) und die Konkavität (Concavity), wie zusammenfassend in [DY05] erläutert wird. In Abbildung 3.1 wird die Vermessung eines Zellkerns mit morphometrischen Merkmalen dargestellt.

Weitere vielseitig eingesetzte morphometrische Merkmale sind *Fourier-Deskriptoren* [Han09], [MMBG04]. Aufbauend auf den Fourier-Deskriptoren werden in der digitalen histopathologischen Bildanalyse Fourier-Energien [Bou08, GBC⁺09] (siehe Abschnitt 3.1.1) verwendet. Weitere wichtige Vertreter der morphometrischen Merkmale sind Momente [RB03, MMBG04]. In [EMT97] werden z.B. invariante Momente nach [Hu62] (*Moment Invariants*) zur Klassifikation von Brustkrebsgewebebiopsien eingesetzt (siehe Abschnitt 3.1.3). In Rangayyan et. al. [REFDA97] werden sowohl Fourier-Deskriptoren, die Roundness als auch Moment Invariants als Analyse von Formfaktoren erfolgreich (95 % Klassifikationsrate) zur Klassifikation von gutartigen und bösartigen Brustkrebstumoren in Mammographieaufnahmen eingesetzt. Dies unterstreicht den Zusammenhang der Kontur und Form eines Tumors im Hinblick auf seine Malignität.

Objektbezogene morphometrische Merkmale lassen sich nach den Ausführungen in [NA08] in Regionen-Deskriptoren und Kontur-Deskriptoren unterscheiden. Regionen-Deskriptoren beschreiben das komplette Bildobjekt R , während Kontur-Deskriptoren nur die Kontur der Objekte charakterisieren. Regionen-Deskriptoren werden nach [NA08] wiederum in Basis-Deskriptoren und Momente unterteilt.

Implementiert wurden für die vorliegende Arbeit Fourier-Deskriptoren, als Vertreter der Kontur-Deskriptoren (siehe Abschnitt 3.1.1). Verwendete und implementierte *Basis-Deskriptoren* werden in Abschnitt 3.1.2 beschrieben. Die implementierten Momente werden in Abschnitt 3.1.3 dargestellt.

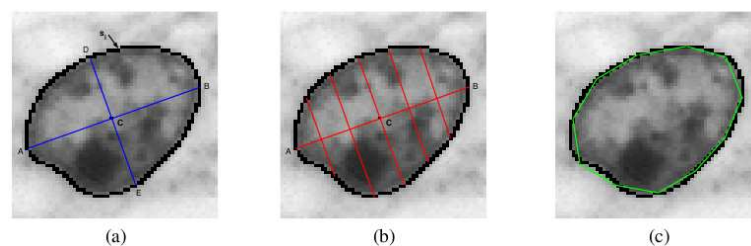


Abbildung 3.1: Beispiel für morphometrische Berechnungen auf Zellebene: (a) Zellkern mit Kontur und Zentroid, b) Liniensegmente zur Symmetriebestimmung, c) Konturabschnitte zur Berechnung der Konkavität (Quelle: [DY05]).

3.1.1 Fourier-Deskriptoren

Fourier-Deskriptoren (kurz FDs) bieten sich als Merkmale an, da sie invariant gegenüber Translation, Rotation und Skalierung sind [MMBG04]. Sie lassen sich schnell implementieren und durch die *DFT* (Diskrete Fourier-Transformation) [BB05] effizient berechnen. Durch die FDs kann die äußere Kontur bzw. die Berandung eines Bildobjektes effektiv beschrieben werden. Die Implementation und folgende Ausführungen basieren auf Grundlage der Studie zur Effizienz von FDs in [ZL02] und unter Zuhilfenahme der Beschreibungen zu FDs in [SKH08]. Zhang et. al. [ZL02] geben in ihrer Studie einen Überblick zu unterschiedlichen Berechnungsmöglichkeiten der FDs und evaluieren diese.

Zunächst wird die äußere Kontur $x(t), y(t), t = 0, 1, \dots, L - 1$ von R mit der Länge L anhand einer 8-Pixelnachbarschaft (siehe Abbildung 2.2) bestimmt. Danach wird eine 1D-Funktion (*Shape signature*) aus der 2D-Konturfunktion entwickelt. Nach den Studien aus [ZL02] ist die Methode der Zentroiden Distanzen $r(t)$ (siehe Formel(3.1)) zur Beschreibung der 2D-Kontur eines R am effektivsten. Dieser Ansatz wird daher für die vorliegenden Arbeit übernommen. Hierzu werden zunächst die Löcher in den Bildobjekten R gefüllt, um eine eindeutige Bestimmung des Zentroiden in Bezug zur äußeren Kontur zu gewährleisten.

$$r(t) = \left([x(t) - \bar{x}_c]^2 + [y(t) - \bar{y}_c]^2 \right)^{\frac{1}{2}} \quad (3.1)$$

Die Funktion $r(t)$ entspricht den euklidischen Distanzen der Konturpunkte zum Schwerpunkt (\bar{x}_c, \bar{y}_c) (siehe Formel (3.2)) der äußeren Kontur.

$$\bar{x}_c = \frac{1}{L} \sum_{t=0}^{L-1} x(t) \quad \bar{y}_c = \frac{1}{L} \sum_{t=0}^{L-1} y(t) \quad (3.2)$$

Bevor die Fourier-Transformation auf $r(t)$ angewendet werden kann, muss für $r(t)$ zunächst eine feste Anzahl an Datenpunkten n bestimmt werden. Unterschiedliche Bildobjekte R haben in der Regel auch unterschiedliche Anzahlen von Konturpunkten. Die Anzahl der Datenpunkte n ist ein zu definierender Parameter. Durch n kann so Einfluss auf die Genauigkeit der Repräsentation von R genommen werden. Je höher n gewählt wird, desto genauer wird die Repräsentation, was gleichzeitig den Berechnungsaufwand erhöht.

Zur Normalisierung der Kontur wird in [ZL02] die Bogenlänge der Kontur als effektivste Methode vorgeschlagen. Dazu werden zu jedem Konturpunkt im Uhrzeigersinn die euklidischen Distanzen $d(t) = \left([x(t) - x(t+1)]^2 + [y(t) - y(t+1)]^2 \right)^{\frac{1}{2}}$ zum jeweiligen Nachbarkonturpunkt bestimmt. Diese Distanzen werden akkumuliert und ergeben die für R spezifische Bogenlänge bzw. den Umfang U . Nun gibt $\frac{n}{U}$ den gleichförmigen Abtastabstand zwischen jedem Datenpunkt an. Durch lineare Interpolation von $r(t)$ anhand $d(t)$ mit dem Abstand $\frac{n}{U}$ zu $r(t)'$ wird die Normalisierung auf genau n Datenpunkte erreicht und kann so auch eine Datenpunktanzahl von n bei $L < n$ garantieren. Dies wird

in der vorliegenden Anwendung aufgrund der sehr heterogenen Größeneigenschaften der unterschiedlichen Bildobjekte notwendig. Gleichzeitig hat dieses Vorgehen einen glättenden Effekt auf die Kontur und entfernt so unerwünschtes Rauschen. Die Anwendung der *Diskreten Fourier-Transformation* [BB05] ergibt

$$\mathbf{u}_n = \frac{1}{N} \sum_{t=0}^{N-1} r(t)' \cdot e^{-\frac{i2\pi nt}{N}}, n = 0, 1, \dots, N - 1 \quad (3.3)$$

als Koeffizientenvektor. Die Koeffizienten \mathbf{u}_n werden als Fourier-Deskriptoren bezeichnet. Sie beschreiben die Amplitude der Basisfunktionen (Kosinus und Sinus) und repräsentieren das Eingangssignal $r(t)'$ und somit die Kontur eindeutig und vollständig [BB05]. Da es sich bei $r(t)$ um eine reellwertige Funktion (Imaginärteil = 0) handelt, sind nur $N/2$ unterschiedliche Frequenzen enthalten weshalb in der vorliegenden Arbeit nur die Hälfte der Werte aus \mathbf{u} zur Beschreibung der Kontur verwendet wird. Die Zentroid-Distanzfunktion (siehe Formel (3.1)) ist invariant gegenüber Translation. Es sind nur wenige Änderungen in Vektor \mathbf{u} notwendig, um Invarianz bzgl. Translation, Rotation und Skalierung auszudrücken. Dies wird durch

$$\mathbf{fd} = \left[\frac{|\mathbf{u}_1|}{|\mathbf{u}_0|}, \dots, \frac{|\mathbf{u}_{N/2}|}{|\mathbf{u}_0|} \right] \quad (3.4)$$

erreicht. Der erste Fourier-Koeffizient \mathbf{u}_0 wird DC-Element (engl. *direct current*) genannt. Er beschreibt die Summe von $r(t)'$ und ist abhängig von der Position. Durch das Ignorieren dieses Elementes wird Translationsinvarianz garantiert. Dass in \mathbf{fd} nur die Beträge der Koeffizienten verwendet werden, entspricht dem Ignorieren der Phaseninformation, womit Rotationsinvarianz erreicht wird. Die Division jedes Koeffizientens durch das DC-Element und damit der skalierungsabhängigen Summe von $r(t)'$ erreicht Skalierungsinvarianz.

Durch die inverse Diskrete-Fouriertransformation DFT^{-1} von \mathbf{fd} ist eine Rekonstruktion der ursprünglichen Form-Funktion $r(t)$ möglich. So lässt sich abschätzen welche Parameter für die Anzahl der Deskriptoren und Datenpunkte sinnvoll sind.

Die Grundidee der Fourier-Deskriptoren ist, dass unterschiedliche Konturen durch weniger Fourier-Koeffizienten und eine Mittlung über alle Konturpunkte (die Fourier-Transformation) repräsentiert werden, welche zur Diskriminierung ausreichen.

Aus FDs lassen sich *Fourier-Energien* bestimmen [Bou08]. Fourier-Energien bzw. *Perimeter Fourier-Energies* [GBC⁺09] werden nach [Bou08] in niedrige, mittlere und hohen Energien unterteilt. Die Energien entsprechen den Summen der vorher berechneten Fourier-Koeffizienten \mathbf{fd} der Länge Nf . In [Bou08] werden diese als

$$\begin{aligned}
feLow & \sum_{n=1}^{Nf/3} |\mathbf{f}d_n|^2 \\
feMid & \sum_{n=Nf/3+1}^{2(Nf/3)} |\mathbf{f}d_n|^2 \\
feHigh & \sum_{n=2(Nf/3)+1}^{Nf} |\mathbf{f}d_n|^2
\end{aligned} \tag{3.5}$$

definiert, wobei dort von einer festen Länge $Nf = 1024$ ausgegangen wird. Die Fourier-Energien bieten eine komprimierte Darstellung der Fourier-Koeffizienten. In der vorliegenden Anwendung orientiert sich die Anzahl Nf an den verwendeten Fourier-Koeffizienten.

3.1.2 Basis-Deskriptoren

Basis-Deskriptoren konzentrieren sich nach [NA08] auf die geometrische Form der Fläche eines Bildobjektes R . In dieser Arbeit eingesetzte Merkmale werden im Folgenden vorgestellt. Die Fläche als Funktion $A(R)$ nach [NA08] wird durch

$$A(R) = \sum_x \sum_y R(x, y) \tag{3.6}$$

bestimmt. Ein weiterer Indikator für unterschiedliche Größeneigenschaften ist der Umfang eines Bildobjektes $P(R)$. Der Umfang eines Bildobjektes wird nach [NA08] als

$$P(R) = \sum_i \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \tag{3.7}$$

bestimmt, wobei die Koordinaten i den Pixeln der äußeren Kontur von R entsprechen. Aus diesen einfachen Messwerten lassen sich weitere Merkmale wie Kompaktheit und Rundheit bestimmen. Die Kompaktheit (engl. *Compactness*) CO ist nach [Bou08] als

$$CO(R) = \frac{4\pi A(R)}{P(R)^2} \tag{3.8}$$

definiert. $CO(R)$ nimmt bei einem perfekten Kreis den Wert 1 an. Die Rundheit CI (engl. *Circularity* oder *Roundness*) untersucht die Kreisförmigkeit eines Objektes und wird in [Han09] als

$$CI(R) = \frac{4A(R)}{\pi D(R)^2} \in [0, 1] \tag{3.9}$$

definiert. Die Rundheit setzt die Fläche $A(R)$ in Relation zu einem Kreis mit maximalem Durchmesser $D(R)$ des Bildobjektes und nimmt wie die Kompaktheit daher für perfekte Kreise den Wert 1 an. Die *Irregularity* [NA08] beschreibt die Dichte eines Objektes R mit den Pixelkoordinaten x_i und y_i und wird in [NA08] als

$$IR(R) = \frac{\pi \max((x_i - \bar{x}_r)^2 + (y_i - \bar{y}_r)^2)}{A(R)} \quad (3.10)$$

definiert. Die Werte \bar{x}_r und \bar{y}_r entsprechen dabei den Koordinaten des Zentroiden von R (siehe Formel (3.2)). Der Zähler definiert die Fläche des Kreises, der das gesamte Bildobjekt R einschließt. Von daher nimmt auch IR für Bildobjekte, die einen perfekten Kreis repräsentieren, den Wert 1 an.

Zur Berechnung der Fläche A und des Umfanges P wird in der vorliegenden Arbeit auf die Funktionen von MATLAB [Matc] zurückgegriffen.

3.1.3 Moment invariants

Moment invariants [BB05] wurden von Hu [Hu62] eingeführt und bieten eine einfache Möglichkeit, verschiedene Bildobjekte R zu beschreiben. In [Hu62] werden sie erfolgreich zur Klassifikation von Schriftzeichen eingesetzt. Sie konzentrieren sich auf die Dichte der binären Bildobjekte und beschreiben unabhängig von der Kontur die räumliche Verteilung der Pixel. Dadurch wird es möglich, z.B. ein Bildobjekt mit zentraler Aushöhlung von dem gleichen Bildobjekt mit gefüllter zentraler Aushöhlung zu unterscheiden, was durch die Fourier-Deskriptoren, welche die äußere Kontur beschreiben, nicht möglich ist.

Die folgenden Beschreibungen zur Definition der Moment invariants basieren auf den Ausführungen in [BB05] und [Hu62], so weit nicht anders angegeben.

An dieser Stelle soll die Herleitung der Moment invariants dargestellt werden. Dabei beziehen sich die Definitionen auf binäre Bildobjekte R . Grundsätzlich ist eine Berechnung auf Pixelintensitäten ebenfalls möglich [BB05].

Die Definition von Momenten kommt aus der Statistik. Sie können als gewichtete Mittelwerte verstanden werden und geben physische Eigenschaften eines Bildobjektes wieder. Ein Moment m_{pq} der Ordnung p, q wird durch

$$m_{pq}(R) = \sum_{(x,y) \in R} R(x,y) \cdot x^p y^q \quad (3.11)$$

definiert. Das Moment mit Ordnung 0,0 entspricht der Fläche A eines Bildobjektes. Der Schwerpunkt (\bar{x}_r, \bar{y}_r) eines Bildobjektes kann durch

$$\bar{x}_r = \frac{m_{010}}{m_{000}} \quad \bar{y}_r = \frac{m_{001}}{m_{000}} \quad (3.12)$$

bestimmt werden. Zur Bestimmung von Merkmalen auf Grundlage der Momente, die unabhängig von der Position eines Bildobjektes sind, werden diese zu *Zentralen Momenten*

$$\mu_{pq}(R) = \sum_{(x,y) \in R} R(x,y) \cdot (x - \bar{x}_r)^p \cdot (y - \bar{y}_r)^q \quad (3.13)$$

erweitert, wodurch der Ursprung des Koordinatensystems an den Schwerpunkt des Bildobjektes verschoben wird. Um nun unabhängig von der Skalierung eines Bildobjektes zu werden, benötigt man *Normalisierte Zentrale Momente*

$$\overline{\mu}_{pq}(R) = \mu_{pq}(R) \cdot \left(\frac{1}{\mu_{00}(R)} \right)^{\frac{p+q+2}{2}} \quad (3.14)$$

, welche den Skalierungsfaktor durch Multiplikation mit dem Kehrwert der potenzierten Fläche für $(p+q) \geq 2$ normalisieren. In [Hu62] werden diese Eigenschaften verwendet, um durch sieben Kombinationen der *Normalisierten Zentralen Momente* eine möglichst orthogonale Basis mit Rotationsinvarianz zu erreichen:

$$\begin{aligned} H_1 &= \bar{\mu}_{20} + \bar{\mu}_{02} \\ H_2 &= (\bar{\mu}_{20} + \bar{\mu}_{02})^2 + 4\bar{\mu}_{11}^2 \\ H_3 &= (\bar{\mu}_{30} - 3\bar{\mu}_{12})^2 + (3\bar{\mu}_{21} - \bar{\mu}_{03})^2 \\ H_4 &= (\bar{\mu}_{30} - \bar{\mu}_{12})^2 + (\bar{\mu}_{21} - \bar{\mu}_{03})^2 \\ H_5 &= (\bar{\mu}_{30} - 3\bar{\mu}_{12}) \cdot (\bar{\mu}_{30} - \bar{\mu}_{12}) \cdot [(\bar{\mu}_{30} - \bar{\mu}_{12})^2 - 3(\bar{\mu}_{21} - \bar{\mu}_{03})^2] + \\ &\quad (3\bar{\mu}_{21} - \bar{\mu}_{03}) \cdot (\bar{\mu}_{21} - \bar{\mu}_{03}) \cdot [3(\bar{\mu}_{30} - \bar{\mu}_{12})^2 - (\bar{\mu}_{21} - \bar{\mu}_{03})^2] \\ H_6 &= (\bar{\mu}_{20} + \bar{\mu}_{02}) \cdot [(\bar{\mu}_{30} - \bar{\mu}_{12})^2 - (\bar{\mu}_{21} - \bar{\mu}_{03})^2] + \\ &\quad 4\bar{\mu}_{11} \cdot (\bar{\mu}_{30} + \bar{\mu}_{12}) \cdot (\bar{\mu}_{21} - \bar{\mu}_{03}) \\ H_7 &= (3\bar{\mu}_{30} - \bar{\mu}_{12}) \cdot (\bar{\mu}_{30} - \bar{\mu}_{12}) \cdot [(\bar{\mu}_{30} - \bar{\mu}_{12})^2 - 3(\bar{\mu}_{21} - \bar{\mu}_{03})^2] + \\ &\quad (3\bar{\mu}_{12} - \bar{\mu}_{30}) \cdot (\bar{\mu}_{21} - \bar{\mu}_{03}) \cdot [3(\bar{\mu}_{30} - \bar{\mu}_{12})^2 - (\bar{\mu}_{21} - \bar{\mu}_{03})^2] \end{aligned} \quad (3.15)$$

Für die Herleitung der rotationsinvarianten Basis sei an dieser Stelle auf [Hu62] verwiesen. Eine Einschränkung der Moment Invariants ist, dass sie für n-fach-symmetrische Objekte die Invarianz-Eigenschaften nicht halten. So zeigte sich bei Tests während dieser Arbeit mit gleichschenkligen Dreiecken und Quadraten, dass die Invarianz-Eigenschaften verloren gehen.

Angemerkt sei an dieser Stelle, dass gezeigt werden konnte, dass die Moment Invariants nach [Hu62] keine eindeutige orthogonale Basis generieren [FSZ09]. In dieser Hinsicht sind z.B. die *Zernike Momente* [NA08] mächtiger. Sie benötigen jedoch zusätzliche Parameter und die Implementierung ist wesentlich aufwändiger, was den Rahmen dieser Arbeit überschritten hätte.

3.2 Strukturelle Merkmale

Strukturelle bzw. räumliche Merkmale werden aufgrund von Graphen berechnet. Sie ermöglichen einen effektiven Weg, um strukturelle Informationen zu repräsentieren. Ein

Graph wird im Folgenden als $G = (V, E)$ mit V Knoten und E Kanten definiert [Die10]. Er ist ungerichtet, ohne Mehrfachkanten und ohne Selbstverweise innerhalb der Knoten V [GBC⁺09]. Die Knoten und Kanten können durch eine gegebene Metrik gewichtet werden. Als Kantengewichte wird in der digitalen Histopathologie unter anderem die euklidische Distanz zwischen zwei Knoten verwendet [DHS⁺07, DAM⁺08].

Graphen zur Beschreibung der räumlichen Anordnungen von Zellen einzusetzen, wurde bereits in den 90ern vorgeschlagen [GBC⁺09]. Ihre Anwendung auf histologische Schnitte ist hingegen relativ neu [GBC⁺09]. Zu jedem Graph werden verschiedene Indizes und statistische Werte bestimmt. Dazu zählen Merkmale auf Basis der Kantenzahlen bzw. Kantengewichtung und Knotengrade (siehe Abschnitt 3.2.3) eines Graphen. Die Knotenidentifikation beruht dabei in den meisten Applikationen auf den Zentroiden der segmentierten Zellstrukturen [GBC⁺09]. Anwendung finden in histopathologischen Applikationen folgende Graphstrukturen [GBC⁺09]: *Delaunay-Graph* (kurz DG, siehe Abschnitt 3.2.2) [ACSGD10, DHS⁺07, DAM⁺08, Bou08] *Voronoi-Diagramm* [DHS⁺07, DAM⁺08], *Minimum-Spanning-Tree* (kurz MST, siehe 3.2.1) [DHS⁺07, DAM⁺08, Bou08], *O'Callaghan Neighbourhood Graph* [Bou08], *Connected Graph*, *Relative Neighbour Graph* und *k-NN Graph*. Definitionen zu den unterschiedlichen Graphstrukturen können in [Bou08] gefunden werden. Auf die für diese Arbeit relevanten Graphstrukturen wird in den folgenden Abschnitten eingegangen. In Abbildung 3.2 werden MST's, DG's und Voronoi-Diagramme auf histopathologischen Gewebeschnitten (HE-Färbung) des Mammakarzinoms auf zellulärer Ebene mit zunehmender Malignitätsstufe abgebildet. Man kann deutlich die zunehmende Strukturichte der Graphen erkennen.

In [ACSGD10] werden als Erweiterung der oben genannten Graphstrukturen nicht nur die Tumorstrukturen in die Erstellung eines Graphen mit einbezogen, sondern auf Grundlage einer Gewebestruktur approximierenden Segmentierung zusätzlich auch Stroma und Lumen¹ in HE-Färbungen von Darmkrebsgewebebiopsien eingesetzt. Dabei wird ein Delaunay-Graph über die Knoten aller Strukturen berechnet. Die Kanten zwischen verschiedenen Gewebeentitäten (Möglichkeiten) werden farblich markiert, wodurch dieser Ansatz zu seinem Namen *Color-Graph* kommt. Für jeden Kanten- und Knotentyp werden einzeln Merkmale berechnet. Dieser Ansatz wird erfolgreich zur automatischen Darmkrebsdiagnose und für ein automatisches Grading eingesetzt. Im Folgenden werden die in der vorliegenden Arbeit verwendeten Graphstrukturen und die als Merkmale implementierten Statistiken erläutert.

¹Hohlräume

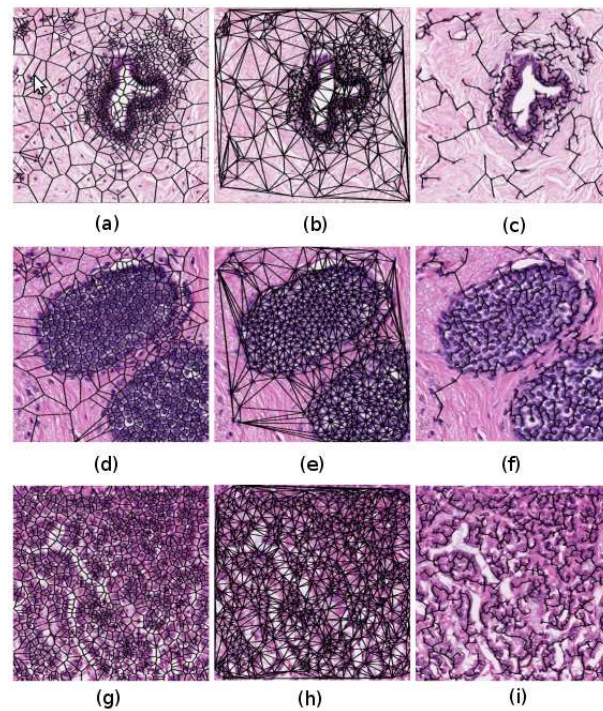


Abbildung 3.2: Beispiel der Graphstrukturen Voronoi-Diagramm (a,d,g), Delaunay-Graph (b,e,h) und Minimum-Spanning-Tree (c,f,g) auf gesundem Gewebe (a-c), niedriggradigem (d-f) und hochgradigem Krebs (g-i) (Quelle: [DAM⁺08]).

3.2.1 Minimum-Spanning-Tree

Ein Baum ist ein Graph G , bei dem zwischen zwei Knoten v genau eine Kante e existiert. Dabei besteht zu allen Knoten V eine Kante e mit $|V| = |E| - 1$ [EP02]. Es existiert in einem Baum demnach eine Kante weniger als Knoten existieren. Knoten, von denen mehr als eine Kante ausgehen, werden als innere Knoten bezeichnet (Grad > 1). Knoten die nur eine Kante enthalten werden Blätter genannt (Grad = 1) [Die10].

Ein Spannbaum (engl. *Spanning-Tree*) ist in der Graphentheorie ein Teilgraph eines ungerichteten Graphen G , der ein ungerichteter Baum ist und alle Knoten V dieses Graphen enthält. Ein *Spanning-Tree* heißt *Minimum-Spanning-Tree*, wenn kein anderer *Spanning-Tree* im selben Graphen G mit geringerem Gewicht, der Summe aller Kantengewichte über die Knoten V , existiert [Bou08]. Dies bedeutet gleichzeitig, dass durchaus unterschiedliche MST's innerhalb eines Graphen vorliegen können, solange sie alle die gleiche minimale Summe der Kantengewichte haben. Von daher wird in der Anwendung dieser Arbeit für die Berechnung des MST's über die selbe Knotenmenge des gleichen Bildmusters immer der gleiche Startknoten (erster in der Liste) gewählt. Zur Veranschaulichung wird in Abbildung 3.3 ein MST als Teilgraph eines *Delaunay-Graphen* (siehe folgender Abschnitt) mit angegebenen Kantengewichten dargestellt.

Grundsätzlich werden in der Literatur zwei unterschiedliche Algorithmen zur automatischen Bestimmung eines MST's, von Kruskal J.B. [Kru57] und Prim R. [Pri57], über die Knoten V beschrieben. In der vorliegenden Arbeit findet der Algorithmus von Prim R. Anwendung, da er mit $O(E * \log(V))$ einen geringeren Aufwand als [Kru57] mit $O(E + X * \log(V))$ hat. X ist hierbei die Anzahl der Kanten, die kürzer als die längste Kante im MST sind. Die Funktionsweise des Algorithmus von Prim R. wird in Algorithmusdarstellung 3 beschrieben. In Abbildung 3.2 wird ein Beispiel des *Minimum-Spanning-Trees* über eine gegebene Zellmenge eines histopathologischen Schnittes dargestellt. Zur Bestimmung des MST's wird in dieser Arbeit auf die Implementation des Algorithmus von Prim R. in der *Bio Informatics -Toolbox* [Mata] zurückgegriffen.

1. Initialisiere die Knoten V , wähle einen initialen Startknoten v_{start} und nenne ihn erreichbar
2. v_{start} wird zu v_a
3. Bestimme den Knoten v_b , der v_a am nächsten liegt (mit geringstem Kantengewicht) und noch nicht erreichbar ist
4. Nenne v_b erreichbar und ziehe eine Kante e zwischen v_a und v_b
5. v_b wird zu v_a
6. Wiederhole Schritte 3 bis 5, bis alle Knoten erreichbar sind.

Algorithmus 3: Algorithmus von Prim R. [Pri57] zur Bestimmung eines MST's (Minimum-Spanning-Tree)

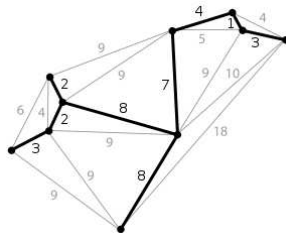


Abbildung 3.3: Beispiel eines Minimum-Spanning-Tree innerhalb eines Delaunay-Graphen mit gegebenen Kantengewichten (Quelle: [MST]).

3.2.2 Delaunay-Graph

Die *Delaunay-Triangulation* bzw. der resultierende *Delaunay-Graph* (im Folgenden kurz als DG bezeichnet) ist der duale Graph des *Voronoi-Diagrammes* [Bou08]. Die *Delaunay-*

Triangulation ist ein allgemein eingesetztes Verfahren zur Bestimmung einer Triangulation über eine gegebene Knotenmenge V [Lam]. Das Verfahren beschreibt ein iteratives Vorgehen, wobei ein erstelltes Dreieck zwischen Knoten V akzeptiert wird, wenn der Umfang des Dreiecks keine anderen Knoten enthält, als die Knoten, die zum erstellten Dreieck gehören [Bou08]. Dies führt dazu, dass die minimalen Winkel der Dreiecke maximiert werden und so ein Graph G mit möglichst minimaler Anzahl von Dreiecken entsteht. In Abbildung 3.4 wird ein Delaunay-Graph von zehn Knoten (schwarze Punkte) mit erfüllter Umfangsbedingung (siehe Kreise) dargestellt. In Abbildung 3.2 wird ein Beispiel der *Delaunay-Triangulierung* über eine gegebene Zellmenge eines histopathologischen Schnittes visualisiert. In dieser Abbildung lässt sich erkennen, dass mit höherer Malignitätsstufe ein dichteres Netz von Dreiecken entsteht.

Die Berechnung des DG erfolgt in der vorliegenden Arbeit durch die Anwendung des in der Standardbibliothek von MATLAB vorhandenen Algorithmus zur Bestimmung des Delaunay-Graphen². Diese Methode verwendet den *QHull-Algorithmus* [BDH96]. Jeder 2D-Knoten des Graphen mit den Koordinaten x und y wird um eine dritte Dimension z mit $z = x^2 + y^2$ erweitert. Um diese 3D-Knoten wird die konvexe Hülle (mit Hilfe des *QHull-Algorithmus*), bestehend aus Dreiecken, erstellt. Dabei werden die nach außen zeigenden Dreiecksnormalen bestimmt. Daraufhin werden alle Dreiecke, deren Normalenvektoren eine negative z -Koordinate haben, wieder in die ursprüngliche xy -Ebene projiziert [Lam]. In dieser Ebene liegt der 2D-Delaunay-Graph vor. Für 2D-Knoten, die auf 3D-Knoten erweitert werden, liegt der Aufwand zur Erstellung eines DG's in der Praxis bei $O(V \log V)$ [Lam].

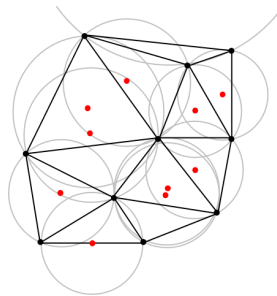


Abbildung 3.4: Beispiel eines Delaunay-Graphens mit erfüllter Umfangsbedingung (siehe Kreise) und Kreismittelpunkten (rote Punkte) (Quelle: [DGW]).

3.2.3 Graphmerkmale

In [GBC⁺09] und [Bou08] wird ein ausführlicher Überblick zu möglichen Merkmalen für Graphstrukturen gegeben. Im Rahmen der vorliegenden Arbeit wird nur eine Auswahl verwendet, welche an dieser Stelle vorgestellt werden soll.

Die Definitionen der im Folgenden beschriebenen und implementierten Merkmale beziehen sich, so weit nicht anders angegeben, auf die Ausführungen aus [Bou08]. Als

²delaunay()

Grundlage für die Bestimmung der Merkmale werden zunächst die Adjazenzmatrix \mathbf{A} und die Distanzmatrix \mathbf{B} berechnet. Ein Eintrag a_{ij} in der Adjazenzmatrix (auch Nachbarschaftsmatrix genannt) \mathbf{A} eines Graphen ist die Anzahl der Kanten $|E|$, die zwischen v_i und v_j bestehen. Da es sich in dieser Anwendung um ungerichtete Graphen ohne Mehrfachkanten und Selbstverweise handelt, besteht \mathbf{A} ausschließlich aus binären Einträgen mit 0 (keine Verbindung) oder 1 (Verbindung vorhanden).

Als Gewicht der Kanten E wird in dieser Anwendung der euklidische Abstand (siehe Formel(2.8)) zwischen den Koordinaten der verbundenen Knoten verwendet. Auf Grundlage dessen lässt sich eine weitere Matrix \mathbf{B} (Distanzmatrix) mit Einträgen b_{ij} definieren, welche die euklidischen Distanzen zwischen den Knoten v_i und v_j enthält.

Beide Matrizen sind symmetrisch und die Diagonalen sind 0. Auf Grundlage dieser Matrizen lassen sich verschiedene Merkmale einfach bestimmen. Merkmale wie die Anzahl der Kanten N_e und die Anzahl der Knoten N_v lassen sich aus den Dimensionen der Distanzmatrix bzw. Adjazenzmatrix berechnen.

Am häufigsten werden verschiedene Statistiken zu den Kantengewichtungen und dem Grad (engl. degree) der Knoten für die Graphen in der Histopathologie berechnet [GBC⁺09]. Die durchschnittliche Kantenlänge e_μ wird durch

$$e_\mu = \frac{\sum_{ij} b_{ij}}{\sum_{ij} a_{ij}} \quad (3.16)$$

bestimmt. Zusätzlich werden die Standardabweichung der Kantenlängen e_σ und das Verhältnis der minimalen Kantenlänge zur maximalen Kantenlänge e_{mima} berechnet, wie es in [DAM⁺08] der Fall ist, um die Spannweite der Distanzen zu erfassen.

Zusätzlich wird in dieser Anwendung außerdem der Variationskoeffizient e_{vc}

$$e_{vc} = \frac{e_\sigma}{e_\mu} \quad (3.17)$$

bestimmt, um ein zusätzliches, von den Größeninformationen der Tumore unabhängigeres Maß zur Bestimmung der Variabilität in den Strukturen zu erhalten. Die gleiche deskriptive Statistik wird für den Grad der Knoten berechnet. Der Grad eines Knotens gibt die Anzahl der Kanten mit ihm verbundener Knoten an. Zur Bestimmung der Grad-Statistiken der Graphen wird die Adjazenzmatrix \mathbf{A} verwendet. Der Grad deg jedes Knotens wird durch die Summe der Zeilen bzw. Spalten von \mathbf{A} bestimmt.

$$deg(v_i) = \sum_{j=1}^{N_v} a_{ij}, i = 1 \cdots N_v \quad (3.18)$$

Dabei gibt N_v die Anzahl der Knoten an. Dies resultiert in einen Vektor

$$\mathbf{deg} = [deg(v_1), \cdots, deg(v_{N_v})] \quad (3.19)$$

zu dem der durchschnittliche Grad deg_μ , die Standardabweichung deg_σ und das Verhältnis von Minimum zu Maximum deg_{mima} bestimmt werden können. Die Anzahl der Knoten N_v , die sich einfach aus den Zeilen bzw. Spalten der Adjazenzmatrix bestimmen lassen, wird ebenfalls als Merkmal verwendet. In [Bou08] wird ein weiteres Maß eingeführt, das die Information aus Grad und Kantenlänge zusammenbringt. Der gewichtete Grad deg_w wird als

$$deg_w(v_i) = \sum_{j=1}^{N_v} b_{ij}, i = 1 \cdots N_v \quad (3.20)$$

definiert und berechnet die Summe der Kantenlänge eines Knotens zu den mit ihm verbundenen Knoten. Aus dem Vektor

$$\mathbf{deg}_w = [deg_w(v_1), \dots, deg_w(v_{N_v})] \quad (3.21)$$

wird der durchschnittliche gewichtete Grad $deg_{w\mu}$ des Graphen berechnet. Neben den Merkmalen, welche den Grad oder die Kantenlänge der Graphen beschreiben, gibt es noch weitere Indizes wie die zyklische Nummer (engl. cyclomatic number) und den *Randić Index* [GBC⁺09]. Die zyklische Nummer wird als

$$C = N_e + N_v + 1 \quad (3.22)$$

definiert, wobei N_e die Anzahl der Kanten ist. Die *cyclomatic number* C gibt an, wie viele Kanten entfernt werden müssen, damit der Graph frei von Zyklen wird. C wird durch die Anzahl der Kanten zu $C_n = \frac{C}{N_e}$ normalisiert. Die zyklische Nummer macht nur bei Graphen Sinn, die intrinsisch in ihrer Struktur Zyklen enthalten. Von daher wird dieses Merkmal nur für den Delaunay-Graph berechnet.

Der *Randić-Index* wird in der Chemie als topologischer Deskriptor von Molekülverbindungen verwendet. Er ist sensitiv für das Aussehen eines Graphes. Er wird als

$$RI = \sum_{i \neq j, a_{ij} \neq 0} \frac{1}{\sqrt{deg(v_i) \cdot deg(v_j)}} \quad (3.23)$$

definiert und beschreibt das Verhältnis des Knotengrades in der Nachbarschaft eines Knotens. Durch $RI_n = \frac{RI}{(N_e/deg_\mu)}$ wird eine Normalisierung des Wertes erreicht.

3.3 Räumliche Relationen - RCC8-Histogramm

Aus Segmentierung hervorgehende Bildobjekte R erlauben es räumliche Beziehungen/Relationen als Merkmal einzusetzen. Dies wird in der medizinischen Bildverarbeitung oft vorgeschlagen, wobei es gleichzeitig an konkret umsetzbaren Implementationen mangelt

[MMBG04]. In [Blo05] werden zusammenfassend verschieden Ansätze zur räumlichen Analyse von Bildobjekten vorgestellt, welche primär unintuitiven Interpretationen unterliegen. Nach Recherchen wurde ein Konzept, das eine intuitive Quantifizierung von räumlichen Relationen im Bezug auf die Bildanalyse von Gewebebildern ermöglicht und zudem bereits in der digitalen Histopathologie eingesetzt wurde, in [Bou08] gefunden. Aufgrund der erfolgreichen Anwendung in der digitalen Histopathologie wird dieser Ansatz den Methoden aus [Blo05] vorgezogen.

Boucheron setzt das *Region Connection Calculus 8* (RCC8) [RCC92] aus dem *Qualitative Spatial Reasoning* (QSR) ein. QSR geht in diesem Bezug der Frage nach: Wenn Bildobjekte R^a, R^b und R^c mit Relation $Rel1(R^a, R^b)$ und Relation $Rel2(R^a, R^c)$ vorliegen, was kann aus dieser Information über die Relation $Rel3$ zwischen R^a und R^c abgeleitet werden? In [Bou08] werden paarweise Relationen zwischen zwei Bildobjekten betrachtet. Die folgenden Definitionen bezüglich der RCC8-Relationen basieren auf den Angaben aus [Bou08] und [RCC92]. Das RCC8 definiert acht Relationen

- DC (disconnected)
- EC (equally connected)
- PO (partial overlapping)
- TPP (tangential proper part)
- TPPI (tangential proper part inverse)
- NTPP (nontangential proper part)
- NTPPI (nontangential proper part inverse)
- EQ (equal)

zwischen zwei Bildobjekten R^a und R^b , welche in Abbildung 3.5 visualisiert werden. DC beschreibt getrennte Bildobjekte, EC am äußeren Rand liegende Bildobjekte, PO sich überlappende Bildobjekte, TPP bzw. TPPI am inneren Rand liegende Bildobjekte, NTPP bzw. NTPPI im inneren liegen Bildobjekte und EQ Bildobjekte mit gleicher Ausdehnung und Position. Zur Bestimmung dieser Relationen wird ein Konnektionsoperator CON , der individuell definiert werden kann, verwendet. Dieser muss sowohl reflexiv $\forall[CON(R^a, R^a)]$ als auch symmetrisch $\forall R^a \forall R^b [CON(R^a, R^b) \rightarrow CON(R^b, R^a)]$ sein.

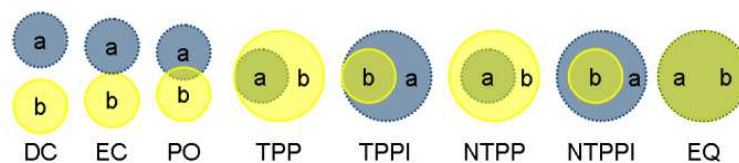


Abbildung 3.5: Relationen des Region Connection Calculus (RCC8) zwischen zwei Objekte (Quelle: [Bou08]).

Das RCC8 basiert auf einer mathematischen Grundlage und wird nach empirischen Studien als kognitiv adäquat in Bezug auf das räumliche konzeptuelle Verständnis von Menschen beschrieben [Bou08]. Boucheron [Bou08] verwendet das RCC8 als Erste auf

Pixelebene im Kontext von histopathologischen Schnitten von Brustkrebs zur Charakterisierung der zellulären Ebene.

Dabei wird zunächst für jedes Bildobjekt R die konvexe Hülle bestimmt. Die Implementation wird in [Bou08] auf Grundlage von Mengenoperationen zwischen den Pixeln der einzelnen Bildobjekte definiert. Der in [Bou08] angegebene Algorithmus zur Bestimmung der Relationen wird in Algorithmusdarstellung 4 beschrieben.

```

IF  $R^a \subset R^b$ 
    IF  $p(R^a) \cap p(R^b) = \emptyset \Rightarrow$  NTPP( $R^a, R^b$ )
    ELSE  $\Rightarrow$  TPP( $R^a, R^b$ )
ELSE IF  $R^b \subset R^a$ 
    IF  $p(R^a) \cap p(R^b) = \emptyset \Rightarrow$  NTPPI( $R^a, R^b$ )
    ELSE  $\Rightarrow$  TPPI( $R^a, R^b$ )
ELSE IF  $R^a \equiv R^b \Rightarrow$  EQ( $R^a, R^b$ )
ELSE IF  $R^a \cap R^b = \emptyset$ 
    IF  $\min(d(R^a, R^b)) > 1 \Rightarrow$  DC( $R^a, R^b$ )
    ELSE  $\Rightarrow$  EC( $R^a, R^b$ )
ELSE IF  $R^a \cap R^b \neq \emptyset \Rightarrow$  PO( $R^a, R^b$ )

```

Algorithmus 4: Bestimmung der RCC8-Relationen nach Boucheron [Bou08].

Die Kontur eines Bildobjektes wird durch $p(\cdot)$ repräsentiert. Der Term $d(\cdot)$ definiert den euklidischen Abstand. Liegt Bildobjekt R^a innerhalb von Bildobjekt R^b ($R^a \subset R^b$) wird zunächst geprüft, ob sich die Konturen der beiden Bildobjekte berühren ($p(R^a) \cap p(R^b)$). Ist dies nicht der Fall, wird die Relation NTPP festgelegt. Andernfalls wird angenommen, dass sich das Bildobjekt R^a am inneren Rand von Bildobjekt R^b befindet womit die Relation TPP festgelegt wird. Gleiches gilt wenn Bildobjekt R^b innerhalb von Bildobjekt R^a liegt ($R^b \subset R^a$). Für diesen Fall werden die invertierten Relationen NTPPI und TPPI angewendet.

Liegen die vorher genannten Relationen nicht vor, wird geprüft, ob sich die beiden Bildobjekte in Position und Größe gleichen ($R^a \equiv R^b$). Tritt diese Bedingung ein, wird die Relation EQ festgelegt. Andernfalls wird getestet ob sich die beiden betrachteten Bildobjekte nicht überlagern ($R^a \cap R^b = \emptyset$). Ist dies der Fall, wird geprüft ob der euklidische Abstand zwischen den beiden Bildobjekten größer als 1 ist und die Objekte dementsprechend räumlich voneinander getrennt sind womit die Relation DC vorliegt. Andernfalls wird die Relation EC angenommen.

Sollten alle vorher genannten Relationen nicht zutreffen und R^a und R^b sich überlagern, wird die Relation PO bestimmt.

In [Bou08] wird zur Unterscheidung zwischen PO und TPPI bzw. TPP ein Schwellwert von 5 Pixeln angegeben, um eine Verzerrung bei schlechter Segmentierung zu vermeiden. Außerdem wird die Einschränkung gemacht, dass die Relationen in der Reihenfolge, in der sie dem Algorithmus präsentiert werden, bestimmt werden. Daraus resultiert, dass bei Vertauschung von R^a und R^b in den Relationen ein unterschiedliches Ergebnis auftritt. Die Anzahlen der berechneten Relationen werden in einem Histogramm (siehe Abbildung 3.6), mit jeweils einem Bin für jede Relation, zusammengefasst. In [Bou08] wurde dieses Histogramm auf zellulärer Ebene beispielsweise zwischen den Gewebeatitäten Zellen(Zellkern und Cytoplasma) und Stroma berechnet. Der an dieser Stelle vorgestellte Algorithmus zur Berechnung des RCC8-Histogramms wird für die Anwendung der vorliegenden Arbeit angepasst und erweitert, worauf in Abschnitt 5.5.1 näher eingegangen wird.

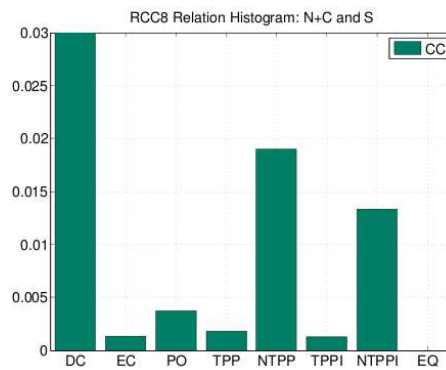


Abbildung 3.6: RCC8-Histogramm zwischen Zellen(Zellkern und Cytoplasma) und Stroma auf Grundlage von *Connected-Components* (Quelle: [Bou08]).

Kapitel 4

ExPrimage als Anwendungsszenario

In diesem Kapitel soll das Projekt ExPrimage als Anwendungsszenario der vorliegenden Arbeit erläutert werden. Zunächst wird die grundlegende Motivation und das Ziel des Projektes dargestellt. Daraufhin wird auf die Entstehung der Bilddaten und die Datenlage im Projekt ExPrimage eingegangen (siehe Abschnitt 4.1). Die Analyseschritte und Prozessabläufe, an welche die vorliegende Arbeit anknüpft, werden in Abschnitt 4.2 beschrieben. Somit soll verständlich werden, in welchem Kontext die in Kapitel 3 beschriebenen und implementierten Merkmale sowie die in Kapitel 5 erläuterten Analysen angewendet werden.

Der momentane Ablauf in der klinischen Praxis zur Erstellung einer Prognose bei der Diagnose eines bösartigen Mammakarzinoms gestaltet sich folgendermaßen: Nach der Entfernung des Tumors durch einen chirurgischen Eingriff muss zur dauerhaften Bekämpfung des Krebses unter anderem die Frage nach der Tumorausbreitung bzw. der Ist-Zustand des Tumors geklärt werden [RP05]. Die dazu nötigen Untersuchungen des Gewebes beruhen auf unterschiedlichen Stichproben (ca. 5) auf zellulärer Ebene des histopathologischen Gewebeschnittes (siehe Abschnitt 4.1).

Anhand eines Grading-Scores¹ erstellt die Pathologie eine Prognose. Dabei gibt ein höherer Grading-Score eine höhere Malignitätsstufe des Tumors an [RP05]. Das Grading ist unter anderem abhängig von den Erfahrungen des Pathologen und nicht immer eindeutig. So kommt es zu einer nicht zu unterschätzenden *Interobserver-Variabilität*², die zeigt, dass der Grading-Score nur mäßig reproduziert werden kann [RPK⁺95, al.05]. Oft basieren die Diagnosen zwar auf geschulten, aber auch subjektiven Eindrücken des einzelnen Domänenexperten. Abhängig von den Prognosen der Pathologie werden weitere Therapiemaßnahmen eingeleitet.

¹Zur Beurteilung der Malignität des Tumors wird das sogenannte Bloom-Richardson-Grading-System bzw. dessen Erweiterung von Lässer et. al. [RP05] eingesetzt. Das Grading setzt sich aus verschiedenen, primär morphologischen, Aspekten zusammen und wird zusammenfassend in G1 (gut differenziert), G2 (mäßig differenziert) und G3 (schlecht differenziert) unterteilt. „Gut differenziert“ bedeutet in diesem Zusammenhang, dass trennbare Zellstrukturen zu erkennen sind. Im Gegensatz dazu steht „schlecht differenziert“ für eher chaotische Zellstrukturen.

²unterschiedliche Ergebnisse eines Beobachtungsverfahrens bei Einsatz verschiedener Beobachter

Am FIT wird mit einer Auswahl von 93 Patientenfällen gearbeitet. Die Prognosen stimmen dabei nicht immer mit dem Krankheitsverläufen der Patienten überein (so genannte *Matched-Pairs*³). Die durch Stichproben der zellulären Ebene untersuchten Faktoren waren an diesen Stellen offensichtlich nicht ausreichend.

Vorhandene Daten, wie Informationen auf der Gewebeebene, die für den Menschen nicht unmittelbar fassbar sind, bleiben weitestgehend ungenutzt. Um diese Daten diagnostisch nutzbar zu machen, kann der Computer dem Arzt durch Quantifizierungen und Vorschläge von möglichen Zusammenhängen innerhalb der Datenbasis durch Maschinelles Lernen (siehe Abschnitt 2) als Unterstützung dienen. Hypothesen der Domänenexperten können durch Quantifizierungen und Korrelationsanalysen der Daten unterstützt oder entkräftet werden.

Im Rahmen des Projekts ExPrimage sollen am FIT Strategien entwickelt werden, die den Domänenexperten durch maschinelle Verarbeitung und Korrelationsanalysen mit dem Krankheitsverlauf in der Gewinnung neuer Erkenntnisse unterstützen. Die Tumore sollen durch computergestützte Analysen auf der Gewebeebene quantitativ beschrieben werden. Dadurch soll eine bessere postoperative Diagnose und Prognose für die Patienten und damit zusammenhängend individuelle *adjuvante Therapien*⁴ ermöglicht werden.

Die Durchführung des Projektes verläuft in enger Zusammenarbeit mit den Domänenexperten (Pathologen und Mikrobiologen) der *Pathologie West* in Hamburg. Durch das kontinuierliche Feedback der Domänenexperten zu den erzielten Zwischenergebnissen werden die angewendeten Arbeitsschritte und Methoden stets - in Form eines qualitätsgesicherten Prozesses - reflektiert.

Neben Bildanalysen von Gewebeschnitten am FIT werden im Verbundprojekt ExPrimage unter anderem Methoden zur Genexpressionsanalyse mittels Real-Time-PCR und Raman-Spektroskopie angewendet.

4.1 Datenlage

Es stehen insgesamt 93 Patientenfallbeispiele zur Verfügung, die in Textform und Bilddaten erfasst sind. Für die Analysen der vorliegenden Arbeit spielen Bilddaten eine zentrale Rolle. An dieser Stelle soll die Entstehung und Bedeutung der verschiedenen Bilddaten dargestellt werden, um ein Verständnis für die Zusammenhänge der eingesetzten Bildmerkmale zu schaffen. Anhand der Abbildung 4.1 wird der Verlauf zur Entstehung der Bilddaten zusammengefasst und soll im Folgenden erläutert werden.

Zunächst wird der Tumor in einem chirurgischen Eingriff aus der Brust entfernt. Danach steht ein OP-Präparat zur Verfügung. In der Pathologie erfolgt daraufhin der Zuschnitt des OP-Präparates in abgetrennte Schichten des organischen Gewebes mit Hilfe eines Mikrotoms⁵.

³Eine Form der Analyse, in dem jeder Fall einer Studiengruppe mit einem Fall aus einer Vergleichsgruppe (Kontrollgruppe) auf Basis eines Faktors (matching-factor) verglichen wird (z.B. der Grading-Score)

⁴ergänzende bzw. unterstützende Therapiemaßnahmen (z.B. Chemotherapie oder Hormontherapie)

⁵Schneidegerät, mit dem sehr dünne Schnittpräparate erstellt werden

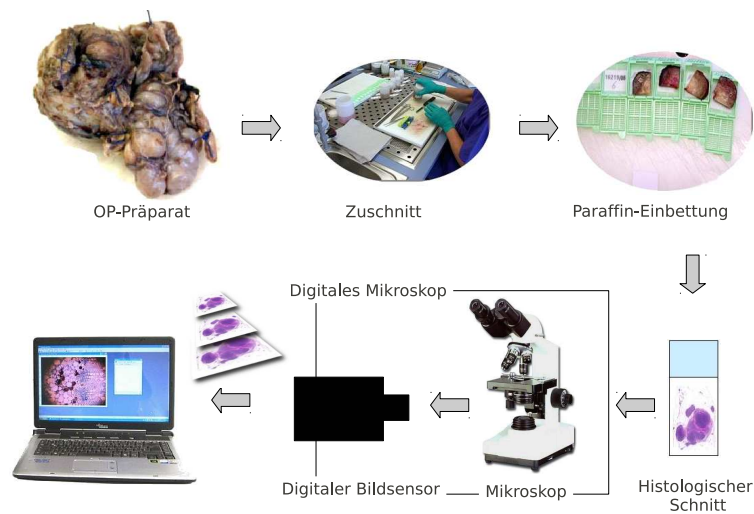


Abbildung 4.1: Entstehung der digitalen histologischen Schnitte im Projekt ExPrimage (Quelle der einzelnen Bilder: [IMG]).

Danach werden die Gewebeschnitte in Paraffin eingebettet, wodurch sie haltbar gemacht werden. Die verschiedenen Gewebeschnitte werden mit speziellen Farbstoffen bzw. Markern angefärbt (Färbungen). Je nach Auswahl des Markers werden anatomische Strukturen oder funktionell relevante Strukturen im Gewebe hervorgehoben.

Der histologische Schnitt wird in ExPrimage vom Pathologen unter einem digitalen Mikroskop untersucht. Ein Computer steuert das Mikroskop und den Scanvorgang des Bildes. Die Färbungen werden in regelmäßigen Abständen gekachelt, in unterschiedlichen Auflösungsstufen aufgenommen und als Farbbild f_{rgb} der Färbung gespeichert.

Während zellorientierte Analysen eine 20- [ACSGD10] oder 40-fache [DHS⁺07], [DAM⁺08] Vergrößerung ausgewählter Bereiche des Gewebes verwenden, wird in ExPrimage eine 10-fache-Vergrößerung des gesamten histologischen Schnittes eingesetzt (siehe Abbildung 1.1). Die digitalisierten Gewebeschnitte bilden die Datenbasis der Bildanalyse im Projekt ExPrimage.

Die Gewebeschnitte stammen aus unterschiedlichen Bereichen des Tumolvolumens. Für jeden Fall stehen drei strukturelle Marker und drei verschiedene funktionelle Marker zur Verfügung. Strukturelle Marker heben die anatomische Struktur des Tumors bzw. der Tumorzellen hervor. Dazu gehören Vimentin (VIM, f_{rgb}^{vim}), Keratin (AE1AE3, f_{rgb}^{ae1ae3}) und Hämatoxylin-Eosin (HE, f_{rgb}^{HE}). In Abbildung 4.2 sind Beispiele korrespondierender struktureller Färbungen eines Falles mit annotierter Färbung bzw. Gegenfärbung abgebildet. In Abbildung 4.2 wird deutlich, dass AE1AE3 die Tumorzellstruktur (braun) mit hohem Kontrast hervorhebt, was in Kapitel 5 im Hinblick auf die Erweiterung der vorverarbeitenden Schritte relevant wird.

HE ist die am weitesten verbreitete und verwendete Färbung in der digitalen Histopathologie [DY05]

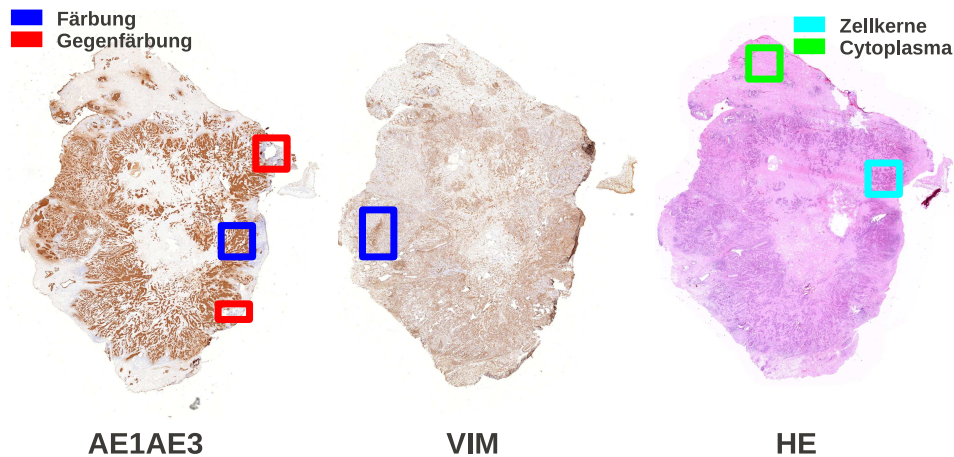


Abbildung 4.2: Zur Verfügung stehende strukturelle Färbungen (links: AE1AE3, mitte: VIM, rechts: HE).

[GBC⁺09]. In ihr werden Zellkerne bläulich (Hämatoxylin) und das Zellplasma rötlich (Eosin) angefärbt. Sie enthält im Verhältnis zu den anderen Färbungen die umfangreichste pathologisch relevante Information.

Abgesehen von der HE-Färbung ist die Färbung braun und die Gegenfärbung lila. Dies gilt ebenfalls für die funktionellen Marker, welche funktionell relevante Strukturen hervorheben.

Der funktionelle Marker CD45 (f_{rgb}^{cd45}) (benannt nach dem CD45-Antigen) reagiert mit bzw. markiert, in erster Linie Leukozyten⁶. Durch ihn werden Strukturen, in denen Entzündungsprozesse ablaufen, sichtbar.

Die Färbungen von Östrogen (ER, f_{rgb}^{er}) und Progesteron (PR, f_{rgb}^{pr}) markieren die Expression der Hormonrezeptoren⁷, welche eine wichtige Rolle im Krankheitsverlauf spielen. Sie geben Aufschluss über die Aggressivität und Therapiemöglichkeiten des Tumors. In [RBH⁺08] konnte durch computerbasierte quantitative Analysen ein Schwellwert ermittelt werden, der bestimmt, ab wann die Expression der Hormonrezeptoren für eine Therapie von Relevanz ist.

In dieser Arbeit wird in einem weiteren Schritt die Verteilung (siehe Abschnitt 5.5), sowohl der Expression der Hormonrezeptoren, als auch des Entzündungsgeschehens innerhalb des Tumors untersucht.

Abbildung 4.3 zeigt korrespondierende Beispielbilder zu funktionellen Färbungen mit markierter Färbung bzw. Gegenfärbung.

Eine Fusion der verschiedenen Marker bzw. Färbungen (engl. *stainings*), ein *Multi-Staining-Ansatz* zu Korrespondenzanalysen auf Grundlage von räumlichen Assoziationen, bietet die Möglichkeit einer wesentlich umfangreicheren Basis für die Gewebeanalyse.

⁶Zellen, die Teil der Immunabwehr sind.

⁷Die Wirkung eines spezifischen Hormons wird durch den Hormonrezeptor der Zellen übermittelt.

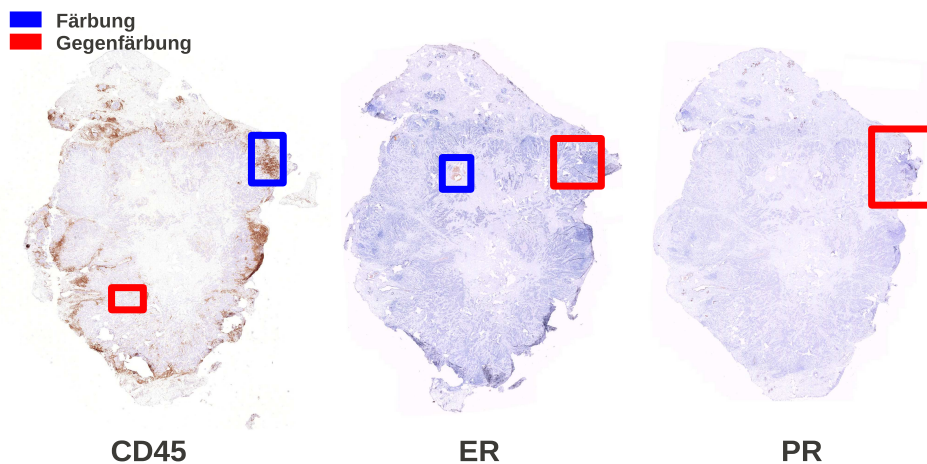


Abbildung 4.3: Zur Verfügung stehende funktionale Färbungen (links: CD45, mitte: ER, rechts: PR).

Aufgrund der Informationsdetails und -menge ist diese Leistung human kognitiv nicht mehr zu bewältigen. Diese färbungsübergreifende Gewebeanalyse kann nur maschinell sinnvoll bearbeitet werden (siehe dazu folgenden Abschnitt 4.2).

Zusätzlich zu den vorhandenen Bilddaten sind in Textform wichtige Diagnosen und personenbezogene Daten zu dem klinischen Verlauf jedes Patientenfalles in tabellarischer Form dokumentiert (klinische Daten⁸). Dazu gehören neben personenbezogenen Daten wie „Alter bei OP“ auch diagnostische:

- „HER2/neu-positive⁹“,
- das Grading,
- der Hormonrezeptorstatus (Detektion ER > 10% des Gewebes, Detektion PR > 10% des Gewebes),
- TN (Tumorgöße und Tumorart) und
- LVR (Grad der Lymphknoteninvasion, Grad der Veneninvasion und Vorhandensein von verbleibendem Tumor).

Diese Daten bilden eigene Merkmalsgruppen im überwachten Lernverfahren *HLVQ* (siehe Abschnitt 2.1), das im Rahmen von ExPrimage entwickelt wurde. Durch Zusammenführung von klinischen Parametern und Bildanalysen sollen Korrespondenzen zum Krankheitsverlauf gefunden werden.

Zur Überprüfung der Einflüsse steht der *Follow-Up-Status* (gesund (*healthy*), wiedererkrankt (*relapse*), verstorben (*deceased*)) für jeden Patientenfall zur Verfügung, der als Klassenzugehörigkeit im *HLVQ*-Verfahren und der *Variation of Information* (siehe Ab-

⁸diese Faktoren wurden auf der internationalen Brustkrebskonferenz in Sankt Gallen identifiziert (<http://www.oncoconferences.ch/dynasite.cfm?dsmid=98911>)

⁹hemmt den programmierten Zelltod (Apoptose), was für eine schlechte Überlebensprognose spricht

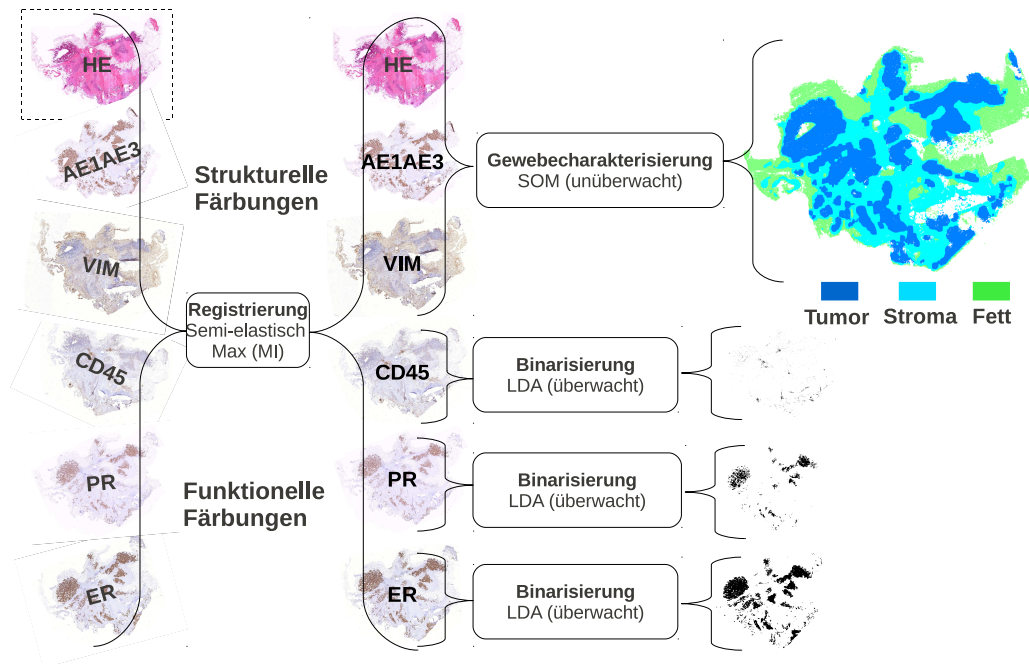


Abbildung 4.4: Schematische Abbildung der bereits integrierten Bildanalysen im Projekt ExPrimage: Zunächst werden die unterschiedlichen Färbungen durch einen Registrierungsprozess zueinander ausgerichtet. Danach erfolgt auf Grundlage der strukturellen Färbungen eine Gewebecharakterisierung durch die SOM. Die Färbungen der funktionellen Marker werden durch die LDA binarisiert.

schnitt 5.6) eingesetzt wird. Der Datensatz teilt sich auf in 50 gesunde, 7 wiedererkrankte und 36 verstorbene Patientenfälle.

Des Weiteren stehen Beobachtungen/Annotationen des Domänenexperten (siehe beiliegende DVD-ROM (Anhang E)) bezüglich verschiedener Ausbreitungsmuster der Tumore in tabellarischer Form zur Verfügung. Aus ihnen wurden unter anderem die Anforderungen für die entwickelnden Merkmale in Zusammenarbeit mit den Domänenexperten erstellt (siehe Abschnitt 5.1).

4.2 Vorhandene Module und Abläufe

Diese Arbeit knüpft an verschiedene bereits implementierte und integrierte Arbeitsabläufe und Bildanalyse-schritte an. Anhand der schematischen Abbildung 4.4 sollen die bereits implementierten Module in ExPrimage erläutert werden.

Registrierung Die histologischen Schnitte werden aus einem Tumolvolumen entnommen (siehe voriger Abschnitt). Die verschiedenen 2D-Färbungen aus unterschiedlichen Ebenen des Tumolvolumens sind nicht korrekt zueinander ausgerichtet. Um färbungs-

übergreifende Bildanalysen im Sinne eines *Multi-Staining-Ansatzes*, durchführen zu können, müssen diese Unterschiede in den geometrischen Ausrichtungen aufgehoben werden. Somit soll eine anatomisch korrekte Überlagerung der Färbungen erreicht werden. Dieser Prozess wird Registrierung genannt [Han09]. Er ist Teil der vorverarbeitenden Schritte (siehe Abschnitt 2.3).

Bei einer Registrierung werden grundsätzlich zwei Bilddatensätze betrachtet: Der Referenzdatensatz *Ref* und der auf *Ref* auszurichtende / zu registrierende *Temp*, Templatdatensatz. Gesucht wird nach einer möglichst optimalen Transformation $T : Temp \rightarrow Ref$, sodass der transformierte Templatdatensatz der anatomischen Position des Referenzdatensatzes entspricht. Die Transformation T wird dementsprechend optimiert.

Im Projekt ExPrimage wird die HE-Färbung aufgrund des höchsten Informationsgehaltes als Referenz *Ref* verwendet. Alle Färbungen werden zunächst in Grauwertbilder transformiert, auf denen danach eine Histogrammspreizung [BB05] zur Kontrastverstärkung angewendet wird.

$$MI(f_{gray}^a, f_{gray}^b) = H(f_{gray}^a) + H(f_{gray}^b) - H(f_{gray}^a, f_{gray}^b) \quad (4.1)$$

Zur Registrierung der Gewebeschnitte wird die Maximierung der *Mutual Information* (kurz MI, siehe Formel (4.1)) [JPWPV03] eingesetzt. Die MI basiert auf der Entropie $H(\cdot)$ und Verbundentropie $H(\cdot, \cdot)$ (siehe Abschnitt 2.4.5) der Pixelintensitäten der Grauwertbilder (f_{gray}^a und f_{gray}^b). Eine hohe Entropie der Bilder deutet darauf hin, dass grundsätzlich genügend Information in beiden Bildern vorhanden ist, während eine geringe Verbundentropie aussagt, dass beide Bilder eine hohe Informationsabhängigkeit haben. Hohe Werte der MI deuten dementsprechend auf eine gute Registrierung der beiden Bilder hin. Mittels eines Gradientenabstiegverfahrens wird die maximale MI zwischen den beiden Bildern ermittelt. Alle Gewebeschnitte liegen in gleicher Vergrößerungsstufe vor. Die Suche nach der korrekten Transformation T beschränkt sich daher auf die Parameter für die Translation und Rotation. Dabei handelt es sich bei der im ExPrimage-Projekt realisierten Registrierung um ein semi-elastisches Verfahren¹⁰. Jede Färbung wird zunächst in Bildkacheln aufgeteilt. Jede Kachel wird einzeln mit Hilfe der *MI* rigide registriert. Daraufhin werden die unterschiedlich ausgerichteten bzw. registrierten Bildkacheln durch Annäherung eines Polynoms wieder zu einem Gesamtbild verbunden.

Die Ergebnisse der Registrierung wurden im Rahmen der vorliegenden Arbeit human untersucht. Dabei wurden vier Kategorien („perfekt“, „gut“, „kritisch“ und „unbrauchbar“) eingeführt. Die Analyse zeigt, dass die angewendete Registrierung überwiegend gute Ergebnisse erzielt und somit für weitere, darauf aufbauende Analysen, ausreichend ist.

Die zueinander registrierten Färbungen werden für eine färbungsübergreifende Gewebecharakterisierung (*Mult-Staining*) eingesetzt, wodurch relationale (siehe Abschnitt 5.5) Analysen zwischen unterschiedlichen Färbungen ermöglicht werden.

¹⁰Man unterscheidet zwischen rigiden (nur affine Transformationen), perspektivischen (erlaubt perspektivische Verzerrungen) und elastischen (erlaubt Deformationen) Transformationen [Han09]

Anwendung der SOM zur strukturellen Gewebecharakterisierung Zur Gewebecharakterisierung der strukturellen Marker wird ein unüberwachtes Lernverfahren (siehe Abschnitt 2.2), die *Self-Organizing-Map* [KSH01] (kurz SOM), verwendet. Die strukturellen Färbungen (f_{rgb}^{AE1AE3} , f_{rgb}^{HE} , f_{rgb}^{VIM}) werden bei diesem Verfahren zunächst in Grauwertbilder transformiert. Daraufhin werden für jede Färbung und für jeden Gewebepixel im Umkreis der Größe 19×19 Pixel der Mittelwert μ und die Standardabweichung σ berechnet, sodass für jeden Gewebepixel ein 6-dimensionaler-Merkmalsvektor entsteht. Durch Anwendung aller strukturellen Marker sollen sich zum einen Artefakte (Färbungsverläufe, Scanfehler, unterschiedliche Beleuchtungsverhältnisse, usw.) einzelner Färbungen durch Mittelung weniger auswirken und zum anderen durch Informationsverdichtung über alle strukturellen Färbungen ein vollständigeres Bild der anatomischen Strukturen geben, als es mit nur einer Färbung möglich wäre.

Die SOM ermöglicht ein Clustering (siehe Abschnitt 2.2) der vorher extrahierten Merkmalsvektoren. Die Clusteranzahl K wird auf drei festgelegt, welche sich als Fett- (grün), Stroma- (cyan) und Krebsgewebe (blau) im Ergebnisbild f_{rgb}^{som}

$$f_{rgb}^{ae1ae3} + f_{rgb}^{he} + f_{rgb}^{vim} \xrightarrow[SOM]{} f_{rgb}^{som} \quad (4.2)$$

darstellen. Ein Beispielergebnis wird in Abbildung 4.4 (oben rechts) dargestellt. Durch den Domänenexperten wurde bestätigt, dass die Ergebnisse einerseits stabil und andererseits aussagekräftig genug für die darauf aufbauenden Analysen der vorliegenden Arbeit sind.

Anwendung der LDA zur Binarisierung der funktionellen Marker Die *Lineare Diskriminanz-Analyse* (kurz LDA) [Run10], ein überwachtes Lernverfahren (siehe Abschnitt 2.1), wird eingesetzt, um die funktionellen Färbungen (f_{rgb}^{cd45} , f_{rgb}^{er} , f_{rgb}^{pr}) zu extrahieren. In diesem Fall müssen nur Vordergrund (Färbung) und Hintergrund (kein Gewebe oder Gegenfärbung), zwei Klassen, in den Bildern detektiert werden. Hierfür wird zunächst eine Lernstichprobe (siehe Abschnitt 2) durch Annotationen in verschiedenen Gewebeschnitten des funktionellen Markers aufgebaut (Färbung und Hintergrund bzw. Gegenfärbung). Die annotierten RGB-Pixelwerte bilden einen 3D-Merkmalraum. Für beide Klassen werden die Mittelwerte $\mu_{Vordergrund}$ und $\mu_{Hintergrund}$ bestimmt. Anhand dieser maximiert die LDA die Kovarianzen zwischen den beiden Klassen und minimiert die Kovarianzen innerhalb der beiden Klassen [Run10].

Ausgegeben wird eine *Diskriminanzhyperebene*. Durch Anwendung des Skalarproduktes zwischen der Normalen der *Diskriminanzhyperebene* und der RGB-Pixelwerte der Gewebeschnitte dieser Färbung kann entschieden werden, ob der Pixelwert vor bzw. hinter der Ebene liegt und somit Vordergrund oder Hintergrund ist. Die funktionellen Marker werden dadurch in die binären Bilder

$$\begin{aligned}
f_{rgb}^{cd45} &\xrightarrow{LDA} f_{bin}^{cd45} \in \{0, 1\} \\
f_{rgb}^{er} &\xrightarrow{LDA} f_{bin}^{er} \in \{0, 1\} \\
f_{rgb}^{pr} &\xrightarrow{LDA} f_{bin}^{pr} \in \{0, 1\}
\end{aligned} \tag{4.3}$$

überführt, welche als Eingabe für die in Abschnitt 5.5 beschriebenen Analysen verwendet werden. Abbildung 4.4 rechts unten zeigt Beispielergebnisse, wobei schwarz die Klasse der Färbung repräsentiert und weiß als Hintergrund bzw. Gegenfärbung klassifiziert wurde.

Kapitel 5

Merkmale und Bildanalysen zur Bestimmung von Ausbreitungsmustern auf Gewebeebene

Dieses Kapitel beschreibt die entwickelten Bildanalysen und Merkmale zur Detektion von Ausbreitungsmustern auf der Gewebeebene. Zunächst wird in Abschnitt 5.1 auf die erhobenen Anforderungen eingegangen. Um die Anforderungen erfüllen zu können, ist eine Erweiterung der vorverarbeitenden Schritte (siehe Abschnitt 4.2) notwendig. Deren Umsetzung wird in Abschnitt 5.2 erläutert. Die Umsetzungen und Erweiterungen der Analysen bezüglich Morphometrie, Struktur und Relationen werden in den Abschnitten 5.3, 5.4 und 5.5 vorgestellt. Klar zu differenzierende Tumorprototypen werden herangezogen, um die Merkmale auf ihre Sinnhaftigkeit für die realen Tumorbilddaten zu testen. Nach den Erläuterungen der Analysemethoden werden die Ergebnisse der Tumorprototypen veranschaulicht. Um die eingesetzten Analysen zur Bestimmung von Ausbreitungsmustern auf Gewebeebene zu validieren, wurde ein Konzept entwickelt und umgesetzt, welches in Abschnitt 5.6 vorgestellt wird. Das implementierte Datenmodell wird in Abschnitt 5.7 erläutert.

5.1 Anforderungen

In der vorliegenden Arbeit wurden in Zusammenarbeit mit den Domänenexperten verschiedene Anforderungen an die zu erfassenden Bildmerkmale aufgestellt. Grundsätzlich geben die Domänenexperten als relevante Kriterien auf Gewebeebene *Heterogenität* und *Homogenität* der Tumorstrukturen an. In diesem Kontext sind, im Sinne von Ausbreitungsmustern:

- Morphologische Eigenschaften des Gewebes,

- die räumliche Verteilung des Gewebes,
- und Kolokalisationen bzw. Relationen zwischen verschiedenen Bildmodalitäten¹

relevant für die Prognose des Patienten. Diese Eigenschaften zu quantifizieren ist Aufgabe der extrahierten Merkmale. Hierzu wird auf Bildobjekten (siehe Abschnitt 2.3) der einzelnen Bildmodalitäten gearbeitet. Morphometrisch relevante Faktoren der Tumorge-webeausbreitung (im Folgenden als Tumorareale bezeichnet) sind nach den Angaben der Domänenexperten:

- Größe
- Berandung (glatt oder rau)
- zentrale Aushöhlungen (als Anzeichen von zentralen nekrotischen Prozessen²)
- Kompaktheit (homogenes kompaktes Erscheinungsbild eines funktionierenden Tumors)
- Dissoziierung (heterogenes filamentöses Erscheinungsbild als mögliche Folge von peripheren und zentralen nekrotischen Prozessen)

Die räumliche Verteilung des Tumorgewebes wird durch die Kategorien

- Kompaktheit (homogene räumliche Verteilung)
- Dissoziierung (heterogene räumliche Verteilung)
- Multifokal³, mit jeweils einheitlichen Verteilungen
- Multifokal, mit jeweils unterschiedlichen Verteilungen

beschrieben.

In Abbildung 5.2 werden drei typische Fälle dargestellt. Tumor (a) ist insgesamt kompakt (markiert durch einen grünen Kreis) und enthält kompakte rundliche Tumorareale (blau markiert). Dieser Tumor wurde vom Domänenexperten insgesamt als „kompakt“ beschrieben. Der Tumor (b) enthält kompakte Tumorareale (blau markiert) und dissoziierte Tumorareale (rot markiert) mit unregelmäßiger Berandung, während das räumliche Verteilungsbild der Tumorareale insgesamt ebenfalls einen zunehmend dissoziierenden Charakter hat (markiert durch mehrere überlappende grüne Kreise). Dieser Fall wurde vom Domänenexperten mit „dissoziiert, Kontinuität erhalten“ annotiert. In Tumor (c) ist deutlich eine starke räumliche Dissoziierung der Tumorareale ohne klaren Zusammenhang (markiert durch mehrere grüne Kreise) zu erkennen. Überwiegend werden kleine unregelmäßig verteilte Tumorareale deutlich.

In der größeren Gewebecharakterisierung f_{rgb}^{som} werden im Tumorgewebe primär morphologische Unterschiede und Eigenschaften sichtbar. So unterscheiden sich die beiden dargestellten größeren Tumorareale (blau) in Abbildung 5.1 durch die Berandung, die Kompaktheit und zentralen Aushöhlungen. Tumor (a) hat beispielsweise eine große zen-

¹beispielsweise zwischen Tumorgewebe und entzündetem Gewebe

²pathologischer Untergang einzelner oder mehrerer Zellen

³mehrere voneinander getrennte Tumorareale

trale Aushöhlung und einen unregelmäßigen Rand, während sich Tumor (b) eher kompakt mit glattem Rand und ohne zentrale Aushöhlungen darstellt.

Kolokalisation und Relationen im Gewebe beschreiben das gemeinsame lokale Auftreten von funktionell markierten Gewebestrukturen mit Tumorgewebe. Kolokalisationen und Relationen im Gewebe können die folgenden Erscheinungsformen haben:

- zentral (im Zentrum der Tumorareale gelegen)
- peripheral (am Rand der Tumorareale gelegen)
- holohedral (sowohl im Zentrum als auch am Rand der Tumorareale gelegen)

In Abbildung 5.3 werden drei typische Fälle unterschiedlicher Relationen zwischen Tumorgewebe und entzündetem Gewebe dargestellt. Hierfür wurde f_{rgb}^{som} mit f_{bin}^{cd45} , welches rot markiert wurde, überlagert. Bei Tumor (a) befindet sich entzündetes Gewebe primär an der Peripherie des Tumors, in Tumor (b) eher zentral und in Tumor (c) sowohl am Rand als auch zentral.

Aus mathematischer Sicht sollten die Merkmale nach [Han09, NA08] zur eindeutigen Beschreibung von Bildobjekten folgende Eigenschaften erfüllen:

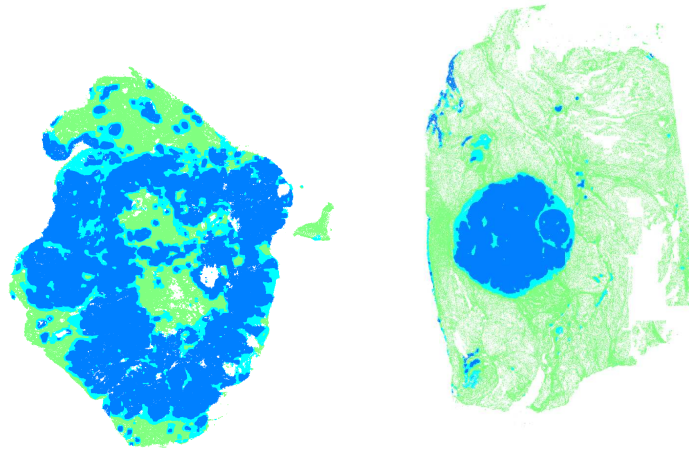
- Invarianz gegenüber Rotation
- Invarianz gegenüber Translation
- Invarianz gegenüber Skalierung
- Kongruent (ähnlich für ähnliche Objekte)
- Kompakt (nur Information verwenden, die notwendig ist)

Invarianz gegenüber Translation und Rotation wird benötigt, um ähnliche Eigenschaften der Bildobjekte unabhängig von ihrer Position und Ausrichtung im Tumor bestimmen zu können. Die Invarianz gegenüber der Skalierung ist in dieser Anwendung nicht immer notwendig bzw. nicht erwünscht, wenn es beispielsweise um die Erfassung von Größe und unterschiedlichen Entfernungen innerhalb eines Tumors geht. Zur Erfassung dieser Eigenschaften müssen gesonderte Merkmale bzw. Merkmalskombinationen verwendet werden. Idealerweise werden für die Berechnung der Merkmale keine Parameter benötigt, um so eine aufwändige Parameteranpassung zu vermeiden.

Das Ziel ist Merkmale zu finden, welche die genannten Eigenschaften möglichst gut erkennen und differenzieren. Somit sollen Merkmalsvektoren mit relevanter und kondensierter Information für jeden Patientenfall generiert werden.

5.2 Erweiterung der vorverarbeitenden Analysen

In Gesprächen mit den Domänenexperten hat sich herausgestellt, dass die Gewebecharakterisierung der SOM (f_{rgb}^{som} , siehe Abschnitt 4.2) das Tumorareal und die Gewebeklassen ausreichend gut beschreibt, um beispielsweise Kolokalisationsanalysen bzw. Relationsanalysen durchzuführen und die generelle morphometrische Ausbreitung und Form des Tumors zu erfassen. Für differenziertere Analysen, zur Bestimmung dissoziierter oder



(a) Tumor (unregelmäßiger Rand, zentrale Aushöhlung)
 (b) Tumor (glatter Rand, kompakt)

Abbildung 5.1: Beispiel von Tumorarealen (blau) aus der SOM-Gewebecharakterisierung: Tumor (a) hat eine große zentrale Aushöhlung und einen unregelmäßigen Rand, während sich Tumor (b) eher kompakt mit glattem Rand und ohne zentrale Aushöhlungen darstellt.

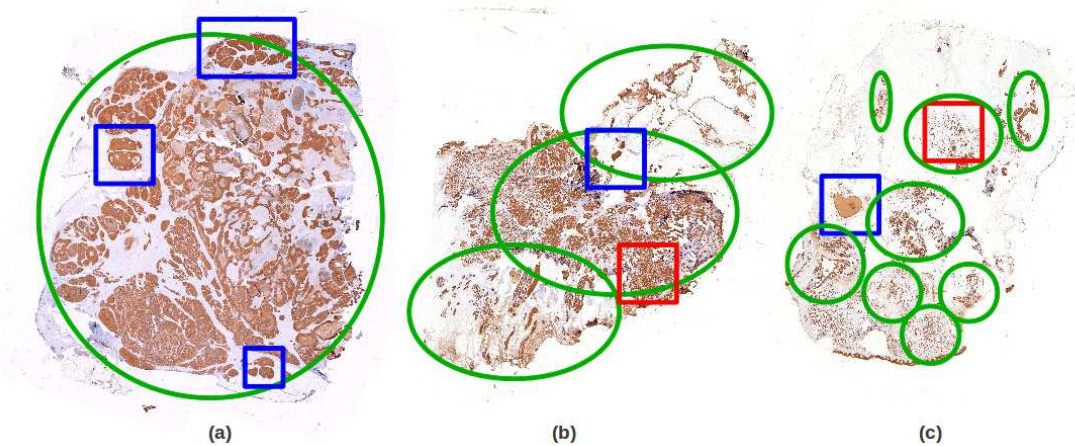


Abbildung 5.2: Unterschiedliche morphologische Eigenschaften und räumliche Verteilungen von Tumorarealen in der AE1/AE3-Färbung: (a) Kompakt, multifokal; (b) dissoziiert, Kontinuität erhalten; (c) stark dissoziiert.

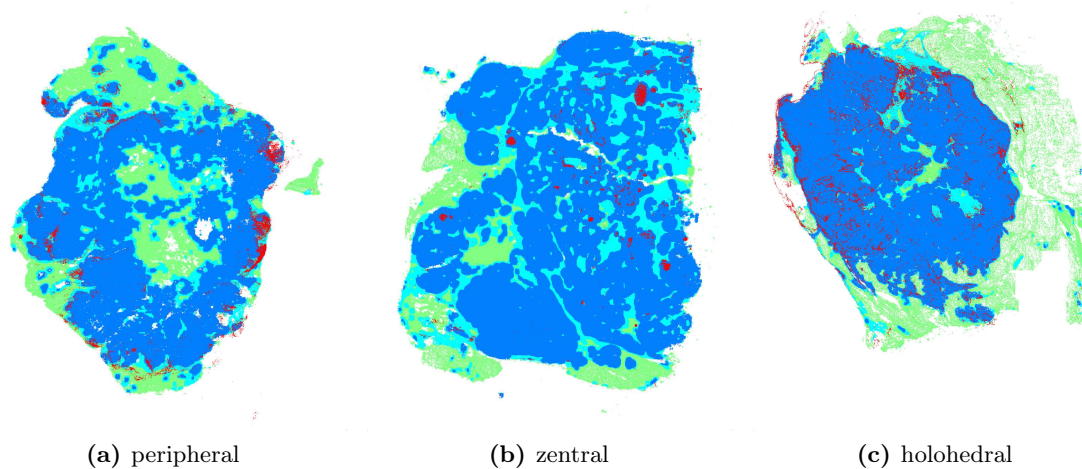


Abbildung 5.3: Unterschiedliche relationale Eigenschaften von funktionellen Markern (hier CD45, rot) im Bezug zum Tumorgewebe (blau) in der SOM-Gewebecharakterisierung (Fettgewebe (grün), Stromagewebe (cyan), Tumorgewebe (blau)): (a) fokale Peripherie (peripheral), (b) fokales Zentrum (zentral), (c) diffuses Peripherie und diffuses Zentrum (holohedral).

kompakter Ausbreitungs- und Verteilungsmuster, bietet die Gewebecharakterisierung der SOM hingegen nicht die notwendige Auflösung feinerer Tumorstrukturen.

Aus diesem Grund muss eine feinere Gewebecharakterisierung für das Tumorgewebe erfolgen. Dabei konzentriert sich die Analyse im Rahmen eines Segmentierungsprozesses (siehe Abschnitt 2.3) auf f_{rgb}^{ae1ae3} (die AE1AE3 Färbung), da sich dort das Tumorgewebe (braun) mit dem höchsten Kontrast (siehe Abschnitt 4.1) darstellt.

Die Gewebecharakterisierung von f_{rgb}^{ae1ae3} beinhaltet verschiedene Schwierigkeiten. Ein ausgleichender Effekt (beispielsweise bei Artefakten in Gewebeschnitten) über mehrere Gewebefärbungen, wie es beispielsweise bei der SOM - Gewebecharakterisierung (siehe Abschnitt 4.1) der Fall ist, entfällt. Es liegen unterschiedliche Beleuchtungsverhältnisse lokal (innerhalb der einzelnen Bildern) vor. Zudem unterscheiden sich die Bilder f_{rgb}^{ae1ae3} untereinander im Kontrastverhältnis (siehe Abbildung 5.4 (a), hoher Kontrast und Abbildung 5.5 (a) niedriger Kontrast). Gleichzeitig sind verschiedene Aufnahmeartefakte (Rauschen und überlagerte Signalartefakte) in den Bilddaten vorhanden. Die Färbung kann sich am Rand gesammelt haben oder ist hinter das Präparat gelaufen. Schwächeres Braun kann daher ein Artefakt sein. Verunreinigung am Rand des Präparates oder auf der Linse des Mikroskops zur Zeit der Aufnahme werden in den Gewebebildern dargestellt.

Diese Störeinflüsse können im Rahmen der vorliegenden Arbeit nur begrenzt in der Vorverarbeitung berücksichtigt werden. Die Korrektur der unterschiedlichen Beleuchtungs- und Kontrastverhältnisse sowie die deutliche Differenzierung zwischen Färbung und Gegenfärbung stehen im Hinblick auf die Bildanalysen zur Merkmalsextraktion im Vordergrund.

In [OJJS96, DY05] wird das Schwellwertverfahren von Otsu (kurz Otsu-Verfahren)(siehe Abschnitt 2.3.1) zur Segmentierung von Grauwertgewebesektionen aufgeführt. In [DY05] wird zusätzlich das *K-Means*-Verfahren(siehe Abschnitt 2.2) zur Erstellung von Segmenten in der digitalen Histopathologie genannt. In [ACSGD10] und [BDNY07] wird es zur Segmentierung von Gewebeausschnitten im RGB-Farbraum eingesetzt.

Anlehnend an diese Anwendungen wurde f_{rgb}^{ae1ae3} zunächst in ein Grauwertbild f_{gray}^{ae1ae3} konvertiert⁴. Daraufhin wurde auf f_{gray}^{ae1ae3} eine Histogrammspreizung [BB05] zur Erhöhung des Kontrastes durchgeführt⁵. Zur Trennung von Färbung und Gegenfärbung wird das Otsu-Verfahren angewendet. In Abbildung 5.4 werden die unterschiedlichen Ergebnisse der angewendeten Verfahren anhand eines vergrößerten Bildausschnittes abgebildet. Dabei wird die Gewebedetektion der Tumoreale schwarz dargestellt. In (b) ist zu sehen, dass die Verwendung des Otsu-Verfahrens auf Basis von Grauwertbildern zur Gewebecharakterisierung, Gegenfärbung (lila) enthält.

Aufgrund dessen wurde das Otsu-Verfahren auf die einzelnen RGB-Farbkanäle angewendet, um mit Hilfe der Farbinformation eine bessere Differenzierung der Färbung und Gegenfärbung zu erreichen. Das Ergebnis konnte dadurch verbessert werden (siehe (c)). Trotzdem bleibt viel Gegenfärbung und verlaufene Färbung (schwaches Braun) in der Gewebedetektion enthalten. Die unterschiedliche Helligkeit der Färbung und die Gegenfärbung kann nicht hinreichend unterschieden werden.

Um eine bessere Trennung und eine bessere Unabhängigkeit von der Beleuchtung bzw. Helligkeit zu erlangen, wird der *Lab*-Farbraum verwendet. Dies findet auch in [ACSGD10] zur Segmentierung von Gewebe Anwendung. Der *Lab*-Farbraum basiert auf der *CIE-Normfarbtafel*[BB05] und trennt die Farbigkeit (Kanal a und b) von der Helligkeit (Kanal L) [BB05].

In [Matb] wird zur Segmentierung einer HE-Färbung unter Verwendung der a und b Kanäle das *K-Means*-Verfahren vorgeschlagen. Dieser Ansatz wurde ebenfalls getestet. Jedes f_{rgb}^{ae1ae3} wird dazu in f_{lab}^{ae1ae3} transformiert⁶. Das *K-Means*-Verfahren wird unter Verwendung von 3 Clustern (starkes Braun, schwaches Braun, Gegenfärbung) und der 2D-Pixelwerte (a-Komponente, b-Komponente) durchgeführt. Das gesuchte Tumorgewebe (starkes Braun) wird anhand des geringsten durchschnittlichen Farbabstandes [BB05] zu einem zuvor festgelegten repräsentativen Farbvektor [$a = 155, b = 164$] (starkes Braun) für jede Ergebnispartition innerhalb eines f_{lab}^{ae1ae3} bestimmt. Dieser Ansatz liefert gute Ergebnisse (siehe Abbildung 5.4 (d)) in Bezug auf die Detektion des Tumorgewebes. Gleichzeitig führt dieser Ansatz jedoch zu viel Detektion von Artefakten außerhalb der AE1AE3-Färbung (siehe Abbildung 5.5 (b)). Die zusätzliche Verwendung der Helligkeitsinformation *L* ergab insgesamt schlechtere Ergebnisse.

Als Alternative kann das Verfahren von Otsu auf jede Komponente des *Lab* Farbraumes einzeln angewendet werden (im Folgenden als Otsu3Lab bezeichnet). Dieser einfache,

⁴mit Hilfe der MATLAB-Funktion `rgb2gray`

⁵mit Hilfe der MATLAB-Funktion `histeq`

⁶eingesetzt werden hierzu die MATLAB-Funktionen: `makecform('srgb2lab')`, und `lab = applycform(rgb, cform)`

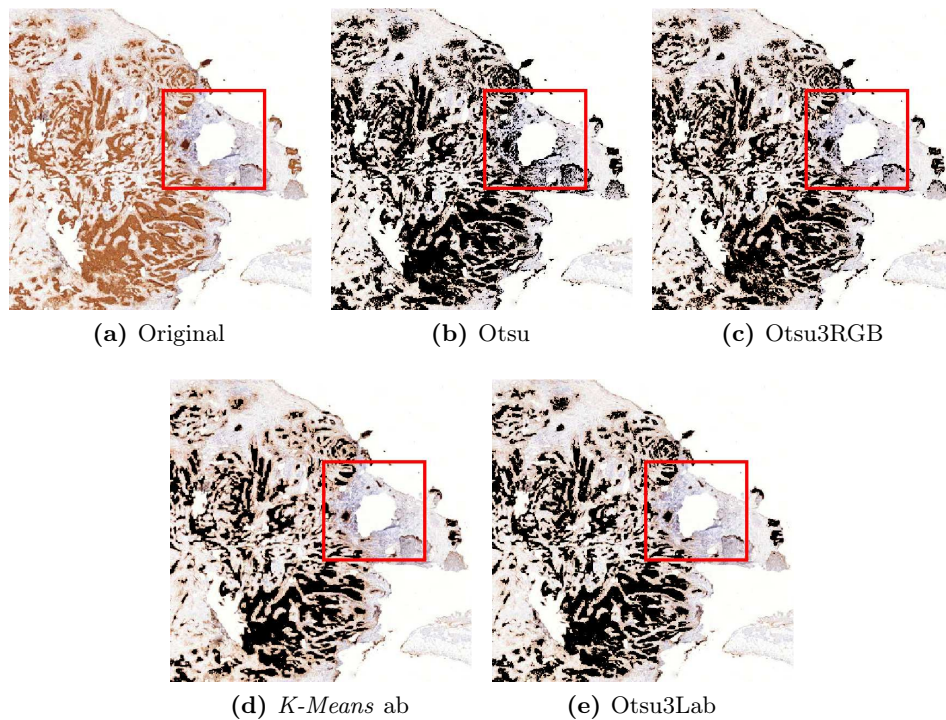


Abbildung 5.4: Beispielergebnisse unterschiedlicher Ansätze zur Tumorgewebedetektion in AE1AE3.

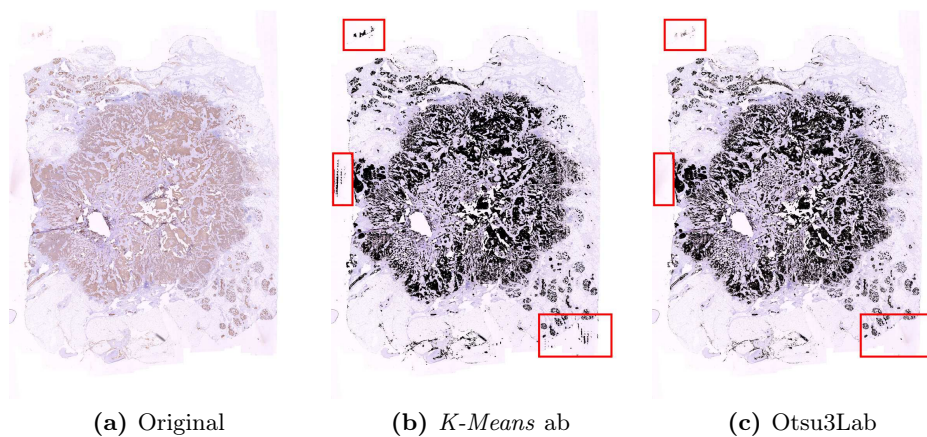


Abbildung 5.5: Beispielergebnisse unterschiedlicher Ansätze zur Segmentierung von AE1AE3: (a) Original AE1AE3-Färbung mit geringem Kontrast, (b) Ergebnis *K-Means*-Verfahren (a-Kombinente und b-Komponente) mit markierten Artefakten (rote Kästen), (c) Otsu3Lab-Ergebnis ohne Artefakte an Randstrukturen.

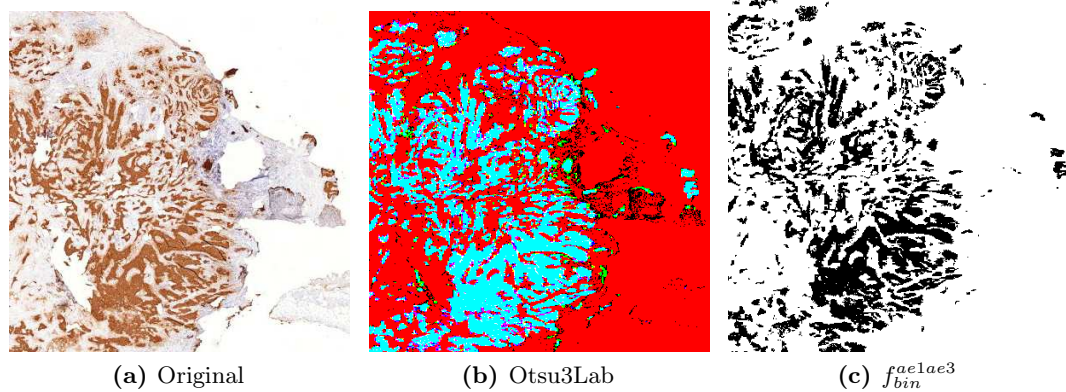


Abbildung 5.6: Ergebnis des Otsu3Lab in zusammengesetzten RGB-Bild (b) und des detektierten Tumorgewebes (c).

aber in der vorliegenden Anwendung effektive Ansatz, wird im Folgenden näher beschrieben⁷.

5.2.1 Anwendung des Otsu-Verfahrens auf den Lab-Farbraum

Das Bildmuster f_{rgb}^{ae1ae3} wird hierzu wiederum zunächst in den *Lab*-Farbraum zu f_{lab}^{ae1ae3} transformiert. Jeder Kanal (L , a und b) wird als einzelnes Bildmuster betrachtet. Für jeden Kanal wird durch das Otsu-Verfahren (siehe Abschnitt 2.3.1) der optimale Schwellwert bestimmt. Dadurch entstehen drei Binärbilder f_{bin}^L , f_{bin}^a und f_{bin}^b . Diese drei Binärbilder werden zu einem RGB-Bild $f_{rgb}^{otsu3Lab}$ fusioniert, wobei die L -Komponente dem Rotkanal, die a -Komponente dem Grünkanal und die die b -Komponente dem Blaukanal zugeordnet wird. Dadurch entstehen 2^3 Farbkombinationen, welche sich als unterschiedlich gefärbte Pixelregionen darstellen (siehe Abbildung 5.6 (b)).

In Abbildung 5.6 (b) kann man erkennen, wie durch f_{bin}^L zunächst der Hintergrund (weiß und schwache Färbung) getrennt wird (rot). Die beiden Binärbilder f_{bin}^a und f_{bin}^b trennen die Gegenfärbung (schwarz) und Färbung (cyan) gut. Die schwächere Braunfärbung, die für verlaufene Regionen spricht, wird durch Magenta, Blau und Grün gut von Cyan getrennt.

Im Vergleich zu diesem Ergebnis liefert das *K-Means*-Verfahren etwas kompaktere Tumorregionen. Dennoch überwiegt der Vorteil des Otsu3Lab-Verfahrens gegenüber dem *K-Means*-Verfahren in dieser Anwendung, da wesentlich weniger Artefakte detektiert werden (siehe Abbildung 5.5), welche zu falschen Tumorobjekten (siehe Abschnitt 5.2.2) führen würden. Außerdem werden keine Parameter, wie beispielsweise die Clusteranzahl benötigt. Das detektierte Tumorgewebe wird als Binärbild $f_{bin}^{ae1ae3} \in \{0, 1\}$ durch Selektion

⁷Getestet wurden ebenfalls die SOM-Analysen und ein lokales Schwellwertverfahren (unter Verwendung des lokalen Mittelwertes) zur Detektion des Tumorgewebes in AE1AE3. Die Ergebnisse waren insgesamt schlechter.

tion der cyan-farbenden-Pixel (siehe Formel (5.1)) als Vordergrund (siehe (c)) für jeden Fall erstellt und ist Grundlage verschiedener im Folgenden beschriebener Analyseschritte.

Anhand von Stichproben wurden die Ergebnisse mit dem Domänenexperten evaluiert. Die Stichproben wurden als ausreichend zur Bestimmung der Feinstrukturen bezüglich der Tumorausbreitung eingeschätzt. Das Verfahren von Otsu wird aus der *ImageProcessing-Toolbox* [Matc] verwendet⁸.

$$f_{rgb}^{otsu3Lab} \xrightarrow{==(0,255,255)} f_{bin}^{ae1ae3} \in \{0,1\} \quad (5.1)$$

5.2.2 Generierung der Bildobjekte

Grundlage für alle nachfolgend beschriebenen Analysen ist zunächst das Segmentieren und Lokalisieren der verschiedenen Bildobjekte R . Dies geschieht in allen Analyseschritten, unter Verwendung unterschiedlicher Bildmodalitäten, prinzipiell gleich. Es stehen die charakterisierten Bilddaten für Fettgewebe, Stromagewebe und Tumorgewebe in f_{rgb}^{som} zur Verfügung. Diese werden durch Selektion der entsprechenden Farbwerte zunächst in binäre Bilddaten f_{bin}^{fat} , f_{bin}^{stroma} und f_{bin}^{tumor} transformiert (siehe Formel (5.2)).

$$\begin{aligned} f_{rgb}^{som} &\xrightarrow{==(128,255,128)} f_{bin}^{fat} \in \{0,1\} \\ f_{rgb}^{som} &\xrightarrow{==(0,255,255)} f_{bin}^{stroma} \in \{0,1\} \\ f_{rgb}^{som} &\xrightarrow{==(0,128,255)} f_{bin}^{tumor} \in \{0,1\} \end{aligned} \quad (5.2)$$

Die funktionellen Marker liegen nach der LDA (siehe Abschnitt 4.2) als binäre Bilddaten f_{bin}^{cd45} , f_{bin}^{er} und f_{bin}^{pr} vor. Die Feinstrukturen des Tumors liegen nach der Anwendung des Otsu3Lab-Verfahrens (siehe voriger Abschnitt) ebenfalls als binäre Bilddaten f_{bin}^{ae1ae3} vor. In Abhängigkeit der vorgesehenen Analysen werden unterschiedliche Bildmodalitäten ϱ verwendet. Dementsprechend werden unterschiedliche Bildobjekte R^ϱ extrahiert. Diese Bildobjekte werden in der vorliegenden Arbeit als Connected-Components R mit 4-Pixelnachbarschaft des Vordergrundes der binären Bilddaten unter Verwendung eines *Flood-Fill-Algorithmus* bestimmt⁹ (siehe Abschnitt 2.3.2). Für jede Bildmodalität ϱ erfolgt vor den folgenden Analysen eine Aufteilung in eine Menge von separierten Bildobjekten

$$f_{bin}^\varrho \xrightarrow{Connected-Components} \mathbf{R}^\varrho = \{R_1^\varrho, \dots, R_n^\varrho | A(R^\varrho) > t^\varrho\} \quad (5.3)$$

, wobei t^ϱ einen Schwellwert für die Größe (A) der verwendeten Bildobjekte R^ϱ festlegt und n die Anzahl der extrahierten Bildobjekte für f_{bin}^ϱ angibt. Der Schwellwert t^ϱ wird

⁸graythresh()

⁹mit Hilfe der MATLAB-Funktion `bwconncomp`

Analysen	f_{bin}^{ae1ae3}	f_{bin}^{tumor}	f_{bin}^{cd45}	f_{bin}^{er}	f_{bin}^{pr}
Morphometrie	✓	✓	-	-	-
DG	✓	-	-	-	-
MST	✓	-	-	-	-
RCC8	-	✓	✓	✓	✓
LDQ	-	✓	✓	✓	✓

Tabelle 5.1: Überblick zu verwendete Bildmodalitäten für die unterschiedlichen Analysen: Morphometrie, Struktur (DG, MST) und Relation (RCC8 und LDQ).

eingesetzt um Rauschen und Artefakte zu minimieren, so dass nur relevante Bildobjekte (und beispielsweise nicht einzelne Pixel) identifiziert werden. Abbildung 5.7 visualisiert die unterschiedliche Bildobjekte am Beispiel von f_{bin}^{tumor} . Für f_{bin}^{ae1ae3} wird t^{ae1ae3} auf 15 Pixel festgelegt.

Jedes Bildobjekt wird unabhängig als einzelnes Objekt behandelt und gespeichert. Dies ermöglicht die unabhängige Rekonstruktion und Behandlung jedes Bildobjektes innerhalb des ursprünglichen globalen Bildmusters. Dadurch kann jedes Bildobjekt einerseits unabhängig und andererseits durch Rekonstruktion zusammen mit anderen Bildobjekten analysiert werden (beispielsweise für die Relationsanalysen des RCC8-Histogramms).

Die Bildobjekte werden unterschieden zwischen Tumorobjekten und Relationsobjekten. Tumorobjekte repräsentieren das Tumorgewebe. Zu ihnen zählen Bildobjekte aus den Bildfunktionen f_{bin}^{ae1ae3} (R^{ae1ae3}) und f_{bin}^{tumor} (R^{tumor}). Bildobjekte der Bildfunktionen f_{bin}^{cd45} , f_{bin}^{er} und f_{bin}^{pr} repräsentieren Relationsobjekte (R^{cd45} , R^{er} und R^{pr}). Relationsobjekte werden im Gegensatz zu Tumorobjekten nicht morphometrisch und strukturell untersucht. Sie dienen dazu Relationen und Koloalisationen zwischen Tumorobjekten und Relationsobjekten zu quantifizieren. Tabelle 5.1 gibt einen Überblick zu den eingesetzten Bildmodalitäten der verschiedenen Analyseanwendungen, welche im Folgenden vorgestellt werden.

5.3 Morphometrische Analysen

Im Folgenden wird zunächst erläutert, welche Merkmale für die Erfüllung der morphometrischen Anforderungen ausgewählt werden. Auf zellulärer Ebene werden, wie bereits in Abschnitt 3.1 erwähnt, hauptsächlich kompakte Basis-Deskriptoren verwendet, welche die primär rundlichen Zellen ausreichend gut beschreiben. Die zu untersuchenden Tumorobjekte (R^{ae1ae3} und R^{tumor}) auf Gewebeebene liegen in wesentlich komplexerer und heterogener Form vor. Um diesen Unterschieden gerecht zu werden, sind umfangreichere Informationen und dementsprechend komplexere Merkmale notwendig. In Tabelle 5.2 wird ein Überblick zu den implementierten bzw. eingesetzten Merkmalen bzgl. der Erfüllung der erstellten Anforderungen gegeben.

Die Größe eines Tumorobjektes R kann durch die Fläche A und den äußeren Umfang P bestimmt werden (siehe Abschnitt 3.1.2). Zur Beschreibung der Berandung eines Tumo-

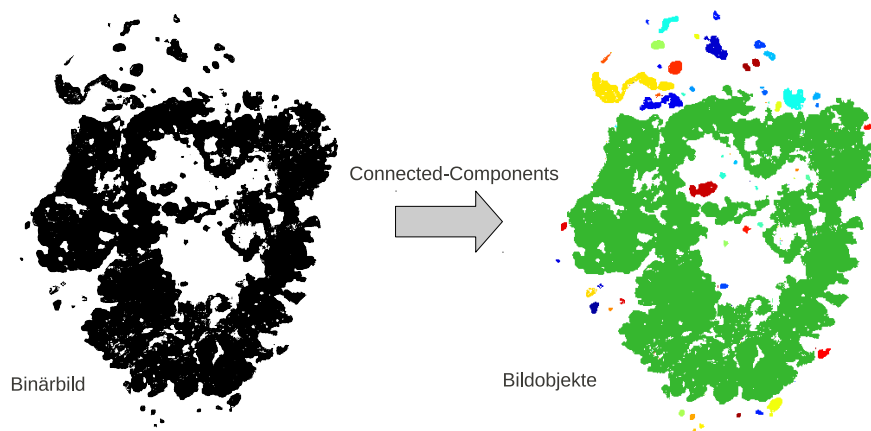


Abbildung 5.7: Beispielergebnis unterschiedlicher Bildobjekte innerhalb eines f_{bin}^{tumor} (farblich und räumlich differenziert) der Bildobjektsegmentierung mittels Connected-Components.

Merkmal	Par	T	R	S	Kt	Kp	Anforderungen
Fourier-Deskriptoren	ja	✓	✓	✓	✓	-	Beschreibung des Tumorrandes
Fourier-Energien	ja	✓	✓	✓	-	✓	Beschreibung des Tumorrandes
Moment Invariants	nein	✓	✓	✓	✓	✓	zentrale Aushöhlungen, Kompaktheit, Dissoziierung
Irregularity	nein	✓	✓	✓	-	✓	Dissoziierung
Compactness	nein	✓	✓	✓	-	✓	Kompaktheit
Roundness	nein	✓	✓	✓	-	✓	Kompaktheit, Beschreibung des Tumorrandes
Area	nein	✓	✓	-	-	✓	Größe
Perimeter	nein	✓	✓	-	-	✓	Größe

Tabelle 5.2: Merkmalsauswahl zur Erfüllung der morphometrischen Anforderungen (Par=Parameter notwendig, T=Translationsinvarianz, R=Rotationinvarianz, S=Skalierungsinvarianz, Kt=Kongruent, Kp=Kompakt).

robjektes werden die Fourier-Deskriptoren (FDs) und Fourier-Energien (siehe Abschnitt 3.1.1) auf Grundlage der äußeren Kontur eines Tumorobjektes eingesetzt.

Die FDs und Fourier-Energien sollten für unregelmäßig berandete Tumorobjekte höhere Werte in den höheren Frequenzbereichen annehmen als für glatt berandete kompakte Tumore. Entscheidend ist in diesem Fall eine geeignete Konturrepräsentation. Ausgewählt wurden hierfür die Zentroiden Distanzen, welche nach den Studien aus [ZL02] die äußere Kontur am effektivsten repräsentieren (siehe Abschnitt 3.1.1). Da in den einzelnen Tumorobjekten zentrale Aushöhlungen vorhanden sind, werden diese zunächst gefüllt¹⁰. Der Zentroid eines Tumorobjektes wird anhand der äußeren Konturpixel des gefüllten Tumorobjektes bestimmt. FDs sind invariant gegenüber Skalierung, Translation und Rotation. Entscheidend ist an dieser Stelle die Auswahl geeigneter Parameter (Anzahl Sampling-Punkte und Anzahl verwendeter FDs). Für diese Anwendung orientiert sich die Auswahl der Parameter an den Ausführungen aus [SKH08], welche eine Anzahl von 15 Fourier-Deskriptoren und 265 Abtastungspunkten als grundsätzlich ausreichend identifiziert haben.

Abbildung 5.8 zeigt die *Shape-Funktion* (Zentroiden Distanzen, blau) sowie die ersten 15 Werte der FDs eines kompakten homogenen Tumorobjektes (a) und eines dissoziierten heterogenen Tumorobjektes (b) aus f_{bin}^{ae1ae3} . Die aus den 15 FDs rekonstruierte *Shape-Funktion* wird rot dargestellt. Für (a) ist deutlich zu erkennen, dass die *Shape-Funktion* gleichmäßig mit niedriger Frequenz verläuft, was für eine glatte und kompakte Kontur spricht. Bei (b) sind deutlich höhere Frequenzen und Amplituden zu erkennen. Dementsprechend sind die Werte der FDs für (a) wesentlich niedriger als für (b) (siehe rote Markierung des Maximalwertes der Y-Achse). Anhand dieser Daten wird von einer hinreichenden Repräsentation der äußeren Kontur durch 15 FDs ausgegangen.

Die Kompaktheit eines Tumorobjektes kann durch die Merkmale Compactness, Roundness und als Gegenstück durch die Irregularity beschrieben (siehe Abschnitt 3.1.2) werden. Diese Merkmale sind zwar prinzipiell invariant allerdings nicht kongruent. Sie können für deutlich unterschiedliche Formen ähnliche Werte annehmen [Han09].

Aus diesem Grund, und um zentrale Aushöhlungen sowie Dissoziierung von Tumorobjekten zu beschreiben, werden zusätzlich die ebenfalls invarianten Moment Invariants (siehe Abschnitt 3.1.3) eingesetzt (siehe Abbildung 5.9). Das siebte Moment H_7 (siehe Formel (3.15)) kann zur Bestimmung von Spiegelungen verwendet werden, die sich durch positive bzw. negative Werte äußern. Da in der vorliegenden Anwendung Spiegelungen nicht von Belang sind, kann dieser Faktor ignoriert werden. Um den sehr großen Wertebereich der Moment Invariants zu reduzieren, wird der Betrag der logarithmierten Momente $|\log(H_k)|$ verwendet.

Die Berechnung der Merkmale ist grundsätzlich abhängig vom eingesetzten Schwellwert t^e (siehe voriger Abschnitt), welcher die minimale Größe der durch die Merkmale zu beschreibenden Tumorobjekte festlegt. Davon abgesehen werden, außer für die Fourier-Deskriptoren und den darauf aufbauenden Fourier-Energien, keine Parameter zur Berechnung der Merkmale benötigt.

¹⁰mit Hilfe der MATLAB-Funktion `imfill`

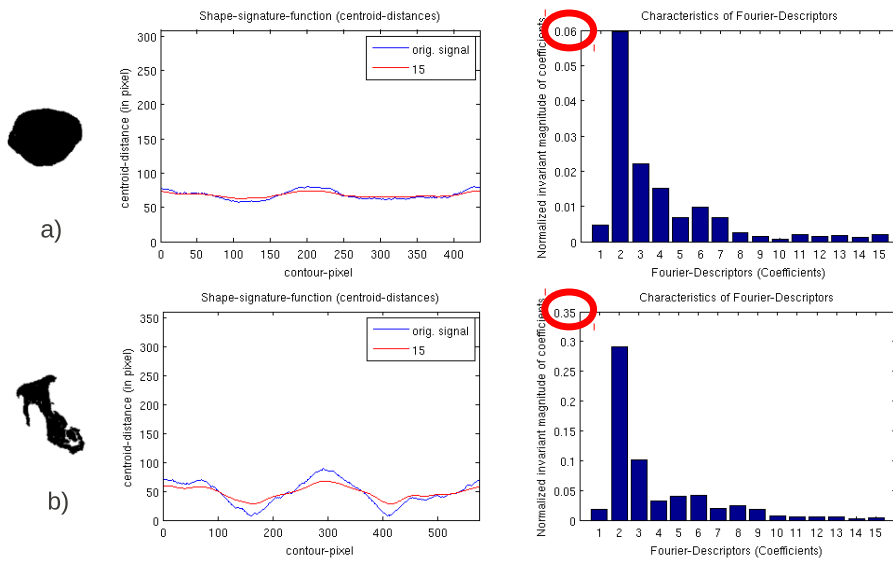


Abbildung 5.8: Diese Abbildung zeigt die *Shape-Funktion* (Zentroiden Distanzen) und ersten 15 Werte der Fourier-Deskriptoren eines kompakten homogenen Bildobjektes und eines dissoziierten heterogenen Bildobjektes aus f_{bin}^{ae1ae3} . Die aus den 15 Fourier-Deskriptoren und 256 Abtastpunkten rekonstruierte *Shape-Funktion* wird rot dargestellt.

Eingesetzt werden diese Merkmale in der Bildmodalität f_{bin}^{ae1ae3} zur Analyse der feinen Tumorstrukturen und in f_{bin}^{tumor} zur Analyse der größeren Tumorstrukturen. Morphologische Operationen, wie beispielsweise ein *Opening* und *Closing* [BB05] zur besseren Separierung der einzelnen Tumorobjekte in f_{bin}^{ae1ae3} wurden getestet. Es zeigte sich, dass dabei relevante Feinstrukturen verloren gehen (wie beispielsweise zentrale Aushöhlungen als Anzeichen für nekrotische Prozesse oder dissoziierte Randstrukturen). Aus diesem Grund wurde beschlossen diese vorverarbeitende Methoden zunächst nicht einzusetzen.

Zur Bestimmung der morphometrischen Eigenschaften werden zwei unterschiedliche Ansätze verwendet. Der erste Ansatz berechnet für jeden Fall deskriptive Statistiken der morphometrischen Merkmale. Als weiterer Ansatz wird versucht mittels eines Clusterings der einzelnen Tumorobjekte sowohl eine Visualisierung als auch eine Dimensionsreduktion und Informationskondensierung der eingesetzten Merkmale, ähnlich wie in [DAM⁺08], zu erreichen. Die beiden Ansätze werden im Folgenden näher beschrieben.

5.3.1 Deskriptive morphometrische Statistiken

In [GBC⁺09] wird vorgeschlagen, die morphometrischen Basis-Deskriptoren für jedes Tumorobjekt R in \mathbf{R} durch deskriptive Statistik, wie beispielsweise den Mittelwert μ und die Standardabweichung σ , zu berechnen. Dieses Vorgehen wurde zunächst übernommen.

Für alle eingesetzten morphometrischen Merkmale werden sowohl der Mittelwert als auch die Standardabweichung für jedes Bildmuster f bzw. die Tumorobjekte \mathbf{R} berechnet. Dabei werden keine einzelnen Fourier-Deskriptoren, sondern lediglich die zusam-

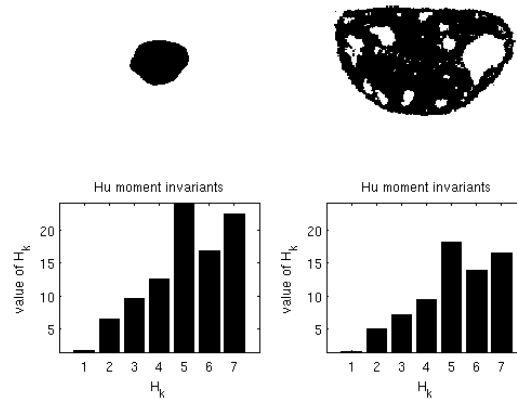


Abbildung 5.9: Werte der sieben Moment Invarianten (H_k) zur Differenzierung von Formen ohne und mit zentralen Aushöhlungen. Die Formen stammen aus f_{bin}^{ae1ae3} . Die Werte der Momente zeigen deutliche Unterschiede, obwohl beide Formen eine ähnliche Kontur und Ausdehnung haben.

menfassenden Fourier-Energien verwendet, um die Anzahl der Merkmale für einen Fall einzuschränken und dennoch die nötige Information in verdichteter Form zu integrieren.

Außerdem wird die Anzahl der Tumorobjekte N_R sowie das Verhältnis $A_{ratio} = \frac{A_\mu}{A_{total}}$ der durchschnittlichen Größe eines Tumorobjektes A_μ zur Größe des gesamten Tumorgewebes A_{total} bestimmt. Dies führt insgesamt zu 32 Merkmalen für jedes Bildmuster. Dieser Ansatz wurde jedoch aufgrund des vorgegebenen Zeitrahmens und dem Vorteil des im folgenden Abschnitt beschriebenen Ansatzes nur zur Bestimmung von Größeneigenschaften (N_R , A_{ratio} , A_μ , A_σ , P_μ , P_σ) in f_{bin}^{ae1ae3} eingesetzt.

5.3.2 Clustering der morphometrischen Merkmale

Der zuvor beschriebene Ansatz ergibt eine hohe Anzahl von Merkmalen für jedes Muster. Da nur 93 Patientenfälle (siehe Abschnitt 4.1) vorliegen, kann eine zu hohe Anzahl von Merkmalen zu einem Überanpassen (engl. *Overfitting*) der Trainingsdaten bzw. Stichproben führen, so dass keine Generalisierung mehr möglich ist [Nie03]. Außerdem wird bei hochdimensionalen Merkmalsräumen die sogenannte *Curse-Of-Dimensionality*¹¹ [Bis06] zu einem Problem. In [DAM⁺08] wird aufgrund dieser Problematik zur Klassifizierung von Gewebeschnitten auf Zellebene und zur Visualisierung des Datenraumes durch *Spectral Clustering* die Dimension der Merkmale unter Berücksichtigung zusammengehörender Merkmalsgruppierungen reduziert. In [KCK⁺11] werden Krebszellen unter Einsatz des *K-Means*-Verfahrens auf morphometrische und textuelle Eigenschaften und Korrelationen untersucht. Aufbauend auf diesen Ideen wird ein weiterer Ansatz zur Bestimmung

¹¹Bei hoch-dimensionalen Merkmalsräumen steigt das Volumen exponentiell. Dabei wird angenommen, dass der Abstand zwischen den kleinsten und größten Abständen gemessen an einer Distanzfunktion beliebig klein wird und somit keine Aussagekraft mehr besteht.

der morphometrischen Eigenschaften eingesetzt, welcher anhand der schematischen Abbildung 5.10 erläutert werden soll.

In diesem Ansatz wird nicht zu jedem Bildmuster ein Merkmalsvektor bestimmt, sondern jedes Tumorobjekt R aller Bildmuster f wird als einzelnes Muster betrachtet. Dies führt zur Menge $\mathbf{R}_{all} = \{\mathbf{R}_1, \dots, \mathbf{R}_n\}$ von unterschiedlichen Tumorobjekten über alle Bildmuster hinweg. Der Schwellwert t^{ae1ae3} zur Bestimmung der Mindestgröße wurde hierfür auf 15 Pixel festgelegt. Dieser Ansatz erhöht die Anzahl unterschiedlicher Muster für f_{bin}^{ae1ae3} von 93 auf 216196. Für jedes Tumorobjekt R werden die Merkmale aus Tabelle 5.2 berechnet und als einzelne Merkmalsvektoren \mathbf{m}_R gespeichert. Die höhere Anzahl von Mustern macht es möglich einzelne Fourier-Deskriptoren einzusetzen, ohne dabei ein *Overfitting* zu riskieren.

Für das Clustering der Muster \mathbf{R}_{all} wird das *K-Means*-Verfahren (siehe Abschnitt 2.2) eingesetzt. Zunächst werden die Merkmale normalisiert, um einen einheitlichen und vergleichbaren Wertebereich für alle Merkmale zu erreichen. Die Normalisierung und die Bestimmung der Clusteranzahl verläuft dabei wie die unüberwachte Evaluation der Merkmale, welche in Abschnitt 5.6 beschrieben wird.

Jedem R_i aus \mathbf{R}_{all} wird nach der Clusteranalyse durch Rückprojektion das zugehörige Cluster zugeordnet. In Abbildung 5.10 wird ein Beispiel mit Clusteranzahl 3 dargestellt. Dort ist zu sehen, wie kompakte Strukturen blau, dissoziierte Strukturen mit zentralen Aushöhlungen cyan und elongierte kompakte Strukturen gelb markiert werden. Daraufhin wird ein neuer Merkmalsvektor für jedes Gesamtbild erstellt. Dabei wird die Prozentzahl der Fläche A jedes Clusters als Merkmalsvektor für f bestimmt. Dies bedeutet beispielsweise, dass für Clusteranzahl $K = 3$ ein 3D-Merkmalsvektor entsteht, welcher dann für die Merkmalsvalidierung (siehe Abschnitt 5.6) in einem weiteren Clustering der Patientenfälle eingesetzt wird.

Da für f_{bin}^{ae1ae3} die Anzahl der Tumorobjekte sehr hoch ist, die Berechnung der Cluster und Clustervalidierung viel Zeit in Anspruch nimmt und eine Generalisierung überprüft werden soll, besteht die Möglichkeit, nur eine geringere Anzahl von Bildmustern f und somit eine geringere Anzahl \mathbf{R}_{all}^S als Stichprobe zu verwenden. Dabei wird unüberwacht mit dem *K-Means*-Verfahren ein Modell der Daten in Form der Clusterzentren $\boldsymbol{\mu}_k$ erstellt. Nachdem das Modell zur Verfügung steht, können alle Tumorobjekte \mathbf{R}_{all} aus allen Bildmustern f mittels des minimalen quadratischen euklidischen Abstandes

$$\|\boldsymbol{\mu}_k - \mathbf{m}_R\| = \min_{\{1, \dots, K\}} \|\boldsymbol{\mu}_k - \mathbf{m}_{R_n}\| \quad (5.4)$$

einem Cluster zugeordnet werden. Durch die Clusterzentren und die quadratischen euklidischen Distanzen wird $\boldsymbol{\delta}(\mathbf{m})$ (siehe Abschnitt 2) definiert. Eine Auswertung der Ergebnisse, insbesondere der Clusteranzahl, erfolgt in den Experimenten in Kapitel 6. Vorteil dieses Vorgehens, im Vergleich zur deskriptiven morphometrischen Statistik (siehe voriger Abschnitt), ist neben der Dimensionsreduktion eine unmittelbare Visualisierung der maschinell als ähnlich betrachteten Tumorobjekte. So kann eine Plausibilitätsüberprüfung durch den Domänenexperten relativ einfach mit Hilfe des unmittelbaren visuellen Feedback's durchgeführt werden.

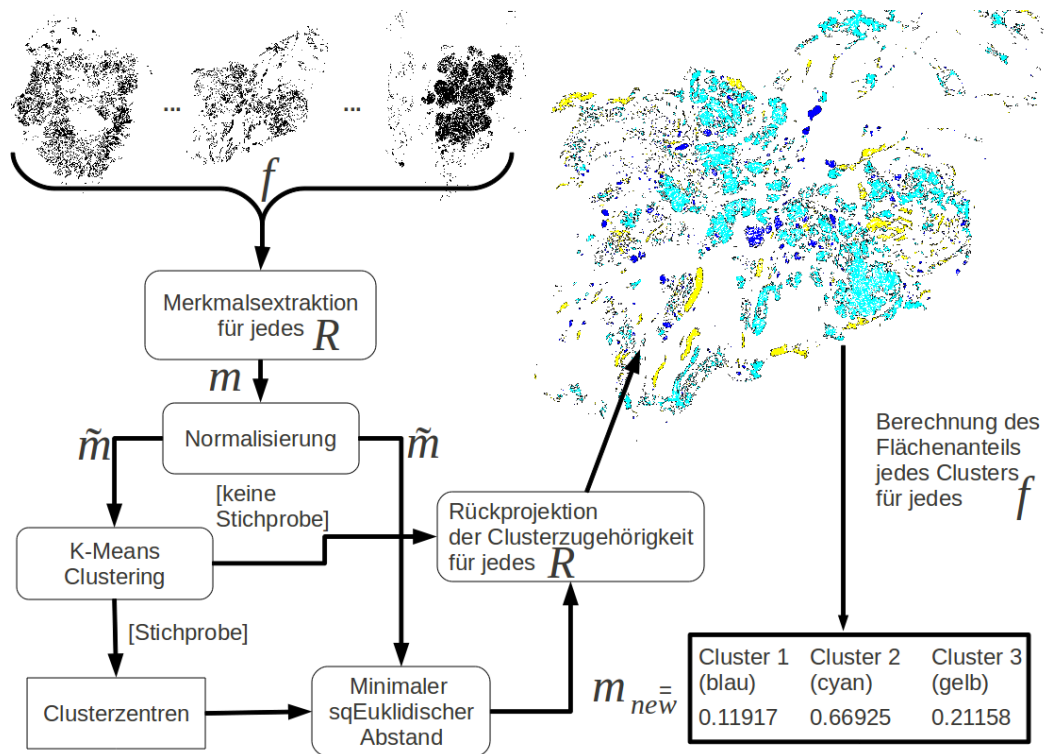


Abbildung 5.10: Schematischer Ablauf des Clusterings der morphometrischen Merkmale: Zunächst werden alle Merkmale für jedes Bildobjekt aus den gegebenen Bildmustern berechnet. Daraufhin werden die Merkmale normalisiert. Durch das *K-Means*-Verfahren werden gemeinsame Eigenschaften der Tumorobjekte zusammengefasst. Hierbei kann eine Stichprobe von Tumorobjekten zur Erstellung eines Modells verwendet werden. Durch Rückprojektion der Clusterzugehörigkeiten jedes Tumorobjektes werden gemeinsame Eigenschaften der Tumorobjekte visualisiert. Für jedes Bildmuster wird ein neuer Merkmalsvektor anhand der prozentualen Fläche jedes Clusters bestimmt.

Unter Verwendung der deskriptiven Statistiken ist eine klare Interpretation nicht möglich, da in einem Bild mehrere Tumorobjekte vorliegen. Aufgrund der Heterogenität in der Größe der Tumorobjekte würde so, beispielsweise ein kleines Tumorobjekt, in seinen morphometrischen Eigenschaften genau so stark gewichtet in die Statistik eingehen, wie ein wesentlich größeres. In diesem Ansatz werden zunächst ähnliche Typen von Tumorobjekten aufgrund ihrer morphologischen Eigenschaft identifiziert und dann aufgrund ihrer Größe gewichtet. Eingesetzt wird diese Methode für f_{bin}^{tumor} und f_{bin}^{ae1ae3} zur morphometrischen Analyse auf zwei Detailstufen.

Ergebnisse ausgewählter Tumorprototypen Anhand der drei Beispielbilder in Abbildung 5.11 sollen die Ergebnisse visualisiert werden. Als Beispiel wurde Clusteranzahl $K = 3$ gewählt. Die Ergebnisse der Rückprojektion beziehen sich auf die Bilder aus Abbildung 5.2.

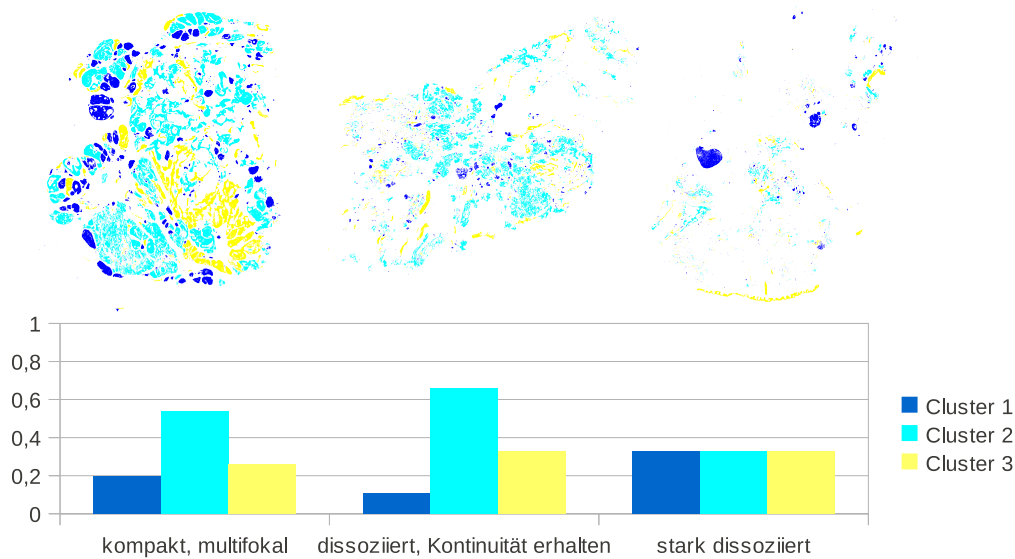


Abbildung 5.11: Beispielergebnis morphometrischer Analysen mit Clusteranzahl 3 und den entsprechenden Clusteranteil der jeweiligen Fälle. Blau markiert kompakte Runde Tumorobjekte, Cyan dissoziierte Tumorobjekte mit unregelmäßiger Berandung und Gelb markiert elongierte Tumorobjekte.

Anhand der Abbildungen 5.11 wird deutlich, dass durch das blaue Cluster (Cluster 1) kompakte abgerundete Tumorobjekte bestimmt werden. Das gelbe Cluster (Cluster 3) markiert elongierte Tumorobjekte und das cyan-farbene Cluster (Cluster 2) dissoziierte Tumorobjekte. Anhand dieser Beispielergebnisse werden erste Einschränkungen deutlich. Viele kompakt erscheinende Regionen werden nicht als solche erkannt. Die Tumorobjekte hängen zum Teil über dünne Verbindungen zusammen, was zu anderen Formen führt. Obwohl die Clusterverteilung auf Grund der unterschiedlichen Fläche jedes Clusters eine Differenzierung zwischen den Fällen erlaubt, entspricht dies nicht den gesuchten Eigenschaften zwischen kompakten und differenzierten Tumoren. Weitere Analysen bezüglich der Clusteranzahl und Bewertung werden in Kapitel 6 dargestellt. Ausgewählte Beispielergebnisse zur morphometrischen Bildanalyse von f_{bin}^{tumor} befinden sich im Anhang (siehe Abbildung C.3).

5.4 Strukturelle Analysen

Um die räumlichen Verteilungsmuster über den gesamten Fall zu erfassen eignen sich Graphmerkmale (siehe Abschnitt 3.2). Als Kantengewichte werden die euklidischen Distanzen $d_e(v_a, v_b)$ (siehe Formel (2.8)) zwischen zwei 2D-Knoten $v_a = (x_a, y_a)$ und $v_b = (x_b, y_b)$ eingesetzt. Die angewendeten Merkmale (siehe Abschnitt 3.2.3) basieren grundsätzlich auf den Kantengewichten zwischen den Knoten und dem Grad eines Knotens.

Merkmal	Par	T	R	S	Kt	Kp	Anforderungen
MST	ja	✓	✓	-	-	✓	räumliche Kompaktheit, räumliche Dissoziierung, Multifokal(mit jeweils ein- heitlichen Verteilungen oder mit jeweils unterschiedlichen Verteilungen)
DG	ja	✓	✓	-	-	✓	räumliche Kompaktheit, räumliche Dissoziierung, Multifokal(mit jeweils ein- heitlichen Verteilungen oder mit jeweils unterschiedlichen Verteilungen)

Tabelle 5.3: Merkmalsauswahl zur Erfüllung der relationalen Anforderungen (Par=Parameter notwendig, T=Translationsinvarianz, R=Rotationinvarianz, S=Skalierungsinvarianz, Kt=Kongruent, Kp=Kompakt).

Die Motivation für den Einsatz auf Gewebeebene liegt darin, dass für räumlich kompakte homogene Tumore geringe, gleichmäßige euklidische Distanzen zwischen den Knoten und ein regelmäßiger Knotengrad zu erwarten sind, während bei dissoziierten räumlich heterogen verteilten Tumorobjekten höhere, ungleichmäßige euklidische Distanzen und Unregelmäßigkeiten im Knotengrad erwartet werden.

Eingesetzt werden im Rahmen dieser Arbeit die Graphstrukturen MST (*Minimum-Spanning-Tree* siehe Abschnitt 3.2.1) und DG (*Delaunay-Graph*, siehe Abschnitt 3.2.2). Die Verwendung des MST's ist durch Gespräche mit den Domänenexperten motiviert, welche diese Graphstruktur als repräsentativ für die Gewebeverteilung sehen. Eingesetzt wird der Algorithmus von Prim [Pri57] (siehe Abschnitt 3.2.1).

Die Berechnung des DG's ist aus der Literatur motiviert. In [ACSGD10] wird dessen vielfache erfolgreiche Anwendung auf Zellebene zur Differenzierung verschiedener Krebseigenschaften hervorgehoben. Auch in [DHS⁺07] und [DAM⁺08] wird der *Delaunay-Graph* zur Erkennung von Brustkrebs bzw. Prostatakrebs eingesetzt. Weitere Anwendungen werden in Kapitel 3.2 aufgezählt.

Da es sich beim MST um einen Baum handelt (siehe Abschnitt 3.2.1), werden unmittelbare Nachbarschaften zwischen den Knoten nicht so gut erfasst, wie es beispielsweise durch den DG möglich sein sollte. Beide Graphstrukturen haben den Vorteil, dass keine Parameter zur Berechnung der Graphen bei gegebener Knotenmenge benötigt werden. Zur Berechnung der Knotenmenge wiederum sind Parameter notwendig (siehe folgender Abschnitt).

Die in diesem Kontext implementierten und angewendeten Merkmale werden in Abschnitt 3.2.3 dargestellt. Die Eigenschaften und abzudeckenden Anforderungen durch die Merkmale werden in Tabelle 5.3 zusammenfassend beschrieben.

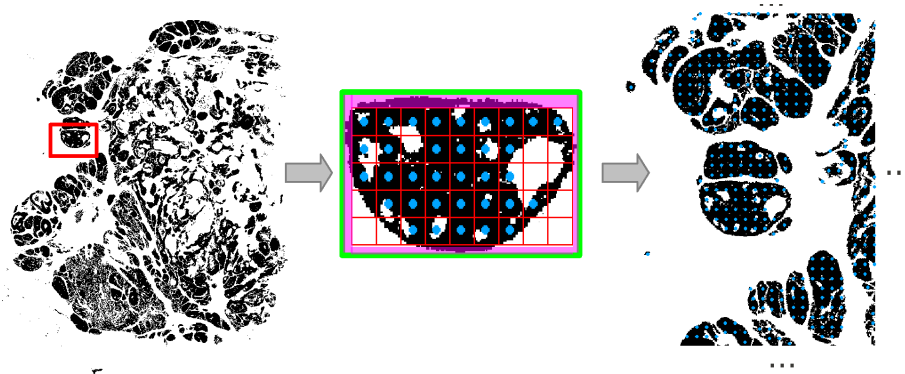


Abbildung 5.12: Knotenidentifikation für Graphberechnungen. Jedes Tumorobjekt wird einzeln durch ein regelmäßiges Raster (rot) auf Knoten (blaue Punkte) untersucht.

Die Graphmerkmale sind aufgrund der eingesetzten Knotenidentifikation (siehe folgender Abschnitt) weitestgehend invariant gegenüber Translation und Rotation. Skalierungsinvarianz wird für die Graphmerkmale nicht benötigt, da zur Erfassung der räumlichen Kompaktheit und Dissoziierung die euklidischen Distanzen entscheidend sind. Eingesetzt werden diese Analysen für f_{bin}^{ae1ae3} , da nach Ansicht der Domänenexperten dort die nötigen Feinstrukturen des Tumors zu erkennen sind. Entscheidend für die Analysen ist eine Knotenidentifikation, welche das Gewebe adäquat repräsentiert. Die angewendete Methode wird im folgenden Abschnitt beschrieben.

5.4.1 Knotenidentifikation

Die Knotenidentifikation ist der zentrale Bestandteil zur Bestimmung der individuellen Graphstrukturen. Zur Identifikation von Knoten wird eine Segmentierung des gesuchten Gewebes benötigt [ACSGD10]. In dieser Anwendung wird hierfür der Vordergrund des nach den Ausführungen in Abschnitt 5.2.1 beschriebenen Binärbildes f_{bin}^{ae1ae3} verwendet.

In den meisten Anwendungen auf Zellebene werden die Zentroide der segmentierten Krebszellen als Knoten eingesetzt (siehe Abschnitt 3.2). Zunächst wurden dementsprechend die Zentroide (\bar{x}_r, \bar{y}_r) (siehe Formel 3.12) der einzelnen Tumorobjekte R^{ae1ae3} aus f_{bin}^{ae1ae3} als Knoten $v(\bar{x}_r, \bar{y}_r)$ verwendet. Aufbauend auf ersten Ergebnissen aus den zuvor beschriebenen morphometrischen Analysen wurde deutlich, dass die Unterschiede in Größe und Form der Tumorobjekte durch die Zentroide als Knoten keine geeignete Repräsentation der Gewebeverteilung erlauben.

Des Weiteren wurde versucht durch einen *Skeletierungsalgorithmus* [Soi98] das Gewebe zu approximieren und daraufhin mittels des *Harris-Corner-Detectors* [BB05] Knoten als Ecken an den Abzweigungen des Bild-Skelettes zu definieren. Dies führte jedoch zu einer Unterrepräsentation des Gewebes bei elongierten Strukturen.

Stattdessen wird der Ansatz aus Bilgin et. al. [BDNY07] auf die vorliegende Problemstellung angepasst. In [BDNY07] werden Zellgraphen von Gewebeproben aus Brustkrebs auf Grundlage eines regelmäßigen Pixelrasters über das gesamte Bild erstellt. Dieser An-

satz wird in der vorliegenden Arbeit übernommen und durch eine differenzierte Analyse der einzelnen Tumorobjekte hinsichtlich der Knotenverteilung erweitert.

Das Vorgehen soll anhand von Abbildung 5.12 erläutert werden. Für jedes Tumorobjekt R wird die Bounding-Box [BB05] aus f_{bin}^{ae1ae3} bestimmt (grünes Quadrat). Daraufhin wird entsprechend der Bounding-Box ein Teilbild f_R^{box} extrahiert, welches R enthält. Durch eine regelmäßig definierte Rasterstruktur mit quadratischer Pixelgröße $n \times n$ (wobei n ungerade sein muss) wird f_R^{box} in i Teilregionen f_R^{box} zerlegt (siehe rotes Raster). Sollte die Rasterstruktur der betrachteten f_R^{box} aufgrund unterschiedlicher Höhen- bzw. Breitenverhältnisse sich nicht eindeutig in $n \times n$ große Teilregionen aufteilen lassen, wird das Raster durch einen entsprechenden Offset (magentafarben) verkleinert zentriert in f_R^{box} dargestellt. Für jede dieser Teilregionen wird der prozentuale Anteil an Tumorgewebe $amount = \frac{Foreground}{n \times n}$ bestimmt. *Foreground* entspricht der Pixelanzahl von Tumorgewebe. Liegt dieser über einem gegebenen Schwellwert $amount > t_{node}$, wird in das Zentrum der Teilregion ein Knoten v (blaue Kreise) gesetzt. Sollte f_R^{box} insgesamt kleiner $n \times n$ bzw. als ein Element der Rasterstruktur sein, wird der Zentroid (\bar{x}_r, \bar{y}_r) (siehe Formel 3.12) des Tumorobjektes R als einziger Knoten v verwendet. So kann sichergestellt werden, dass eine ausreichend genaue Approximation der Gewebestruktur erreicht wird. Der Schwellwert t_{node} wird im Rahmen dieser Anwendung auf 0.5 gesetzt. Die Pixelrastergröße s wird auf 15 Pixel festgelegt. Nachdem die Knoten V festgelegt sind, können die einzelnen Graphstrukturen G berechnet werden. Im Folgenden werden die Ergebnisse ausgewählter Tumorprototypen vorgestellt. Größere Grafiken zu den berechneten Graphstrukturen über f_{bin}^{ae1ae3} befinden sich im Anhang C.5.

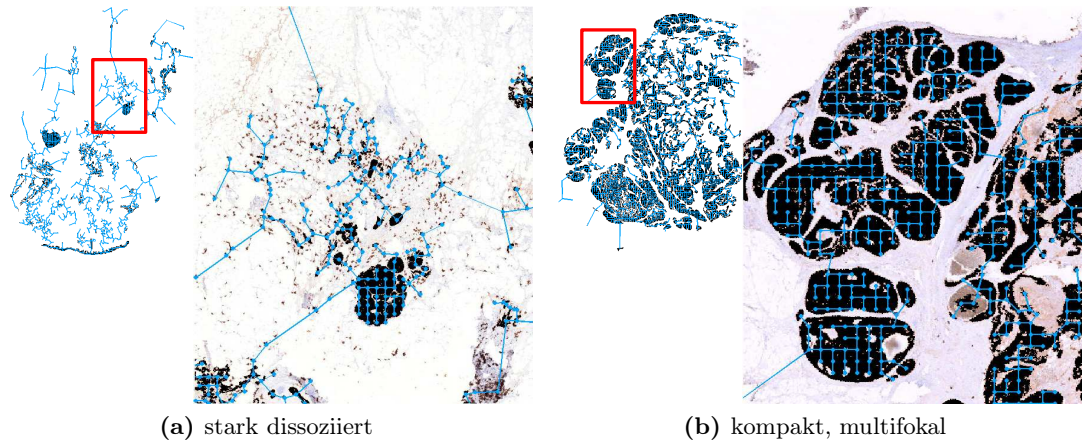


Abbildung 5.13: Berechnete MST's zu Tumorprototypen, welche für Tests auf den realen Daten herangezogen wurden ((a) stark dissoziiert (b) kompakt, multifokal). Dargestellt wird die AE1AE3-Färbung überlagert mit dem detektierten Tumorgewebe (schwarz) und dem berechneten MST-Strukturen (blau).

Fall	N_v	e_μ	e_σ	e_{mima}	e_{vc}	C_n
MST(a)	1829	39.4226	33.4923	0.0140	0.8496	-
MST(b)	3204	32.1810	11.4473	0.0278	0.3557	-
Diff(%)	27.32	10.11	49.05	33.01	40.97	-
	deg_μ	deg_σ	deg_{mima}	deg_{vc}	deg_w	RI
MST(a)	1.9989	0.7161	0.2500	0.3583	39.4010	1.1517
MST(b)	1.9994	0.8294	0.2500	0.4148	32.1709	1.1162
Diff(%)	00.00	07.33	00.00	07.30	10.10	01.56

Tabelle 5.4: MST-Merkmalwerte der Tumorprototypen (stark dissoziierter Tumor (a) und kompakter Tumor (b)). Die Abkürzungen der Merkmalswerte werden in Abschnitt 3.2.3 erläutert.

Ergebnisse ausgewählter Tumorprototypen In Abbildung 5.13 werden zwei vom Domänenexperten deutlich unterschiedene Ausbreitungsmuster von Tumoren ((a) und (b)) und der korrespondierende MST dargestellt. Diese Bildmuster werden als grundsätzlich zu differenzierende Tumorprototypen herangezogen. In der Abbildung ist deutlich die Heterogenität der Kantenlängen in Fall (a) und die homogene Struktur in Fall (b) zu erkennen. Anhand dieser Beispiele soll die grundsätzliche Plausibilität der Merkmale geprüft werden. Die berechneten Merkmale und Differenzen beider Fälle werden in Tabelle 5.4 dargestellt.

Als Vergleich zum MST und DG werden die beiden repräsentativen Fälle in Abbildung 5.13 bzw. 5.14 und die entsprechenden Merkmale und Differenzen in Tabelle 5.4 bzw. 5.5 dargestellt. Auch hier ist die Heterogenität (a) bzw. Homogenität (b) der Graphstruktur zu erkennen.

Sowohl die MST-Merkmale als auch die DG-Merkmale auf Basis der Kantengewichte zeigen deutlich differenzierte Werte. Wie zu erwarten sind beispielsweise das durchschnittliche Kantengewicht (e_μ), die Standardabweichung (e_σ) und der Variationskoeffizient (e_{vc}) für den dissoziierten Fall (a) deutlich höher als für den kompakten Fall (b). Die Merkmalswerte auf Grundlage des Knotengrades zeigen für den MST nur wenig Unterscheidungskraft. Der Wert deg_{mima} als Abschätzung der Spannweite unterscheidet sich nicht. Ähnliches gilt für die DG-Merkmalswerte. Lediglich die Standardabweichung deg_σ lässt auf eine heterogene Struktur in (a) im Verhältnis zu (b) schließen. In Kapitel 6 werden die Merkmale im Hinblick auf ihre Differenzierungsfähigkeit weiter untersucht.

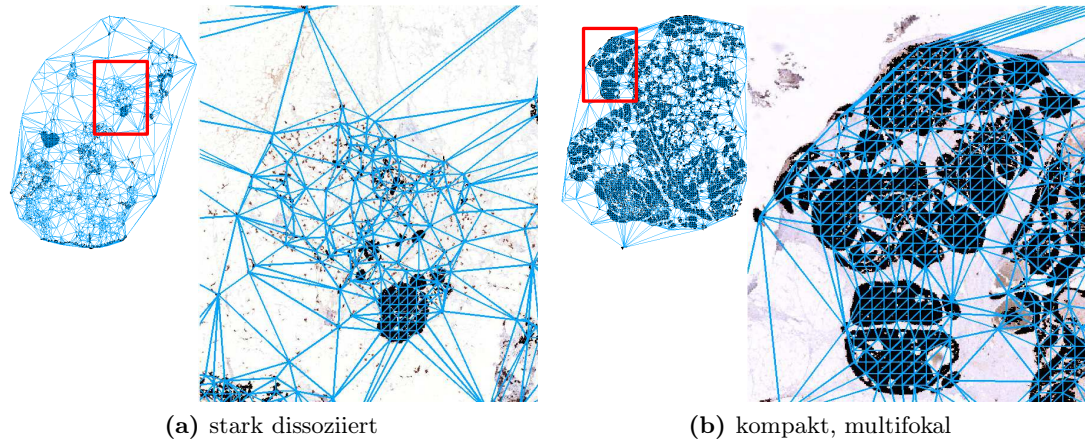


Abbildung 5.14: Berechnete DG's (blau) zu Tumorprototypen, welche für Tests auf den realen Daten herangezogen wurden ((a) stark dissoziiert (b) kompakt, multifokal). Dargestellt wird die AE1AE3-Färbung überlagert mit dem detektierte Tumorgewebe (schwarz) und Delaunay-Graph (blau).

Fall	N_v	e_μ	e_σ	e_{mima}	e_{vc}	C_n
DG(a)	1829	41.37	54.63	0.0023	1.3204	1.3358
DG(b)	3204	26.51	27.64	0.0064	1.0429	1.3344
Diff(%)	27.32	21.89	32.80	47.12	11.74	00.00
	deg_μ	deg_σ	deg_{mima}	deg_{vc}	deg_w	RI
DG(a)	5.9595	1.3628	0.2308	0.228	246.58	1.0389
DG(b)	5.9825	1.0771	0.2143	0.180	158.60	1.0232
Diff(%)	00.00	11.70	03.70	11.76	21.71	00.76

Tabelle 5.5: DG-Merkmalwerte der Prototypen (stark dissoziierter Tumor (a) und kompakter Tumor (b)). Die Abkürzungen der Merkmalswerte werden in Abschnitt 3.2.3 erläutert.

5.5 Relationale Analysen

Um Relationen und Koloalisationen (zentral, holohedral und peripheral) zwischen verschiedenen Bildmodalitäten zu erfassen, wird das RCC8-Histogramm (siehe Abschnitt 3.3) eingesetzt. Für zentrale Verteilungen sollte das RCC8-Histogramm viele Relationen in NTTP/NTPPI (innen) und TPP/TPPI (äußerer Rand innen) anzeigen. Bei peripheralen Verteilungen werden hohe Werte in den Relationen EC, PO (außerhalb) und TPP/TPPI (äußerer Rand innen) erwartet. Holohedrale Verteilungen sollten gleich verteilte Werte in allen Relationen haben.

Für die Anwendung auf Gewebeebene werden verschiedene Anpassungen und Erweiterungen zur Berechnung des RCC8-Histogramms nötig, welche in Abschnitt 5.5.1

Merkmal	Par	T	R	S	Kt	Kp	Anforderungen
RCC8-Histogramm	ja	✓	✓	✓	✓	✓	peripheral, zentral, holohedral
LDQ	ja	✓	✓	✓	✓	✓	peripheral, zentral, holohedral

Tabelle 5.6: Merkmalsauswahl zur Erfüllung der relationalen Anforderungen (Par=Parameter notwendig, T=Translationsinvarianz, R=Rotationinvarianz, S=Skalierungsinvarianz, Kt=Kongruent, Kp=Kompakt).

näher erläutert werden. Aufgrund der im Folgenden beschriebenen Einschränkungen des RCC8-Histogramms wird zusätzlich ein selbstständig entwickelter Ansatz, die *Lineare Distanz-Quantifizierung* (kurz LDQ), zur Quantifizierung der Verteilungen von funktionellen Markern eingesetzt, welcher in Abschnitt 5.5.2 näher beschrieben wird. Eingesetzt werden diese Analysen zwischen f_{bin}^{tumor} und den funktionellen Markern (f_{bin}^{cd45} , f_{bin}^{er} , f_{bin}^{pr}). Tabelle 5.6 gibt einen zusammenfassenden Überblick zu den Eigenschaften der Merkmale. Solange die Positionsverhältnisse zwischen den Tumorobjekten und Relationsobjekten (siehe Abschnitt 5.2.2) erhalten bleiben, sind die eingesetzten Verfahren zur Bestimmung der unterschiedlichen Relationen grundsätzlich invariant gegenüber Translation, Rotation und Skalierung. Für beide Analysen werden Parameter benötigt (siehe folgender Abschnitt).

5.5.1 Erweiterung des RCC8-Histogramms

Das RCC8-Histogramm eignet sich aufgrund seines intuitiven Verständnisses der erfassten Relationen zur Bestimmung der gesuchten Eigenschaften (holohedral, zentral und peripheral) und gibt zusätzliche Informationen, wie die Relationen DC¹² und EQ¹³.

Durch die intuitiv verständlichen Relationsdefinitionen (siehe Abbildung 3.5) sollte eine Verringerung der Semantischen Lücke möglich sein.

Das RCC8-Histogramm wird zwischen den Tumorobjekten R^{tumor} und den Relationsobjekten R^{cd45} , R^{er} und R^{pr} angewendet (siehe Abschnitt 5.2.2). Die Tumorobjekte R^{tumor} bieten die nötige Vergrößerung des Tumorgewebes, um Abschätzungen für die Relationen zu erlauben. Da die Berechnungen für die Relationsobjekte im RCC8-Histogramm gleich ablaufen, werden diese im Folgenden für die Erläuterungen der Berechnungen des RCC8-Histogramms zusammenfassend als R^{rel} bezeichnet.

In den Beschreibungen des RCC8-Histogramms (siehe Abschnitt 3.3) werden Einschränkungen deutlich, auf Grund derer für die Anwendung auf Gewebeebene eine Anpassung der Algorithmik durchgeführt wird. Es ist in der Anwendung auf Gewebeebene nicht ausreichend die konvexe Hülle der Regionen zu verwenden, da in f_{bin}^{tumor} sehr konvexe Bildobjekte vorliegen können (siehe Abbildung 5.15). In der vorliegenden Arbeit werden daher stattdessen alle zentralen Aushöhlungen eines Tumorobjekts und Relati-

¹²disconnected (getrennt)

¹³equal (gleiche Position und Überlagerung)

onsobjekts R zunächst gefüllt. Dafür wird auf die entsprechende Funktionalität¹⁴ aus [Matc] zurückgegriffen. So bleibt die grundlegende Forminformation erhalten.

Im RCC8-Histogramm nach [Bou08] werden keine Größeninformationen der Relationen berücksichtigt. Entscheidend in der vorliegenden Arbeit ist die Frage wie viel, beispielsweise entzündetes Gewebe, sich zentral oder am Rand des Tumors befindet. Es ist ein Unterschied, ob beispielsweise ein Relationsobjekt R^{cd45} , welches entzündetes Gewebe repräsentiert, 300 Pixel groß ist oder nur 100 Pixel. Eine Integration von Größeninformation in das RCC8-Histogramm muss daher erfolgen.

In [Bou08] wird nicht klar, ob berücksichtigt wird, dass ein Bildobjekt der Modalität a auch mehrere Relationen mit Bildobjekten der Modalität b (siehe Abschnitt 3.3) und umgekehrt haben kann. Im Rahmen der vorliegenden Arbeit wird daher die Relation jedes Tumorobjektes R^{tumor} zu jedem Relationsobjekt R^{rel} bestimmt.

Der Schwellwert von 5 Pixeln zur Unterscheidung zwischen sich überlappenden Bildobjekten (PO) und am inneren Rand liegenden Regionen (TPPI bzw. TPP) (siehe Abschnitt 3.3) ist bei äußerst unterschiedlichen Größen von Bildobjekten nicht ausreichend, um den Bias der zugrunde liegenden Segmentierung und Registrierung zu berücksichtigen. Aus diesem Grund werden in der vorliegenden Arbeit adaptive Schwellwerte auf Grundlage der Größeninformation der Bildobjekte eingeführt.

Die implementierte und angepasste Algorithmik zur Bestimmung der Relationen wird im Folgenden dargestellt. Der angepasste Algorithmus wird formell unter Verwendung von Mengenoperationen zwischen den Pixeln der einzelnen Bildobjekte in Algorithmusdarstellung 5 beschrieben.

¹⁴mit Hilfe der MATLAB-Funktion `imfill`

```

FOREACH  $R^{rel}$ 
  FOREACH  $R^{tum}$ 
    IF ( $R^{tum} \cap R^{rel} = \emptyset$  &  $\min(d(p(R^{tum}), p(R^{rel}))) > t_{dDCToEC}$ )
       $\Rightarrow$   $RCC8 = DC(R^{tum}, R^{rel})$ 
    ELSE IF ( $\min(d(p(R^{tum}), p(R^{rel}))) \leq t_{dDCToEC}$  &  $\#(R^{tum} \cap R^{rel}) / \# \min(R^{tum}, R^{rel}) < t_{jDCToEC}$ )
       $\Rightarrow$   $RCC8 = EC(R^{tum}, R^{rel})$ 
    ELSE IF ( $\#(R^{tum} \cap R^{rel}) / \# \min(R^{tum}, R^{rel}) \geq t_{jDCToEC}$  &  $\#(R^{tum} \cap R^{rel}) / \# \min(R^{tum}, R^{rel}) < t_{jECToPO}$ )
       $\Rightarrow$   $RCC8 = PO(R^{tum}, R^{rel})$ 
    ELSE IF ( $R^{tum} \subset R^{rel}$  &  $\#(R^{tum} \cap R^{rel}) / \# \min(R^{tum}, R^{rel}) \geq t_{jECToPO}$  &  $\min(d(p(R^{tum}), p(R^{rel}))) \leq t_{dTPPToNTPP}$  &  $\#((R^{tum} \cup R^{rel}) \setminus R^{tum} \cap R^{rel}) / \#(R^{tum} \cup R^{rel}) > t_{errToEQ}$ )
       $\Rightarrow$   $RCC8 = TPP(R^{tum}, R^{rel})$ 
    ELSE IF ( $R^{rel} \subset R^{tum}$  &  $\#(R^{tum} \cap R^{rel}) / \# \min(R^{tum}, R^{rel}) \geq t_{jECToPO}$  &  $\min(d(p(R^{tum}), p(R^{rel}))) \leq t_{dTPPToNTPP}$  &  $\#((R^{tum} \cup R^{rel}) \setminus R^{tum} \cap R^{rel}) / \#(R^{tum} \cup R^{rel}) > t_{errToEQ}$ )
       $\Rightarrow$   $RCC8 = TPPI(R^{tum}, R^{rel})$ 
    ELSE IF ( $R^{tum} \subset R^{rel}$  &  $\#(R^{tum} \cap R^{rel}) / \# \min(R^{tum}, R^{rel}) \equiv \min(R^{tum}, R^{rel})$  &  $\min(d(p(R^{tum}), p(R^{rel}))) > t_{dTPPToNTPP}$ )
       $\Rightarrow$   $RCC8 = NTPP(R^{tum}, R^{rel})$ 
    ELSE IF ( $R^{rel} \subset R^{tum}$  &  $\#(R^{tum} \cap R^{rel}) / \# \min(R^{tum}, R^{rel}) \equiv \min(R^{tum}, R^{rel})$  &  $\min(d(p(R^{tum}), p(R^{rel}))) > t_{dTPPToNTPP}$ )
       $\Rightarrow$   $RCC8 = NTPPI(R^{tum}, R^{rel})$ 
    ELSE IF ( $\#((R^{tum} \cup R^{rel}) \setminus R^{tum} \cap R^{rel}) / \#(R^{tum} \cup R^{rel}) \leq t_{errToEQ}$ )
       $\Rightarrow$   $RCC8 = EQ(R^{tum}, R^{rel})$ 
    IF  $RCC8 \neq DC(R^{tum}, R^{rel}) \Rightarrow RCC8Hist(RCC8) += A^{(uf R^{rel})} //$ 
      oder (combined)  $A^{(uf R^{tum})} + A^{(uf R^{rel})}$ 
    IF ALL  $RCC8(R^{tum}, R^{rel}) \equiv DC(R^{tum}, R^{rel}) \Rightarrow RCC8Hist(DC) +=$ 
       $A^{(uf R^{rel})} //$  oder (combined)  $A^{(uf R^{tum})} + A^{(uf R^{rel})}$ 
    END
  END
  Normalisiere RCC8Hist-Bins durch Gesamtsumme der flächengewichteten Relationen

```

Algorithmus 5: Erweiterung des Algorithmus zur Bestimmung des RCC8-Histogramms um Größeninformationen, Mehrfachrelationen und Schwellwerte. Die einzelnen Relationen werden in Abschnitt 3.3 erläutert und visualisiert (R^{tumor} wird in dieser Darstellung zur besseren Übersicht als R^{tum} bezeichnet).

Schwellwert	Bedeutung	Wert
$t_{dDCToEC}$	Euklidischer Abstand zwischen äußeren Konturen der Bildobjekte zur Differenzierung zwischen den Relationen DC und EC	1
$t_{dTPPToNTPP}$	Euklidischer Abstand zwischen äußeren Konturen (bei Überlagerung) der Bildobjekte zur Differenzierung zwischen den Relationen TPP und NTPP bzw. TPPI und NTPPI	1
$t_{jDCToEC}$	Minimale Überlagerung (in %) der Bildobjekte zur Bestimmung der Relation PO	0.1
$t_{jECToPO}$	Minimale Überlagerung (in %) der Bildobjekte zur Bestimmung der Relation TPP	0.9
$t_{errToEQ}$	Erlaubte maximal Abweichung (in %) zwischen zwei Bildobjekten zur Bestimmung der Relation EQ	0.05

Tabelle 5.7: Erläuterungen zu verwendeten Schwellwerten zur Bestimmung der RCC8-Relationen.

In einer doppelten Schleife wird jedes Relationsobjekt R^{rel} mit den Tumorobjekten R^{tumor} auf die jeweilig vorliegende Relation untersucht. Dies führt zu einer sehr hohen Anzahl von DC-Relationen, wodurch die anderen Relationen überlagert würden. Aus diesem Grund wird eine Relation zwischen zwei Bildobjekten R nur als von einander getrennt (DC) betrachtet, wenn zwischen einem Relationsobjekt R^{rel} und allen Tumorobjekten R^{tumor} keine andere Relation vorliegt. Der Term $\min(R^{tumor}, R^{rel})$ bestimmt das kleinere Bildobjekt. Die Anzahl an Pixelelementen eines Bildobjektes wird durch $\#(R)$ definiert. Die Kontur wird durch p angegeben. Die euklidische Distanz zwischen den Konturen von R wird durch d beschrieben.

Die eingesetzten Schwellwerte wurden empirisch anhand synthetischer und ausgewählter realer Beispiele getestet. In Tabelle 5.7 werden die Bedeutungen und die verwendeten Parameter der Schwellwerte aus Algorithmusdarstellung 5 dargestellt. Der Schwellwert $t_{dDCToEC}$ definiert den euklidischen Abstand der zwischen den beiden Konturen der Bildobjekte besteht, um zwischen DC und EC zu unterscheiden. Er wird, wie in [Bou08], auf 1 gesetzt. Der Schwellwert $t_{dTPPToNTPP}$ beschreibt den Übergang von TPP zu NTPP (bzw. TPPI zu NTPPI) und wird ebenfalls auf 1 gesetzt. Die Schwellwerte $t_{jDCToEC}$ und $t_{jECToPO}$ mit $t_{jDCToEC} + t_{jECToPO} = 1$ definieren den prozentualen Anteil, den das kleinere Bildobjekt an der überlappenden Fläche zwischen den beiden Bildobjekten R^{tumor} und R^{rel} haben muss, um zwischen den Relationen EC und PO sowie PO und TPP bzw. TPPI in Abhängigkeit von der Größe der Bildobjekte, adaptiv unterscheiden zu können. Der Schwellwert $t_{jDCToEC}$ wird auf 0.1 festgelegt. Erst wenn 10% der kleineren Region ($\min(R^{tumor}, R^{rel})$) mit der größeren Region überlappen, wird die Relation PO bestimmt. Wenn es weniger als 10% sind, wird EC festgelegt. Der Schwellwert $t_{jECToPO}$ wird dementsprechend auf 0.9 gesetzt. Sobald mehr als 90% der kleineren Regi-

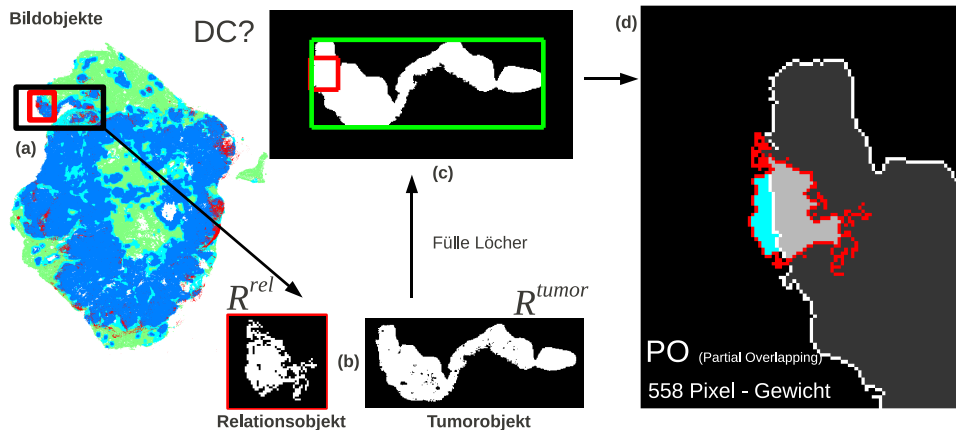


Abbildung 5.15: Visualisierung der einzelnen Schritte zur Bestimmung einer RCC8-Relation: Dargestellt werden das Tumorgewebe (blau) mit überlagertem entzündetem Gewebe (rot) (siehe (a)). Als Relation wird in diesem Fall PO (partial overlapping) bestimmt und mit 558 Pixeln gewichtet.

on ($\min(R^{tumor}, R^{rel})$) mit der größeren Region überlappen, liegt kein Relation PO mehr vor, sondern die Relation TPP bzw. TPPI. Diese Schwellwerte erlauben eine intuitive Repräsentation der Relationen in Abhängigkeit der Größe der Bildobjekte. Verzerrungen auf Grund der Segmentierung und Registrierung kann so besser entgegengewirkt werden. Die Größeninformation $A(R)$ der ungefüllten Bildobjekte ${}^{uf}R$ gewichtet jede Relation. Dabei wird zwischen zwei Varianten unterschieden.

Je nach Auswahl kann eine kombinierte Gewichtung (*combined*) aus den Größen beider Bildobjekte R^{tumor} und R^{rel} erfolgen oder nur durch die Größe von R^{rel} . Die Relationsobjekte werden einzeln gewichtet, ohne die Größeninformation der Tumorobjekte, da ansonsten ein verfälschtes Bild entstehen würde.

Bei kombinierter Gewichtung würde eine Relation zwischen einem sehr kleinen Relationsobjekt und einem sehr großen Tumorobjekt aus f_{bin}^{tumor} ein sehr hohes Gewicht bekommen. Wesentlich größere Relationsobjekte des funktionellen Markers eines kleineren Tumorobjekts, welche eine andere Relation haben, würden deutlich weniger gewichtet. Verwendet man die nicht-kombinierte Gewichtung verliert das RCC8 seine symmetrische Eigenschaft. Bei kombinierter Gewichtung bleibt die Symmetrie erhalten.

Das normalisierte RCC8-Histogramm (RCC8Hist) wird nicht bezüglich der Anzahl der Relationen bestimmt (so wie in [Bou08]), sondern durch die Größe $A(R^{rel})$ (bzw. $A(R^{rel}) + A(R^{tumor})$ bei kombinierter Gewichtung), die eine Relation unter Verwendung von R^{rel} hat. Durch die Normalisierung anhand der Gesamtsumme aller Bins des Histogramms wird Skalierungsinvarianz erreicht. Durch die doppelte Schleife wird jedes Relationsobjekt R^{rel} mit jedem Tumorobjekt R^{tumor} verglichen, wodurch Mehrfachrelationen ermöglicht werden.

Anhand von Abbildung 5.15 soll beispielhaft die Bestimmung einer Relation PO (Partial Overlapping) dargestellt werden. Abgebildet wird f_{rgb}^{som} mit überlagertem f_{bin}^{cd45} (rot)

(siehe (a)). Es werden zwei Bildobjekte, das Relationsobjekt aus f_{bin}^{cd45} ($R^{rel} = R^{cd45}$) und das Tumorobjekt aus f_{bin}^{tumor} (R^{tumor}) extrahiert (siehe (b)). Zunächst werden die Löcher beider Bildobjekte gefüllt (siehe (c)). Danach wird durch die Bounding-Box (BB) beider Bildobjekte (rot und grün) abgeschätzt, ob eine DC-Relation vorliegt (siehe (c)). Ist dies der Fall, brauchen keine weiteren Berechnungen durchgeführt werden. In diesem Beispiel überschneiden sich die BB (siehe (c)), was zu weiteren Analysen führt. Zunächst wird die Kontur beider Bildobjekte extrahiert (rot für R^{rel} aus f_{bin}^{cd45} und weiß für R^{tumor} aus f_{bin}^{tumor} , siehe (d)). Anhand der Konturen wird mittels des minimalen euklidischen Abstandes der Pixel wiederum auf Relation DC getestet. Trotz überlagerter BB könnte bei sehr konkaven Bildobjekten weiterhin DC vorliegen. Die hellgraue Pixelregion entspricht $(R^{tumor} \cap R^{rel})$. Die kleinere Region $min(R^{tumor}, R^{rel})$ ist in diesem Fall R^{rel} (rot) und entspricht der vereinigten Pixelregion zwischen cyan-farbener Region und hellgrauer Region (siehe (d)). Da der Anteil der kleineren Region R^{rel} (rot) zu $(R^{tumor} \cap R^{rel})$ ($\#(R^{tumor} \cap R^{rel}) / \#min(R^{tumor}, R^{rel})$) über 10% und unter 90% liegt, wird die Relation PO bestimmt und durch die ungefüllte Fläche von $A(R^{rel})$ (siehe (b)) mit 558 Pixeln gewichtet und dem Histogramm-Bin der Relation PO zugeordnet. Ein weiteres erläuterndes Beispiel befindet sich im Anhang (siehe Abbildung C.7).

Der Aufwand des Algorithmus ist $O(n_{tumor} \cdot n_{rel})$, wobei n_{tumor} die Anzahl der Tumorobjekte R^{tumor} angibt und n_{rel} die Anzahl der Relationsobjekte R^{rel} .

Bei der Anwendung auf die realen Daten kommt es auf Grund des nahezu quadratischen Aufwandes, der verhältnismäßig langsamen Zielsprache MATLAB und der großen Bilddaten zu sehr langen Laufzeiten (> 1 Woche) für die einzelnen Berechnungsdurchläufe des RCC8-Histogrammes. Verschiedene Versuche der Parallelisierung innerhalb von MATLAB oder Vorberechnungen einzelner Ergebnisse führten zu erhöhten Speicheraufkommen (> 4 GB). Erst durch die Abschätzung der DC-Relation anhand der Bounding-Box beider Tumorobjekte konnte ein entscheidender Geschwindigkeitsvorteil ohne zu hohes Speicheraufkommen erzielt werden.

Dennoch musste der Schwellwert t^θ (siehe Abschnitt 5.2.2) für die Berechnung auf 100 Pixel für die eingesetzten Bildmodalitäten festgelegt werden, um die Anzahl der Bildobjekte zu verringern und so eine akzeptable Laufzeit zu erreichen.

An dieser Stelle muss angemerkt werden, dass durch die eingeführten adaptiven Schwellwerte eine klare mathematische Definition nach [RCC92] nicht mehr möglich ist.

Es kann durchaus sein, dass es Fälle gibt, in denen auf Grundlage dieses Algorithmus und der gewählten Schwellwerte eine Relation nicht eindeutig bestimmt werden kann. Aus diesem Grund wurde ebenfalls eine Rückweisungsrelation NOTHING implementiert. Eine Relation dieser Art fließt nicht in die Gewichtung ein. Tests anhand der synthetischen Daten (siehe Abschnitt 5.6.2) und ausgewählter realer Daten haben gezeigt, dass dieser Fall vernachlässigt werden kann.

Ergebnisse ausgewählter Tumorprototypen An dieser Stelle soll ebenfalls anhand von ausgewählten Tumorprototypen zunächst die grundlegende Sinnhaftigkeit für die Anwendung getestet werden. Die Ergebnisse beziehen sich auf die Fälle in Abbildung 5.3 (a), (b) und (c).

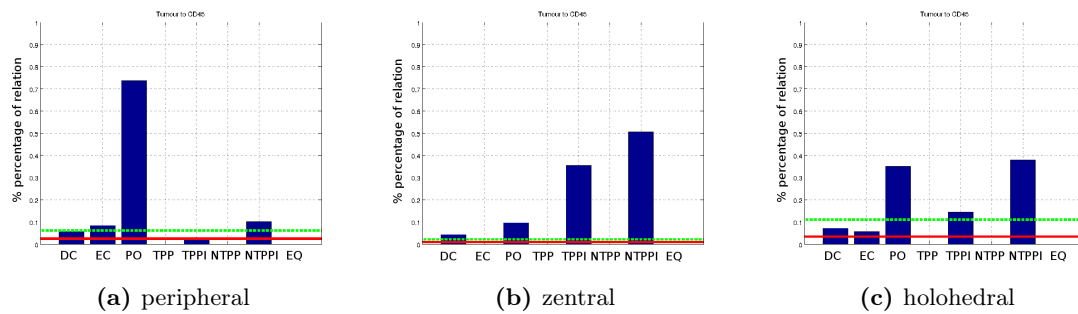


Abbildung 5.16: Beispielergebnisse der RCC8-Analysen zu den Tumorprototypen (peripheral (a), zentral (b) und holohedral (c)) aus Abbildung 5.3). Der rote Balken gibt die verwendete Information im Verhältnis zur vorhandenen Information (grün gestrichelter Balken) an.

In (a) überwiegt die Relation PO. Dies spricht für eine periphere Verteilung des funktionellen Markers CD45. In (b) überwiegen die Relationen TPPI und NTPPI, was für eine zentrale Verteilung spricht. Die Verteilung in Diagramm (c) zeigt die Relationen PO und NTPPI in gleicher Ausprägung, während die anderen Relationen deutlich unterhalb der beiden Relationen liegen. Somit kann von einer holohedralen Verteilung ausgegangen werden. Die Ergebnisse der Tumorprototypen stimmen mit den Annotationen des Domänenexperten überein. Deutlich wird in (c) die Problematik des Schwellwertes. Für die Analysen in (c) wurde lediglich die Hälfte der vorhandenen Information verwendet. Zu diesem Wert kommt es, da in Abbildung 5.3 c) sehr viele kleine Tumorobjekte (teilweise einzelne Pixel) vorliegen, welche auf Grund des Schwellwertes nicht berücksichtigt werden.

5.5.2 Lineare Distanz-Quantifizierung

Da wegen des Zeitaufwandes zur Berechnung der RCC8-Merkmale ein hoher Schwellwert ($t^{\ell} = 100$) verwendet werden muss, wurde ein zusätzlicher, selbst entwickelter Ansatz zur Quantifizierung der Verteilungen von funktionellen Marken umgesetzt. Mit diesem Ansatz sollen die RCC8-Merkmale unterstützt werden und zusätzliche Informationen verwendet werden. Auf Grundlage der *Linearen Distanztransformation* (siehe [Soi98]) können für jedes Tumorobjekt R^{tumor} ausgehend von der Kontur die euklidischen Distanzen jedes Pixels zum nächsten Konturpunkt bestimmt werden. Durch diese Entfernungsinformation kann eine Einteilung in unterschiedliche räumliche Areale erfolgen. Zur Bestimmung der *Linearen Distanztransformation* wird auf die entsprechende Funktionalität aus [Matc] zurückgegriffen¹⁵. Anhand von Abbildung 5.17 soll das Vorgehen verdeutlicht werden.

Abbildung 5.17 zeigt f_{rgb}^{som} mit überlagertem f_{bin}^{cd45} (rot). Zunächst wird anhand der Bounding-Box (siehe [BB05]) eines Tumorobjektes in f_{bin}^{tumor} ein Teilbild f_r^{box} , welches das Tumorobjekt R^{tumor} enthält, extrahiert (siehe (a)). Entsprechend dieser Bounding-Box

¹⁵mit Hilfe der MATLAB-Funktion `bwdist`

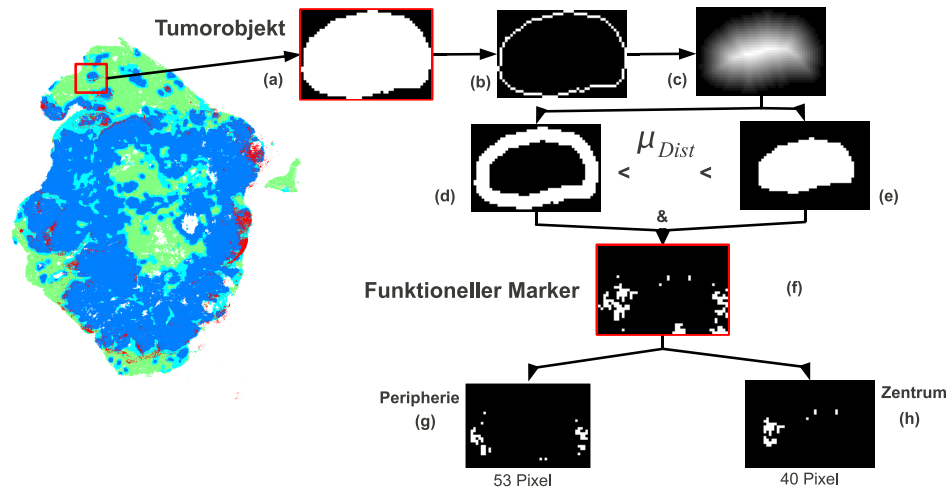


Abbildung 5.17: Visualisierung der einzelnen Schritte der Linearen-Distanz-Quantifizierung.

wird der korrespondierende Bildausschnitt des funktionellen Markers bestimmt (siehe (f)).

Daraufhin werden die Konturen von R^{tumor} in f_r^{box} ermittelt (siehe (b)). Anhand der Konturen werden die euklidischen Distanzen d jedes Pixels mittels der *Linearen Distanztransformation* (siehe [Soi98]) zum nächstgelegenen Konturpunkt bestimmt (siehe (c)). Dabei werden nur die Distanzen innerhalb von R^{tumor} betrachtet. In Folge dessen wird der Mittelwert der Distanzen μ_{Dist} berechnet.

Als zentrale Region reg_z werden alle Pixel bestimmt, deren Distanzwerte d größer als μ_{Dist} sind (siehe (e)). Alle Distanzwerte d kleiner gleich μ_{Dist} werden als periphere Region reg_p betrachtet (siehe (d)). Durch logische Verknüpfung (&), der als Binärmasken eingesetzten Regionen zwischen reg_z und R^{tumor} bzw. reg_p und R^{tumor} , wird die Menge an Pixeln in der zentralen bzw. peripheralen Region von R^{tumor} bestimmt (siehe (g) bzw. (h)). Dieser Vorgang wird für jedes Tumorbild innerhalb von f_{bin}^{tumor} ausgeführt.

Die Menge an, nach dieser Definition, peripher LDQ_p gelegenen und zentral LDQ_z gelegenen Pixeln für jedes Tumorbild R^{tumor} wird aufsummiert und gespeichert. Durch Normalisierung der beiden Werte durch die Gesamtsumme entsteht eine Verteilung LDQ (genannt Lineare Distanz-Quantisierung) mit $LDQ_p + LDQ_z = 1$. Somit sollen innerhalb der Tumorbilder die in den Anforderungen betrachteten Fälle peripheral (hoher LDQ_p -Wert), zentral (hoher LDQ_z -Wert) und holohedral (LDQ_p und LDQ_z relativ gleich verteilt) bestimmt werden. Zusätzlich wird an dieser Stelle der %-Anteil LDQ_{amount} von funktionellen Markern zu Tumorgewebe in f_{bin}^{tumor} bestimmt, um beispielsweise zwischen Fällen mit viel und mit wenig entzündetem Gewebe zu unterscheiden.

Der Vorteil dieses Vorgehens liegt darin, dass die funktionellen Marker nicht als Bildobjekte (bzw. Relationsobjekt) interpretiert werden (wie beim RCC8-Histogramm), sondern jeder einzelne Pixelwert betrachtet wird. Außerdem ist die Berechnung aufgrund der einfachen Bildarithmetik sehr effizient, so dass keine Einschränkung der Datenmenge



Abbildung 5.18: Periphere Regionen innerhalb zentraler Aushöhlungen bei Anwendung der LDQ-Analysen: a) Tumorobjekt (R^{tumor}), b) zentrale Region (reg_z), c) periphere Region (reg_p).

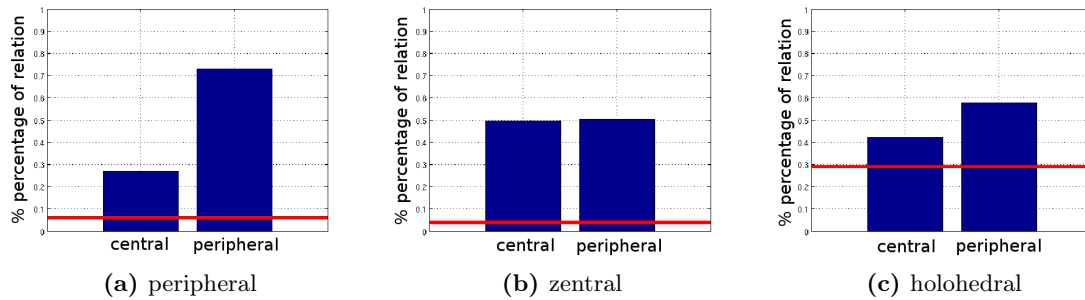


Abbildung 5.19: Beispielergebnisse der LDQ-Analysen zu den Tumorprototypen (peripheral (a), zentral (b) und holohedral (c)) aus Abbildung 5.3). Der rote Balken gibt LDQ_{amount} an.

erfolgen muss. Der Nachteil dieses Vorgehens liegt im Vergleich zum RCC8-Histogramm darin, dass keine Informationen außerhalb der Tumorobjekte erfasst werden.

Ein wesentlicher Unterschied zum RCC8-Histogramm ist, dass die Tumorobjekte vor den Berechnungen nicht gefüllt werden, so dass auch periphere Regionen innerhalb zentraler Aushöhlungen von Tumorobjekten bestimmt werden (siehe Abbildung 5.18). Durch Anwendung beider Ansätze soll ein besseres Gesamtbild der Verteilungen entstehen. Durch Multiplikation eines Schwellwertes $t_{ldq} \in \{0, 1\}$ mit μ_{Dist} kann die Grenze der peripheren und zentralen Region variiert werden. In Tests an synthetischen Daten erwies sich der Mittelwert der Distanzen als am effektivsten. Der Aufwand des Algorithmus beträgt $O(n)$ wobei n für die Anzahl der Tumorobjekte steht.

Ergebnisse ausgewählter Tumorprototypen Die Ergebnisse werden in den Diagrammen in Abbildung 5.19 zusammenfassend dargestellt. Als Tumorprototypen dienen die drei Fälle aus Abbildung 5.3 bezüglich der räumlichen Verteilungen von f_{bin}^{cd45} in f_{bin}^{tumor} .

Diagramm (a) deutet mit über 70% peripheral und unter 30% zentral auf eine primär periphere Verteilung hin. Die Ergebnisse aus (c) zeigen weniger periphere (unter 60%) und mehr zentrale (über 40%) Verteilungen des funktionellen Markers. Die zentrale Verteilung wird durch 50% zentral und 50% peripheral im Verhältnis zu den anderen beiden Tumorprototypen plausibel differenziert dargestellt. Die Ergebnisse der LDQ-Relationen lassen sich den Beobachtungen des Domänenexperten zuordnen. Von einer sinnhaften

Differenzierung durch die LDQ-Verteilungsmerkmale wird auf Grund der Ergebnisse bei den Tumorprototypen ausgegangen.

5.6 System zur Validierung und Evaluierung der Merkmale

Um zu bestimmen, inwieweit sich durch die zuvor beschriebenen Bildanalysen generalisierende relevante Unterscheidungen von Mustern abzeichnen, muss eine Evaluation dieser Verfahren durchgeführt werden. Die Schwierigkeit liegt dabei in der fehlenden Ground-Truth bezüglich der beobachteten Ausbreitungsmuster jedes individuellen Patientenfalles (siehe Kapitel 1).

Im Rahmen des Projektes ExPrimage hat sich gezeigt, dass durch kategorielle Annotation (Zuweisung eindeutiger Klassenzugehörigkeiten) des Domänenexperten keine ausreichend reliable Übereinstimmung über alle Fälle erreicht werden kann. So liegen zwar Annotationen des Domänenexperten vor (siehe Abschnitt 4.1), diese lassen sich allerdings nicht alle eindeutigen Kategorien zuordnen. Die Bewertungen und Beobachtungen der Domänenexperten sind in der Realität mehr auf einzelne Tumorfälle bezogen als generalisierend.

Durch vorgeschaltetes, mehrstufiges validiertes Maschinelles Lernen kann dem Domänenexperten auf Grundlage der Merkmale ein Vorschlag mit hoher Informationsverdichtung gemacht werden, welcher anschließend bewertet wird (Post-Labeling). Im Rahmen der vorliegenden Arbeit wird hierzu ein Konzept vorgeschlagen, welches zunächst an synthetischen Daten getestet und schließlich bei ausgewählten Beispielen auf den realen Daten eingesetzt wird (siehe Kapitel 6).

5.6.1 Konzept

Das entwickelte Konzept soll anhand der schematischen Abbildung 5.20 erläutert werden. Nach der Registrierung der Bildmuster und Gewebecharakterisierungen (siehe Kasten „Vorverarbeitung“) erfolgt zunächst die in den vorigen Abschnitten beschriebenen Bildanalysen zur Merkmalsextraktion. Es werden grundsätzlich zwei Varianten der Validierung unterschieden. Zum einen wird anhand der Klassifikationsrate, der Trefferquote und der Präzision (siehe Abschnitt 2.4.1) geprüft, inwieweit die entwickelten Merkmale die Klassifikation des HLVQ-Verfahrens (siehe Abschnitt 2.1) im Hinblick auf den Verlauf (*Follow-Up-Status*) verbessern (siehe Kasten „Klassifikation“). Zum anderen erfolgt eine gesonderte unüberwachte Evaluation, welche im Rahmen der vorliegenden Arbeit entwickelt wurde. Durch sie soll geprüft werden, inwieweit sich differenzierte Tumorgruppen durch die entwickelten Merkmale darstellen. Dieser Ansatz soll im Folgenden näher vorgestellt werden.

Die extrahierten Merkmalswerte m der Merkmalsvektoren \mathbf{m} werden aufgrund unterschiedlicher Wertebereiche zunächst auf einen einheitlichen Wertebereich $m \in \{0, 1\}$ normiert. Allgemein wird davon ausgegangen, dass Merkmale normal verteilt sind [AH00]. In [Run10] wird zwischen Normalisierung von Gleichverteilung und Normalverteilung unterschieden. Sowohl Aksoy et. al. [AH00] als auch in den weiteren Ausführungen in Runkler

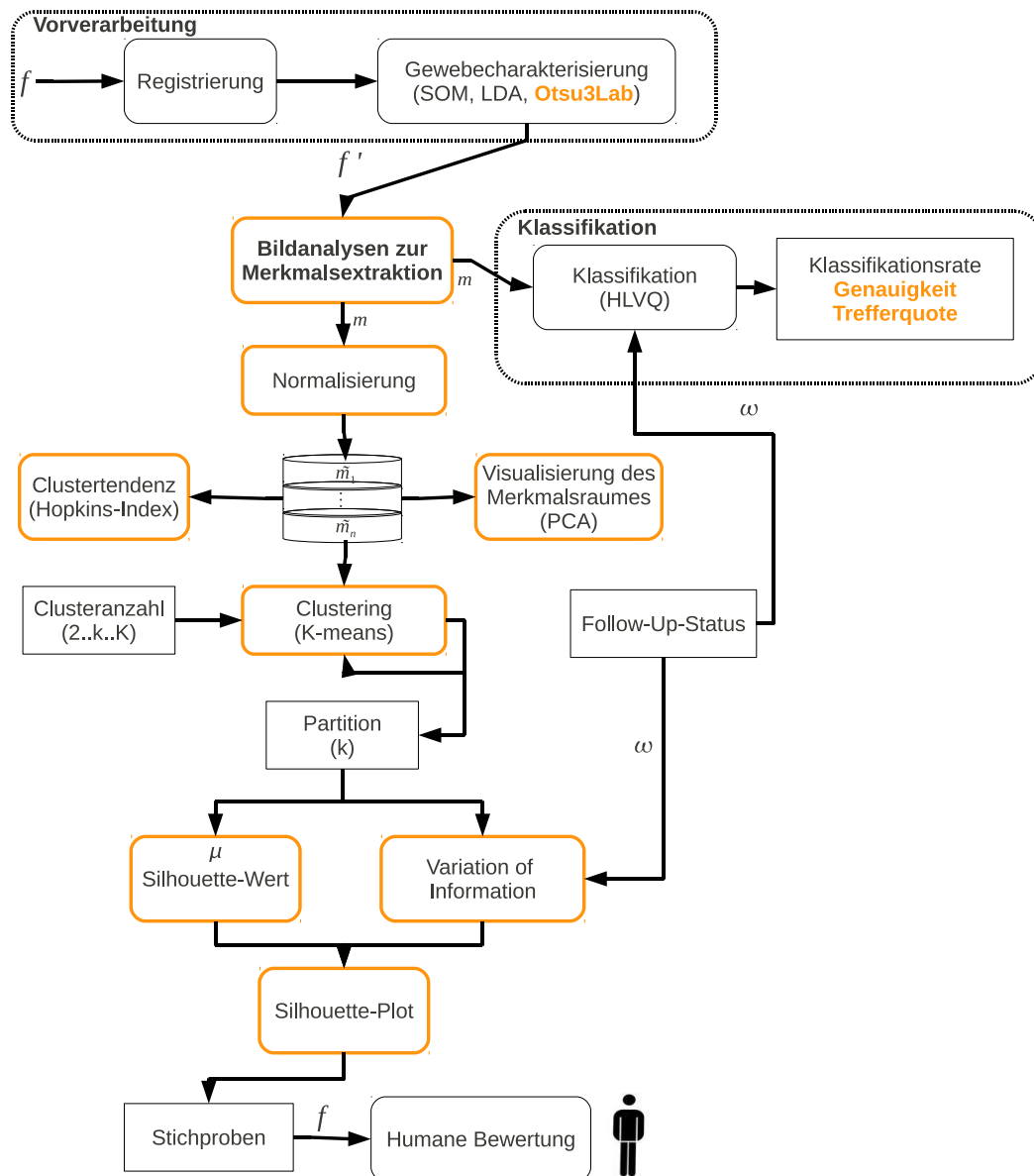


Abbildung 5.20: Konzept zur Validierung der Bildanalysen: Nach der Vorverarbeitung (siehe Kasten „Vorverarbeitung“) erfolgen die Bildanalysen zur Merkmalsextraktion. Die Merkmale werden daraufhin in einem überwachten Klassifikationsschritt getestet (siehe Kasten „Klassifikation“). In einem weiteren unüberwachten Validierungsverfahren sollen die entwickelten Merkmale auf ihre Differenzierungsleistung zwischen verschiedenen Tumoren getestet werden. Die Ergebnisse sollen dem Domänenexperten als Diskussionsgrundlage dienen, welcher anhand von Stichproben einen ersten Vorschlag für Zusammenhänge in den Ausbreitungsmustern erhält (im Rahmen der vorliegenden Arbeit hinzugefügte Komponenten sind orange markiert).

[Run10] wird erklärt, dass solche Annahmen nicht adäquat sind und das Ergebnis verzerren können. Sie schlagen eine Approximation von Funktionen oder Verteilungen vor. Aufgrund der zeitlichen Einschränkung wird in der vorliegenden Arbeit lediglich zwischen Gleichverteilung und Normalverteilung unterschieden. Hierzu werden die Merkmale der einzelnen Merkmalsdimensionen mittels des *Jarque-Bera-Test* [JB87] zunächst auf Normalverteilung geprüft. Ist dies der Fall, erfolgt eine Normalisierung der Merkmalswerte m nach

$$\tilde{m} = \frac{\frac{m - \mu_M}{3\sigma_M} + 1}{2} \quad (5.5)$$

, wobei μ_M der Mittelwert und σ_M die Standardabweichung von Merkmal M der betrachteten Merkmalsvektoren ist [AH00]. Dies garantiert, dass 99% von \tilde{m} innerhalb des Wertebereiches $\{0, 1\}$ liegen [AH00]. Liegt nach dem *Jarque-Bera-Test* keine Normalverteilung vor, wird davon ausgegangen, dass eine Gleichverteilung vorliegt. In diesem Fall werden die Merkmalswerte durch

$$\tilde{m} = \frac{m - m_{max}}{m_{max} - m_{min}} \quad (5.6)$$

auf eine einheitliche Skala normiert, wobei m_{max} den Maximalwert und m_{min} den Minimalwert des Merkmals M repräsentieren.

Anhand der normalisierten Merkmale wird zunächst die Clustertendenz mittels des Hopkins-Index 2.4.2 geprüft. Dies gibt einen ersten Einblick in die Datenlage und prüft inwieweit trennbare homogene Gruppen von Mustern vermutet werden. Der Hopkins-Index wird zehn mal auf den ersten drei PCA-Projektionen (PC1 zu PC2, PC2 zu PC3 und PC1 zu PC3) ausgeführt. Als zufällige Auswahl der Merkmalsvektoren für die Mengen U und V werden entsprechend [Fer00] 10% aller Merkmalsvektoren verwendet. Das Ergebnis wird gemittelt. Die Aussagekraft des Hopkins-Index wird entsprechend den in Abschnitt 2.4.2 beschriebenen Statistiken überprüft.

Eine Visualisierung der Merkmalsvektoren in 3D durch die PCA (siehe Abschnitt 2.4.3) unter Berücksichtigung der abgebildeten Varianzen ermöglicht eine visuelle Kontrolle. Sind auf dieser Ebene bereits Cluster zu sehen, kann davon ausgegangen werden, dass solche vorliegen (siehe Abschnitt 2.4.3).

Als unüberwachtes Lernverfahren wird hierzu das *K-Means*-Verfahren 2.2 unter Verwendung der quadratischen euklidischen Distanz (2.8) eingesetzt, welche nach [XW05], die am häufigsten eingesetzte Metrik darstellt.

An dieser Stelle sei angemerkt, dass mit dieser Metrik im unüberwachten Evaluations-schritt (Clusteranalyse) der vorliegenden Arbeit auch Abstände zwischen Merkmalen bestimmt werden, für welche die euklidische Metrik nicht adäquat ist (beispielsweise zwischen Verteilungen (RCC8-Histogramm und LDQ-Merkmale)).

Aufgrund des eingeschränkten Zeitrahmens konnte dieser Punkt nicht weiter berücksichtigt werden. Initialisiert werden die Clusterzentren anhand einer zufälligen Auswahl

von vorliegenden Merkmalsvektoren. Da das *K-Means*-Verfahren zu lokalen Minima konvergieren kann, wird an dieser Stelle das Clustering 10-mal mit unterschiedlich initialisierten Clusterzentren (Prototypen) durchgeführt. Das Partitionsergebnis mit dem geringsten durchschnittlichen quadratischen euklidischen Abstand der Merkmalsvektoren zu den Clusterzentren wird als Ergebnispartition verwendet.

Da das *K-Means*-Verfahren als Parameter die Clusteranzahl K benötigt, muss die Clusteranzahl geschätzt werden. Hierfür wird das *K-Means*-Verfahren mit unterschiedlichen Clusteranzahlen $2, \dots, k, \dots, K$ ausgeführt, wobei K ein festzulegender Parameter ist. Für jede Partition wird der durchschnittlichen Silhouetten-Werte SV_μ (siehe 2.4.4) bestimmt. Die normalisierte Variation of Information VI_n (siehe Abschnitt 2.4.5) wird im Rahmen der vorliegenden Arbeit für die Bestimmung der Güte der Clusteranalyse im Bezug zu gegebenen Klassenzugehörigkeiten eingesetzt. Verwendet man beispielsweise für P die Klassen Ω mit den gegebenen Klassenzugehörigkeiten ω jedes Merkmalsvektors, kann man durch die VI_n validieren wie groß die Übereinstimmung der Partition des Clusterergebnisses C mit der Klassenzugehörigkeiten der Merkmalsvektoren ist. Ein Vorteil der VI_n ist, dass keine Übereinstimmung in Clusteranzahl und Klassenanzahl der zu vergleichenden Verteilungen benötigt wird. Somit kann beispielsweise die sinnvollste Clusteranzahl, im Sinne der Übereinstimmung der gegebenen Klassenzugehörigkeiten, für die Anwendung des *K-Means*-Verfahrens bestimmt werden.

Mittels externer Clustervalidierung durch die VI_n wird nach einer Partition von Mustern gesucht, die schon in der unüberwachten Validierung auf einen Zusammenhang mit den bekannten Klassenzugehörigkeiten ω des *Follow-Up-Status* hinweisen. Sollte dies der Fall sein, könnte ein unmittelbarer Zusammenhang zwischen Ausbreitungsmuster und Verlauf hergestellt werden.

Welche Partition näher untersucht wird, entscheidet die Signifikanz des berechneten SV_μ oder der VI_n . Liegt ein signifikanter Wert für die VI_n vor, wird die entsprechende Partition aufgrund der Relevanz bzgl. des Verlaufes näher untersucht. Liegt keine Partition mit signifikanten VI_n vor, wird die Partition mit dem höchsten SV_μ -Wert näher untersucht.

Näher untersuchen bedeutet in diesem Zusammenhang, dass in einem Plot die Silhouetten-Werte jedes Musters der zu untersuchenden Partition dargestellt wird. Anhand des höchsten Silhouetten-Wertes (am kompaktesten innerhalb des Cluster und am besten separiert von den anderen Clustern (siehe Abschnitt 2.4.4)), wird das Muster identifiziert, welches jedes Cluster am besten repräsentiert. Zusätzlich werden weitere Muster mit den nächst höchsten Silhouetten-Werten, in Form einer Stichprobe, den Domänenexperten vorgelegt, welche dann eine humane Bewertung der Tumorgruppierungen hinsichtlich ihrer Plausibilität durchführen. Hierfür können Evaluationsbögen eingesetzt werden. In Abbildung C.8 (Anhang) wird ein Beispiel für einen konzipierten Evaluationsbogen eines Testfalles aus Abschnitt 6.2 dargestellt.

Durch die unterschiedlichen Methoden soll ein möglichst objektives Gesamtbild der extrahierten Merkmale hinsichtlich ihrer Differenzierungsleistung gegeben werden. Auch das Auffinden von zu differenzierenden Tumorgruppierungen im Merkmalsraum wird so unterstützt. Sollten beispielsweise sowohl der Hopkins-Index, die Visualisierung durch

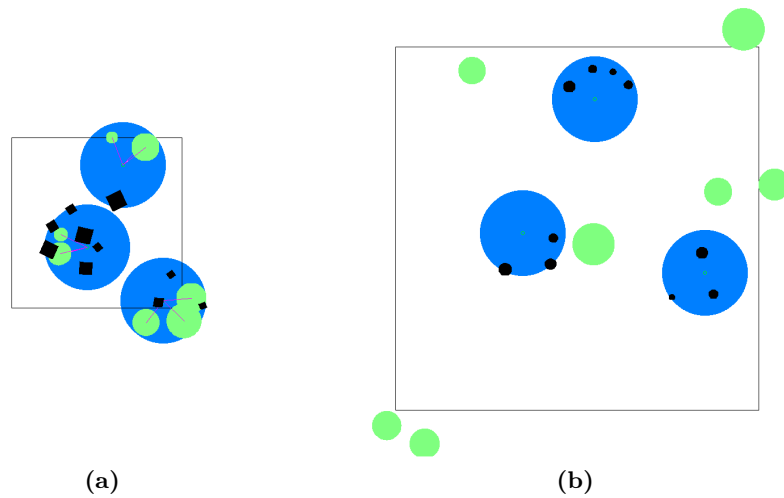


Abbildung 5.21: Beispiel randomisierte synthetische Daten. (a) repräsentiert einen fokalen Fall mit unterschiedlicher Verteilung und der Relation TPP. (b) repräsentiert einen multifokalen Fall mit unterschiedlichen Verteilungen und der Relation DC.

die PCA, die externe Validierung und die interne Validierung klare Hinweise auf Gruppierungen liefern, kann man davon ausgehen, dass sich klar differenzierte Tumorgruppen für diese Daten anhand der Merkmale bestimmen lassen. Dies spricht für ein relevantes Ergebnis, welches dann human kontrolliert werden muss. Dabei dienen die verschiedenen Maße als erste Abschätzung und Informationsverdichtung, so dass vom Domänenexperten nur relevante Ergebnisse begutachtet werden müssen.

5.6.2 Erzeugung randomisierter synthetischer Daten

Um zum einen die implementierten Merkmale und Algorithmen auf ihre Korrektheit und Sinnhaftigkeit zu untersuchen und zum anderen das im vorigen Abschnitt vorgestellte Konzept zur Validierung der Ergebnisse zu testen, werden synthetische Daten herangezogen. Da für die zu Untersuchenden Faktoren Struktur und Relation, so weit bekannt, keine adäquaten öffentlich zugänglichen synthetischen Bilddaten existieren, wurde in Anlehnung an [HW05] ein Algorithmus entwickelt, welcher automatisch randomisierte Bilddaten mit Klassenzugehörigkeiten erstellt.

Anhand von Abbildung 5.21 soll die Funktionsweise des implementierten Algorithmus vorgestellt werden. Ziel ist es mit diesen Daten die gesuchten Eigenschaften im Tumor zu approximieren. Eine fixe Bildgröße (beispielsweise 800x800 Pixel) sei gegeben. Zunächst wird zufällig bestimmt, ob es sich um einen fokalen oder multifokalen (siehe Abschnitt 5.1) Fall handelt. Dies erfolgt anhand eines fixen Schwellwertes, der den Bildbereich für die zu verteilenden Basisobjekte (repräsentativ für Tumorobjekte f_{bin}^{tumor} als blaue Kreissegmente dargestellt) einschränkt. In der Abbildung wird dieser Bereich als schwarzer Kasten angezeigt. Liegt ein fokaler Fall vor, werden Kreise von fester Größe

zufällig innerhalb des kleineren schwarzen Kastens verteilt (siehe (a)). Bei multifokalen Fällen wird der Bereich für die zufällige Verteilung vergrößert (siehe (b)). So kann die Klassenzugehörigkeit fokal bzw. multifokal definiert werden.

Aufbauend auf den Basisobjekten wird zufällig gleichverteilt eine der 8 betrachteten Relationen des RCC8 (siehe Abschnitt 3.3) für das Gesamtbild ausgewählt. Daraufhin wird eine zufällige Anzahl (innerhalb eines gegebenen Wertebereiches) von Relationsobjekten (grüne ausgefüllte Kreissegmente) unterschiedlicher Größe erzeugt. Die Verteilung der Relationsobjekte ist abhängig von der vorliegenden Relation. Liegt die Relation DC vor, werden in Abhängigkeit der Radien und Zentroide der Relationsobjekte und Basisobjekte die Relationsobjekte so zufällig im Bildbereich verteilt, dass kein Basisobjekt berührt wird (siehe (b)).

Liegen Relationen EC, PO, TPP, TPPI, NTPP oder NTPPI vor, wird für jedes Relationsobjekt ein Abstand in Abhängigkeit der Größe eines zufällig ausgewählten Basisobjektes und der zuvor definierten Relation ausgehend vom Zentroiden bestimmt (siehe (a) margentafarbene Linie). Dieser Abstand bestimmt den erlaubten Radius auf dem das Relationsobjekt zufällig verteilt werden darf. Dabei können verschiedene Relationsobjekte übereinander liegen. Die Unterscheidung der größenabhängigen Abstände der Relationen wurde empirisch bestimmt. Bei allen Relationen kann ein Fehlerbereich eingestellt werden in dessen Rahmen die Relationsobjekte von den vorgegebenen relativen Abstand abweichen können. Dieser wurde für die Tests in Kapitel 6 auf 5% festgelegt. Die invertierten Relationen TPPI und NTPPI werden durch vertauschen der Basisobjekte und Relationsobjekte realisiert. Für jeden synthetischen Fall wird die dementsprechende Klassenzugehörigkeit der berechneten Relation ausgegeben. In Abbildung (a) wird beispielsweise eine TPP-Relation dargestellt.

Zur Simulation der Tumorobjekte aus f_{bin}^{ae1ae3} werden innerhalb der Basisobjekte wiederum einfache geometrische Formobjekte verteilt. Dabei wird zufällig zwischen Kreisen, Dreiecken und Quadraten für jeden synthetischen Fall gewählt, welche drei grundlegende unterschiedliche morphometrische Klassenzugehörigkeiten repräsentieren. Um die Invarianzeigenschaften der Merkmale zu testen, können die Formobjekte innerhalb der Basisobjekte zufällig unterschiedlich skaliert, transliert und rotiert (siehe (a) Quadrate) werden. Gleichzeitig werden zwei unterschiedliche Anzahlen von zu verteilenden Formobjekten angegeben. So kann eine Simulation und automatische Bestimmung der Klassenzugehörigkeiten bezüglich der Verteilungstypen „einheitliches Verteilungsmuster“ und „unterschiedliche Verteilungsmuster“ erfolgen (siehe Abschnitt 5.1), da bei höherer Anzahl von Formobjekten in den Basisobjekten eine Dichte einheitliche Besetzung erfolgt und bei niedrigerer eine dünne uneinheitliche (siehe (a) und (b)).

Die Basisobjekte, Relationsobjekte und Formobjekte werden jeweils in unterschiedlichen Bildern gespeichert. Durch einen Parameter kann die Anzahl der zu erstellenden Fälle angegeben werden. Die Klassenzugehörigkeiten zu jedem synthetischen Fall werden automatisch ausgegeben. Getestet werden können an diesen Daten grundlegende morphometrische Merkmale und Abläufe, strukturelle Graphmerkmale und die relationalen Merkmale.

5.7 Implementation

An dieser Stelle wird auf die konkrete Umsetzung der Bildanalysen und das entwickelte Datenmodell (siehe Abschnitt 5.7.2) sowie die eingesetzten Technologien (siehe folgender Abschnitt) eingegangen.

5.7.1 Verwendete Technologien

Zu Beginn des Projektes wurde festgelegt, die einzelnen Aufgaben in der Skript-Sprache *MATLAB* umzusetzen, welche bereits eine Reihe an benötigter Funktionalität in den entsprechenden Bibliotheken (Toolboxes) zur Verfügung stellt. *MATLAB* ist für die Erprobung und schnelle Umsetzung neuer Algorithmen gut geeignet. In der vorliegenden Arbeit wurde *MATLAB* in der Version R2009b eingesetzt. Für Bildverarbeitungsaufgaben konnte dabei auf die *ImageProcessing*-Toolbox [Matc] zurückgegriffen werden, welche grundlegende Funktionalitäten für die Bildverarbeitung bereitstellt. Es wurde darauf geachtet nur dokumentierte, *MATLAB* interne Toolboxes zu verwenden.

Für den implementierten Algorithmus zur Erstellung synthetischer Bildtestdaten (siehe Abschnitt 5.6.2) wurde für den Export der Bilddaten das Matlabscript „export_fig“¹⁶ eingesetzt. Das *K-Means*-Verfahren (siehe Abschnitt 2.2), der Silhouetten-Wert (siehe Abschnitt 2.4.4) und die PCA (siehe Abschnitt 2.4.3) wurden aus der *Statistic-Toolbox* [Matd] verwendet.

Sowohl die Mehrzahl der eingesetzten Merkmale als auch der Hopkins-Index und die normalisierte Variation of Informaiton stehen in den Standard-Toolboxes von *MATLAB* nicht zur Verfügung. Abgesehen von den Merkmalen Area und Perimeter wurden daher alle verwendeten Merkmale eigenständig implementiert.

Der *Hopkins-Index* (siehe Abschnitt 2.4.5) sowie die Umsetzung der normalisierten *Variation of Informaiton* mittels Kontingenztafeln zur Validierung der Merkmale wurden ebenfalls eigenständig implementiert.

5.7.2 Datenmodell

In *MATLAB* wird vorzugsweise prozedural entwickelt, was schnell zu Unübersichtlichkeit im Quelltext und den zusammengesetzten Abläufen führen kann. Außerdem wird es so schwierig, auf sich schnell ändernde Anforderungen zu reagieren.

Um diesen Schwierigkeiten entgegenzuwirken, wurde entschieden die Möglichkeit von *MATLAB* zur objekt-orientierten Programmierung (OOP) einzusetzen. Hierfür wurde ein Datenmodell entwickelt, welches die Hauptkomponenten der angewendeten Analysen in Klassen kapselt und die Polymorphie der OOP ausnutzt.

Dabei wurden die einzelnen entwickelten Komponenten in *Packages* organisiert. Die entwickelten und erweiterten Algorithmen zur Clusteranalyse befinden sich im *Package Clustering*. Die vorverarbeitenden Schritte (siehe Abschnitt 5.2) und die Algorithmik zur Erstellung der synthetischen Daten werden in den *Packages Preprocessing* bzw. *Creating-SyntheticData* gekapselt.

¹⁶Oliver Woodford, <http://www.mathworks.com/matlabcentral/fileexchange/23629>

Das *Package FeatureComputation*, als zentrale Komponente der entwickelten Analysen und Merkmale, soll anhand des Klassendiagrammes in Abbildung 5.22 vorgestellt werden.

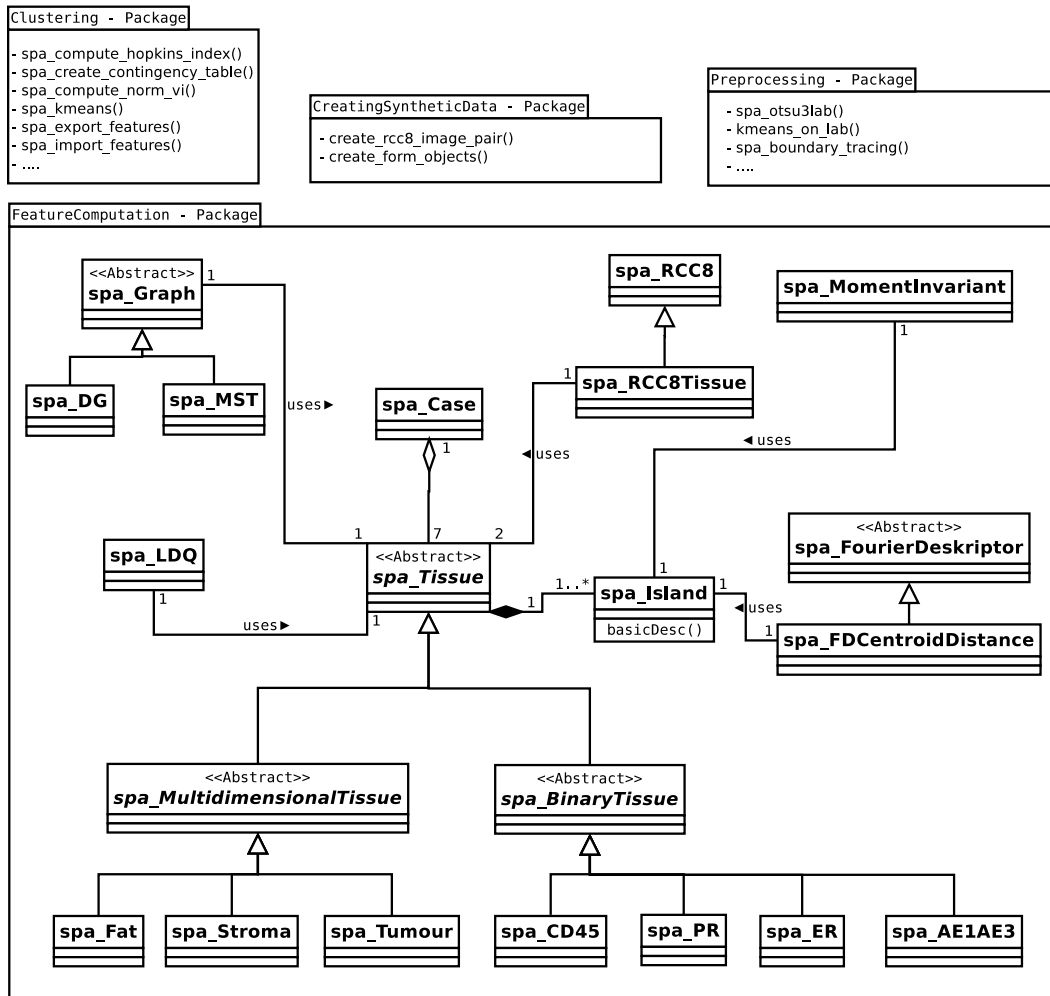


Abbildung 5.22: Klassendiagramm Package: FeatureComputation. Die Methoden und Attribute wurden zur besseren Übersicht

In diesem Klassendiagramm werden die implementierten Klassen¹⁷ und ihre Zusammenhänge untereinander zusammenfassend dargestellt. Auf eine detaillierte Auflistung der Attribute und Methoden wird zur besseren Übersicht verzichtet.

In der Klasse `spa_Case` werden alle berechneten Merkmale für einen Patientenfall in Form von `spa_Tissue`-Instanzen aggregiert. Die abstrakte Klasse `spa_Tissue` repräsentiert eine Gewebeentitäten. Sie wird zu den abstrakten Klassen `spa_BinaryTissue` und `spa_MultidimensionalTissue` spezialisiert. Die Klasse `spa_BinaryTissue` repräsentiert

¹⁷das Kürzel „spa“ steht für Spatial-Pattern-Analysis

die als Binärbilder bzw. Bildmodalitäten vorliegenden Gewebeentitäten der funktionellen Marker ($f_{bin}^{cd45}, f_{bin}^{er}, f_{bin}^{pr}$) und AE1AE3 (f_{bin}^{ae1ae3}). Die spezifischen Eigenschaften werden in den spezialisierten konkreten Klassen `spa_AE1AE3`, `spa_CD45`, `spa_ER` und `spa_PR` implementiert.

Die abstrakte Klasse `spa_MultidimensionalTissue` kapselt die interne Behandlung von f_{rgb}^{som} , welche die Eigenschaften der drei Gewebeentitäten in den konkretisierten Klassen `spa_Fat` (f_{bin}^{fat}), `spa_Stroma` (f_{bin}^{stroma}) und `spa_Tumor` (f_{bin}^{tumor}) umsetzt.

Die `spa_Tissue`-Klasse enthält die Menge von Tumorobjekte R in Form eines Vektors einer Bildmodalität. Jedes R wird durch die Klasse `spa_Island` repräsentiert. Die Klasse `spa_Island` speichert jegliche benötigte Information eines Tumorobjektes innerhalb einer Bildmodalität. In ihr ist die Funktionalität zur Berechnung der Basis-Deskriptoren umgesetzt (`basicDesc()`). In der Klasse `spa_Tissue` sind die Berechnungen für die deskriptive Statistik (siehe Abschnitt 5.3.1) der morphometrischen Analysen der Tumorobjekte (`spa_Island`) implementiert.

Die zuvor beschriebenen Klassen kapseln die Bildmuster sowie deren berechnete Merkmalswerte und sind primär für das Datenmanagement zuständig. Die Funktionalität der Merkmalsextraktion ist weitestgehend in unabhängigen Klassen gekapselt. Die Klasse `spa_MomentInvariant` verwendet die `spa_Island`-Klasse zur Berechnung der Momente (siehe Abschnitt 3.1.3) jedes Tumorobjektes. Gleiches gilt für die Berechnung der Fourier-Deskriptoren, deren Funktionalität in der Klasse `spa_FDCentroidDistance` gekapselt ist. Die Klasse `spa_FDCentroidDistance` spezialisiert und konkretisiert die Klasse `spa_FourierDescriptors`. Dabei ist die Klasse `spa_FourierDescriptors` als *Template-Method* implementiert. So können unterschiedliche Shape-Repräsentationsfunktionen (siehe Abschnitt 3.1.1) ohne Veränderung der Algorithmik hinzugefügt werden.

Die Berechnungen der einzelnen RCC8-Relationen zwischen zwei Tumorobjekten in Form der Klasse `spa_Island` werden in der Klasse `spa_RCC8` realisiert. Durch Spezialisierung zur Klasse `RCC8Tissue` wird die Funktionalität auf die Bestimmung des normalisierten RCC8-Histogramms (siehe Abschnitt 5.5.1) zwischen zwei `spa_Tissue`-Instanzen erweitert. Die abstrakte Klasse `spa_Graph` implementiert die Berechnung zur Knotenidentifikation (siehe Abschnitt 5.4.1) und definiert die gemeinsame Schnittstelle zur Berechnung der Graphmerkmale sowie die Distanz- und Adjazenzmatrix (siehe Abschnitt 3.2.3). In den spezialisierten und konkretisierten Klassen `spa_DG` (für Delaunay-Graph) und `spa_MST` (für Minimum-Spanning-Tree) ist die Implementation für die Berechnung der Graphstrukturen, Graphmerkmale und Visualisierungen enthalten. Diese Funktionalität verwendet eine `spa_Tissue`-Instanz.

Durch das Datenmodell ist es möglich, die Ergebnisse der Merkmalsextraktion in Form von serialisierten `spa_Case`-Instanzen zu speichern. Für darauf aufbauende Berechnungen (beispielsweise Evaluationen) können die Objekte mit allen Informationen wieder in den Hauptspeicher geladen werden. Die Funktionalität zur Berechnung der Merkmale ist weitestgehend von der Datenhaltung gekapselt, wodurch Veränderungen und Erweiterungen an der Funktionalität und Algorithmik effizient umgesetzt werden können.

Die implementierte Klassenhierarchie und Algorithmen können als generelle *Toolbox* zur Bestimmung von Ausbreitungsmustern auf binären Bilddaten eingesetzt werden. In

Verbindung mit der integrierten Clustervalidierung hat der Anwender die Möglichkeit die Merkmalsdaten hinsichtlich der erstellten Gruppierungen näher zu untersuchen.

Kapitel 6

Experimente und Ergebnisse

Dieses Kapitel beschreibt die durchgeführten Experimente und deren Ergebnisse. Die Auswahl der getesteten Merkmalskombinationen erfolgte anhand von semantisch zusammenhängenden Merkmalsgruppen.

In Abschnitt 6.1 werden die Tests und Ergebnisse auf synthetischen Daten zur grundlegenden Validierung der entwickelten Merkmale und Clusteranalysen vorgestellt. Im Rahmen dieser Tests werden erste Einschränkungen der Merkmale deutlich.

Abschnitt 6.2 beschreibt die Anwendung der Merkmale und Clusteranalysen auf die realen Tumorbilddaten. In Abschnitt 6.3 werden die Klassifikationsergebnisse des überwachten HLVQ-Verfahrens unter Verwendung der entwickelten Merkmale vorgestellt.

Die Merkmalsgruppen wurden wiederum zunächst einzeln und die aussichtsreichsten schließlich mit den klinischen Daten zusammen getestet. Die eingesetzten Merkmalskombinationen werden unter Angabe des Testfalls in jedem Experiment anhand der im Abschnitt B aufgelisteten Merkmalskonfigurationen angegeben.

Angemerkt sei an dieser Stelle, dass aufgrund des eingeschränkten Zeitrahmens nur eine Auswahl von Tests und Merkmalskombinationen durchgeführt werden konnte. Lediglich eine Auswahl von Testdurchläufen wurde im Rahmen der vorliegenden Arbeit dokumentiert. Die dokumentierten Experimente der realen Daten beschränken sich auf die morphometrischen Analysen von f_{bin}^{ae1ae3} , die strukturellen Analysen von f_{bin}^{ae1ae3} und die relationalen Analysen zwischen f_{bin}^{tumor} und f_{bin}^{cd45} bzw. f_{bin}^{er} .

6.1 Synthetische Daten - Clusteranalysen

Um die implementierten Algorithmen, Merkmale und Abläufe zunächst auf ihre grundlegende Korrektheit zu prüfen, wurden Tests auf synthetischen Daten durchgeführt. Dabei wurden Daten verwendet, bei denen die Klassenzugehörigkeiten bekannt sind. Zusätzlich soll an dieser Stelle überprüft werden, inwieweit sich durch die implementierten Merkmale und Clustervalidierungsmethoden korrekte Klassenzugehörigkeiten identifizieren lassen. Angewendet wird hierfür das in Abschnitt 5.6 vorgestellte Konzept.

Für diese Experimente wurde, in Abhängigkeit der zu testenden Merkmalsgruppen, zunächst eine individuelle Datenbasis mit Hilfe des in Abschnitt 5.6.2 beschriebenen

Algorithmus erzeugt. Zudem wurde auf eine öffentlich zugängliche Form-Datenbank zurückgegriffen. Die getesteten Clusteranzahlen können dabei je nach Testfall variieren.

Für jeden Test werden die Ergebnisse des Hopkins-Index zwischen den ersten drei PCA-Hauptachsen dargestellt („PC1 to PC2“, „PC1 to PC3“, „PC2 to PC3“, siehe Abschnitt 2.4.2). Die normalisierte Variation of Information (VI_n) (siehe Abschnitt 2.4.5) und der durchschnittliche Silhouetten-Wert (SV) (siehe Abschnitt 2.4.4) werden in einem Diagramm zusammengefasst.

Aufbauend auf das morphometrische Clustering (siehe Abschnitt 5.3) wird in einer weiteren (zweiten) Clusteranalyse (in den Diagrammen als *SecCl* bezeichnet) die Clusterzugehörigkeit des gesamten Bildmusters bestimmt. Hierfür werden die generierten Merkmalsvektoren (%-Fläche jedes Clusters) der morphometrischen Clusteranalysen genutzt. Die Ergebnisse für jede Clusteranzahl des morphometrischen Clusterings werden durch Annotation der Clusteranzahl mit den besten Werten des zweiten Clusterings in Form eines grünen Kästchens (siehe Abbildung 6.1) dargestellt¹.

Das beste Clusterergebnis wurde bei den synthetischen Daten primär an der externen Validierung durch die VI_n festgelegt. Angemerkt sei an dieser Stelle, dass zur besseren Interpretation die VI_n in den Ergebnisvisualisierungen durch $1 - VI_n$ invertiert wurden, sodass hohe Werte von $1 - VI_n$ ebenso wie hohe Werte des *Silhouetten-Wertes* für ein gutes Clusterergebnis sprechen. Die jeweils signifikantesten Ergebnisse werden durch eine Rückprojektion der Clusterzugehörigkeit in die PCA-Projektion der Merkmalsdaten visualisiert.

Da auf den synthetischen Bildern wesentlich weniger Bildobjekte vorhanden sind und die Bildgröße insgesamt kleiner ist, wurde t^θ (siehe Abschnitt 5.2.2) für jede Bildmodalität in allen synthetischen Testfälle auf 15 Pixel festgelegt.

6.1.1 Morphometrische Analyse - Form-Datenbank

Zum Testen der morphometrischen Clusteranalysen und zur Überprüfung der Funktionsfähigkeit der morphometrischen Merkmale zur Bestimmung komplexerer Formen wurde die Algorithmik anhand einer öffentlich zugänglichen Form-Datenbank (siehe [SKK01]) getestet. Diese Datenbank wurde zum Testen von Image-Retrieval Anwendungen zur Indexierung ähnlicher Formen entwickelt. In der Anwendung von Sebastian et. al. [SKK01] wird hierfür ein Graph-Matching-Algorithmus verwendet. Die Datenbank enthält 216 unterschiedliche binäre synthetisch erzeugte Bilder in jeweils 18 Kategorien bzw. Klassen (siehe Abbildung C.1). Die Klassenzugehörigkeit ist bekannt, sodass die VI_n genutzt werden konnte.

Eingesetzt wurden für diesen Test (MTF2, siehe Tabelle B.4) die Merkmale Compactness, Roundness, Irregularity, 15 Fourier-Deskriptoren (FDs, mit 256 Abtastungspunkten) und die Moment Invariants (insgesamt 25 Merkmale). Als Clusteranzahl für das morphometrische Clustering sowie für das Clustering auf Grundlage der Flächenverteilungen

¹Diese Zusammenfassung wurde notwendig, da $K_1 \cdot K_2$ (K_1 (Clusteranzahl morphometrisches Clustering), K_2 (Clusteranzahl für Clustering auf Grundlage des generierten Merkmalsvektors (%-Fläche jedes Clusters)) Partitionen entstehen und nur die besten Ergebnisse von Relevanz sind.

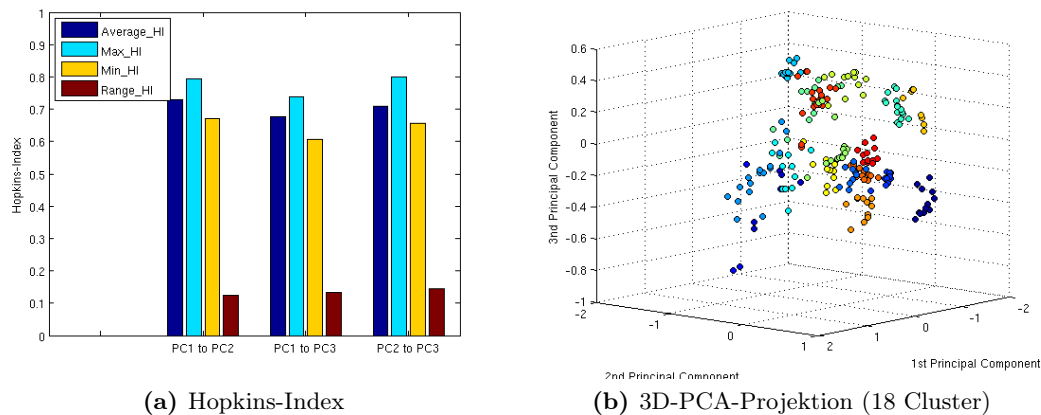


Abbildung 6.1: Ergebnisse Form-Datenbank: Hopkins-Index (a) und 3D-PCA-Projektion (b).

(*SecCl*) wurde K auf 2-22 festgelegt. In diesem Fall wurde das Clustering aufgrund der geringen Anzahl von Mustern komplett über alle Fälle ausgeführt ohne Stichproben bzw. Lernen eines Modells (siehe Abschnitt 5.3).

Der Hopkins-Index (siehe Abbildung 6.1(a)) des Tests liegt für die einzelnen Hauptkomponenten bei ca. 0.7. Dies deutet auf eine Clustertendenz hin. Die ersten drei Hauptkomponenten decken ca. 75% der Gesamtvarianz ab, sodass die Abbildung der PCA's (siehe Abbildung 6.1) eine visuelle Kontrolle erlaubt. In dieser Abbildung wurde bereits das beste Ergebnis (18 Cluster) nach der internen und externen Clusteranalyse eingefärbt. In der Projektion sind Gruppierungen zu erkennen. Die Cluster liegen relativ kompakt im Datenraum der PCA-Projektionen.

Der durchschnittliche Silhouetten-Wert liefert drei weniger signifikante Spitzenwerte bei Clusteranzahl $K = 2$, $K = 18$ und $K = 22$, welche bei ca. 0.5 liegen, was für mittel stark kompakte und separierte Clusterstrukturen spricht. Das beste Ergebnis der VI_n liegt bei Clusteranzahl $K = 18$ im ersten Clustering und 22 im zweiten. Mit diesen Werten wird eine VI_n von ca. 0.1 (bzw. $1 - VI_n \approx 0.9$) erreicht. Dies spricht für eine sehr gute Übereinstimmung mit den korrekten Klassenzugehörigkeiten.

Die Ergebnisse der Clustervalidierung deuten auf ein gutes Clusterergebnis unter Verwendung von 18 Clustern im morphometrischen Clustering hin. Dies entspricht der korrekten Klassenanzahl. Auch die humane visuelle Kontrolle zeigt gute Ergebnisse bei Clusteranzahl $K = 18$ (siehe Abbildung C.2), welche unter Berücksichtigung der Clusterzugehörigkeit eingefärbt wurden.

Getestet wurde außerdem, ob die Ergebnisse sich verbessern, wenn nur die FDs bzw. Momente eingesetzt werden. In beiden Fällen zeigten sich schlechtere Ergebnisse, was für die Kombination der Merkmale spricht.

Betrachtet man die Ergebnisse aus diesem Test, kann von einer korrekten Funktionalität der Implementation ausgegangen werden. Die guten Ergebnisse und Differenzierungen

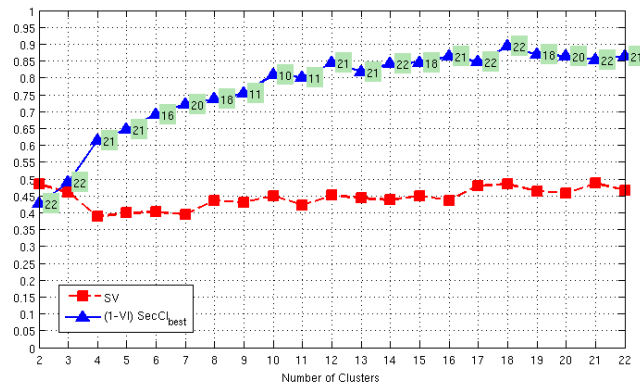


Abbildung 6.2: Durchschnittlicher Silhouetten-Wert SV und das beste Ergebnis der $VI_n(1 - VI)SecCl_{best}$ zweiten Clusterings für jedes Ergebnis des morphometrischen Clustering.

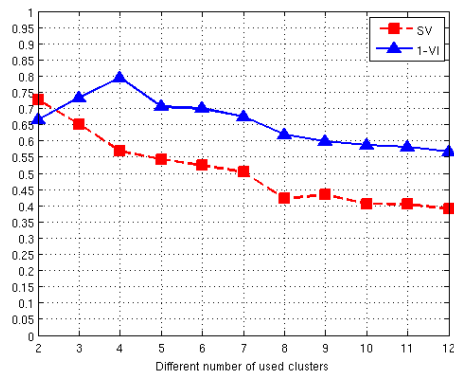
auf der Form-Datenbank mit deutlich unterschiedlich geformten Bildmustern sprechen für eine Anwendung auf die realen Daten.

6.1.2 Graphanalyse - DG

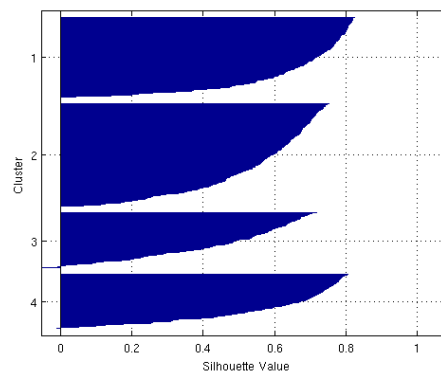
Als Test zur Funktionalität der Graphmerkmale wurde eine synthetische Datenbasis (*syntheticImageDatabase04-1000*) mit 1000 Fällen generiert. Verwendet wurden ausschließlich Kreisobjekte gleicher Größe, die entsprechend der in Abschnitt 5.6.2 beschriebenen Algorithmik, unterschiedlich verteilt wurden (minimal 10 Formobjekte (dissoziiert) und maximal 35 Formobjekte (kompakt)).

Der Fokus dieses Tests liegt darin, ob vier Fälle (kompakt, dissoziiert, multifokal mit gleicher Verteilung und multifokal mit unterschiedlicher Verteilung) aus den Verteilungsanforderungen (siehe Abschnitt 5.1) mit den Merkmalen unterschieden werden können. Die zu testende Clusteranzahl wurde auf $K = 2$ bis $K = 12$ Cluster festgelegt.

Getestet wurden zunächst die Merkmale auf Grundlage des *Delaunay-Graphens* (DG). Als Merkmale wurden alle implementierten Graph-Merkmale (siehe Abschnitt 3.2.3) ohne Anzahl der Knoten verwendet (DGTF1, siehe B.1). Die Anzahl der Knoten wurde bewusst außer Acht gelassen, da ansonsten anhand der Knotenanzahl zwei Klassen getrennt werden könnten.

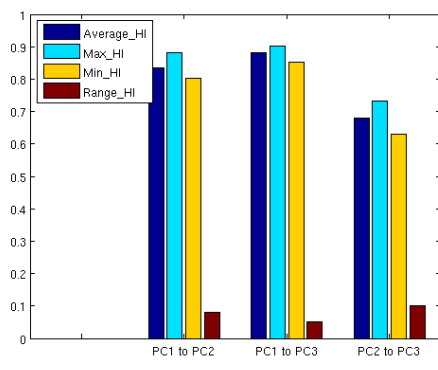


(a) Interne und externe Clustervalidierung

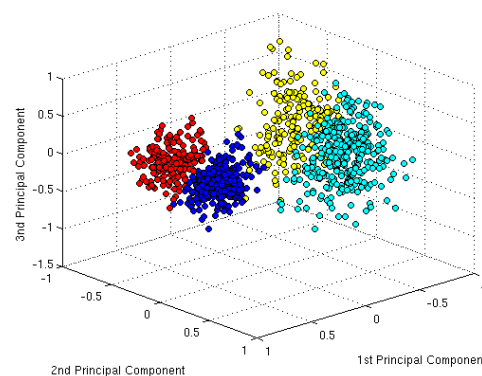


(b) Silhouetten-Plot (4 Cluster)

Abbildung 6.4: Testergebnisse auf synthetischen Daten (DG): Durchschnittlicher Silhouettenwert (SV) und normalisierte Variation of Information (VI_n) (a) sowie der Silhouetten-Plot bei Clusteranzahl 4 (b).



(a) Hopkins-Index



(b) 3D-PCA-Plot (4 Cluster)

Abbildung 6.3: Testergebnisse auf synthetischen Daten (DG): Hopkins-Index (a) und PCA-Projektion (b).

Der Hopkins-Index (siehe Abbildung 6.3) zeigt stabile Werte. Die *Range* ist gering und man kann anhand der Werte von einer Clustertendenz ausgehen. Die 3-D Visualisierung der PCA zeigt zwei deutlich getrennte Datenwolken, die sich jeweils in zwei weitere weniger stark getrennte Datenwolken spalten. Die Visualisierung der drei ersten PCA's enthält ca. 95% der Gesamtvarianz. Die Datenpunkte wurden entsprechend des besten Ergebnisses der externen Clustervalidierung (siehe Abbildung 6.4) mit der korrekten Clusteranzahl $K = 4$ eingefärbt.

Die interne und externe Clustervalidierung zeigen unterschiedliche Ergebnisse. Während der maximale SV -Wert bei Clusteranzahl $K = 2$ mit ca. 0.75 erreicht wird und

mit steigender Clusteranzahl stetig absinkt, liegt bei der externen Validierung ein deutlich signifikanter Wert bei Clusteranzahl $K = 4$ vor. Die VI_n erreicht einen Wert von ca. 0.2 (bzw. $1 - VI_n \approx 0.8$). Sie zeigt eine klare Übereinstimmung mit den Klassenzugehörigkeiten. Der SV -Wert liegt bei $K = 4$ mit ca. 0.6 über 0.5 und deutet auf eine mäßige Strukturierung hin. In 6.4 (b) werden die Silhouetten-Werte jedes Falles für Clusteranzahl $K = 4$ dargestellt. Cluster 1 und 2 zeigen eine starke Strukturierung. Das dritte Cluster ist im Verhältnis zu den anderen weniger gut differenziert. Nur sehr wenige Werte liegen bei Null oder unter Null.

Die Ergebnisse der Analysen zeigen, dass bei vier Clustern ein gutes Ergebnis erreicht wurde. Eine humane Stichprobenanalyse bestätigte dies. In Abbildung 6.5 werden die repräsentativen Fälle anhand des maximalen Silhouetten-Werte des jeweiligen Clusters abgebildet. Die vier gesuchten Klassen konnten durch das Clustering korrekt erfasst werden. Die Ergebnisse sprechen für eine Anwendung der Merkmale auf die realen Daten.

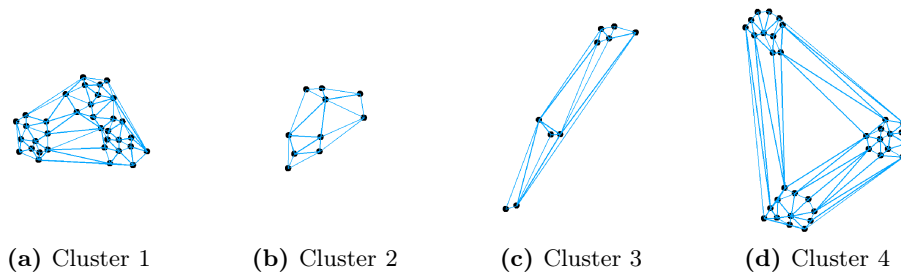


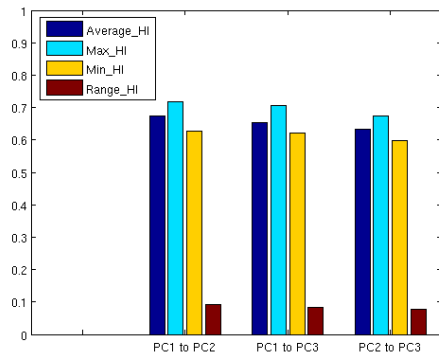
Abbildung 6.5: Cluster 1 = Kompakt, Cluster 2 = Dissoziiert, Cluster 3 = Multifokal mit unterschiedlichem Muster, Cluster 4 = Multifokal mit einheitlichem Muster.

6.1.3 Graphanalyse - MST

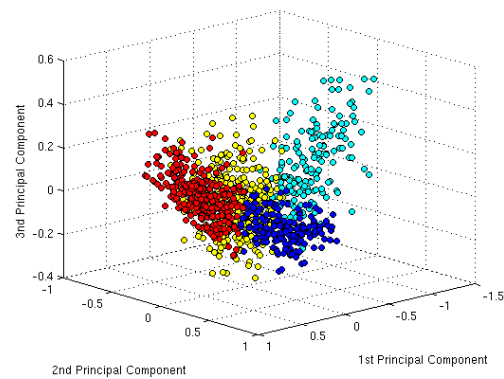
Zum Testen der Graphanalysen des MST wurde die gleiche synthetische Bilddatenbank *syntheticImageDatabase04-1000*, wie im vorigen Test, (siehe 6.1.2) verwendet. Die Clusteranzahl wurde auf 2 bis 10 beschränkt. Gesucht wurde nach den gleichen vier Clustern, wie im vorigen Test. Das beste Ergebnis unter Verwendung des MST's wurde mit den Merkmalen e_μ , e_σ , e_{vc} und e_{mima} erzielt (MSTTF1, siehe B.2). Die Verwendung der Merkmale, welche auf dem Grad der Knoten als Information basieren, zeigten deutlich schlechtere Ergebnisse.

In Abbildung 6.6 (a) wird nur eine schwache Clustertendenz sichtbar. Die Werte sind stabil. Die PCA-Varianzen der drei Hauptkomponenten in 6.6 (b) erreichen ca. 97% der Varianzen, sodass diese Projektionen ein glaubhaftes Bild des Datenraumes wiedergeben. Die Visualisierung der Datenpunkte zeigt, dass hier deutlich weniger kompakte Datenbereiche im Vergleich zu den Ergebnissen des Delaunay-Graphen vorliegen.

Der durchschnittliche Silhouetten-Wert (siehe Abbildung 6.7 (a)) erreicht bei der korrekten Clusteranzahl vier mit ca. 0.65 sein Maximum. Die VI_n erreicht ebenfalls einen deutlich signifikanten Wert bei der korrekten Clusteranzahl vier mit ca. 0.3 (bzw. $1 - VI_n \approx 0.7$). Der Silhouetten-Plot von Clusteranzahl vier deutet auf mäßig separierte

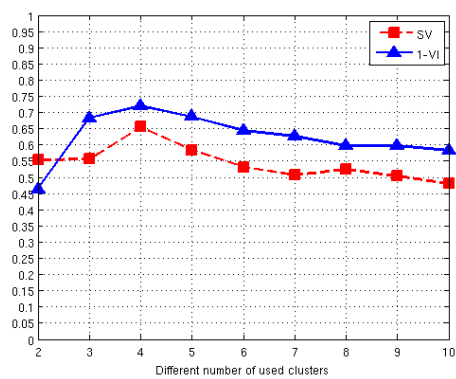


(a) Hopkins-Index

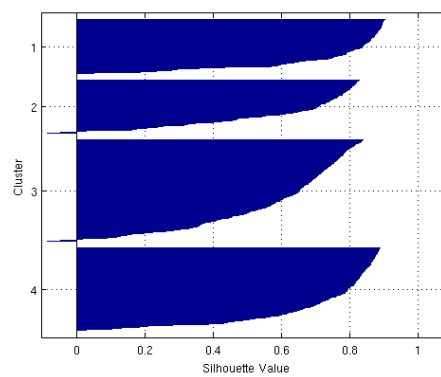


(b) 3D-PCA-Projektion (4 Cluster)

Abbildung 6.6: Testergebnisse auf synthetischen Daten (MST): Hopkins-Index (a) und PCA-Projektion (b).



(a) Interne und externe Clustervalidierung



(b) Silhouette-Plot (4 Cluster)

Abbildung 6.7: Testergebnisse auf synthetischen Daten (MST): Durchschnittlicher Silhouettenwert (SV) und normalisierte Variation of Information (VI_n) (a) sowie der Silhouetten-Plot bei Clusteranzahl 4 (b).

Cluster hin (siehe Abbildung 6.7 (b)). Eine humane Stichprobenanalyse ergab, dass die meisten Fälle korrekt zugeordnet wurden. Abbildung 6.8 zeigt die repräsentativen Fälle mit den höchsten Silhouetten-Werten der vier Cluster. Dieser Test zeigt, dass die korrekte Clusteranzahl und somit die unterschiedlichen Verteilungen mit Hilfe der MST-Merkmale bestimmt werden konnten. Lediglich der Hopkins-Index spricht gegen ein plausibles Ergebnis.

Zusammenfassend, im Vergleich zu den Werten des Tests mit DG auf der gleichen Datenbasis, zeigen die MST-Merkmale eine schlechtere Performance im Sinne der korrekten Klassenzugehörigkeiten.

Zudem konnte beobachtet werden, dass alle Merkmale, die auf dem Knotengrad basieren, für die synthetischen Fälle nahezu identisch sind und keine Diskriminierung der Daten ermöglichen. Dies bestätigt die ersten Eindrücke der Tests mit den realen Tumorprototypen (siehe Abschnitt 5.4).

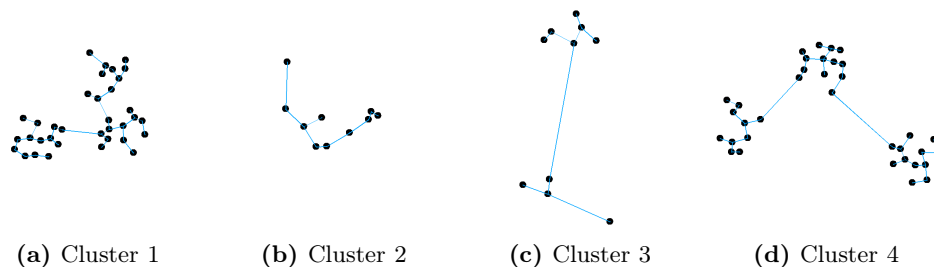


Abbildung 6.8: Cluster 1 = Kompakt, Cluster 2 = Dissoziiert, Cluster 3 = Multifokal mit unterschiedlichen Muster, Cluster 4 = Multifokal mit einheitlichem Muster.

6.1.4 Relationale Analyse - RCC8

Dieser Test (RTF1, siehe Tabelle B.3) soll die grundlegende Funktionalität der implementierten und im Rahmen dieser Arbeit erweiterten RCC8-Merkmale bzw. des RCC8-Histogramms (siehe Abschnitt 5.5.1) prüfen. Hierfür wurde eine Datenbasis von 1000 unterschiedlichen Verteilungen (*syntheticImageDatabase01-1000*) nach den Ausführungen in Abschnitt 5.6.2 erstellt. Es wurde nach acht Clustern gesucht, welche jeweils eine der acht RCC8-Relationen (siehe Abschnitt 3.3) repräsentieren. Angewendet wurde hierfür das nicht-kombinierte Relationsgewicht (siehe Abschnitt 5.5.1).

Eine Clustertendenz ist nach dem Hopkins-Index (siehe Abbildung 6.9 (a)) deutlich und stabil zu erkennen. Die PCA-Visualisierung (siehe Abbildung 6.9 (b)) stellt nur ca. 60% der Varianz in den Daten dar und ist daher weniger für eine gute Abschätzung geeignet. Es lassen sich keine Clusterstrukturen erkennen.

Erst in der internen und externen Clustervalidierung werden Clusterstrukturen (siehe 6.10 (a)) deutlich. Der durchschnittliche Silhouetten-Wert SV erreicht das Maximum (ca. 0.9) bei der korrekten Clusteranzahl 8. Ein Wert ca. 0.9 deutet auf eine starke Strukturierung der Muster hin. Die VI_n erreicht ebenfalls einen signifikanten Wert bei Clusteranzahl 8 mit ca. 0.15 (bzw. $1 - VI_n \approx 0.85$). Die Untersuchung der einzelnen

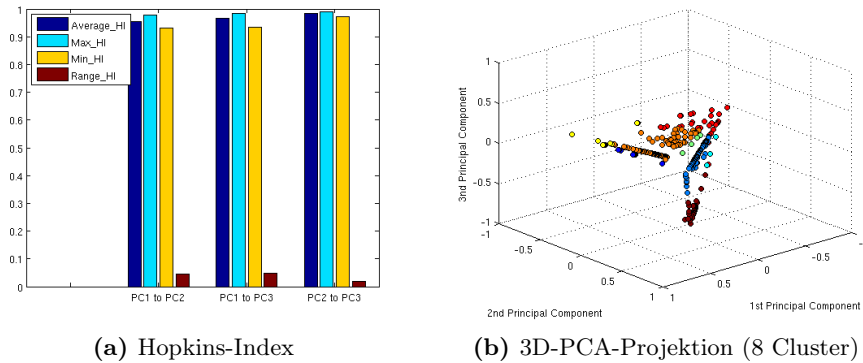


Abbildung 6.9: Testergebnisse auf synthetischen Daten (RCC8): Hopkins-Index (a) und PCA-Projektion (b).

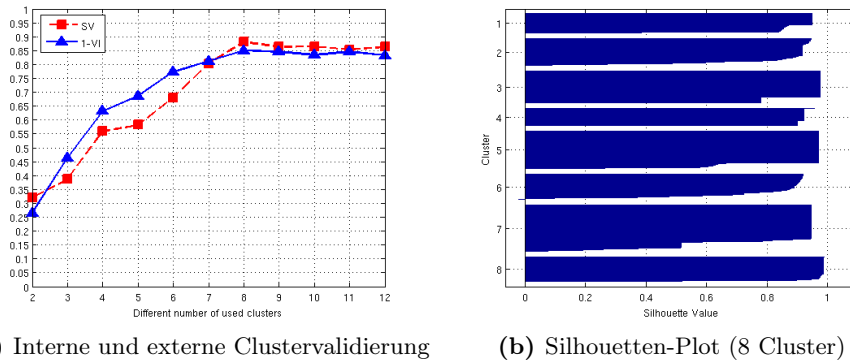


Abbildung 6.10: Testergebnisse auf synthetischen Daten (RCC8): Durchschnittlicher Silhouetten-Wert (SV) und normalisierte Variation of Information (VI_n) (a) sowie der Silhouetten-Plot bei Clusteranzahl 8 (b).

Silhouetten-Werte in Abbildung 6.10 (b) zeigt die gute Kompaktheit und Separierung des Clusterergebnisses. In Abbildung 6.11 werden die Fälle mit den höchsten Silhouetten-Werten der einzelnen Cluster dargestellt.

Sowohl der Hopkins-Index, als auch die interne und externe Clustervalidierung lassen auf ein sehr gutes Clusterergebnis schließen. Die Übereinstimmung mit den korrekten Klassenzugehörigkeiten ist hoch. Unter Verwendung dieser Testdaten erzielt die Erweiterung des RCC8-Histogramms korrekte Ergebnisse.

6.1.5 Relationale Analyse - LDQ

Um die Funktionalität der entwickelten LDQ-Merkmale zu prüfen (RTF2, siehe Tabelle B.3), wurden diese auf den gleichen synthetischen Datensatz wie die RCC8-Merkmale angewendet (*syntheticImageDatabase01-1000*).

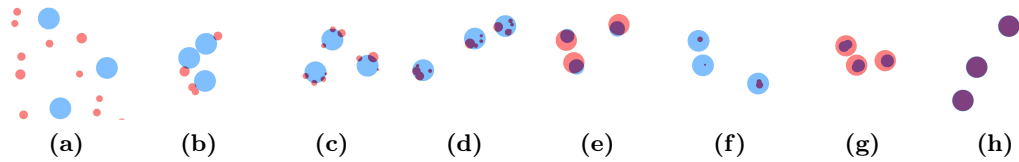
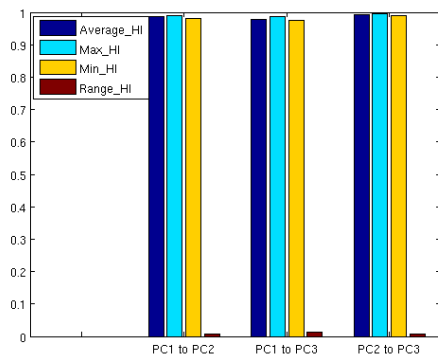
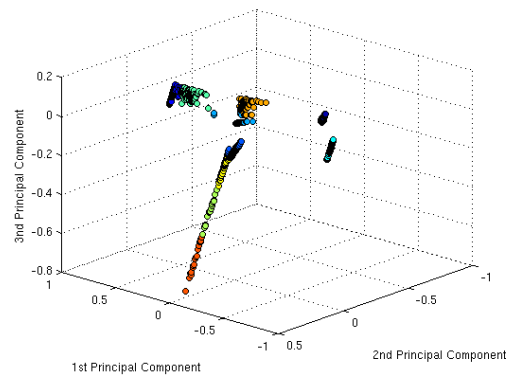


Abbildung 6.11: (a) DC \rightarrow Cluster 8, (b) EC \rightarrow Cluster 2, (c) PO \rightarrow Cluster 7, (d) TPP \rightarrow Cluster 6, (e) TPPI \rightarrow Cluster 4, (f) NTPP \rightarrow Cluster 5, (g) NTPPI \rightarrow Cluster 3, (h) EQ \rightarrow Cluster 1

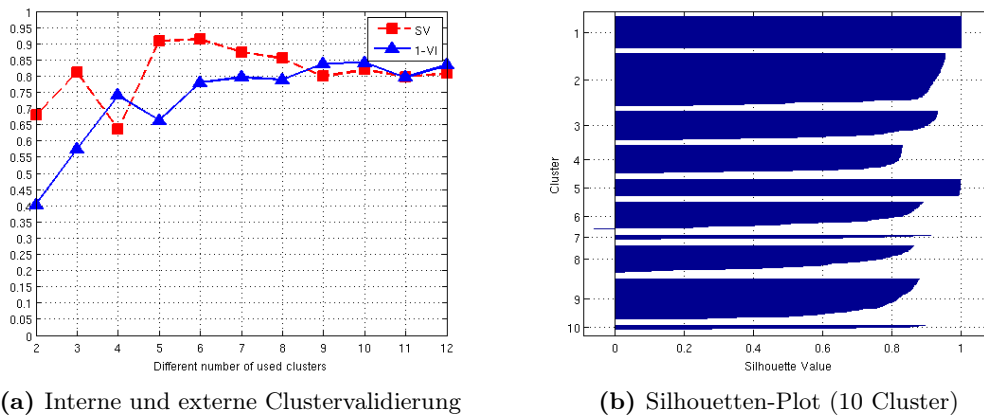


(a) Hopkins-Index



(b) 3D-PCA-Projektion (10 Cluster)

Abbildung 6.12: Testergebnisse auf synthetischen Daten (LDQ): Hopkins-Index (a) und PCA-Projektion (b).



(a) Interne und externe Clustervalidierung

(b) Silhouetten-Plot (10 Cluster)

Abbildung 6.13: Testergebnisse auf synthetischen Daten (RCC8): Durchschnittlicher Silhouetten-Wert (SV) und normalisierte Variation of Information (VI_n) (a) sowie der Silhouetten-Plot bei Clusteranzahl 10 (b).

Dem Hopkins-Index zur Folge lässt sich in den Merkmalsdaten eine stabile und deutliche Clustertendenz erkennen (siehe Abbildung 6.12 (a)). Da die LDQ-Merkmale eine Gruppe von drei Merkmalen sind, zeigt die PCA-Visualisierung (siehe Abbildung 6.12 (b)) 100% der Varianzen. Gruppen lassen sich hier zwar identifizieren, stellen sich allerdings teilweise ähnlich linear verteilt dar wie die Ergebnisse der RCC8-Merkmale. Dies kann dafür sprechen, dass die zufällige Verteilung der synthetischen Daten eine zu geringe Varianz in sich trägt.

Die interne und externe Clustervalidierung zeigen für diesen Test sehr unterschiedliche Ergebnisse (siehe Abbildung 6.13). Während der durchschnittliche Silhouetten-Wert sein Maximum (ca. 0.9) bei fünf Clustern hat, hat die VI_n mit ca. 0.15 (bzw. $1 - VI_n \approx 0.85$) ihr Minimum bei zehn Clustern.

Im Silhouetten-Plot ist zu sehen, dass bei Clusteranzahl $K = 10$ zwei sehr dünn besetzte Cluster (7 und 10) vorliegen. Auf Grund der niedrigen VI_n kann davon ausgegangen werden, dass diese beiden Cluster irrelevant sind. Somit kann auf die korrekte Anzahl von 8 natürlichen Clustern entsprechend den Klassenzugehörigkeiten geschlossen werden. Der durchschnittliche Silhouetten-Wert zeigt bei 8 Clustern mit ca. 0.85 eine starke Strukturierung in den Merkmalswerten.

Dieser Test zeigt, dass die LDQ-Merkmale in der Lage sind die unterschiedlichen Relationen kompakt und separiert zu differenzieren. Der niedrige Wert der VI_n entspricht einer hohen Korrelation mit den Klassenzugehörigkeiten. Im Vergleich zu den RCC8-Merkmalen sind die Ergebnisse jedoch nicht so eindeutig. Von einem sinnvollen Einsatz für die realen Daten wird dennoch ausgegangen.

6.2 Reale Daten - Clusteranalysen

Grundsätzlich entspricht der Ablauf und die Präsentation der Clusteranalyse der realen Tumordaten dem Vorgehen bei den Experimenten zu den synthetischen Daten (siehe Abschnitt 6.1). Jede Merkmalsgruppe wurde in den Clusteranalysen unabhängig voneinander getestet. In jedem Experiment wird auf die verwendeten Merkmalskombination im Anhang (siehe Anhang B) verwiesen.

Als Klassenzugehörigkeit für die VI_n wurde für die realen Daten der *Follow-Up-Status* (siehe Abschnitt 5.6) eingesetzt, um zu prüfen, inwieweit sich relevante Gruppen für den Verlauf bestimmen lassen. Das beste Clusterergebnis wird bei den Tests an den realen Daten hingegen primär durch den durchschnittlichen Silhouetten-Wert SV bestimmt. Nur wenn ein im Verhältnis zu den anderen Clusteranzahlen signifikanter Wert der VI_n auftritt, wird diese Partition näher untersucht.

Da das Projekt ExPrimage im Laufe der Erstellung der vorliegenden Arbeit beendet wurde und der zeitliche Rahmen für eine humane Evaluation unterschätzt wurde, konnte eine humane Bewertung durch den Domänenexperten, wie sie in Abschnitt 5.6 vorgestellt wurde, nicht mehr bzw. nur teilweise erfolgen.

Aus diesem Grund wurde ein Vergleich mit den dokumentierten Beobachtungen (siehe Abschnitt 4.1), sofern entsprechende vorlagen, des Domänenexperten verwendet. Dies konnte im Rahmen der vorliegenden Arbeit nur in Form von Stichproben erfolgen. Weitere Analysen waren aufgrund des begrenzten Zeitrahmens nicht möglich. Lagen für jeden der 93 betrachteten Fälle Beobachtungen vor, wurde anhand des Silhouetten-Plots das kompakteste Cluster ausgesucht. Lagen mehr als 10 Fälle in diesem Cluster vor, wurden nur die zehn Fälle mit den höchsten Silhouetten-Werten als Stichprobe herangezogen.

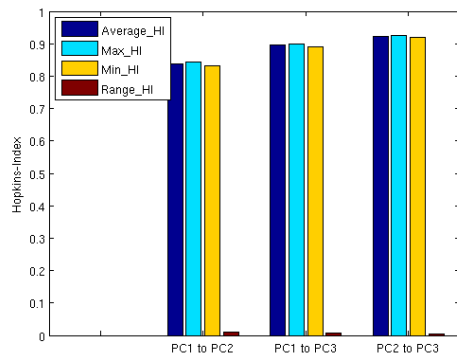
Entsprechend der Silhouetten-Analyse sollten die Muster mit den höchsten Silhouetten-Werten das Cluster am besten repräsentieren (siehe Abschnitt 5.6). Besaß nicht jedes Bildmuster eine eindeutige Annotation des Domänenexperten, wurden ähnliche eindeutig vergleichbare Annotationen ausgewählt und diese mit den Clusterzuordnungen verglichen. Die Vergleiche werden in tabellarischer Form präsentiert.

Die Schwellwerte t^{θ} und Parameter wurden entsprechend den Ausführungen zu den einzelnen Analysen in Kapitel 5 gewählt.

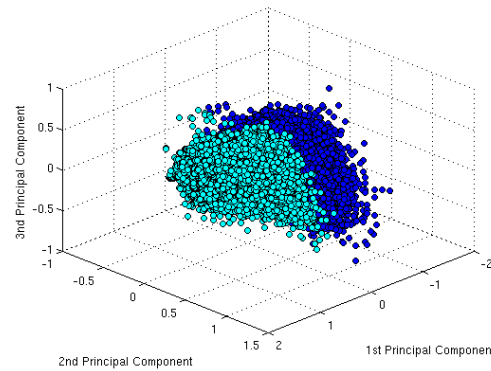
6.2.1 Morphometrische Analysen - AE1AE3

In diesem Testfall wurden die morphometrischen Eigenschaften der Tumorobjekte unter Verwendung der Bildmodalität f_{bin}^{ae1ae3} untersucht. Eingesetzt wurden alle morphometrischen Merkmale (MOTF3, siehe Tabelle B.4). Für die Fourier-Deskriptoren wurden 256 Abtastungspunkte verwendet. Die Clusteranzahl wurde auf $K = 2 - 16$ festgelegt. Zum Lernen des Modells (siehe Abschnitt 5.3) wurden 31 der 93 Bildmuster f_{bin}^{ae1ae3} als Stichprobe ausgewählt, welche insgesamt 62 111 Tumorobjekte enthalten.

Der Hopkins-Index, die PCA-Projektion und der durchschnittliche Silhouetten-Wert SV (rot) beziehen sich auf die Clusterergebnisse der 62 111 Tumorobjekte (morphometrisches Clustering). Die Clusteranzahl des morphometrischen Clusterings zur Bestimmung unterschiedlicher Tumorobjekte wurde auf $K = 2 - 16$ beschränkt. Die Clusteranzahl zur



(a) Hopkins-Index



(b) 3D-PCA-Projektion (2 Cluster)

Abbildung 6.14: Testergebnisse auf realen Daten (Morphometrisch AE1AE3): Hopkins-Index (a) und PCA-Projektion (b).

Bestimmung der Verteilungsmuster auf Grundlage der %-Cluster-Flächenverteilungen (zweites Clustering, *SecCl*) wurde auf $K = 2 - 10$ eingestellt.

In Abbildung 6.15 werden lediglich die besten Werte des zweiten Clusterings unter Verwendung der unterschiedlichen Clusteranzahlen des morphometrischen Clusterings dargestellt. Die Clusteranzahl mit dem höchsten *SV* ($SV_{SecCl_{best}}$) bzw. der niedrigsten VI_n ($(1 - VI)_{SecCl_{best}}$) des zweiten Clusterings wird in grünen Kästchen annotiert.

Der Hopkins-Index zeigt durchgängig stabile Werte und eine eindeutige Clustertendenz (siehe Abbildung 6.14 (a)). Die Werte liegen zwischen 0.85 und 0.95. Die 3D-PCA-Projektion deckt ca. 70% der Gesamtvarianz ab. Sie erlaubt einen Einblick in den Merkmalsraum. In ihr lassen sich keine Clusterstrukturen erkennen. Die Datenpunkte liegen in einer undifferenzierten Datenwolke vor.

Die VI_n (*SecCl* VI) liegt für alle Clusteranzahlen unter 0.1 (bzw. $1 - VI_n \approx 0.85$). Ein Zusammenhang mit dem *Follow-Up-Status* kann nicht hergestellt werden. Der durchschnittliche Silhouetten-Wert (*SV*) hat sein Maximum bei 2 Clustern mit ca. 0.45, was für eine schwache Strukturierung spricht. Lediglich der durchschnittliche Silhouetten-Wert ($SV_{SecCl_{best}}$) des zweiten Clusterings auf Grundlage der %-Flächenverteilung des morphometrischen Clusterings zeigt mit ca. 0.87 bei Clusteranzahl $K = 2$ eine deutliche Tendenz für gut strukturierte Clusteraufteilungen.

Dem Domänenexperten wurden jeweils zehn Stichproben mit Clusteranzahl $K = 2 - 5$ der visualisierten Rückprojektion des morphometrischen Clusterings (siehe Abschnitt 5.3) vorgelegt. Jeder Fall wurde durch den Domänenexperten als „plausibel“ (1), „teilweise plausibel“ (0) und „nicht plausibel“ (-1) hinsichtlich der Clusterrückprojektionen der einzelnen Tumorobjekte bewertet (siehe Tabelle 6.1).

Entsprechend den Analysen des Domänenexperten ist Clusteranzahl $K = 3$ am plausibelsten, gefolgt von Clusteranzahl $K = 2$. Die höheren Clusteranzahlen $K = 4$ und $K = 5$ sind überwiegend nicht plausibel.

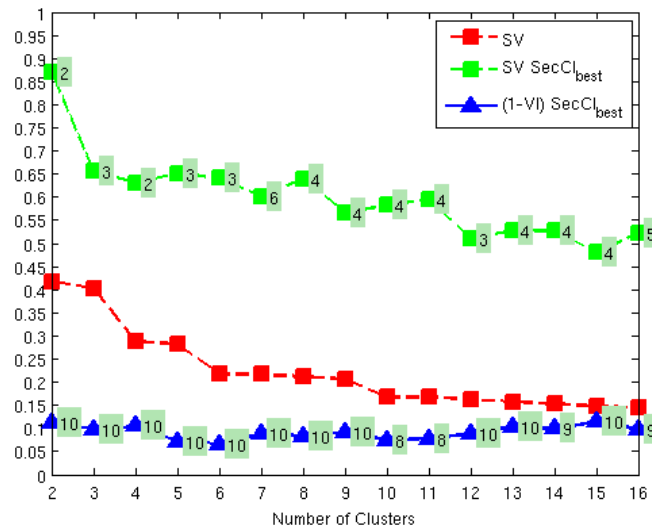


Abbildung 6.15: Testergebnisse der internen und externen Clustervalidierung auf den realen Daten (Morphometrisch): Durchschnittlicher Silhouetten-Wert (SV) des morphometrischen Clustering der Tumorobjekte, beste Ergebnisse des zweiten Clustering ($SV SecCl_{best}$) und normalisierte Variation of Information ($1 - VI_n SecCl_{best}$).

Patient-ID	2 Cluster	3 Cluster	4 Cluster	5 Cluster
002-11916	0	1	-1	-1
004-08068	0	1	-1	0
004-14248	1	0	-1	-1
004-20862	0	0	1	-1
011-05730	0	1	-1	-1
012-18864	1	-1	-1	1
031-09500	-1	1	0	-1
034-21893	0	1	-1	-1
971-14011	1	1	-1	1
971-18370	0	-1	0	-1
Σ	2	4	-7	-5

Tabelle 6.1: Beurteilungen der Stichprobe zu den Clusteranalysen durch den Domänenexperten(-1 = nicht plausibel, 0 = teilweise plausibel, 1 = plausibel).

Insgesamt wurde vom Domänenexperten angemerkt, dass diese Analysen nicht die Faktoren erfassen können, welche unter den Anforderungen (siehe Abschnitt 5.1) für die morphometrischen Analysen angesetzt wurden. Der Faktor Größe wird nicht genügend berücksichtigt. Eine Differenzierung der gesuchten Strukturen kann nur in Einzelfällen dargestellt werden. Eine allgemeine, plausible und fallübergreifende Differenzierung wird mit diesem Ansatz nicht erreicht.

6.2.2 Strukturelle Analysen - AE1AE3 - DG

Dieser Test basiert auf der Berechnung des DG's unter Verwendung der Bildmodalität f_{bin}^{ae1ae3} (siehe Abschnitt 5.4). Getestet wurden zunächst alle Merkmale des DG's, nur Merkmale auf Grundlage der Kantengewichte und nur N_v (Anzahl der Knoten), e_μ (durchschnittliche Kantengewichte) und e_σ (Standardabweichung der Kantengewichte).

Für die Graphanalysen wurde der höchste Silhouetten-Wert mit der Merkmalskonfiguration DGTF2 (siehe Tabelle B.1) erzielt. Der Hopkins-Index zeigt instabile geringe Werte (siehe Abbildung 6.18 (a)) von ca. 0.6. Da für diesen Testfall nur drei Merkmalswerte verwendet wurden, deckt die 3D-PCA-Projektion 100% der Varianzen ab. Auch der PCA-Plot lässt keine eindeutigen Cluster erkennen (siehe Abbildung 6.18 (b)).

Die VI_n liegt bei jeder getesteten Clusteranzahl über 0.9 bzw. $(1 - VI_n)$ unter 0.1 (siehe Abbildung 6.19 (a)). Es liegt kein signifikanter Wert vor. Der maximale durchschnittliche Silhouetten-Wert wird bei Clusteranzahl $K = 3$ mit ca. 0.65 erreicht (siehe Abbildung 6.19 (a)).

Im Silhouetten-Plot (siehe Abbildung 6.19 (b)) der Partition mit Clusteranzahl $K = 3$ ist zu sehen, dass die ersten beiden Cluster hohe Silhouetten-Werte erreichen und dementsprechend relativ kompakt und gut separiert sind.

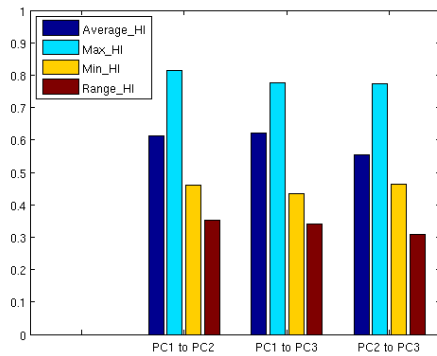
Das dritte Cluster ist weniger gut separiert. In der PCA-Projektion wird das dritte Cluster durch die gelb markierten Datenpunkte repräsentiert. Diese Datenpunkte sind wesentlich weitläufiger im Datenraum verteilt, als die Datenpunkte der beiden anderen Cluster (blau und cyan).

Aufgrund des höchsten durchschnittlichen Silhouetten-Wertes wurde die Partition mit Clusteranzahl $K = 3$ für die Stichprobenanalyse mit den Beobachtungen des Domänenexperten herangezogen (siehe Tabelle 6.2). Da für diesen Test keine eindeutigen Beobachtungen für jeden Patientenfall vorlagen, wurden zufällig eindeutige Beobachtungen herausgesucht und mit den Clusterzugehörigkeiten des Falles aus der Clusteranalyse verglichen.

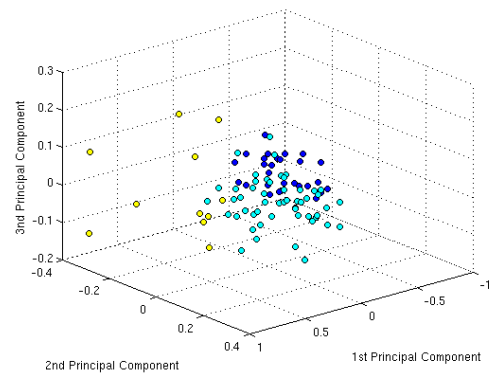
Sieben von neun eindeutigen Beobachtungen liegen ebenfalls im gleichen Cluster. Für diese Stichproben kann davon ausgegangen werden, dass die Anwendung der Merkmale ähnliche Zusammenhänge erkennt, wie der Domänenexperte in diesen Fällen.

6.2.3 Strukturelle Analysen - AE1AE3 - MST

Die Analysen dieses Testfalls basieren ebenfalls auf der Bildmodalität f_{bin}^{ae1ae3} . Getestet wurden zunächst alle Merkmale des MST's (siehe 3.2.3), nur Merkmale auf Grundlage

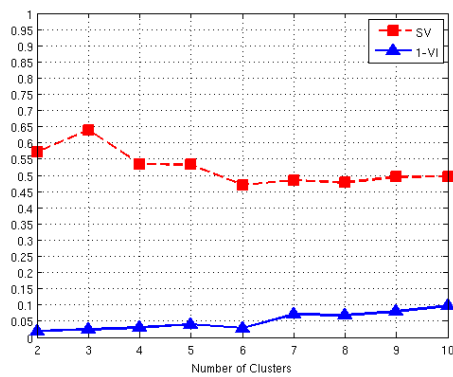


(a) Hopkins-Index

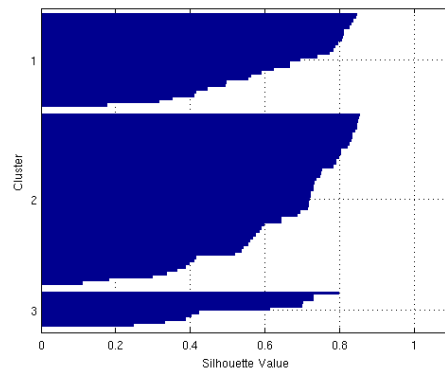


(b) 3D-PCA-Projektion (3 Cluster)

Abbildung 6.16: Testergebnisse auf realen Daten (DG - AE1AE3): Hopkins-Index (a) und PCA-Projektion (b).



(a) Interne und externe Clustervalidierung



(b) Silhouetten-Plot (3 Cluster)

Abbildung 6.17: Testergebnisse auf realen Daten (DG - AE1AE3): Durchschnittlicher Silhouetten-Wert (SV) und normalisierte Variation of Information (VI_n) (a) sowie der Silhouetten-Plot bei Clusteranzahl 3 (b).

Patienten-ID	Annotation	Cluster(1-3)
044-15571	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	2
032-00918	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	2
013-13832	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	2
014-17283	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	2
054-00865	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	2
013-13832	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	2
022-22356	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	3
031-02224	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	3
014-14612	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	2

Tabelle 6.2: Stichprobenanalyse: Vergleich von Fällen mit gleicher eindeutiger Beobachtung des Domänenexperten mit den berechneten Clusterzugehörigkeiten (DG AE1AE3).

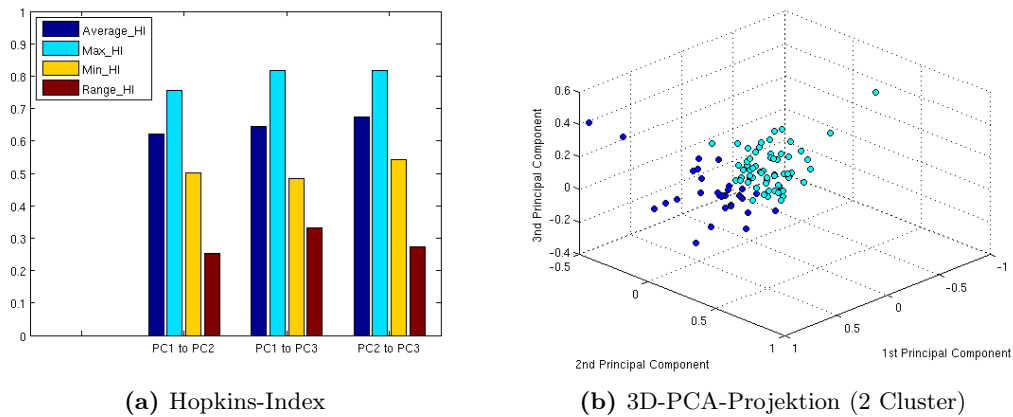


Abbildung 6.18: Testergebnisse auf realen Daten (AE1AE3 - MST): Hopkins-Index (a) und PCA- Projektion (b).

der Kantengewichte und nur N_v (Anzahl der Knoten), e_μ (durchschnittliche Kantengewichte) und e_σ (Standardabweichung der Kantengewichte). Der höchste durchschnittliche Silhouetten-Wert wurde mit den Merkmalen auf Grundlage der Kantengewichte (MSTTF2, siehe B.2) erreicht, welcher an dieser Stelle genauer vorgestellt wird.

In diesem Testfall wird lediglich eine geringe Clustertendenz mit instabilen Werten vorausgesagt (siehe Abbildung 6.18 (a)). Auch in der PCA-Projektion (siehe Abbildung 6.18 (b)) lässt sich keine deutliche Clusterstruktur erkennen, wobei die Visualisierung 100% der Varianzen abdeckt. Der maximale durchschnittliche Silhouetten-Wert wird bei 2 Clustern mit ca. 0.65 erreicht, was auf eine gut separierte Clusterstruktur hinweist (siehe Abbildung 6.19). Keine der unterschiedlichen Clusteranzahlen zeigt einen Zusammenhang mit dem *Follow-Up-Status*. Alle Werte der VI_n liegen oberhalb von 0.9 bzw. $1 - VI_n$ unterhalb von 0.1.

Der Silhouetten-Plot lässt erkennen, dass Cluster 1 wesentlich weniger kompakt und gut separiert ist als Cluster 2. In Cluster 1 gibt es Fälle mit negativen Silhouetten-Werten. Als Vergleich zwischen den Merkmalen von MST und DG wurden die gleichen Beobachtungen des Domänenexperten herangezogen (siehe Tabelle 6.3) wie bei den Ergebnissen des MST's und die Clusterzugehörigkeiten der Patientenfälle eingetragen.

Von neun Fällen wurden sechs Cluster 1 zugeordnet und drei Cluster 2. Es besteht also ein Drittel Abweichung, während im Testfall von DG lediglich zwei Fälle abweichen. Da im Testfall des MST's nur zwei Cluster für die Stichproben verwendet werden, ist die Wahrscheinlichkeit für zufällige Übereinstimmungen im Vergleich zum DG größer. Obwohl beide Partitionen (DG und MST) einen ähnlichen durchschnittlichen Silhouetten-Wert haben, wird aufgrund der Stichproben von einem besseren Ergebnis der Merkmale des DG's ausgegangen.

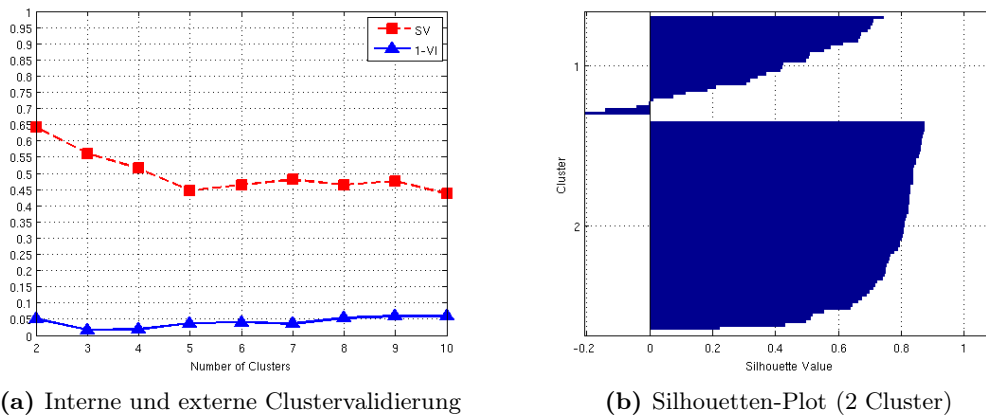


Abbildung 6.19: Testergebnisse auf realen Daten (DG - AE1AE3): Durchschnittlicher Silhouetten-Wert (SV) und normalisierte Variation of Information (VI_n) (a) sowie der Silhouetten-Plot bei Clusteranzahl 2 (b).

Patienten-ID	Annotation	Cluster(1-2)
044-15571	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	1
032-00918	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	2
013-13832	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	1
014-17283	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	2
054-00865	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	2
013-13832	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	1
022-22356	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	1
031-02224	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	1
014-14612	dissoziierter Tumor, Kontinuität der einzelnen Anteile zueinander erhalten	1

Tabelle 6.3: Stichprobenanalyse: Vergleich von zehn Fällen mit gleicher eindeutiger Beobachtung des Domänenexperten mit den berechneten Clusterzugehörigkeiten (MST AE1AE3).

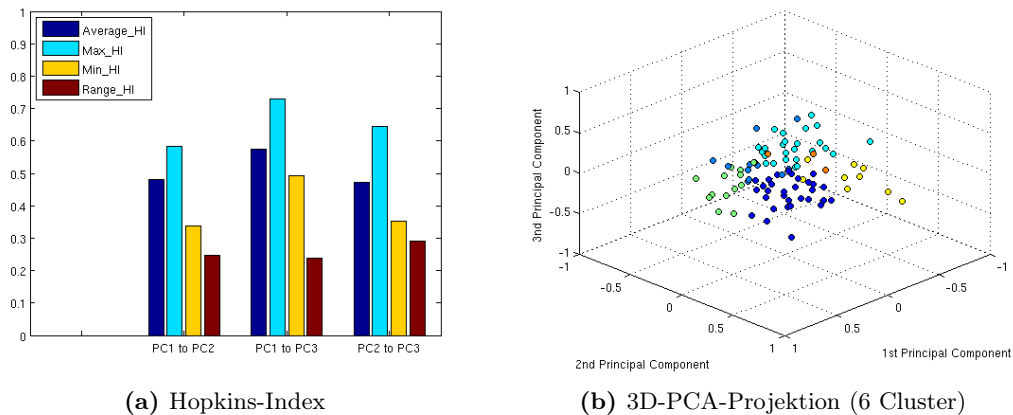


Abbildung 6.20: Testergebnisse auf realen Daten (Relationen - CD45): Hopkins-Index (a) und PCA-Projektion (b).

6.2.4 Relationale Analysen - CD45

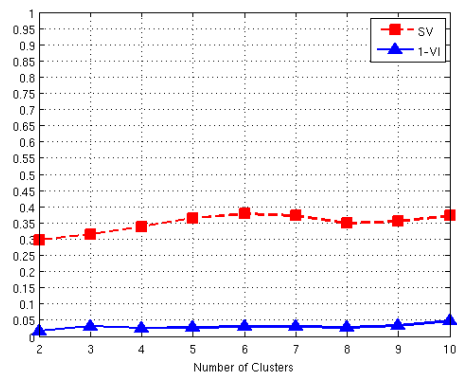
Bei diesem Test soll die Ausbreitung des entzündeten Gewebes f_{bin}^{cd45} in Relation zum Tumorgewebe f_{bin}^{tumor} untersucht werden. Dazu wurden die RCC8-Merkmale bzw. das RCC8-Histogramm und die LDQ-Merkmale kombiniert eingesetzt (RTF3, siehe Tabelle B.3). Die zu testende Clusteranzahl wurde auf $K = 2 - 10$ festgelegt.

Der Hopkins-Index zeigt keine Clustertendenz (siehe Abbildung 6.20 (a)) bei Werten zwischen 0.4 und 0.6. Obwohl der Hopkins-Index auch in diesem Fall eher instabile Werte liefert, liegen selbst die Maximalwerte Max_HI unter den Werten der anderen durchgeführten Clusteranalysen auf den realen Daten. Die PCA-Projektion (siehe Abbildung 6.20 (a)) lässt ebenfalls keine Clusterstrukturen erkennen.

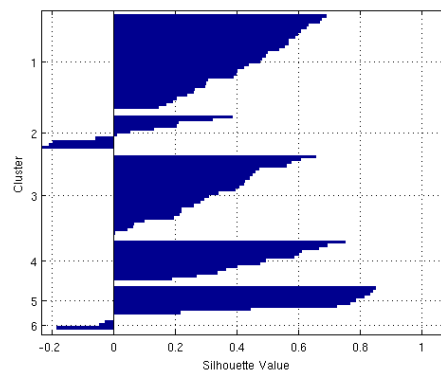
Die interne Clustervalidierung mittels des durchschnittlichen Silhouetten-Wertes erreicht ihr Maximum bei Clusteranzahl 6 mit ca. 0.38 (siehe Abbildung 6.21), was für schwach kompakte und separierte Cluster spricht. Dies bestätigt ebenfalls die Rückprojektion der Clusterzugehörigkeit in der PCA-Projektion der Merkmalsdaten. Es werden keine räumlich trennbaren Gruppen sichtbar. Die VI_n -Werte bleiben bei diesem Test über 0.95 bzw. unter $1 - VI_n = 0.05$. Auf Grundlage des unüberwachten Lernens wird kein Zusammenhang mit dem *Follow-Up-Status* deutlich.

Der Silhouetten-Plot der Partition mit Clusteranzahl $K = 6$ bestätigt das zu erwartende Ergebnis des Clusterings. Die Cluster haben im Vergleich zu den vorigen Clusteranalysen sehr geringe Silhouetten-Werte. Cluster 6 ist falsch zugeordnet worden und enthält keine Strukturierung (nur negative Werte).

Für den Stichprobenvergleich der Beobachtungen des Domänenexperten wird das Cluster mit den höchsten Silhouetten-Werten, Cluster 5, ausgewählt. Für diesen Test liegen eindeutig vergleichbare Beobachtungen für jeden Fall vor. In Tabelle 6.4 werden die Beobachtungen des Domänenexperten bzgl. der Verteilung von entzündetem Gewebe zu den Fällen mit den Silhouetten-Werten aus Cluster 5 zusammengefasst.



(a) Interne und externe Clustervalidierung



(b) Silhouetten-Plot (6 Cluster)

Abbildung 6.21: Testergebnisse auf realen Daten (Relationen - CD45): Durchschnittlicher Silhouetten-Wert (SV) und normalisierte Variation of Information (VI_n) (a) sowie der Silhouetten-Plot bei Clusteranzahl 6 (b).

Patienten-ID	Annotation	Silhouette-Wert	Cluster (1-6)
011-08250	fokal Zentrum	0.2139	5
013-15178	fokal Zentrum	0.4420	5
004-08068	fokal Peripherie, fokal Zentrum	0.7225	5
011-13642	Peripherie (von Tumor, und Nekrose)	0.7674	5
032-08016	fokal Peripherie	0.7839	5
054-00097	diffus (Tu-Septen)	0.8118	5
021-23131	fokal Peripherie, fokal Zentrum	0.8295	5
022-09100	fokal Zentrum	0.8401	5
971-14011	fokal Peripherie	0.8504	5

Tabelle 6.4: Stichprobenanalyse: Vergleich der kompaktesten Fälle mit Beobachtungen des Domänenexperten (CD45).

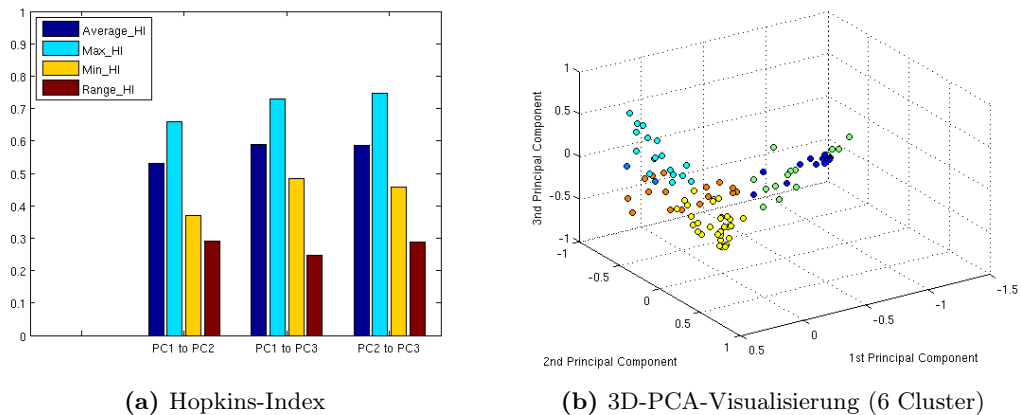


Abbildung 6.22: Testergebnisse auf realen Daten (Relationen - ER): Hopkins-Index (a) und PCA-Projektion (b).

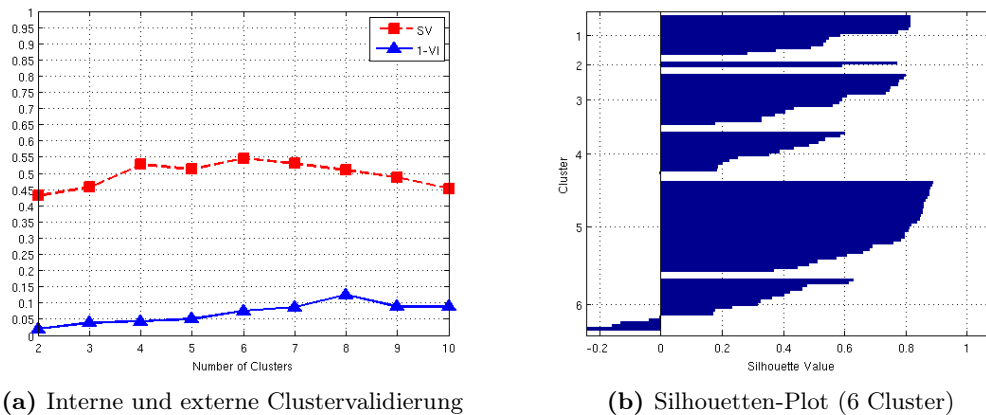
Die Stichprobenanalyse zeigt, dass in diesem Fall keine Übereinstimmung zwischen der Clusterzugehörigkeit und den Beobachtungen des Domänenexperten besteht. Selbst die Fälle mit den drei höchsten Silhouetten-Werten (971-14011, 022-09100 und 021-23131), welche nach den eingesetzten Merkmalen ähnlich sein sollten, unterscheiden sich in den Beobachtungen des Domänenexperten grundsätzlich. Sowohl die schlechten Ergebnisse der Clustervalidierung als auch der Stichprobenvergleich mit den humanen Beobachtungen des Domänenexperten zeigen, dass durch die eingesetzten Merkmale in der unüberwachten Validierung keine nachvollziehbare Unterscheidung der Verteilung und Ausbreitung von entzündetem Gewebe möglich ist.

6.2.5 Relationale Analysen - ER

Anhand dieses Tests soll die Ausbreitung des funktionellen Markers ER (f_{bin}^{er}) in Relation zum Tumorgewebe (f_{bin}^{tumor}) untersucht werden. Dazu wurden die RCC8-Merkmale (bzw. das RCC8-Histogramm) und die LDQ-Merkmale kombiniert eingesetzt (RTF4, siehe Tabelle B.3). Die zu testende Clusteranzahl wurde auf $K = 2 - 10$ festgelegt. In Abbildung 6.22 (a) zeigt der Hopkins-Index nur eine schwache Clustertendenz. Die Stabilität des Hopkins-Index ist auch in diesem Fall sehr gering, was an den hohen Schwankung zwischen Minimum und Maximum zu sehen ist.

Die PCA-Projektion (siehe Abbildung 6.22 (b)) bildet ca. 75% der totalen Varianz der Merkmalsdaten ab. Somit kann von einer plausiblen visuellen Abschätzung der Verteilung der Merkmalsdaten ausgegangen werden. In der PCA-Projektion mit Rückprojektion der Partition mit Clusteranzahl $K = 6$ (maximaler durchschnittlicher Silhouetten-Wert, siehe Abbildung 6.23 (a)) sind kompakte Gruppierungen zu erkennen (siehe gelbe und blaue Datenpunkte). Differenzierungen durch die Merkmale werden deutlich.

Die externe Clustervalidierung durch die VI_n zeigt lediglich einen minimal signifikanten Wert bei Clusteranzahl $K = 8$ mit ca. 0.15 (siehe Abbildung 6.23 (a)). Alle anderen



(a) Interne und externe Clustervalidierung

(b) Silhouetten-Plot (6 Cluster)

Abbildung 6.23: Testergebnisse auf realen Daten (Relationen - ER): Durchschnittlicher Silhouetten-Wert (SV) und normalisierte Variation of Information (VI_n) (a) sowie der Silhouetten-Plot bei Clusteranzahl 6 (b).

Werte liegen unter 0.10. Ein Zusammenhang mit dem *Follow-Up-Status* kann aufgrund der Partitionen nicht hergestellt werden. Der durchschnittliche Silhouetten-Wert erreicht sein Maximum bei Clusteranzahl $K = 6$ mit ca. 0.55, was für eine mittlere Strukturierung der Daten spricht. Der Silhouetten-Plot (siehe Abbildung 6.23 (b)) der Partition mit Clusteranzahl $K = 6$ zeigt bis auf Cluster 4 und 6 überwiegend Silhouetten-Werte höher 0.6. Besonders hohe Silhouetten-Werte (> 0.8) und somit besonders kompakt und gut separiert sind die Merkmalsvektoren in Cluster 5. Die Merkmalsvektoren dieses Clusters scheinen den Analysen entsprechend eine hohe Ähnlichkeit zu haben. Aus diesem Cluster werden die 10 Fälle mit den höchsten Silhouetten-Werten für einen Stichprobenvergleich mit den Beobachtungen des Domänenexperten herangezogen.

Für diesen Test lagen eindeutig vergleichbare Beobachtungen für jeden Fall vor. In Tabelle 6.5 sind die Beobachtungen des Domänenexperten bezüglich der Verteilung von ER-positivem Gewebe zu den Fällen mit den Silhouetten-Werten aus Cluster 5 eingetragen.

Die Übereinstimmung mit den Beobachtungen des Domänenexperten in der Stichprobenanalyse sind signifikant. Die sieben Beobachtungen mit den höchsten Silhouetten-Werten beschreiben die gleiche Verteilung: „diffus Zentrum“. Lediglich Fall „032-00918“ und „024-21887“ werden unterschiedlich beschrieben. Acht von zehn Beobachtungen liegen mit hohen Silhouetten-Werten im gleichen Cluster. Der PCA-Plot und der durchschnittliche Silhouetten-Wert deuten auf eine gute Differenzierung der Fälle durch die Merkmalsvektoren hin. Der Stichprobenvergleich zeigt, dass sich (zumindest für diese Auswahl und die verwendeten Methoden) die Beobachtungen des Domänenexperten in den extrahierten Merkmalen widerspiegeln.

Patienten-ID	Annotation	Silhouette-Wert	Cluster (1-6)
024-21887	diffus Peripherie nach Zentrum zunehmend	0.8549	5
032-00694	diffus Zentrum	0.8571	5
032-00918	diffus, alle Tumor-Anteile	0.8583	5
022-09100	diffus Zentrum	0.8621	5
021-23131	diffus Zentrum	0.8695	5
044-15571	diffus Zentrum	0.8739	5
043-18071	diffus Zentrum	0.8769	5
004-08068	diffus Zentrum	0.8783	5
021-00098	diffus Zentrum	0.8856	5
032-26200_1	diffus Zentrum	0.8866	5

Tabelle 6.5: Stichprobenanalyse: Vergleich der kompaktesten Fälle mit Beobachtungen des Domänenexperten (ER)

6.3 Reale Daten - Überwachtes Lernen

Nach der Clusteranalyse erfolgte die Analyse mittels des HLVQ-Lernverfahrens (siehe Abschnitt 2.1), als Teil der überwachten Evaluation der Merkmale (siehe Abschnitt 5.6). Getestet wurden zunächst alle Merkmalsgruppen einzeln.

Als initiale Lernrate wurde 0.03 gewählt. Die initiale Prototypenanzahl wurde auf 7 festgelegt. Es wurden 100 unterschiedliche Lerndurchläufe durchgeführt. Die Trainingsdaten (78 Muster) und Testdaten (15 Muster) wurden bei jedem Durchlauf zufällig und disjunkt bestimmt.

Im *HLVQ*-Verfahren werden Trainings- und Testdatensatz prozentual gleich mit entsprechenden Mustern der unterschiedlichen Klassenzugehörigkeit gewählt. Die Anzahl der Prototypen für jede Klasse wird entsprechend der Anzahl von Mustern unterschiedlicher Klassenzugehörigkeiten aufgeteilt. Für jede Klasse werden vor dem Training aus den Trainingsdaten entsprechend viele Merkmalsvektoren zufällig ausgewählt und als Prototypen initialisiert. Daraufhin erfolgt eine Klassifikation vor dem Trainingsschritt anhand der zufällig generierten Prototypen. Dieses Ergebnis wird im Folgenden als *Random* bezeichnet. Es wird als Vergleich herangezogen, um zu bestimmen, inwieweit die eingesetzten Merkmale im Rahmen des Trainings das Klassifikationsergebnis verbessern.

Nachdem alle Merkmalsgruppen einzeln getestet wurden, werden in Abschnitt 6.3.6 die Testergebnisse mit den aussichtsreichsten Merkmalskombinationen der Merkmalsgruppen (gekennzeichnet durch \otimes) zusammen mit den klinischen Daten anhand der Klassifikationsrate verglichen.

Entscheidend für die Evaluation der Merkmale auf dieser Stufe ist, inwieweit sich durch das Lernen der Merkmale ein positiver Einfluss im Vergleich zu den Random-Daten abzeichnet. Die Auswahlen für eine nähere Untersuchung und die Kombination

mit den klinischen Daten orientieren sich von daher an der höchsten Klassifikationsrate in den Trainingsdaten.

Eine nähere Untersuchung bedeutet, dass die Klassifikationsergebnisse zusätzlich durch Diagramme visualisiert werden, welche neben Statistiken zur Klassifikationsrate die Genauigkeit (Precision) und die Trefferquote (Recall) jeder Klasse angeben. Somit wird eine differenzierte Auswertung zwischen den einzelnen Klassen möglich. Zusätzlich werden die Standardabweichungen der jeweiligen Ergebnisse mit angegeben (Fehlerbalken in Diagrammen).

Die genauere Untersuchung wird aufgrund des unbalancierten Datensatzes (50 gesund (healthy), 7 wiedererkrankt (relapse) und 36 tot (dead)) (siehe Abschnitt 4.1) notwendig, um die unterschiedliche Differenzierungsleistung bezüglich der verschiedenen Klassen abschätzen zu können.

6.3.1 Morphometrische Analysen - AE1AE3

In diesem Experiment wurde der komprimierte Merkmalsvektor des morphometrischen Clusterings (kurz MO) (siehe Abschnitt 5.3.2) von f_{bin}^{ae1ae3} eingesetzt. Die eingesetzten morphometrischen Merkmale zur Bestimmung der Clusterverteilungen basieren auf den Clusteranalysen in Abschnitt 6.2.1.

Als Clusteranzahl wurde $K = 2, 3$ und 4 verwendet, welche nach den Clusteranalysen die höchsten durchschnittlichen Silhouetten-Werte im morphometrischen Clustering und im zweiten Clustering erreicht haben. Außerdem bezieht sich die Auswahl auf die vom Domänenexperten als am plausibelsten bewertete Clusteranzahl ($K = 2$ und $K = 3$, siehe Abschnitt 6.2.1). In Tabelle 6.6 werden die durchschnittlichen Klassifikationsraten und die Standardabweichungen bezüglich der unterschiedlichen Testfälle angegeben.

Merkmale	Random		Training		Test	
	μ	σ	μ	σ	μ	σ
MO2(2 Cluster)	0.4590	0.0633	0.5064	0.0163	0.5027	0.0459
MO3(3 Cluster)⊗	0.4388	0.0547	0.5349	0.0181	0.5293	0.0229
MO4(4 Cluster)⊗	0.4300	0.0505	0.5373	0.0153	0.5287	0.0332

Tabelle 6.6: Klassifikationsergebnisse morphometrisch AE1AE3.

Die höchste durchschnittliche Klassifikationsrate auf den Trainingsdaten wurde in diesem Test mit MO4 (vier morphometrische Cluster) erreicht, welcher vom Domänenexperten im Verhältnis zu den beiden anderen Clusteranzahlen als wenig plausibel beschrieben wurde.

Die höchste Klassifikationsrate auf den Testdaten wurde mit MO3 erreicht. Da MO3 sehr ähnliche Werte wie MO4 hat und vom Domänenexperten als am plausibelsten gesehen wird, wurde es ebenfalls für die weiteren Analysen in Kombination mit den klinischen Daten eingesetzt. Die näheren Untersuchungen in Abbildung 6.24 beziehen sich auf MO4.

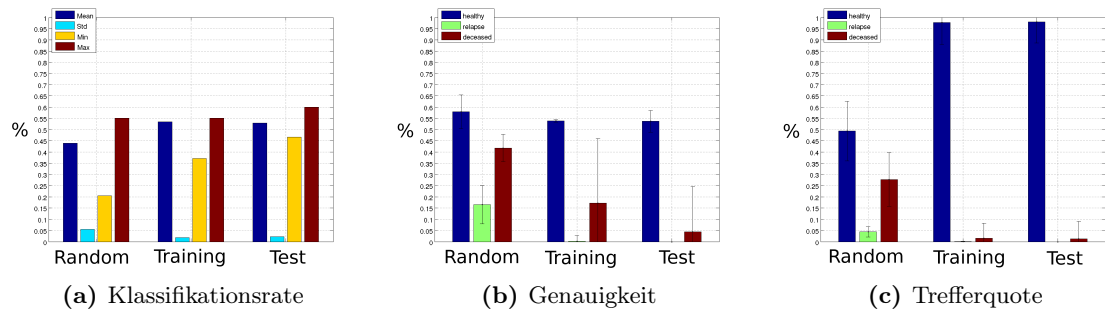


Abbildung 6.24: Bestes Klassifikationsergebnis der morphometrischen Analysen von AE1AE3: Klassifikationsrate (a) (blau: durchschnittliche Klassifikationsrate, cyan: Standardabweichung, gelb: minimale Klassifikationsrate, rot: maximale Klassifikationsrate), Genauigkeit (b) und Trefferquote (c) (blau: Klasse *healthy*, grün: Klasse *relapse*, rot: Klasse *deceased*).

Die Klassifikationsrate erreicht ihr Maximum bei ca. 55% auf den Trainingsdaten und ca. 60% auf den Testdaten und liegt somit nur geringfügig über den Ergebnissen des Randomdatensatzes.

Die Genauigkeit (b) liegt sowohl bei den Trainingsdaten als auch bei den Testdaten für jede Klasse unterhalb des Random-Datensatzes. Kein Fall aus Klasse *relapse* wurde korrekt erkannt.

In der Übersicht zu den Trefferquoten (c) der einzelnen Klassen wird deutlich, wie eine Klassifikationsrate von ca. 53% erreicht wurde. Nahezu alle Fälle wurden aus Klasse *healthy* sowohl bei den Trainingsdaten als auch bei den Testdaten korrekt erkannt. Demgegenüber wurden nahezu keine Fälle aus Klasse *relapse* und *dead* erkannt. Die Klasse *healthy* entspricht 50% des Datensatzes, deren alleinige Erkennung eine Klassifikationsrate von ca. 53% erklärt. Nahezu alle Fälle wurden der Klasse *healthy* zugeordnet. Eine Differenzierung zwischen den Klassen konnte mit Hilfe dieser Merkmale nicht erreicht werden. Insgesamt wurde durch den Lernprozess mit diesen Merkmalen die Klassifikationsleistung im Vergleich zum Zufallsdatensatz verschlechtert.

6.3.2 Strukturelle Analysen - AE1AE3 - DG

In diesem Experiment wurden die Graphanalysen auf f_{bin}^{ae1ae3} unter Verwendung der Merkmale des Delaunay-Graphen (DG's) eingesetzt. Es wurden drei verschiedene Merkmalskombinationen DGTF1, DGTF2 und DGTF3 (siehe Tabelle B.1) getestet.

Merkmale	Random		Training		Test	
	μ	σ	μ	σ	μ	σ
DGTF1	0.4322	0.0695	0.5427	0.0440	0.4460	0.0912
DGTF2	0.4485	0.0628	0.5342	0.0307	0.4573	0.1007
DGTF3 [⊗]	0.4408	0.0573	0.5735	0.0478	0.4233	0.1248

Tabelle 6.7: Klassifikationsergebnisse AE1AE3 - DG.

Testfall DGTF2 erzielte die höchste Klassifikationsrate auf den Trainingsdaten. Die Analysen der Testdaten zeigen in allen Testfällen eine hohe Standardabweichung in den Ergebnissen. Die Klassifikationsrate der Trainingsdaten von Testfall DGTF2 liegt ca. 13% über denen des Random-Datensatzes, was für einen positiven Einfluss der Merkmale bezüglich des Lernens spricht. Auf den Testdaten wurde bei diesen Tests lediglich eine Klassifikationsrate von 0.4573 in Testfall DGTF2 erreicht.

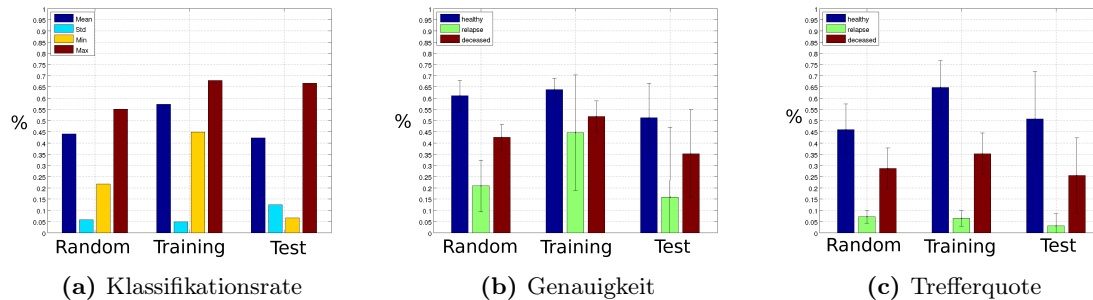


Abbildung 6.25: Bestes Klassifikationsergebnis der strukturellen Analysen von AE1AE3 - DG: Klassifikationsrate (a) (blau: durchschnittliche Klassifikationsrate, cyan: Standardabweichung, gelb: minimale Klassifikationsrate, rot: maximale Klassifikationsrate), Genauigkeit (b) und Trefferquote (c) (blau: Klasse *healthy*, grün: Klasse *relapse*, rot: Klasse *deceased*).

Die genauere Untersuchung von Testfall DGTF3 (siehe Abbildungen 6.25) zeigt, dass mit dieser Merkmalskombination in der Klassifikation ein Maximum von ca. 67% auf den Trainingsdaten und Testdaten erreicht wurde.

Die Genauigkeit konnte durch das Lernen dieser Merkmalskombination im Verhältnis zum Random-Datensatz auf den Trainingsdaten verbessert werden. Für die Klasse *relapse* wurde im Trainingsdatensatz ein Anstieg von ca. 25% erreicht. Dabei muss die hohe Standardabweichung (siehe Fehlerbalken) berücksichtigt werden. Auch die Genauigkeit der beiden anderen Klassen verbesserte sich bei niedrigerer Standardabweichung. Die Genauigkeit des Testdatensatzes ist für Klasse *healthy* am höchsten. Im Vergleich zum vorigen Test wurden die beiden anderen Klassen wesentlich besser erkannt.

Die Trefferquote liegt im Trainingsdatensatz für die Klassen *healthy* und *dead* über den Random-Datensatz. Für die Klasse *healthy* stieg die Genauigkeit um ca. 20% bei ähnlicher Standardabweichung. Für die Klasse *relapse* wiederum sank die Trefferquote

minimal. Für den Testdatensatz liegen die Werte für jede Klasse unter 50%. Auch hier ist das Ergebnis für Klasse *relapse* am schlechtesten.

Insgesamt liegen die Ergebnisse nur für die Trainingsdaten über den Zufallswerten. Die schlechte Trefferquote zeigt, dass viele Fälle nicht korrekt erkannt wurden.

6.3.3 Strukturelle Analysen - AE1AE3 - MST

In diesem Experiment wurden die Graphanalysen auf f_{bin}^{ae1ae3} unter Verwendung der Merkmale des Minimum-Spanning-Trees (MST's) eingesetzt. Sowohl aus den zunächst getesteten Tumorprototypen (siehe Abschnitt 5.4), als auch aus den synthetischen und realen Clusteranalysen, ging hervor, dass der Grad beim MST keine differenzierende Rolle spielt. Aus diesem Grund wurde auf die Verwendung der Gradbasierenden Merkmale in diesem Experiment verzichtet.

Zudem hat sich in den Clusteranalysen herausgestellt, dass der deg_w beim MST mit den durchschnittlichen Kantengewichten e_μ korreliert. Aus diesem Grund wurde auch dieses Merkmal für die weiteren Analysen in diesem Experiment nicht berücksichtigt.

Merkmale	Random		Training		Test	
	μ	σ	μ	σ	μ	σ
MSTTF1 \otimes	0.4550	0.0648	0.5823	0.0327	0.4727	0.1025
MSTTF3	0.4615	0.0598	0.5795	0.0377	0.4920	0.1116
MSTTF4	0.4572	0.0647	0.5392	0.0328	0.4640	0.0938

Tabelle 6.8: Klassifikationsergebnisse AE1AE3 - MST.

Getestet wurden drei verschiedene Merkmalskombinationen MSTF1, MSTF3 und MSTF4 welche in Tabelle B.2 dokumentiert sind. Die Klassifikationsraten für den Trainingsdatensatz aller Testfälle liegen über denen des Random-Datensatzes. Die höchste Klassifikationsrate mit ca. 58% auf den Trainingsdaten wurde unter Verwendung des Testfalls MSTF1 erreicht. Auf den Testdaten bleiben alle Testfälle unter 50% Klassifikationsrate. Die Standardabweichung ist auch bei diesen Testfällen auf dem Testdatensatz sehr hoch. Zur genaueren Untersuchung wurde Testfall MSTF1 herangezogen (siehe Abbildung 6.26).

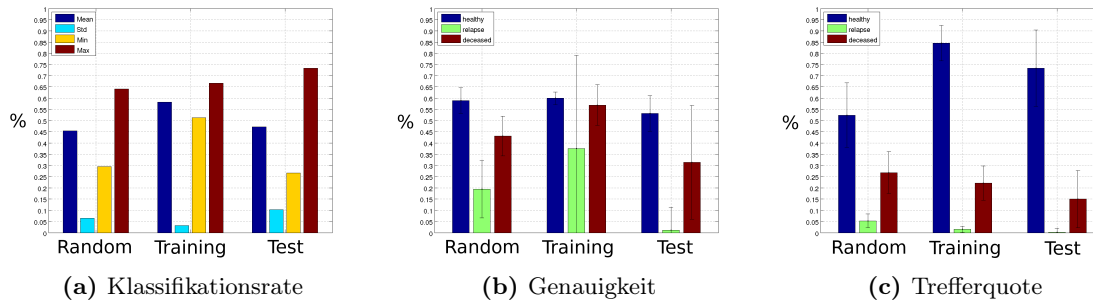


Abbildung 6.26: Bestes Klassifikationsergebnis der strukturellen Analysen von AE1AE3 - MST: Klassifikationsrate (a) (blau: durchschnittliche Klassifikationsrate, cyan: Standardabweichung, gelb: minimale Klassifikationsrate, rot: maximale Klassifikationsrate), Genauigkeit (b) und Trefferquote (c) (blau: Klasse *healthy*, grün: Klasse *relapse*, rot: Klasse *deceased*).

Dieser Testfall zeigt, dass eine maximale Klassifikationsrate von 65% erreicht wurde (siehe (a)). Die Genauigkeit (siehe (b)) auf den Trainingsdaten wurde nur für Klasse *deceased* im Vergleich zum Random-Datensatz erhöht. Für die Klasse *relapse* ergibt sich durch das Lernen auf den Testdaten eine wesentlich höhere Standardabweichung. Die Genauigkeit verbesserte sich durch das Lernen für alle Klassen. Die Genauigkeit für die Klasse *relapse* wurde um ca. 15% erhöht, wobei eine hohe Standardabweichung vorliegt. Für die Klasse *deceased* verbesserte sich die Genauigkeit um ca. 10%.

Die Trefferquoten wurde durch das Lernen im Vergleich zum Random-Datensatz für die Klassen *healthy* und *deceased* verbessert (siehe (c)). Für die Klasse *healthy* wurde eine Steigerung um ca. 30% im Vergleich zum Random-Datensatz erreicht. Die Trefferquote der Klasse *relapse* verschlechterte sich durch den Lernprozess. Auf den Testdaten wurde nur für die Klasse *healthy* mit knapp 75% eine höhere Trefferquote als 50% erreicht. Die Trefferquoten der beiden anderen Klassen verschlechterten sich im Vergleich zum Random-Datensatz.

6.3.4 Relationale Analysen - CD45

In diesem Experiment wurden die relationalen Analysen zwischen f_{bin}^{tumor} und f_{bin}^{cd45} im Hinblick auf ihren Einfluss auf die Klassifikationsergebnisse untersucht. Getestet wurden drei Merkmalskombinationen RTF2, RTF3 und RTF4 (siehe Tabelle B.3). In allen Testfällen wurde eine höhere Klassifikationsrate auf den Trainingsdaten als auf dem Random-Datensatz erreicht. Die höchste Klassifikationsrate auf den Trainingsdaten wurde unter Verwendung der Merkmalskombination RTF4 erzielt (siehe Tabelle 6.9).

Merkmale	Random		Training		Test	
	μ	σ	μ	σ	μ	σ
RTF2	0.4446	0.0583	0.5441	0.0195	0.4947	0.0671
RTF3	0.4708	0.0693	0.5877	0.0475	0.5120	0.1024
RTF4 [⊗]	0.4658	0.0612	0.5903	0.0423	0.4893	0.0833

Tabelle 6.9: Klassifikationsergebnisse CD45 - (LDQ+RCC8).

Zur näheren Untersuchung wurde Merkmalskombination RTF4 herangezogen. Die Ergebnisse werden in Abbildung 6.27 dargestellt. Die Genauigkeit hat sich im Bezug auf die Trainingsdaten für die Klassen *healthy* und *deceased* verbessert. Für Klasse *relapse* verschlechterte sich die Genauigkeit. Auf den Testdaten bleibt die Genauigkeit für alle Klassen unter 50%. Die Trefferquote verschlechterte sich durch das Lernen für die Klassen *relapse* und *deceased* minimal (um. ca 2%), während sie sich für die Klasse *healthy* um knapp 20% verbesserte. Auf den Testdaten wurden nahezu keine Fälle aus Klasse *relapse* erkannt. Alle Ergebnisse auf den Testdaten weisen hohe Standardabweichungen auf.

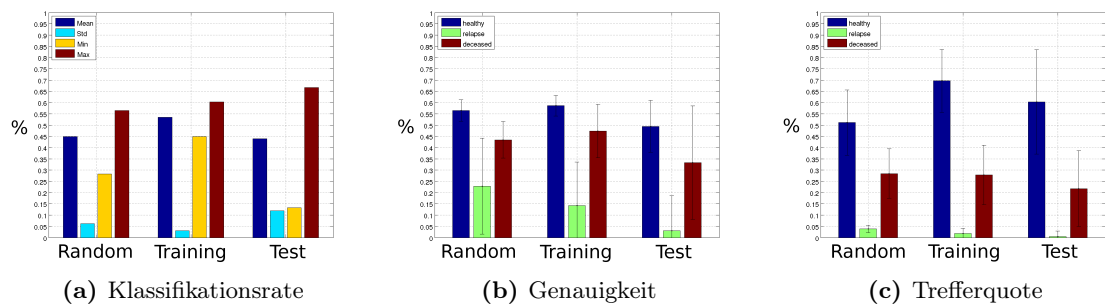


Abbildung 6.27: Bestes Klassifikationsergebnis der relationalen Analysen von CD45: Klassifikationsrate (a) (blau: durchschnittliche Klassifikationsrate, cyan: Standardabweichung, gelb: minimale Klassifikationsrate, rot: maximale Klassifikationsrate), Genauigkeit (b) und Trefferquote (c) (blau: Klasse *healthy*, grün: Klasse *relapse*, rot: Klasse *deceased*).

6.3.5 Relationale Analysen - ER

In diesem Experiment wurden die relationalen Analysen zwischen f_{bin}^{tumor} und f_{bin}^{er} im Hinblick auf ihren Einfluss auf die Klassifikationsergebnisse untersucht. Die Ergebnisse der eingesetzten Merkmalskombinationen werden in Tabelle 6.10 dargestellt. Alle Merkmalskombinationen (RTF) ergeben auf den Trainingsdaten eine höhere Klassifikationsrate, als auf dem Random-Datensatz. Dies spricht für eine Relevanz der Merkmale im Bezug auf den Verlauf. Auf den Testdaten ist im Vergleich zu den anderen Experimenten eine niedrigere Standardabweichung zu beobachten.

Merkmale	Random		Training		Test	
	μ	σ	μ	σ	μ	σ
RTF2	0.4549	0.0623	0.5469	0.0206	0.5127	0.0713
RTF3	0.4601	0.0632	0.5456	0.0134	0.5127	0.0508
RTF4 \otimes	0.4694	0.0615	0.5622	0.0241	0.4827	0.0735

Tabelle 6.10: Klassifikationsergebnisse ER (LDQ+RCC8).

Die höchste Klassifikationsrate auf den Trainingsdaten wurde unter Merkmalskombination RTF4 erreicht, welche näher untersucht wurde. Die Genauigkeit wurde für die Klassen *relapse* und *deceased* durch den Lernprozess jeweils um ca. 10% verbessert, wobei sich die Standardabweichung gleichzeitig erhöht. Auf den Testdaten wurde durchschnittlich nur eine geringe Genauigkeit (unter 50% in allen Klassen) erreicht.

Die Trefferquote wurde durch die Merkmalskombination lediglich für die Klasse *healthy* verbessert. Für die beiden anderen Klassen verschlechterte sich die Trefferquote um ca. 5%. Auf den Testdaten liegt lediglich, ähnlich wie in den vorangegangenen Experimenten, die Trefferquote für die Klasse *healthy* mit 75% über 50%. Sowohl auf den Trainingsdaten, als auch auf den Testdaten konnte keine Verbesserung der Trefferquote in den Klassen *relapse* und *deceased* erreicht werden.

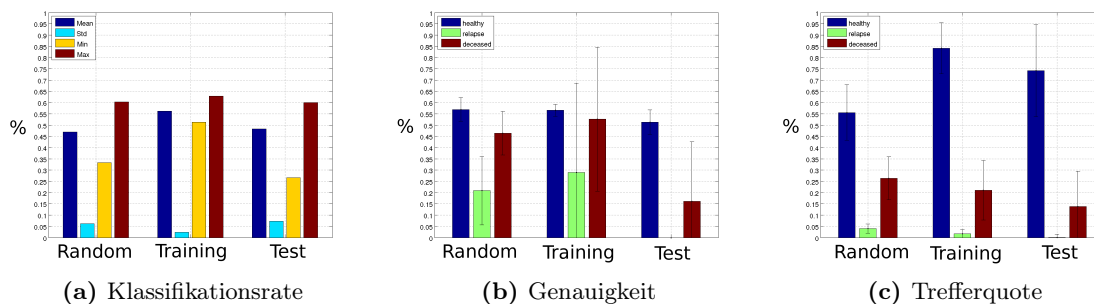


Abbildung 6.28: Bestes Klassifikationsergebnis der relationalen Analysen von ER: Klassifikationsrate (a) (blau: durchschnittliche Klassifikationsrate, cyan: Standardabweichung, gelb: minimale Klassifikationsrate, rot: maximale Klassifikationsrate), Genauigkeit (b) und Trefferquote (c) (blau: Klasse *healthy*, grün: Klasse *relapse*, rot: Klasse *deceased*).

6.3.6 Kombinationen mit klinischen Daten

Bei diesen Experimenten wurden die berechneten Merkmalskombinationen mit den besten Ergebnissen (\otimes) der vorigen Experimente mit den klinischen Daten (siehe Abschnitt 4.1) zusammengebracht und verglichen.

Um die Auswahl der besten klinischen Merkmale zu bestimmen, wurden zwei Merkmalskombinationen getestet (siehe Tabelle 6.11, Test der klinischen Daten). Die höchste Klassifikationsrate auf den Trainingsdaten wurde unter Verwendung aller klinischen Da-

ten (kID) erreicht. Die Kombination aus den Merkmalen des Gradings und des Alters der Patienten erreichte zwar höhere Werte auf den Testdaten, jedoch geringere Werte auf den Trainingsdaten. Aus diesem Grund werden alle klinischen Merkmale in die Kombinationstests integriert.

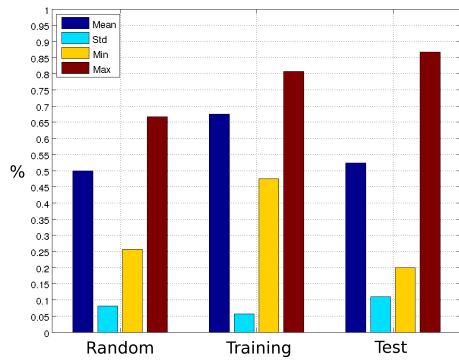
Die besten Testfälle der einzelnen Merkmalsgruppen wurden zunächst mit den klinischen Daten einzeln getestet (siehe Tabelle 6.11, Kombinationstests A). Die Testfälle mit den höchsten Klassifikationsraten auf den Trainingsdaten der Kombinationstests A, wurden zusammen mit Merkmalsgruppen, unter denen in den Kombinationstests A die höchste Klassifikationsrate erreicht wurde, getestet. Diese Tests werden als Kombinationstests B bezeichnet. Die einzelnen Ergebnisse werden in Tabelle 6.11 abgebildet.

Auf den Trainingsdaten wurde bei den Kombinationstests A die maximale Klassifikationsrate (0.6690) unter Verwendung der Merkmalskombination CD45 RTF4 erreicht. Somit liegt die Klassifikationsrate knapp 2% über dem Maximum bei Verwendung der klinischen Daten. Auf den Testdaten wurde die höchste Klassifikationsrate (0.5433) mit Merkmalskombination DG DGTF3 erzielt. Sie liegt 2% unter dem Maximalwert bei den klinischen Daten.

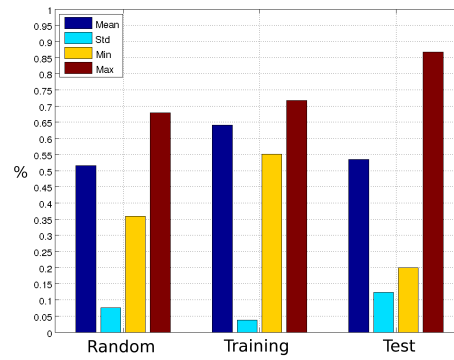
In den Kombinationstests B wurde auf dem Trainingsdatensatz eine Klassifikationsrate von 0.6747 erreicht. Somit wurde eine 3% höhere Klassifikationsrate, als bei ausschließlicher Verwendung der klinischen Daten erzielt. Gleichzeitig kann ein Anstieg der Standardabweichung sowohl bei den Kombinationstests A als auch bei den Kombinationstests B im Vergleich zur ausschließlichen Verwendung der klinischen Daten beobachtet werden. In den Kombinationstests B wurde hingegen die Standardabweichung für die Ergebnisse der Testdaten minimal gesenkt.

Für die nähere Untersuchung und einen direkten Vergleich werden die Ergebnisse der klinischen Daten mit Kombinationstest B „ER RTF4[⊗], CD45 RTF4[⊗], MO4, kID“ (im Folgenden kurz kID+[⊗]) verwendet. In den Diagrammen in Abbildung 6.29 wird eine Gegenüberstellung der beiden Ergebnisse dargestellt.

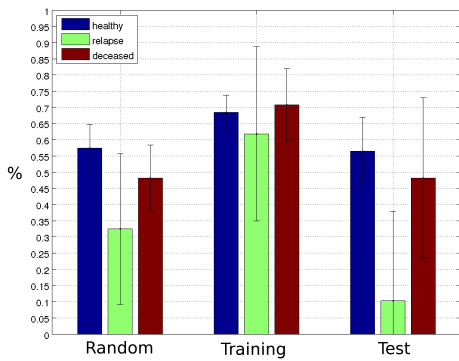
Unter Einsatz von (kID+[⊗]) wurde auf den Trainingsdaten ein Maximum von über 80% (+6%) und auf den Testdaten von über 90% (+0%) erreicht. Die Genauigkeit wurde für Klasse *relapse* (+ ca. 10%) und Klasse *deceased* (ca. 5%) erhöht. Für die Testdaten wurde die Genauigkeit für alle Klassen verschlechtert. Die Trefferquote wurde lediglich für die Klasse *healthy* auf den Trainingsdaten und Testdaten verbessert. Für die anderen Klassen hat sich die Trefferquote sowohl in den Trainings- als auch in den Testdaten verschlechtert.



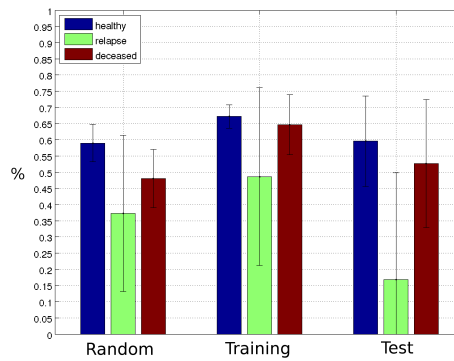
(a) Klassifikationsrate (kID+⊗)



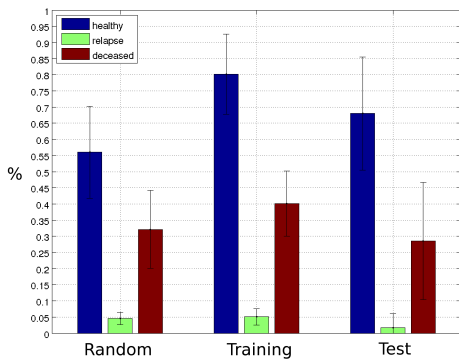
(b) Klassifikationsrate (kID)



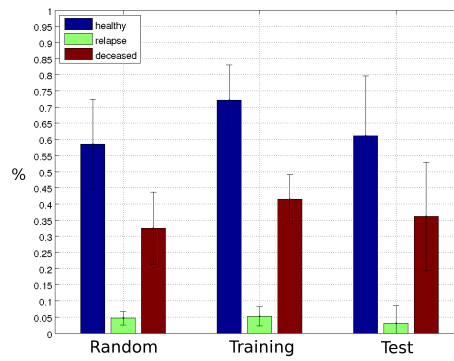
(c) Genauigkeit (kID+⊗)



(d) Genauigkeit (kID)



(e) Trefferquote (kID+⊗)



(f) Trefferquote (kID)

Abbildung 6.29: Vergleich der Klassifikationsergebnisse unter Verwendung der klinischen Daten (kID) und in Kombination mit den implementierten Merkmalen (kID+⊗): Klassifikationsrate (a) (blau: durchschnittliche Klassifikationsrate, cyan: Standardabweichung, gelb: minimale Klassifikationsrate, rot: maximale Klassifikationsrate), Genauigkeit (b) und Trefferquote (c) (blau: Klasse *healthy*, grün: Klasse *relapse*, rot: Klasse *deceased*).

Merkmale	Random		Training		Test	
	μ	σ	μ	σ	μ	σ
Test der klinischen Daten						
alle klinischen Daten (klD)	0.5150	0.0751	0.6404	0.0374	0.5353	0.1233
Grading,Alter	0.4976	0.0787	0.5815	0.0452	0.5633	0.1150
Kombinationstests A						
MO3 \otimes ,klD	0.5221	0.0759	0.6460	0.0413	0.5140	0.1102
MO4 \otimes ,klD	0.5085	0.0749	0.6567	0.0422	0.5273	0.1221
DG DGTF3 \otimes ,klD	0.5271	0.0678	0.6237	0.0470	0.5433	0.1204
MST MSTTF1 \otimes ,klD	0.5192	0.0787	0.6346	0.0479	0.5013	0.1119
CD45 RTF4 \otimes ,klD	0.4985	0.0801	0.6690	0.0441	0.5233	0.1086
ER RTF4 \otimes ,klD	0.5149	0.0624	0.6594	0.0450	0.5027	0.1219
Kombinationstests B						
ER RTF4 \otimes , CD45 RTF4 \otimes , klD	0.4931	0.0730	0.6673	0.0504	0.5273	0.1077
ER RTF4 \otimes , CD45 RTF4 \otimes , MO4, klD	0.4991	0.0806	0.6747	0.0571	0.5240	0.1093

Tabelle 6.11: Klassifikationsergebnisse in Kombination mit den klinischen Daten.

Kapitel 7

Diskussion

Die eingesetzten Merkmale und das Konzept zur unüberwachten Merkmalsvalidierung liefern auf den synthetischen Daten plausible Ergebnisse. Die Kombination der implementierten Basis-Deskriptoren, Moment invariants, der Fourier-Deskriptoren und die Anwendung des *K-Means*-Verfahrens zur Informationsverdichtung ermöglichen eine gute Differenzierung zwischen den unterschiedlichen Objekten der Form-Datenbank. Dies hat die Anwendung auf die realen Daten motiviert.

Die strukturellen Analysen auf den synthetischen Daten ergeben, dass der Delaunay-Graph (DG) für die Repräsentation der gesuchten Strukturen besser als der Minimum-Spanning-Tree (MST) geeignet ist. Er erreicht eine höhere Übereinstimmung mit den gesuchten Klassenzugehörigkeiten. Außerdem wird bei den Tests deutlich, dass die Merkmale für den MST, welche den Grad der Knoten miteinbeziehen, keine bessere Diskriminierung erlauben. Sowohl durch die LDQ-Merkmale, als auch durch die RCC8-Merkmale, wird eine hohe Übereinstimmung in den Klassenzugehörigkeiten erreicht. Gleichzeitig ergeben sich sehr kompakte und gut differenzierte Cluster.

Der durchschnittliche Silhouetten-Wert korrespondiert nicht immer mit den Werten der normalisierten Variation of Information. Kompakte gut separierte Cluster müssen demnach nicht einer besseren Klassenzugehörigkeit entsprechen.

Die eingesetzten synthetischen Daten geben nur eine eingeschränkte Aussage über die Sinnhaftigkeit der Merkmale für die wesentlich komplexeren realen Daten. Diese erste Testreihe gibt jedoch Aufschluss darüber, dass die Implementation und die selbstständig erweiterten Ansätze, zumindest unter eingeschränkten Rahmenbedingungen, funktionieren. Dies ist Grundvoraussetzung für die Anwendung der Bildmerkmale und Analyse-schritte auf den realen Daten.

Die Clusteranalysen auf den realen Daten stellen sich wesentlich schwieriger dar. Das Projekt ExPrimage ist während der Anfertigung der vorliegenden Arbeit offiziell beendet worden, was den Zugriff auf den Domänenexperten einschränkte. Zudem wurde der Zeitaufwand für eine humane Evaluation und Diskussion der Ergebnisse unterschätzt. Aus diesen Gründen konnten nur einzelne Ergebnisse begutachtet werden, weswegen auf einen Stichprobenvergleich mit dokumentierten Beobachtungen zurückgegriffen wurde.

In keiner der getesteten Merkmalskombinationen kann eine Übereinstimmung mit dem *Follow-Up-Status* der Patienten erkannt werden. Alle Werte der VI_n bleiben über 0.85. Aus diesem Grund orientierte sich die Auswahl der näher zu untersuchenden Partitionen primär an dem durchschnittlichen Silhouetten-Wert. In den Tests wird offensichtlich, dass der Hopkins-Index für die Analysen, unter Verwendung niedrigerer Anzahlen von Merkmalsvektoren (z.B. 93), an Stabilität und Aussagekraft verliert. Er ist somit für Clusteranalysen mit geringer Datenzahl kaum geeignet.

Die morphometrischen Analysen der Bildmodalität f_{bin}^{ae1ae3} , unter Anwendung des *K-Means*-Verfahrens zur Informationsverdichtung und Visualisierung (siehe Abschnitt 5.3), ergeben geringfügig kompakte Cluster im ersten Clustering. Im zweiten Clustering (*SecCl*) werden hohe durchschnittliche Silhouetten-Werte erreicht. Durch diesen Ansatz werden zwar ähnliche Formen erkannt, die gesuchten Eigenschaften des Tumors können dadurch jedoch nicht repräsentiert werden. Nur teilweise werden relevante Ergebnisse in vereinzelt Beispielen unter Clusteranzahl $K = 3$ erkannt. Über mehrere Tumorfälle wird keine ausreichende Übereinstimmung erreicht. Die Anforderungen Größe, Berandung, zentrale Aushöhlungen, Kompaktheit und Dissoziierungen können auf dieser Detailstufe nicht ausreichend mit den eingesetzten Methoden erfasst werden. Dies bestätigt die Rückmeldung des Domänenexperten, welcher die Ergebnisse als teilweise plausibel ansieht.

Eine Erklärung hierfür ist die nicht ausreichend differenzierte Detektion der Tumorobjekte. Es bestehen sehr heterogene Übergänge zwischen dem Tumorgewebe, welche durch das eingesetzte Verfahren Otsu3Lab nicht ausreichend getrennt werden. Die Pixelklassifikation durch das Verfahren Otsu3Lab und die darauf folgende Detektion der Tumorobjekte mittels Connected-Components ist zur Analyse von morphometrischen Eigenschaften auf Gewebeebene nicht ausreichend. Diese Aufgabe könnte durch den Einsatz einer markerkontrollierten Segmentierung, unter Verwendung der Wasserscheide-Transformation [Soi98], gelöst werden.

Ungeachtet dessen bleibt eine morphometrische Analyse von f_{rgb}^{ae1ae3} auf Gewebeebene schwierig, da sich oft keine eindeutigen Entitäten, in Form von Tumorobjekten, auf dieser Auflösungsstufe der Gewebeschnitte unterscheiden lassen. Morphometrischen Analysen von f_{rgb}^{ae1ae3} werden erst auf einer höheren Auflösungsstufe der Gewebeschnitte, bei denen sich die Grenzen des Tumorgewebes eindeutig identifizieren lassen, aussagekräftig. Die gesuchten Eigenschaften in f_{rgb}^{ae1ae3} können durch Texturanalysen auf der Gewebeebene gut differenziert werden, wie die Ergebnisse aus [Kha11] demonstrieren.

Die Ergebnisse des überwachten Lernens zeigen, dass die Klassifikationsrate durch das Training gesteigert werden konnte. In der näheren Untersuchung hingegen wird jedoch deutlich, dass sich die Genauigkeit verschlechtert hat und die Trefferquote lediglich für die Klasse *healthy* gesteigert werden konnte. Erst in den Kombinationstests B wird eine geringfügige Relevanz der morphometrischen Analysen ersichtlich. Insgesamt wird bei den eingesetzten morphometrischen Methoden auf Gewebeebene daher von einer geringen Bedeutung im Bezug auf die Differenzierung von Tumoren ausgegangen.

Für die Anwendung der morphometrischen Analysen auf f_{bin}^{tumor} ist die Datenlage besser, da dort wesentlich klarer getrennte Tumorobjekte vorliegen. Aufgrund des eingeschränkten Zeitrahmens konnten dazu nur wenige Tests durchgeführt werden, zu denen

kein Feedback des Domänenexperten vorliegt. Auf eine Dokumentation im Rahmen der vorliegenden Arbeit wurde daher verzichtet. Aus Sicht des Anwenders werden in f_{bin}^{tumor} relevante Strukturen besser erfasst (siehe Abbildung C.3). Durch diese Analysen kann die grobe Morphometrie des Tumors erfasst werden.

Die strukturellen Analysen mit Hilfe des MST's und DG's zeigen in den Clusteranalysen der realen Daten ähnliche Ergebnisse. Beide erreichen das beste Clusterergebnis ohne Verwendung der Gradbasierten Merkmale. Für den MST ist dies plausibel, da in einem Baum nur in den inneren Knoten Abweichungen vom Grad vorliegen. Der häufige Einsatz dieser Merkmale in der Literatur [GBC⁺09, Bou08] ist daher nicht nachzuvollziehen.

Bei den Analysen des DG's scheinen die Kantengewichte ebenfalls eine differenziertere Rolle als der Knotengrad zu haben. Dies bestätigen die Beispielergebnisse der Tumorprototypen (siehe Abschnitt 5.4.1). In den Stichprobenanalysen zeigen die Ergebnisse des DG's eine etwas bessere Übereinstimmung mit den Annotationen des Domänenexperten.

Das Ergebnis der überwachten Analysen des DG's zeigt, im Vergleich zum Random-Datensatz, eine Relevanz für den Verlauf. Erhöht werden Genauigkeit und Trefferquote. Die Ergebnisse für den MST sind in der Klassifikationsrate minimal (knapp 1%) besser, zeigen aber in der Genauigkeit und Trefferquote für die Klassen *relapse* und *deceased* schlechtere Ergebnisse als der DG.

In den Kombinationstests A wird durch die Merkmale des DG's die höchste Klassifikationsrate (ca. 54%) auf den Testdaten erreicht. Anhand der Ergebnisse kann darauf geschlossen werden, dass die Merkmale, auf Basis der Kantengewichte und Knotenzahl des DG's, eine bessere Diskriminierung auf Gewebeebene und im Hinblick auf die Verlaufsanalysen erlauben, als es mit dem MST möglich ist.

Durch die Abtastung des gesamten detektierten Tumorgewebes sind die Graphanalysen unabhängiger von den Grenzen der Tumorobjekte als die morphometrischen Analysen. Sie eignen sich daher besser für Analysen auf Gewebeebene.

Mit den Graphanalysen werden globale Ausbreitungsmuster im Tumorgewebe erfasst. So kann grundsätzlich zwischen kompakten und stark dissoziierten Tumoren unterschieden werden. Subregionen mit unterschiedlicher Verteilung innerhalb des Tumorgewebes werden nicht erfasst. Im Ausblick wird ein Konzept angesprochen, welches diese Einschränkungen aufheben könnte.

Die relationalen Analysen mittels LDQ- und RCC8-Merkmale ergeben bei ausgewählten Tumorprototypen jeweils sinnvolle Ergebnisse (siehe Abschnitt 5.5). Die LDQ-Merkmale in Kombination mit den RCC8-Merkmalen zeigen bei den Clusteranalysen für CD45 (f_{bin}^{cd45}) mäßige Resultate. Die Stichprobenanalysen zeigen keine Übereinstimmung mit den Beobachtungen des Domänenexperten. Die relationalen Clusteranalysen von ER (f_{bin}^{er}) erreichen bessere Ergebnisse. Im Vergleich zu den Analysen von CD45 werden kompaktere und mehr separierte Cluster generiert. Die Stichprobenanalyse zeigt eine Übereinstimmung mit den Beobachtungen des Domänenexperten.

Ein Grund für diese unterschiedlichen Ergebnisse kann in der Festlegung des Schwellwerts für die RCC8-Merkmale liegen. Die Verteilung von CD45 enthält in vielen Fällen einzelne Pixel und wenig kompakte Tumorobjekte, welche aufgrund des Schwellwertes

nicht weiter berücksichtigt werden. Bei den Verteilungen von ER liegen oftmals größere kompakte Tumorobjekte vor.

Nach weiteren manuellen Untersuchungen der RCC8-Merkmale von CD45 und ER wird deutlich, dass die Analysen auf Grundlage der Connected-Components nur in den Fällen sinnvoll erscheinen, in denen kompakte kleinere Relationsobjekte der funktionellen Marker vorliegen. Liegen sehr große Relationsobjekte vor, wird das gesamte Bildobjekt einer Relation zugeordnet. Den gesuchten Anforderungen entsprechend ist es sinnvoller, dass größere Tumorobjekte nicht als ganzes betrachtet werden sondern gleichzeitig unterschiedlichen Relationen zugeordnet werden können.

Anhand der Analysen der Tumorprototypen durch die LDQ-Merkmale wird deutlich, dass eine Tendenz für einen zu hohen peripheren Anteil vorliegt. Da die Tumorobjekte vorher nicht morphologisch geschlossen werden, entstehen viele kleinere periphere Anteile innerhalb der Tumorobjekte, was bei größeren Aushöhlungen erwünscht ist (die ebenfalls als Peripherie gesehen werden können), bei kleineren Aushöhlungen jedoch für Artefakte der Gewebecharakterisierung der SOM-Analysen (siehe Abschnitt 5.7.2) sprechen kann. Aus diesem Grund sollte vor Berechnung der LDQ-Merkmale ein morphologisches Schließen [BB05] der einzelnen Tumorobjekte in Erwägung gezogen werden. Da die LDQ-Merkmale auf Pixelebene der funktionellen Marker erstellt werden, sind sie nicht, wie die RCC8-Merkmale, abhängig von den Grenzen der Relationsobjekte.

Sowohl die CD45-Analysen, als auch die ER-Analysen erreichen unter jeder getesteten Merkmalskonfiguration eine höhere Klassifikationsrate im Vergleich zum Random-Datensatz. Dies spricht für eine Relevanz im Zusammenhang mit dem Verlauf. Die Kombination von LDQ- mit RCC8-Merkmalen zeigt eine höhere Relevanz gegenüber dem Verlauf, als einzeln verwendet. Daher erscheint eine Kombination der beiden Ansätze sinnvoll. Die relationalen Analysen von ER zeigen, gegenüber den Analysen von CD45, eine höhere Genauigkeit und eine höhere Trefferquote im Bezug zur Differenzierung der drei gegebenen Klassen. Insgesamt bestätigt dies die besseren Ergebnisse der relationalen Clusteranalysen von ER.

Die Kombinationstests mit den klinischen Daten zeigen bei allen getesteten Merkmalskombinationen (außer den strukturellen) auf den Trainingsdatensatz eine Erhöhung der Klassifikationsrate. Auf den Trainingsdaten konnte in den Kombinationstests eine Erhöhung der Klassifikationsrate um ca. 3% erreicht werden. Es wurde ein Maximalwert von 80% beobachtet (Erhöhung um 10%). Auf den Testdaten wurde, im Vergleich zu den klinischen Daten, keine Erhöhung der Klassifikationsrate erzielt.

Da die Anzahl der Merkmale im Verhältnis zu den gegebenen Fällen im letzten Kombinationstest mit der höchsten Klassifikationsrate relativ zur verwendeten Musteranzahl (93) hoch ist (>24) und die Klassifikationsrate der Testdaten sich nicht verbessert hat, kann dies an einem Overfitting der Trainingsdaten liegen.

In allen überwachten Tests wird deutlich, dass die Klasse *relapse* nur in wenigen Fällen korrekt erkannt wird. Dies hängt wahrscheinlich mit dem unbalancierten Datensatz zusammen, da in der Klasse *relapse* lediglich 7 von 93 Fällen vorliegen. In allen Tests wird deutlich, dass auf den Testdaten keine stabilen Ergebnisse mit den Merkmalen erreicht werden. Die Standardabweichung ist durchgängig hoch (nahezu in allen Tests

>0.1). Die eingesetzten Merkmale zeigen zwar prinzipiell eine geringe Relevanz mit den Verlaufsdaten, führen hingegen nicht zu einer Stabilisierung (im Sinne einer geringeren Standardabweichung) der Ergebnisse.

Die Tests konnten nur unter Einschränkung einer geringen Anzahl von Merkmalskombinationen durchgeführt werden. Andere Kombinationen könnten bessere Ergebnisse liefern.

Eine allgemeine Aussage, inwiefern die entwickelten Merkmale für den Menschen nachvollziehbare relevante Eigenschaften erfassen und die aufgestellten Anforderungen erfüllen, kann aufgrund der fehlenden Ground-Truth bezüglich der Ausbreitungsmuster im Rahmen der vorliegenden Arbeit nicht erbracht werden.

Das Graiding der Domänenexperten konnte in diesem Fall nicht verwendet werden. Die sich stetig ändernde Herangehensweisen unterschiedlicher Graiding-Systeme und die *Intra- und Interobserver-Variabilität* ermöglichen keine konsistente Aussage [Bou08].

Der *Follow-Up-Status* ist klarer definiert und verringert die Problematik des Graiding [Bou08]. Im Bezug auf die untersuchten Ausbreitungsmuster stellt er jedoch lediglich ein indirektes Label bzw. eine indirekte Klassenzugehörigkeit zur Verfügung. Eine verlässlichere Aussage kann nur in direkter Zusammenarbeit mit den Domänenexperten geschehen.

Die benötigte Zeit und der Aufwand einer geeigneten Präsentation und Visualisierung für eine umfassende humane Evaluierung, unter Einbeziehung des Domänenexperten, stellt eine Herausforderung dar, die unterschätzt wurde.

Die untersuchten Prototypen, Clusteranalysen und Stichproben geben lediglich Hinweise dazu, inwieweit die Semantische Lücke verringert werden konnte. In Folge dessen sollte eine bessere Visualisierung der Ergebnisse, vor allem der relationalen Analysen der einzelnen Merkmalsgruppen, erfolgen (siehe Abschnitt 8). Anhand der Clusteranalysen müssen zur Validierung der eingesetzten Merkmale, wie in Abschnitt 5.6 beschrieben, Stichproben unmittelbar mit verschiedenen Domänenexperten diskutiert werden. Erst dann kann eine genauere Aussage darüber gemacht werden, inwieweit die eingesetzten Merkmale für den Domänenexperten relevante Informationen enthalten und mit den Ausprägungen des Mammakarzinoms zusammenhängen. Notwendig sind hierfür Reliabilitätstests¹ durch mehrere Domänenexperten.

Anhand der Clusteranalysen erscheinen die Ergebnisse des DG's und die relationalen Analysen von ER am relevantesten für den Domänenexperten. Allerdings müssen dabei die erkannten Einschränkungen der RCC8- und LDQ-Merkmale beachtet werden. Es wird davon ausgegangen, dass sich die LDQ-Merkmale besser für die Analysen der Relationen der funktionellen Marker eignen. Wie bereits erwähnt, agieren sie auf Pixelebene der funktionellen Marker und sind dadurch unabhängiger von den Formen der Connected-Components.

Der Einsatz der Connected-Components stellt eine grundsätzliche Schwierigkeit für alle darauf basierenden Methoden dar. Wie bereits angesprochen, ist nicht immer klar, inwiefern die Connected-Components bzw. die detektierten Bildobjekte, die tatsächlichen Grenzen der unterschiedlichen Gewebeentitäten (Tumorgewebe, entzündete Regio-

¹Zuverlässigkeitstests

nen, usw.) widerspiegeln und korrekt trennen. Zur Lösung dieser Schwierigkeit müssen differenziertere Vorverarbeitungsschritte gefunden und auf Gewebeebene evaluiert werden.

Kapitel 8

Zusammenfassung und Ausblick

8.1 Zusammenfassung

Ziel der vorliegenden Arbeit war es, Merkmale zur Bestimmung von Ausbreitungsmustern in segmentierten Bilddaten von histopathologischen Gewebeschnitten auf Gewebeebene des Mammakarzinoms zu entwickeln. Die Durchführung der vorliegenden Arbeit geschah im Rahmen des Projektes ExPrimage.

Die Merkmale sollten zunächst in Form von synthetischen Daten auf ihre Korrektheit überprüft werden. Aufgrund der fehlenden Ground-Truth musste ein Konzept zur Validierung der Merkmale auf den realen Daten entwickelt werden.

Da sich die vorliegende Arbeit dem Bereich des Maschinellen Lernens zuordnet, wurde einführend ein Überblick zu den Grundlagen dieser Thematik gegeben (siehe Kapitel 2). In diesem Kontext wurden sowohl eingesetzte Verfahren erläutert, als auch eine Einordnung der Arbeit in das Maschinelle Lernen durchgeführt.

Ein Überblick zu eingesetzten Bildmerkmalen in der medizinischen Bildanalyse und der digitalen Histopathologie wurde in Kapitel 3 beschrieben. In diesem Kontext stellen sich die Kategorien Morphometrie (Größe und Form), Struktur (Merkmale auf Basis von Graphen) und Relation (Kokalisationen und Relationen), als für die vorliegende Arbeit relevant heraus. In der Literatur beschriebene Anwendungen von Merkmalen im Bereich der digitalen Krebsanalyse, welche diesen Kategorien zugeordnet werden, wurden aufgeführt. Vorwiegend werden auf zellulärer Ebene Graphen und morphometrische Basis-Deskriptoren eingesetzt. Auf Gewebeebene werden hauptsächlich Texturanalysen angewendet. Lediglich eine Anwendung auf Basis des RCC8-Histogramms konnte im Bereich der relationalen Analysen identifiziert werden.

In Zusammenhang dessen wurden die implementierten und verwendeten Merkmale sowie die benötigten Algorithmen erläutert.

Daraufhin wurden das Anwendungsszenario im Projekt ExPrimage, die Datenlage und die vorausgegangenen Analysen dargestellt (siehe Kapitel 4). Neben klinischen Parametern der Patienten stand Bildmaterial verschiedener Bildmodalitäten zur Verfügung, welche zum einen, die strukturellen Gewebeeigenschaften erfassen, und zum anderen, funktionell relevante Strukturen im Gewebe hervorheben.

In Zusammenarbeit mit dem Domänenexperten wurden Anforderungen an die Merkmale definiert (siehe Abschnitt 5.1). Um diese Anforderungen erfüllen zu können, musste ein weiterer Vorverarbeitungsschritt zur Detektion feinerer Tumorstrukturen umgesetzt werden (siehe Abschnitt 5.2). Getestet wurde das *K-Means*-Verfahren und das Schwellwertverfahren von Otsu. Untersucht wurden außerdem unterschiedliche Farbräume (RGB und Lab). Als am effektivsten stellte sich das Schwellwertverfahren von Otsu heraus, welches auf die einzelnen drei Kanäle des Lab-Farbraumes (zur besseren Beleuchtungsunabhängigkeit) angewendet und zur Detektion der feineren Tumorstrukturen eingesetzt wurde.

Die einzelnen Bildobjekte der verschiedenen Bildmodalitäten wurden als Connected-Components detektiert (siehe Abschnitt 5.2.2). Dabei wurde unterschieden zwischen Tumorobjekten (repräsentativ für Tumorgewebe) und Relationsobjekten (repräsentativ für funktionell relevante Strukturen).

Um die Anforderungen zur Beschreibung der Tumorobjekte in verschiedenen Auflösungsstufen zu erfüllen, wurden in der Kategorie Morphometrie Basis-Deskriptoren (Area, Roundness, Perimeter, Irregularity und Compactness) zur Beschreibung der Größe und Kompaktheit implementiert. Zur Beschreibung der Kontur bzw. Berandung, wurden die Fourier-Deskriptoren und Fourier-Energien unter Verwendung der zentroiden Distanzen eines Tumorobjektes implementiert. Als weitere Merkmale wurden die Moment Invariants eingesetzt, um zusätzlich Aushöhlungen in Tumorobjekten erfassen zu können.

Zunächst wurden diese Merkmale in Form von deskriptiver Statistik (Mittelwert und Standardabweichung) für jedes Bildmuster berechnet (siehe Abschnitt 5.3.1). Da dies zu einer hohen Anzahl von Merkmalen führte und keine unmittelbare Visualisierung bzw. ein Vergleich zwischen den einzelnen Tumorobjekten innerhalb einer Bildmodalität möglich war, wurde zusätzlich der Weg einer Informationsverdichtung, in Form eines Clusterings der Merkmale jedes einzelnen Tumorobjektes durch das *K-Means*-Verfahren, gewählt (siehe Abschnitt 5.3.2). Durch Rückprojektion der Clusterzugehörigkeiten jedes Tumorobjektes wurde auf Grund des %-Flächenanteils jedes Clusters in der Bildmodalität ein neuer Merkmalsvektor mit niedrigerer Dimension erstellt. Zudem stand so eine unmittelbare Visualisierung für den Domänenexperten zur Verfügung.

Zur Erfüllung der Anforderungen der Bestimmung räumlicher Dissoziierung und Kompaktheit wurden im Rahmen der strukturellen Analysen die Graphstrukturen MST (Minimum-Spanning-Tree) und DG (Delaunay-Graph) eingesetzt (siehe Abschnitt 5.4). Als Knotenidentifikation wurde ein Verfahren zur Abtastung der Tumorobjekte, unter Verwendung eines regelmäßigen Pixelrasters, weiterentwickelt. Dadurch wurde eine adäquate Repräsentation des Tumorgewebes möglich. Als Kantengewichte wurden die euklidischen Distanzen zwischen den Knoten eingesetzt. Auf Grundlage der Adjazenz- und Distanzmatrix wurden Merkmale des Knotengrades (Mittelwert, Standardabweichung, Minimum-Zu-Maximum, Variationskoeffizient, Randic Index und gewichteter Knotengrad) und der Kantengewichte (Mittelwert, Standardabweichung, Minimum-Zu-Maximum, Variationskoeffizient) sowie die Anzahl der Knoten und für den DG die Anzahl der Zyklen bestimmt.

Zur Bestimmung der Kolokalisationen und Relationen zwischen funktionellen Markern und Tumorgewebe wurde das RCC8-Histogramm für die Geweebene angepasst.

Dabei wurden Größeninformationen der Tumorobjekte als Gewichtung, adaptive Schwellwerte sowie die Möglichkeit für Mehrfachrelationen integriert (siehe Abschnitt 5.5.1).

Ein neuer Ansatz zur Quantifizierung von Kolo-kalisationen und zur Unterstützung der RCC8-Merkmale wurde entwickelt (siehe Abschnitt 5.5.2). Auf Grundlage der Linearen Distanzanalysen werden mit dem entwickelten LDQ-Verfahren (Lineare Distanz-Quantifizierung) auf Pixelebene für jedes Tumorobjekt periphere und zentrale Regionen definiert. Die Überlagerung mit funktionellen Markern in diesen Regionen kann dadurch quantifiziert werden.

Es wurde ein flexibles objektorientiertes Datenmodell entwickelt, welches die Datenhaltung von den Funktionen der Bildanalysen zur Bestimmung der Merkmale entkoppelt (siehe Abschnitt 5.7). Die Implementation wurde so ausgelegt, dass die entwickelten Algorithmen und Abläufe, im Sinne einer Toolbox, zur Bestimmung von Ausbreitungsmustern auf Grundlage binärer Bilddaten genutzt werden können. Somit konnte ein Beitrag zur Musteranalyse von 2D-Bilddaten geleistet werden.

Die Validierung der eingesetzten Analysen und Merkmale gestaltete sich aufgrund der fehlenden Ground-Truth schwierig. Da der Zugriff auf die Domänenexperten nur noch begrenzt möglich war und sich die Interpretation der numerischen Werte durch den Menschen über mehrere Fälle hinweg als unintuitiv und zeitaufwändig erwiesen hatte, wurde ein mehrstufiger Validierungsprozess entwickelt (siehe Abschnitt 5.6.1).

Hierfür ist ein Konzept auf Grundlage einer Clusteranalyse unter Verwendung des *K-Means*-Verfahrens verwendet worden. Eingesetzt wurden der Hopkins-Index, eine Visualisierung des Merkmalsraumes durch die PCA, die Berechnung der Silhouetten-Werte und die Berechnung der normalisierten Variation Of Information (VI_n). Teil des Konzeptes war die Entwicklung eines Algorithmus zur automatischen Erzeugung einfacher Bilddaten mit Ground-Truth, welche einen ähnlichen Problembereich abbilden, wie die gesuchten Eigenschaften auf den realen Bilddaten.

Zunächst wurden das Konzept und die Merkmale anhand synthetischer Daten getestet (siehe Abschnitt 6.1). Für die Validierung des Ansatzes zur Bestimmung morphometrischer Eigenschaften wurde eine öffentlich zugängliche Form-Datenbank eingesetzt. Unter Verwendung der synthetischen Daten konnte für die morphometrischen, die strukturellen und die relationalen Analysen VI_n -Werte unter 0.15 erreicht werden, was für eine hohe Übereinstimmung mit den Klassenzugehörigkeiten und somit den gesuchten Eigenschaften spricht. Auf den synthetischen Daten konnte die korrekte Funktionsweise der implementierten und entwickelten Analysen dargestellt werden.

Die entwickelten Analysen wurden für die realen Daten zunächst anhand von Tumorprototypen getestet (siehe Kapitel 5). Die morphometrischen Analysen trennen die gesuchten Eigenschaften auf der Detailstufe von f_{bin}^{ae1ae3} nicht ausreichend. Gründe hierfür wurden in der Diskussion (siehe Kapitel 7) erläutert. Die Tumorprototypen konnten durch die strukturellen Analysen auf Grundlage der Kantenlängen getrennt werden. Die Merkmale auf Basis des Grades zeigten wenig Differenzierungspotenzial. Sowohl die LDQ-Merkmale, als auch das RCC8-Histogramm zeigten bei den Tumorprototypen plausible Werte und waren in der Lage diese zu differenzieren. Für diese ausgewählten Fälle konnten die strukturellen und relationalen Anforderungen erfüllt werden.

Die Clusteranalysen auf den realen Daten ergaben für die morphometrischen Analysen von f_{bin}^{ae1ae3} im morphometrischen Clustering nur geringfügig kompakte Cluster. Die Stichprobenanalysen des Domänenexperten zeigten, dass maximal 6 von 10 Fällen plausibel sind.

Die strukturellen Analysen lassen erkennen, dass die Verwendung von Gradbasierten Merkmalen grundsätzlich zu weniger gut strukturierten Clustern führen. Der DG erzielt zusammenfassend etwas bessere Ergebnisse als der MST. Die Stichprobenanalysen zeigten beim DG eine höhere Übereinstimmung.

Die relationalen Analysen ergaben unter Verwendung von f_{bin}^{cd45} (CD45) nur mäßige Clusterergebnisse. Die Stichprobenanalysen zeigten keine Übereinstimmung mit den Beobachtungen des Domänenexperten. Für die relationalen Analysen auf f_{bin}^{er} (ER) wurden bessere Ergebnisse erzielt. Die Stichprobenanalysen zeigten eine hohe Übereinstimmung mit den Beobachtungen des Domänenexperten.

Des Weiteren wurden die Merkmale unter Verwendung des HLVQ-Verfahrens getestet (siehe Abschnitt 6.3). Somit sollte der Einfluss der Merkmale auf die Trainingsdaten in Bezug auf eine zufällige Verteilung der Verlaufsdaten untersucht werden. Alle getesteten Merkmalskombinationen zeigten, dass die Klassifikationsrate durch das Lernen erhöht wurde, was grundsätzlich für eine Relevanz der Merkmale in Bezug auf den *Follow-Up-Status* spricht.

Bei ausgewählten näheren Untersuchungen der Genauigkeit und Trefferquote, bezüglich der einzelnen Klassen, wurden Einschränkungen dieser Beobachtung deutlich. Aufgrund des sehr unbalancierten Datensatzes erhielt Klasse *relapse* (7 Fälle) nur in wenigen Fällen eine hohe Trefferquote. Auf den Testdaten zeichnete sich bei keinem Testfall eine eindeutige Tendenz ab. Die Klassifikationsrate blieb bei allen einzeln getesteten Testfällen unter 58%. Eine Generalisierung konnte nicht beobachtet werden.

Die Merkmalskombinationen mit den aussichtsreichsten Ergebnissen wurden mit den klinischen Daten kombiniert eingesetzt, um zu testen, inwieweit unter diesen Voraussetzungen eine Verbesserung der Klassifikation erreicht wird. Hierzu wurden verschiedene Kombinationstests durchgeführt. Die Klassifikationsrate auf den Testdaten konnte nicht verbessert werden. Die höchste Klassifikationsrate der Kombinationstests lag 2% unter den besten Ergebnissen der getesteten klinischen Daten. Auf den Trainingsdaten konnte die Klassifikationsrate um ca. 3% erhöht werden. Die Genauigkeit konnte für Klasse *relapse* um ca. 12% und für die Klasse *deceased* um ca. 5% gesteigert werden. Für Klasse *healthy* wurde die Genauigkeit um ca. 2% schlechter. Auf den Testdaten wurde die Genauigkeit in allen Klassen um ca. 5% schlechter. Die Trefferquote verbesserte sich lediglich für die Klasse *healthy* auf den Trainings- und Testdaten. Insgesamt ließ sich beobachten, dass die eingesetzten Merkmalskombinationen eine geringfügig verbesserte Klassifikationsleistung auf den Trainingsdaten erreichten. Auf den Testdaten konnte keine Verbesserung beobachtet werden.

Die Ergebnisse und Ansätze der vorliegenden Arbeit wurden in Kapitel 7 diskutiert. Als Schwierigkeit der eingesetzten Bildanalysen und Bildmerkmale wurde die Grundlage der zu wenig differenzierten Grenzen der Bildobjekte durch die Connected-Components auf Gewebeebene identifiziert. Diese Einschränkungen verringern die Aussagekraft des

RCC8-Histogramms und der morphometrischen Analysen. Die strukturellen Analysen auf Grundlage von Graphen und die LDQ-Merkmale werden dadurch weniger negativ beeinflusst, welche aus diesen Gründen als für die Gewebeebene angemessener angesehen werden.

8.2 Ausblick

Der morphometrische Ansatz sollte auf Gewebeebene lediglich für die Tumorobjekte aus den SOM-Analysen (f_{bin}^{tumor}) oder auf einer feineren Auflösungsstufe der Gewebeschnitte eingesetzt werden. In diesem Zusammenhang kann eine unüberwachte Merkmalsselektion durchgeführt werden, um wenig relevante Merkmale zu identifizieren, welche die Differenzierung negativ beeinflussen. In Frage kommt hierfür der B2-Algorithmus [Jol72, Jol73], der Merkmale mit geringer Varianz entfernt.

Um die relationalen Merkmale besser human beurteilen zu können, müssen vor allem für die RCC8- und LDQ-Merkmale geeignete Visualisierungen gefunden werden. So könnte jedes Tumorobjekt bzw. Pixel einer Bildmodalität, im Hinblick auf seine Relationen farblich markiert werden.

Die RCC8-Merkmale können zusätzlich dafür eingesetzt werden, die Relationen zwischen Fettgewebe bzw. Stromagewebe und Tumorgewebe zu quantifizieren. Sowohl die RCC8-Merkmale, als auch die LDQ-Merkmale sollten durch einen Fuzzy-Ansatz erweitert werden, welcher jedem Pixel eine gewisse Wahrscheinlichkeit für seine Relation angibt. Für das Region-Connection-Calculus 8 wurde dies bereits in [ASI09] als theoretisches Konzept vorgeschlagen. Somit könnte man Unterschiede und Abweichungen in Größe und Form der Connected-Components sowie Fehler in der Registrierung sinnvoll ausgleichen.

Für die strukturellen Analysen, basierend auf Graphen, können, wie in [Bou08] vorgeschlagen, einzelne Subregionen bestimmt werden. Entsprechend dieses Ansatzes wurde im Rahmen der vorliegenden Arbeit ein erstes Konzept implementiert. Nach diesem Konzept wird das *Voronoi-Diagramm* zu den gegebenen Knoten berechnet. Aufbauend darauf werden die Knoten als Subregionen bestimmt, deren *Voronoizellen* sich berühren, was zu einer plausiblen räumlichen Nachbarschaft führt. Daraufhin werden die Graphen und Graphmerkmale für jede Subregion berechnet. Für jede Subregion können so in Form eines Clusterings (siehe Abbildung C.6), gemeinsame Eigenschaften zusammengefasst werden und durch eine Rückprojektion visualisiert werden (ähnlich den morphometrischen Analysen). Dies ermöglicht eine differenziertere Analyse und Visualisierung einzelner Tumorbereiche bezüglich ihrer Ausbreitungsmuster.

Zusätzlich kann über Graphen versucht werden relationale Verteilungen der funktionalen Marker in Bezug zum Tumorgewebe, in Form eines *Color-Graph's* [ACSGD10], zu quantifizieren.

Als Verfahren für die Clusteranalyse kann zusätzlich ein hierarchisches Clusterverfahren [XW05] verwendet werden, um mit Hilfe eines Dendogrammes [XW05] ähnliche Fälle zu visualisieren. Als Werkzeug für den Domänenexperten sollte, neben einer Clusteranalyse, ein Image-Retrieval [MMBG04] eingesetzt werden, anhand dessen der Domänenex-

perre auf Grundlage der Merkmale nach ähnlichen Beispielen suchen kann. Entscheidend für das Explorieren der Merkmals- und Bilddaten sind in jedem Fall grafische Evaluationsumgebungen, welche dem Domänenexperten das Nachvollziehen der Berechnungen und Zusammenhänge erleichtern und plausibel machen. Dies kann dazu beitragen die Semantische Lücke zu verringern.

Um eine allgemeine Aussage bezüglich der Relevanz der Merkmale und quantifizierten Eigenschaften zu ermöglichen, muss für die Clusteranalysen und überwachten Analysen eine größerer Patientenfallzahl mit entsprechenden Bilddaten verwendet werden. Entscheidend für den weiteren Einsatz der entwickelten Merkmale ist zudem eine Validierung durch verschiedene Domänenexperten.

Sollten sich die eingesetzten Bildanalysen zur Merkmalsextraktion, unter Berücksichtigung der genannten Einschränkungen, über mehr Testfälle als valide und für den Domänenexperten plausibel darstellen, können die Merkmale dazu beitragen, die unterschiedlichen Eigenschaften des Mammakarzinoms besser zu verstehen, um somit eine individuellere Therapie für die Patienten zu ermöglichen.

Anhang A

Symbolverzeichnis

Symbol	Bedeutung
Ω_k	Klasse
ω	Klassenzugehörigkeit
M	Merkmal
\mathbf{m}	Merkmalsvektor
$\tilde{\mathbf{m}}$	Normalisierter Merkmalsvektor
\mathbf{m}^i	Komponente eines Merkmalsvektors
m	Einzelner Merkmalswert
C_k	Cluster
K	Clusteranzahl
VI_n	Normalisierte Variation of Information
SV	Durchschnittlicher Silhouetten-Wert
f_{ζ}^{ϱ}	Bildfunktion bzw. Bildmuster der Modalität ϱ und vom Typ ζ mit Pixelkoordinaten (x,y)
f_{bin}^{ae1ae3}	Feine Tumorstrukturen (binarisiertes AE1AE3)
f_{bin}^{tumor}	Grobe Tumorstrukturen (Tumorgewebe aus den SOM-Analysen)
f_{bin}^{cd45}	Entzündetes Gewebe (binarisierte CD45-Färbung)
f_{bin}^{er}	Expression von Östrogen (binarisierte ER-Färbung)
f_{bin}^{pr}	Expression von Progesteron (binarisierte PR-Färbung)
R^{ϱ}	Connected-Component, Ein Bildobjekt der Modalität ϱ mit Pixelkoordinaten (x,y)
\mathbf{R}^{ϱ}	Vektor mit Bildobjekten R^{ϱ}
R^{tumor}, R^{ae1ae3}	Tumorobjekte
R^{cd45}, R^{er}, R^{pr}	Relationsobjekte R^{rel} (verwendet zur Bestimmung von Relation zwischen Tumorobjekten R^{tumor} und Relationsobjekten im RCC8-Histogramm)

$G(V, E)$	Ein Graph G mit einer Menge von V Knoten v und E Kanten e .
MST	Minimum-Spanning-Tree
DG	Delaunay-Graph
A	Fläche
DC	disconnected (RCC8-Relation)
EC	equally connected (RCC8-Relation)
PO	partial overlapping (RCC8-Relation)
TPP	tangential proper part (a in b) (RCC8-Relation)
TPPI	tangential proper part inverse (b in a) (RCC8-Relation)
NTPP	nontangential proper part (a in b) (RCC8-Relation)
NTPPI	nontangential proper part inverse (b in a) (RCC8-Relation)
EQ	equal (RCC8-Relation)

Anhang B

Merkmalsauswahl für Testfälle

Die Tabellen beschreiben die eingesetzten Merkmalskonfigurationen für die Experimente aus Kapitel 6.

Testfall	N_v	e_μ	e_σ	e_{mima}	e_{vc}	C_n	deg_μ	deg_σ	deg_{mima}	deg_{vc}	deg_w	RI
DGTF1	-	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DGTF2	✓	✓	✓	-	-	-	-	-	-	-	-	-
DGTF3	-	✓	✓	✓	✓	-	-	-	-	-	-	-

Tabelle B.1: Delaunay-Graph: eingesetzte Merkmale (markiert durch ✓) in dokumentierten Testfällen (TF).

Testfall	N_v	e_μ	e_σ	e_{mima}	e_{vc}	deg_μ	deg_σ	deg_{mima}	deg_{vc}	deg_w	RI
MSTTF1	-	✓	✓	✓	✓	-	-	-	-	-	-
MSTTF2	-	✓	✓	✓	✓	-	-	-	-	✓	-
MSTTF3	✓	✓	✓	✓	✓	-	-	-	-	-	-
MSTTF4	✓	✓	✓	-	-	-	-	-	-	-	-

Tabelle B.2: Minimum-Spanning-Tree: eingesetzte Merkmale (markiert durch ✓) in dokumentierten Testfällen (TF).

Testfall	LDQ_{amount}	LDQ_z	LDQ_p	DC	EC	PO	TPP	TPPI	NTPP	NTPPI	EQ
RTF1	-	-	-	✓	✓	✓	✓	✓	✓	✓	✓
RTF2	✓	✓	✓	-	-	-	-	-	-	-	-
RTF3	-	-	-	✓	✓	✓	✓	✓	✓	✓	-
RTF4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Tabelle B.3: Relationale Merkmale: eingesetzte Merkmale (markiert durch ✓) in dokumentierten Testfällen (TF).

Merkmale	MTF1	MTF2	MTF3
<i>A</i>	-	-	✓
<i>P</i>	-	-	✓
<i>CI</i>	✓	✓	✓
<i>CO</i>	✓	✓	✓
<i>IR</i>	✓	✓	✓
<i>fd</i> ₁	-	✓	✓
<i>fd</i> ₂	-	✓	✓
<i>fd</i> ₃	-	✓	✓
<i>fd</i> ₄	-	✓	✓
<i>fd</i> ₅	-	✓	✓
<i>fd</i> ₆	-	✓	✓
<i>fd</i> ₇	-	✓	✓
<i>fd</i> ₈	-	✓	✓
<i>fd</i> ₉	-	✓	✓
<i>fd</i> ₁₀	-	✓	✓
<i>fd</i> ₁₁	-	✓	✓
<i>fd</i> ₁₂	-	✓	✓
<i>fd</i> ₁₃	-	✓	✓
<i>fd</i> ₁₄	-	✓	✓
<i>fd</i> ₁₅	-	✓	✓
<i>feLow</i>	-	-	✓
<i>feMid</i>	-	-	✓
<i>feHigh</i>	-	-	✓
<i>H1</i>	-	✓	✓
<i>H2</i>	-	✓	✓
<i>H3</i>	-	✓	✓
<i>H4</i>	-	✓	✓
<i>H5</i>	-	✓	✓
<i>H6</i>	-	✓	✓
<i>H7</i>	-	✓	✓

Tabelle B.4: Morphometrische Analysen: eingesetzte Merkmale (markiert durch ✓) in dokumentierten Testfällen (TF).

Anhang C

Bilder

Eine Auswahl von Ergebnisbildern unterschiedlicher Bildanalysen sowie ein Beispiel für einen vorgesehenen Evaluationsbogen für den Domänenexperten.

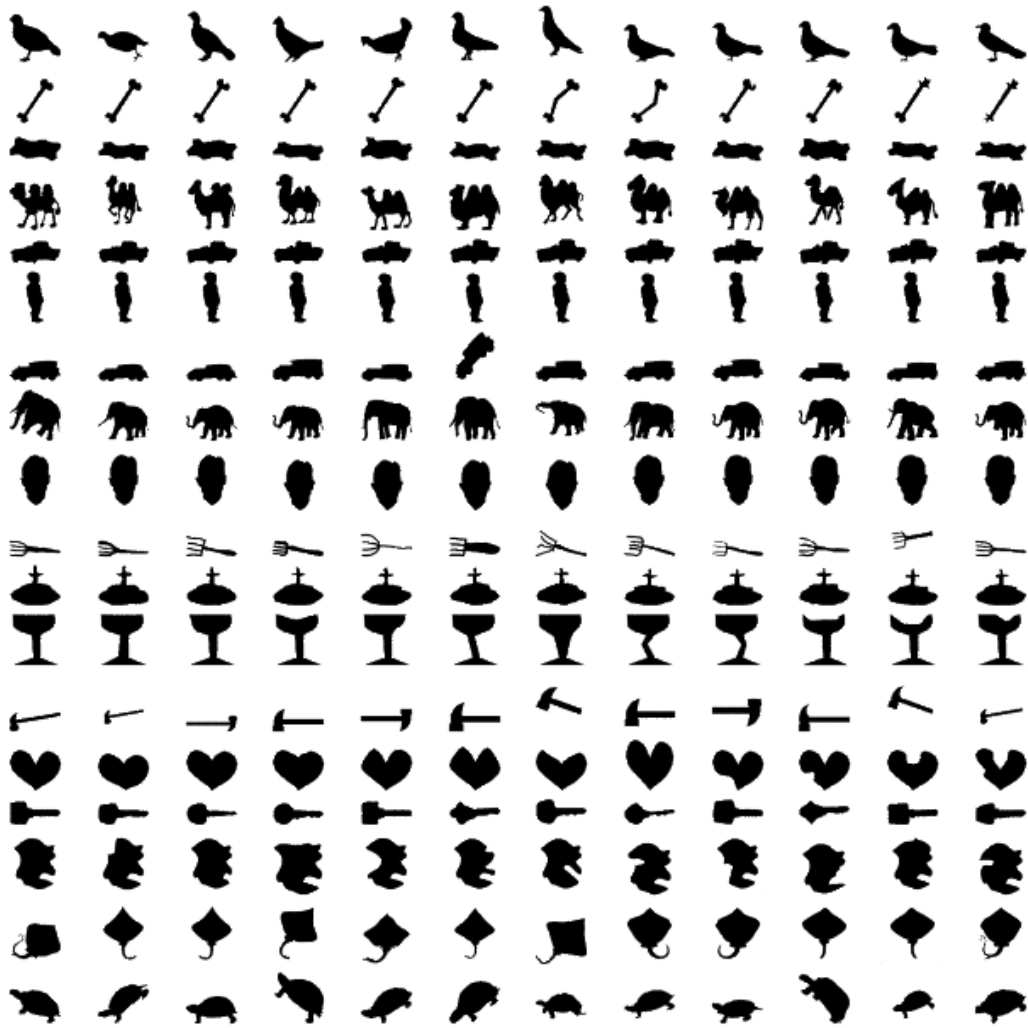


Abbildung C.1: Für morphometrische Tests eingesetzte Form-Datenbank [SKK01].

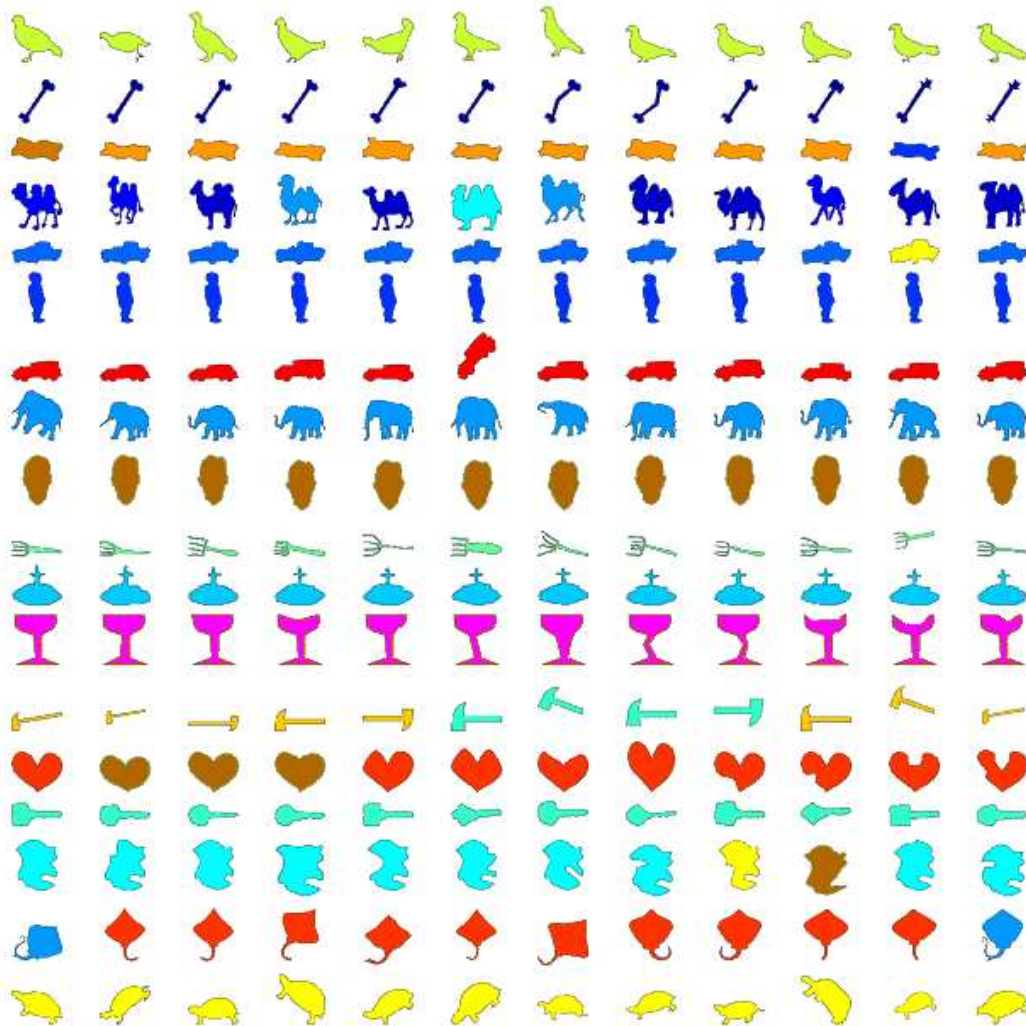
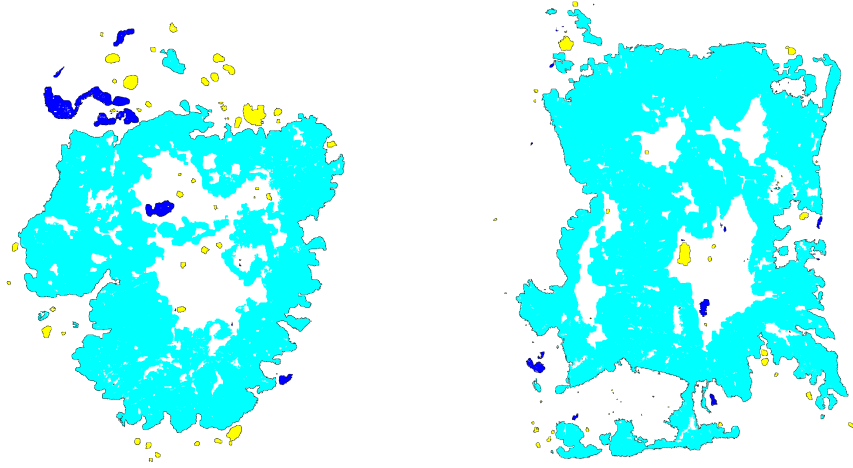
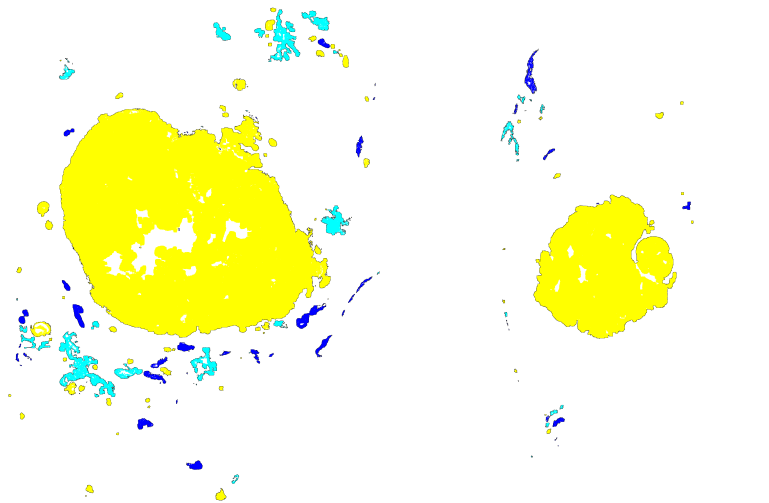


Abbildung C.2: Morphometrisches Clusterergebnis bei Clusteranzahl $K = 18$ auf der Form-Datenbank [SKK01]. Für diese Clusteranzahl wurde eine VI_n von ca. 0.1 erreicht. (Anmerkung: Das Blau in der fünften Zeile unterscheidet sich von dem Blau der sechsten Zeile. Beide Gruppen wurden korrekt getrennt.)



(a) zentrale Nekrose, unregelmäßiger Rand (b) zentrale Nekrose, unregelmäßiger Rand



(c) kompakt, glatter Rand

(d) kompakt, glatter Rand

Abbildung C.3: Beispielergebnisse des morphometrischen Clusterings bei Clusteranzahl 3 von f_{bin}^{tumor} : Das gelbe Cluster repräsentiert kompakte Tumorobjekte mit glatterm Rand. Das cyanfarbene Cluster entspricht Tumorobjekten mit zentraler Nekrose und einem unregelmäßigen Rand.

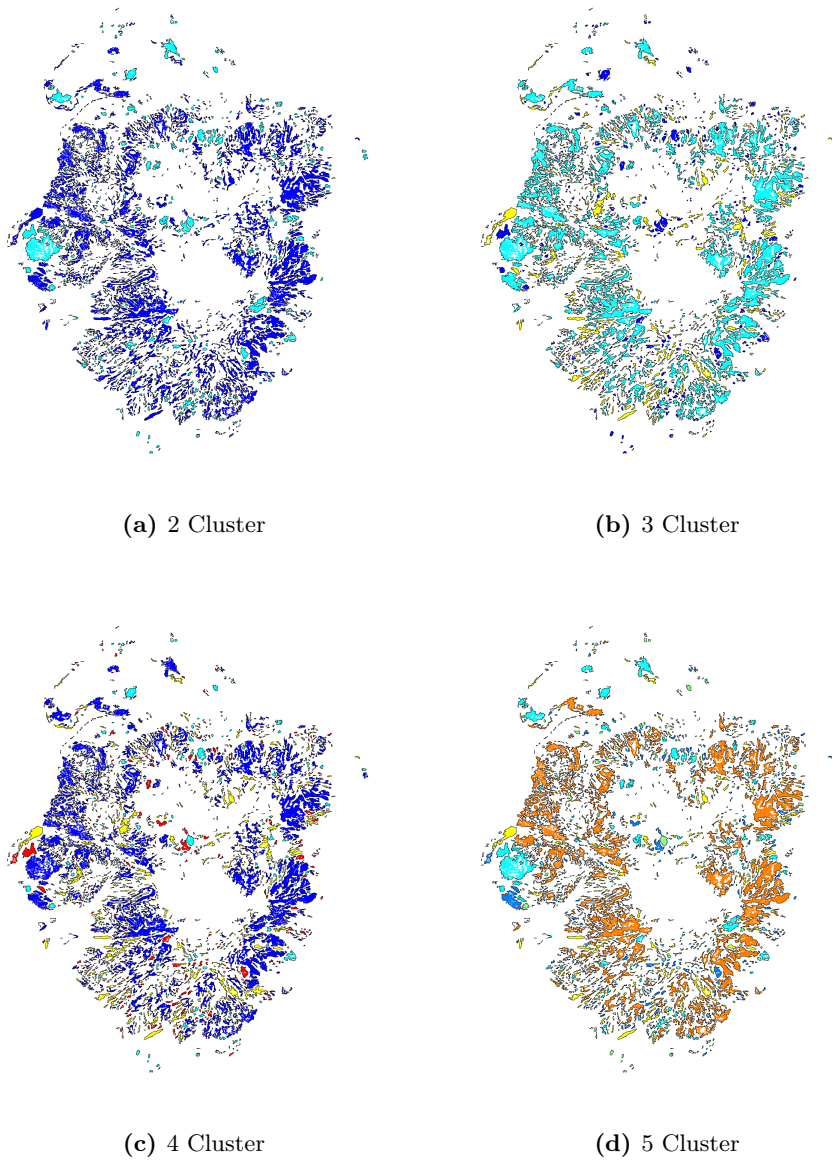


Abbildung C.4: Beispielergebnisse des morphometrischen Clusterings bei Clusteranzahl 2 bis 5 von f_{bin}^{ae1ae3} .

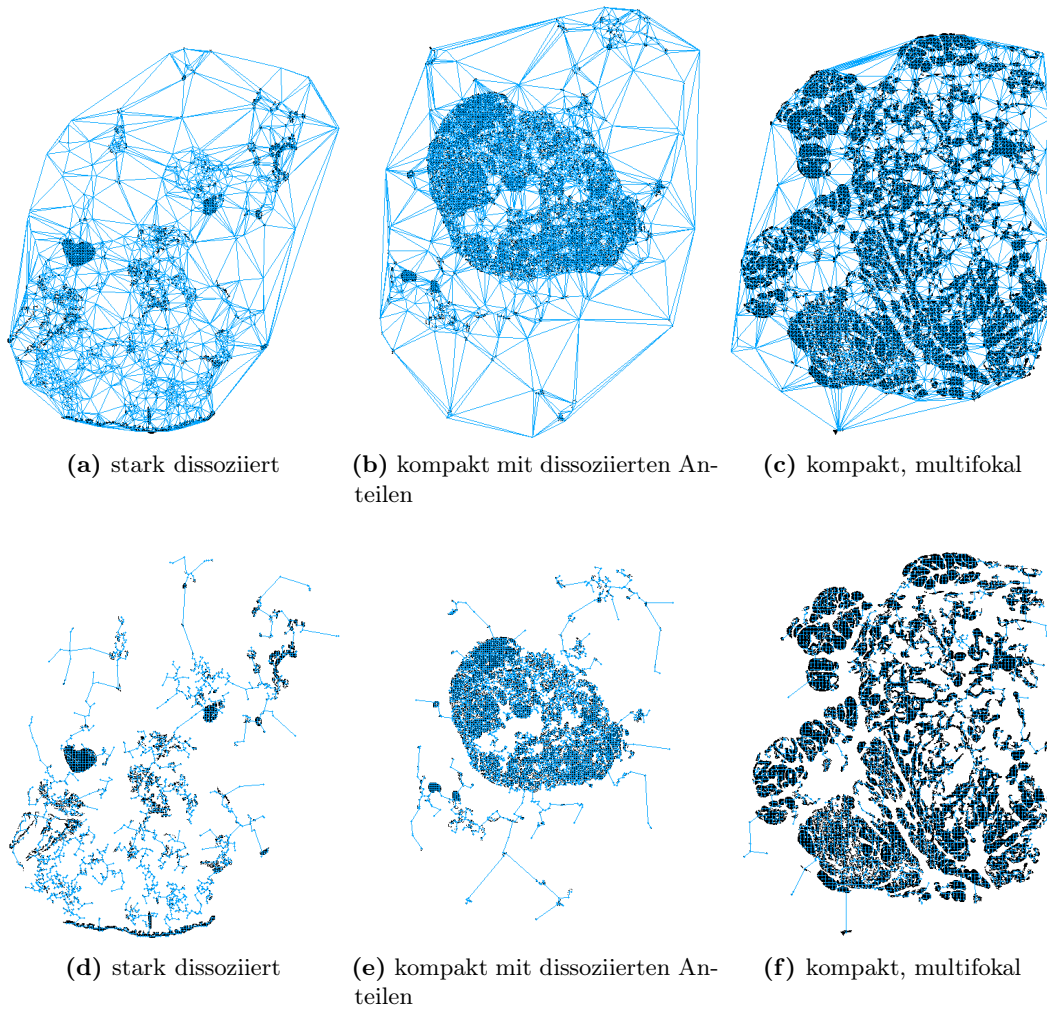
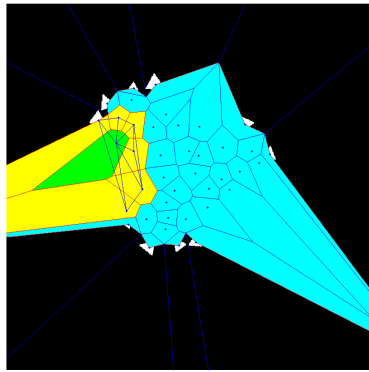
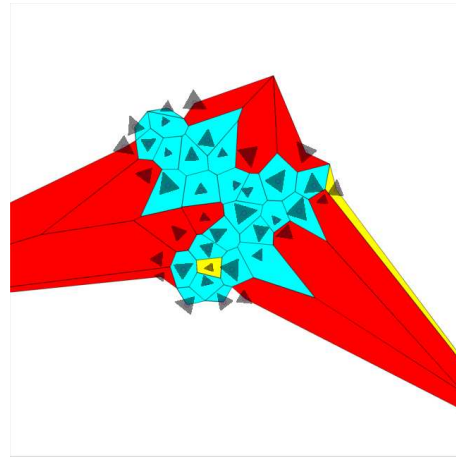


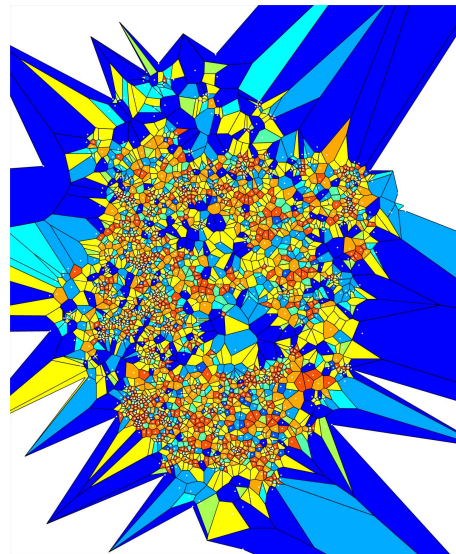
Abbildung C.5: Beispielbilder zu erstellten Graphen von f_{bin}^{ae1ae3} : (a) - (c) Delaunay-Graph, (d) - (f) Minimum-Spanning-Tree.



(a) Subgraph (DG)

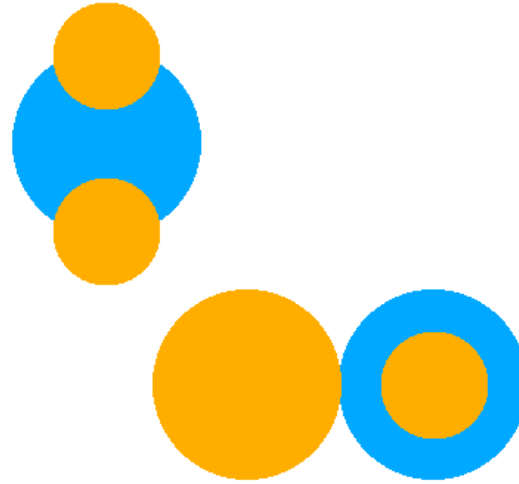


(b) Ergebnis (3 Cluster)

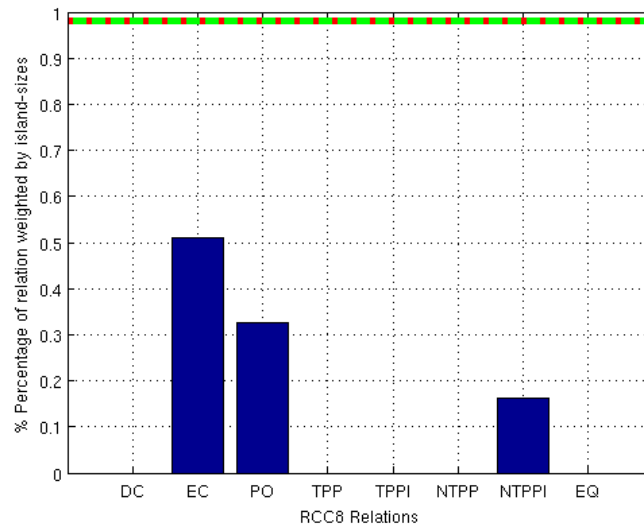
(c) f_{bin}^{ae1ae3} 

(d) Ergebnis (9 Cluster)

Abbildung C.6: Beispielbilder des Clustering der Graphmerkmale (unter Verwendung von Subregionen) auf synthetischen Daten: Delaunay-Subgraph (blau) innerhalb Voronoi-Nachbarzellen (gelb) (a), Rückprojektion des Clusterergebnisses (3 Cluster) (b); und realer Daten am Beispiel von f_{bin}^{ae1ae3} (c) und die Rückprojektion mit 9 Clustern (d).



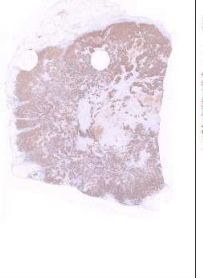
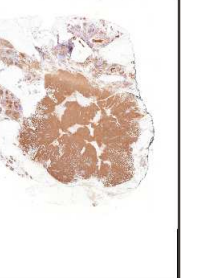


(a) Bildobjekte


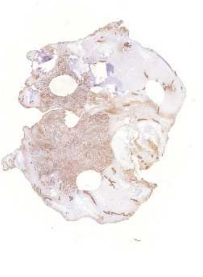
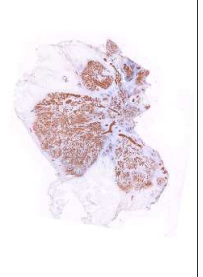
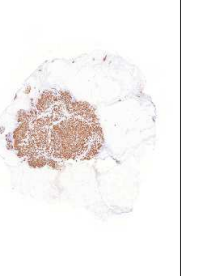


(b) RCC8-Histogramm





Abbildung C.7: Synthetisches Beispiel des erweiterten RCC8-Histogramms: orange Relationsobjekte R^{rel} und blau Tumorobjekte R^{tumor} (a); Ergebnis des erweiterten RCC8-Histogramms (b).

Delaunay-Graph (Bestes Clustering: meanDist, stdDist, nNodes)				
Cluster 1	Am nächsten zu Prototyp	1	2	3
Patienten-ID	043-18071	022-09100	022-23125	014-00815
Tumor				
Silhouetten-Wert	0.8463	0.8479	0.8357	0.8303
Bewertung				
Bemerkung				

(a) Cluster 1

Cluster 2	Am nächsten zu Prototyp	1	2	3
Patienten-ID	013-13832	054-00865	014-17283	024-19369
Tumor				
Silhouetten-Wert	0.8523	0.8493	0.8479	0.8462
Bewertung				
Bemerkung				

(b) Cluster 2

Cluster 3	Am nächsten zu Prototyp	1	2	3
Patienten-ID	022-22356	031-07992	021-21449	013-18006
Tumor				
Silhouetten-Wert	0.7983	0.7285	0.7285	0.7018
Bewertung				
Bemerkung				

(c) Cluster 3

Abbildung C.8: Beispiel für einen Evaluationsbogen mit Stichproben: Clusterergebnis des Delaunay-Graphen mit Merkmalen N_v , e_μ und e_σ und Clusteranzahl 3. Für die Bewertung wurden die Kategorien „plausibel“, „teilweise plausibel“ und „nicht plausibel“ vorgesehen.

Anhang D

PCA

Erläuterungen zu den Berechnungsschritten der PCA anhand von Abbildung D.1 [Shl05, Han09]:

1. Von jedem Merkmalsvektor wird der Mittelwert μ der gegebenen Merkmalsvektoren abgezogen, um die Daten zu zentrieren (Translation in Koordinatenursprung, siehe Abbildung D.1 (b)).
2. Die Kovarianzmatrix \mathbf{C} der Merkmalsvektoren wird bestimmt.
3. Eigenvektoren γ und Eigenwerte λ von \mathbf{C} werden bestimmt. Die Eigenvektoren werden in der Ergebnismatrix $\mathbf{\Gamma}$ normiert angegeben. Die korrespondierenden Eigenwerte werden als $\lambda_1 > \dots > \lambda_n$ notiert. Sie geben die Achsenvarianz der korrespondierenden Eigenvektoren an. Die Eigenvektoren entsprechen den gesuchten Hauptkomponenten.
4. Multiplikation von $\mathbf{\Gamma}$ mit Merkmalsvektoren (Rotation auf Achsenvarianzen, siehe Abbildung D.1 (c)).
5. Durch Entfernen von Eigenvektoren mit niedrigen Eigenwerten (Varianzen) in $\mathbf{\Gamma}$ zu $\mathbf{\Gamma}'$ können die Merkmalsvektoren in einen Datenraum mit geringerer Dimension projiziert werden (siehe Abbildung D.1 (e)).
6. Durch Anwendung von $\mathbf{\Gamma}'^{-1}$ auf die transformierten Merkmalsvektoren werden diese wieder in das ursprüngliche Koordinatensystem rotiert (siehe Abbildung D.1 (f)).
7. Die Addition des Mittelwertes μ transliert die Merkmalsvektoren wieder an ihre ursprüngliche Position (siehe Abbildung D.1 (g)).

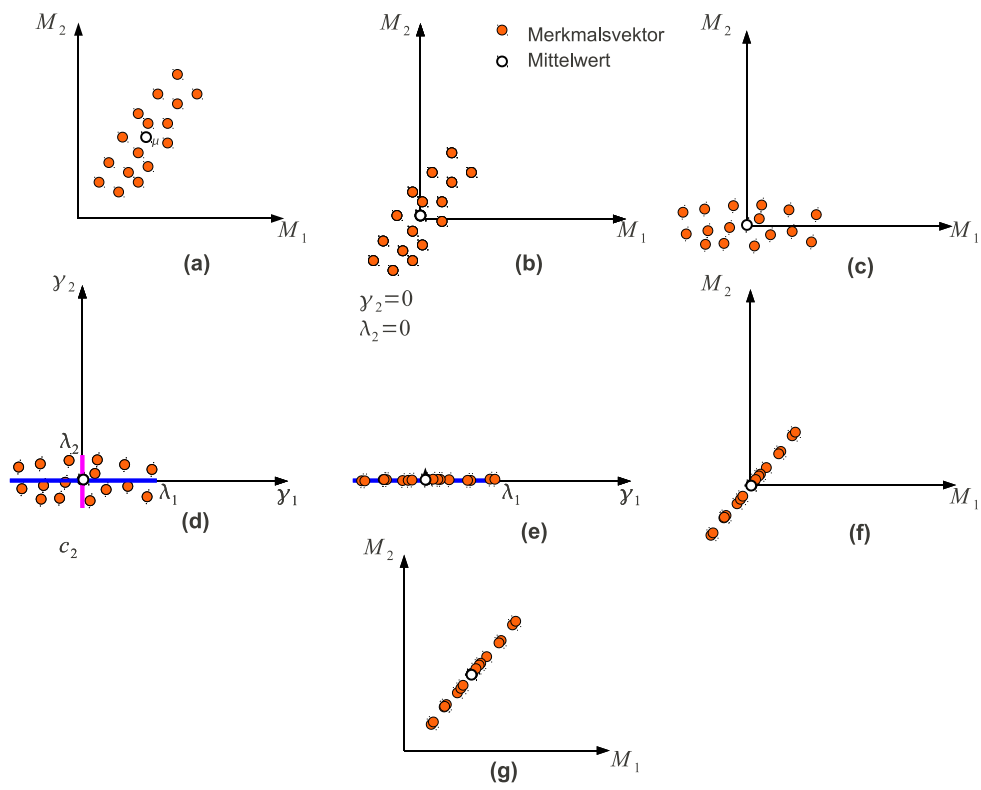


Abbildung D.1: Dimensionsreduktion eines 2D-Merkmalraumes auf eine Dimension durch die PCA.

Anhang E

DVD-ROM

Inhalt der beiliegenden DVD-ROM:

Verzeichnis	Inhalt
<i>Ausarbeitung/</i>	Latex-Quelltext und Bilder
<i>Programmcode/</i>	Implementierter Quellcode der entwickelten Algorithmen, Merkmale und Abläufe sowie angewendete Testkonfigurationen
<i>Literatur/</i>	Referenzierte Literatur sofern in digitaler Form vorliegend
<i>Sonstiges/</i>	
<i>Annotationen/</i>	Annotationen des Domänenexperten
<i>BerechneteMerkmale/</i>	CSV-Dateien mit berechneten Merkmalen für die synthetischen und realen Bilddaten
<i>RealeDaten/</i>	Eine Auswahl verwendeter realer Bilddaten für Testläufe
<i>SynthetischeTestdaten/</i>	Verwendete synthetische Datenbasis und Form-Datenbank für Testläufe

Literaturverzeichnis

- [ACSGD10] ALTUNBAY, Dogan ; CIGIR, Celal ; SOKMENSUER, Cenk ; GUNDUZ-DEMIR, Cigdem: Color graphs for automated cancer diagnosis and grading. In: *IEEE transactions on bio-medical engineering* 57 (2010), März, Nr. 3, 665–74. <http://dx.doi.org/10.1109/TBME.2009.2033804>. – DOI 10.1109/TBME.2009.2033804. – ISSN 1558–2531
- [AH00] AKSOY, Selim ; HARALICK, Robert M.: Effects of Feature Normalization on Image Retrieval. In: *Database* (2000), S. 2–5
- [al.05] AL., J. S. M.: Breast carcinoma malignancy grading by bloom-richardson system vs proliferation index: reproducibility of grade and advantages proliferation index. In: *Modern Pathology* 18 (2005), Nr. 8, S. 1067–1078
- [Ame11] *Learn about cancer - Breast Cancer*. Version: 2011. <http://www.cancer.org/Cancer/BreastCancer/DetailedGuide/breast-cancer-pdf12>, Ab-
ruf: 02. September. 2011
- [ASI09] ALBOODY, Ahed ; SEDES, Florence ; INGLADA, Jordi: Fuzzy Intersection and Difference Model for Topological Relations. In: *Work* (2009), S. 1079–1084
- [BB05] BURGER, Wilhelm ; BURGE, Mark J.: *Digitale Bildverarbeitung: Eine Einführung mit Java und ImageJ (eXamen.press)*. 1. Springer, 2005
- [BD04] BANERJEE, a. ; DAVE, R.N.: Validating clusters using the Hopkins statistic. In: *2004 IEEE International Conference on Fuzzy Systems (IEEE Cat. No.04CH37542)* (2004), 149–153. <http://dx.doi.org/10.1109/FUZZY.2004.1375706>. – DOI 10.1109/FUZZY.2004.1375706. ISBN 0–7803–8353–2
- [BDH96] BARBER, C. B. ; DOBKIN, David P. ; HUHDANPAA, Hannu: The Quickhull algorithm for convex hulls. In: *ACM TRANSACTIONS ON MATHEMATICAL SOFTWARE* 22 (1996), Nr. 4, S. 469–483
- [BDNY07] BILGIN, Cagatay ; DEMIR, Cigdem ; NAGI, Chandandeep ; YENER, Bulent: Cell-graph mining for breast tissue modeling and classification. In: *Confe-*

rence proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference 2007 (2007), Januar, 5311–4. <http://dx.doi.org/10.1109/IEMBS.2007.4353540>. – DOI 10.1109/IEMBS.2007.4353540. – ISSN 1557–170X

- [BEH⁺06] BRAUMANN, Ulf-dietrich ; EINENKEL, Jens ; HORN, Lars-christian ; KUSKA, Jens-peer ; SCHERF, Nico ; WENTZENSEN, Nicolas: Registration of Histologic Colour Images of Different Staining. In: *IEEE Transactions on Medical Imaging* (2006)
- [Bis06] BISHOP, Christopher M.: *Pattern Recognition and Machine Learning*. 1. Springer, 2006. – 703 S.
- [Blo05] BLOCH, I: Fuzzy spatial relationships for image processing and interpretation: a review. In: *Image and Vision Computing* 23 (2005), Februar, Nr. 2, 89–110. <http://dx.doi.org/10.1016/j.imavis.2004.06.013>. – DOI 10.1016/j.imavis.2004.06.013. – ISSN 02628856
- [Bou08] BOUCHERON, Laura E.: *Object- and Spatial-Level Quantitative Analysis of Multispectral Histopathology Images for Detection and Characterization of Cancer*, UNIVERSITY of CALIFORNIA Santa Barbara, Diss., 2008
- [DAB⁺04] DIAMOND, James ; ANDERSON, Neil H. ; BARTELS, Peter H. ; MONTIRONI, Rodolfo ; HAMILTON, Peter W.: The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia. In: *Human Pathology* 35 (2004), September, Nr. 9, 1121–1131. <http://dx.doi.org/10.1016/j.humpath.2004.05.010>. – DOI 10.1016/j.humpath.2004.05.010. – ISSN 00468177
- [DAM⁺08] DOYLE, Scott ; AGNER, Shannon ; MADABHUSHI, Anant ; FELDMAN, Michael ; TOMASZEWSKI, John: Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features. In: *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (2008), Mai, 496–499. <http://dx.doi.org/10.1109/ISBI.2008.4541041>. – DOI 10.1109/ISBI.2008.4541041. ISBN 978–1–4244–2002–5
- [DGW] *Delaunay-Triangulation*. http://en.wikipedia.org/wiki/File:Delaunay_circumcircles_centers.png, Abruf: 03. Dezember. 2011
- [DHS⁺07] DOYLE, Scott ; HWANG, Mark ; SHAH, Kinsuk ; MADABHUSHI, Anant ; FELDMAN, Michael ; TOMASZEWSKI, John: Automated Grading of Prostate Cancer Using Architectural and Textural Image Features. In: *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (2007), April, 1284–1287. <http://dx.doi.org/10.1109/ISBI.2007.357094>. – DOI 10.1109/ISBI.2007.357094. ISBN 1–4244–0671–4

- [Die10] DIESTEL, Reinhard: *Graphentheorie*. 4. Springer Verlag, 2010 <http://diestel-graph-theory.com/basic.html>
- [DY05] DEMIR, Cigdem ; YENER, Bülent: *Automated cancer diagnosis based on histopathological images: a systematic survey*. 2005
- [EMT97] ENDERWICK, C.Y. ; MICHELI-TZANAKOU, E.: Classification of mammographic tissue using shape and texture features. In: *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 'Magnificent Milestones and Emerging Opportunities in Medical Engineering'* (Cat. No.97CH36136) 2 (1997), Nr. C, 810–813. <http://dx.doi.org/10.1109/IEMBS.1997.757772>. – DOI 10.1109/IEMBS.1997.757772. ISBN 0–7803–4262–3
- [EP02] ERK, Katrin ; PRIESE, Lutz: *Theoretische Informatik*. 2. Springer Verlag, 2002. – 467 Seiten, 115 Abbildungen
- [Fer00] FERNÁNDEZ PIERNA, J: Improved algorithm for clustering tendency. In: *Analytica Chimica Acta* 408 (2000), März, Nr. 1-2, 13–20. [http://dx.doi.org/10.1016/S0003-2670\(99\)00879-X](http://dx.doi.org/10.1016/S0003-2670(99)00879-X). – DOI 10.1016/S0003-2670(99)00879-X. – ISSN 00032670
- [FSZ09] FLUSSER, Jan ; SUK, Tom ; ZITOV, Barbara: *Moments and Moment Invariants in Pattern Recognition*. Chichester, UK : John Wiley & Sons, Ltd, 2009. – 312 S. <http://dx.doi.org/10.1002/9780470684757>. <http://dx.doi.org/10.1002/9780470684757>. – ISBN 9780470684757
- [GBC⁺09] GURCAN, Metin N. ; BOUCHERON, Laura ; CAN, Ali ; MADABHUSHI, Anant ; RAJPOOT, Nasir ; YENER, Bulent: Histopathological Image Analysis: A Review. In: *IEEE reviews in biomedical engineering* 2 (2009), Januar, 147–171. <http://dx.doi.org/10.1109/RBME.2009.2034865>. – DOI 10.1109/RBME.2009.2034865. – ISSN 1941–1189
- [GLB⁺10] GATZA, Michael L. ; LUCAS, Joseph E. ; BARRY, William T. ; KIM, Jong W. ; WANG, Quanli ; CRAWFORD, Matthew D. ; DATTO, Michael B. ; KELLEY, Michael ; MATHEY-PREVOT, Bernard ; POTTI, Anil ; NEVINS, Joseph R.: A pathway-based classification of human breast cancer. In: *Proceedings of the National Academy of Sciences of the United States of America* 107 (2010), April, Nr. 15, 6994–9. <http://dx.doi.org/10.1073/pnas.0912708107>. – DOI 10.1073/pnas.0912708107. – ISSN 1091–6490
- [Han09] HANDELS, Heinz: *Medizinische Bildverarbeitung: Bildanalyse, Mustererkennung und Visualisierung für die computergestützte ärztliche Diagnostik und Therapie*. 2. Vieweg+Teubner, 2009

- [HKK05] HANDL, Julia ; KNOWLES, Joshua ; KELL, Douglas B.: Computational cluster validation in post-genomic data analysis. In: *Bioinformatics (Oxford, England)* 21 (2005), August, Nr. 15, 3201–12. <http://dx.doi.org/10.1093/bioinformatics/bti517>. – DOI 10.1093/bioinformatics/bti517. – ISSN 1367–4803
- [Hu62] HU: Visual pattern recognition by moment invariants. In: *IEEE Transactions on Information Theory* 8 (1962), Februar, Nr. 2, 179–187. <http://dx.doi.org/10.1109/TIT.1962.1057692>. – DOI 10.1109/TIT.1962.1057692. – ISSN 0018–9448
- [HW05] HLAOUI, Adel ; WANG, Shengrui: Median graph computation for graph clustering. In: *Soft Computing* 10 (2005), April, Nr. 1, 47–53. <http://dx.doi.org/10.1007/s00500-005-0464-1>. – DOI 10.1007/s00500-005-0464-1. – ISSN 1432–7643
- [IMG] *Klinische Diagnostik - Allgemeines.* <http://www.pathologie-uniklinikum-mannheim.de/Klinische-Diagnostik/Allgemeines/>, Abruf: 27. November. 2011
- [JB87] JARQUE, C M. ; BERA, A K.: A test for normality of observations and regression residuals. In: *International Statistical Review* 55 (1987), Nr. 2, 163–172. <http://www.jstor.org/stable/1403192>
- [JMF99] JAIN, a. K. ; MURTY, M. N. ; FLYNN, P. J.: Data clustering: a review. In: *ACM Computing Surveys* 31 (1999), September, Nr. 3, 264–323. <http://dx.doi.org/10.1145/331499.331504>. – DOI 10.1145/331499.331504. – ISSN 03600300
- [Jol72] JOLLIFFE, I. T.: Discarding Variables in a Principal Component Analysis. I: Artificial Data. In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 21 (1972), Nr. 2, 160–173. <http://dx.doi.org/10.2307/2346488>. – DOI 10.2307/2346488. – ISSN 00359254
- [Jol73] JOLLIFFE, I. T.: Discarding Variables in a Principal Component Analysis. II: Real Data. In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 22 (1973), Nr. 1, 21–31. <http://dx.doi.org/10.2307/2346300>. – DOI 10.2307/2346300. – ISSN 00359254
- [JPWPV03] JOSIEN P. W. PLUIM, J. B. Antoine M. ; VIERGEVER, Max A.: Mutual information based registration of medical images: a survey. In: *IEEE Transactions On Medical Imaging* 20 (2003), 1, Nr. 100, S. 1–11
- [KCK+11] KONG, Jun ; COOPER, Lee ; KURC, Tahsin ; MIKKELSEN, Tom ; SCARPACE, Lisa ; BRAT, Daniel ; SALTZ, Joel: Nuclear Morphometry Analysis

- with Whole-Slide Microscopy Images of Low-grade Glioma. In: *ACM Conference on Bioinformatics, Computational Biology and Biomedicine Conference*, 2011
- [Kha11] KHABEROVA, Eleonora: *Image-processing descriptors for inner tumour growth patterns*, Rheinisch-Westfälische Technische Hochschule Aachen, Diplomarbeit, 2011
- [Kru57] KRUSKAL, J.B.: On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem. In: *Proceedings of the American Mathematical Society* 7 (1957), S. 48–50
- [KSH01] KOHONEN, T. (Hrsg.) ; SCHROEDER, M. R. (Hrsg.) ; HUANG, T. S. (Hrsg.): *Self-Organizing Maps*. 3rd. Secaucus, NJ, USA : Springer-Verlag New York, Inc., 2001. – ISBN 3540679219
- [Lam] LAMBERT, Tim: *Delaunay Triangulation Algorithms*. <http://www.cse.unsw.edu.au/~lambert/java/3d/delaunay.html>, Abruf: 23. Oktober. 2011
- [LZLM07] LIU, Y ; ZHANG, D ; LU, G ; MA, W: A survey of content-based image retrieval with high-level semantics. In: *Pattern Recognition* 40 (2007), Januar, Nr. 1, 262–282. <http://dx.doi.org/10.1016/j.patcog.2006.04.045>. – DOI 10.1016/j.patcog.2006.04.045. – ISSN 00313203
- [Mar02] MARINA, Meila: *Comparing Clustering*. Washington, 2002
- [Mata] *Bioinformatics Toolbox*. <http://www.mathworks.de/products/bioinfo/index.html>, Abruf: 21. Oktober. 2011
- [Matb] *Color-Based Segmentation Using K-Means Clustering*. <http://www.mathworks.de/products/image/demos.html?file=/products/demos/shipping/images/ipexhistology.html>, Abruf: 19. Oktober. 2011
- [Matc] *Image Processing Toolbox*. <http://www.mathworks.de/products/image/index.html>, Abruf: 19. Oktober. 2011
- [Matd] *Statistics Toolbox*. <http://www.mathworks.de/products/statistics/index.html>, Abruf: 21. Oktober. 2011
- [MKSW99] MAKHOUL, John ; KUBALA, Francis ; SCHWARTZ, Richard ; WEISCHEDEL, Ralph: *Performance measures for information extraction*. 1999
- [MMBG04] MÜLLER, Henning ; MICHOUX, Nicolas ; BANDON, David ; GEISSBUHLER, Antoine: A review of content-based image retrieval systems in medical applications-clinical benefits and future directions. In: *International journal of medical informatics* 73 (2004), Februar, Nr. 1, 1–23. <http://dx.doi.org/10.1016/j.ijmedinf.2003.11.024>. – DOI 10.1016/j.ijmedinf.2003.11.024. – ISSN 1386–5056

- [MST] *Minimaler Spannbaum.* http://de.wikipedia.org/wiki/Minimaler_Spannbaum, Abruf: 21. Oktober. 2011
- [NA08] NIXON, Mark S. ; AGUADO, Alberto S.: *Feature Extraction & Image Processing.* 2. Academic Press is an imprint of Elsevier, 2008. – 423 S.
- [Nie03] NIEMANN, H: *Klassifikation von Mustern.* 2. Springer, 2003. – 482 S.
- [OJJS96] ONG ; JIN ; JAYASOORIAH ; SINNIAH: Image Analysis Of Tissue Sections. In: *Compu. Biolog. Med.* 26 (1996), Nr. 3, S. 269–279
- [Pri57] PRIM, R.: Shortest Connection Networks and Some Generalizations. In: *Bell System Technical Journal* 36 (1957), S. 1389–1401
- [RB03] RODENACKER, Karsten ; BENGTTSSON, Ewert: A feature set for cytometry on digitized microscopic images. In: *Analytical cellular pathology : the journal of the European Society for Analytical Cellular Pathology* 25 (2003), Januar, Nr. 1, 1–36. <http://www.ncbi.nlm.nih.gov/pubmed/12590175>. – ISSN 0921–8912
- [RBH⁺08] REXHEPAJ, Elton ; BRENNAN, Donal ; HOLLOWAY, Peter ; KAY, Elaine ; MCCANN, Amanda ; LANDBERG, Goran ; DUFFY, Michael ; JIRSTROM, Karin ; GALLAGHER, William: Novel image analysis approach for quantifying expression of nuclear proteins assessed by immunohistochemistry: application to measurement of oestrogen and progesterone receptor levels in breast cancer. In: *Breast Cancer Research* 10 (2008), Nr. 5, S. R89
- [RCC92] RANDELL, D. A. ; CUI, Z. ; COHN, A. G.: A Spatial Logic based on Regions and Connection. In: *3rd Int. Conf. on Knowledge Representation and Reasoning*, 1992, S. 165–176
- [REFDA97] RANGAYYAN, R M. ; EL-FARAMAWY, N M. ; DESAUTELS, J E. ; ALIM, O a.: Measures of acutance and shape for classification of breast tumors. In: *IEEE transactions on medical imaging* 16 (1997), Dezember, Nr. 6, 799–810. <http://dx.doi.org/10.1109/42.650876>. – DOI 10.1109/42.650876. – ISSN 0278–0062
- [Rou87] ROUSSEEUW, P: Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. In: *Journal of Computational and Applied Mathematics* 20 (1987), November, Nr. 1, 53–65. [http://dx.doi.org/10.1016/0377-0427\(87\)90125-7](http://dx.doi.org/10.1016/0377-0427(87)90125-7). – DOI 10.1016/0377-0427(87)90125-7. – ISSN 03770427
- [RP05] REGIERER, Anne ; POSSINGER, Kurt: *Mammakarzinom - manual Diagnostik und Therapie.* 1. Dieselstraße 2, 50859 Köln : Deutscher Ärzte-Verlag GmbH, 2005

- [RPK⁺95] ROBBINS, P ; PINDER, S ; KLERK, N de ; DAWKINS, H ; HARVEY, J ; STERRETT, G ; ELLIS, I ; ELSTON, C: Histological grading of breast carcinomas: a study of interobserver agreement. In: *Human pathology* 26 (1995), August, Nr. 8, 873–9. <http://www.ncbi.nlm.nih.gov/pubmed/7635449>. – ISSN 0046–8177
- [Run10] RUNKLER, Thomas A.: *Data Mining - Methoden und Algorithmen intelligenter Datenanalyse*. 1. Vieweg+Teubner, 2010. – 169 S.
- [Sch10] SCHNEIDER, Petra: *Advanced methods for prototype-based classification*, University of Groningen, Diss., 2010. – 114 S.
- [Sha48] SHANNON, C E.: A mathematical theory of communication. In: *Bell Syst. Tech. J* vol (1948), S. 27pp379–423/623–656
- [Shl05] SHLENS, Jonathon: *A Tutorial on Principal Component Analysis*. <http://www.sn1.salk.edu/~shlens/pub/notes/pca.pdf>. <http://www.sn1.salk.edu/~shlens/pub/notes/pca.pdf>. Version: December 2005
- [SKH08] SVOBODA, Tomas ; KYBIC, Jan ; HLAVAC, Vaclac: *Image Processing, Analysis, and Machine Vision: a MATLAB companion*. Toronto, Ontario : Thomson Learning, 2008
- [SKK01] SEBASTIAN, Thomas B. ; KLEIN, Philip N. ; KIMIA, Benjamin B.: Recognition of Shapes by Editing Shock Graphs. In: *ICCV*, 2001, Accepted - To appear
- [Soi98] SOILLE, P.: *Morphologische Bildverarbeitung*. Berlin, Heidelberg : Springer-Verlag, 1998. – ISBN 3–540–64323–0
- [SRE⁺79] SMITH, P ; REID, D B. ; ENVIRONMENT, Cluttered ; PALO, Lockheed ; ALTO, Palo ; SMITH, P L.: A Threshold Selection Method from Gray-Level Histograms. In: *IEEE Transactions on Systems, Man, and Cybernetics* 9 (1979), Nr. 1, 62–66. <http://dx.doi.org/10.1109/TSMC.1979.4310076>. – DOI 10.1109/TSMC.1979.4310076. – ISSN 0018–9472
- [SY96] SATO, A ; YAMADA, K: Generalized learning vector quantization. In: *Advances in Neural Information Processing Systems* (1996), S. 423–429
- [WXC09] WU, Junjie ; XIONG, Hui ; CHEN, Jian: Adapting the right measures for K-means clustering. In: *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '09* (2009), 877. <http://dx.doi.org/10.1145/1557019.1557115>. – DOI 10.1145/1557019.1557115. ISBN 9781605584959
- [XW05] XU, Rui ; WUNSCH, Donald: Survey of clustering algorithms. In: *IEEE transactions on neural networks / a publication of the IEEE Neural Networks*

Council 16 (2005), Mai, Nr. 3, 645–78. <http://dx.doi.org/10.1109/TNN.2005.845141>. – DOI 10.1109/TNN.2005.845141. – ISSN 1045–9227

- [ZL02] ZHANG, Dengsheng ; LU, Guojun: Shape Based Image Retrieval Using Generic Fourier Descriptors. In: *Signal Processing: Image Communication* 17, 2002, S. 825–848
- [Züh12] ZÜHLKE, Dietlind: *Learning vector quantization for heterogeneous structured data*, UNIVERSITY of Groningen, Netherlands, Diss., 2012