

Schlussbericht
zum BMBF-Verbundvorhaben

Histoquant Mamma – Bildanalyse zur Molekulardiagnostik von Brustkrebs

Dr. Bettina Winkler (Projektleitung)
Dr. Günter Schmidt
Prof. Dr. Gerd Binnig

Förderkennzeichen: 01IS10021

Laufzeit: 01.10.2010 – 30.11.2012

Zuwendungsempfänger:

Definiens AG, Bernhard-Wicki-Straße 5, 80636 München

DEFINIENS[®]
Understanding Images

Unterauftragnehmer:

Charité Berlin, Institut für Pathologie, Charitéplatz 1, 10117 Berlin

München, Mai 2013

Das diesem Bericht zu Grunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS10021 gefördert.



Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

I. Kurze Darstellung

1. Aufgabenstellung

8-9% aller Frauen erkranken im Laufe ihres Lebens an Brustkrebs, womit dieser bei Frauen eine der häufigsten Krebserkrankungen ist. Weltweit wird jährlich bei über einer Million Frauen neu Brustkrebs diagnostiziert, und nahezu 400 000 Frauen sterben jedes Jahr an dieser Krankheit.

Untersuchungen haben gezeigt, dass bei rund 20–30% aller Brustkrebspatientinnen ein HER2 (Human Epidermal growth factor Receptor 2)-positiver Tumor vorliegt. Bei diesem Typ sind auf der Oberfläche der Krebszellen erhöhte Mengen des HER2-Proteins vorhanden. Man spricht von einer „HER2-Überexpression“. Die Behandlung mit Herceptin, einem humanisierten HER2 Antikörper, verbessert die Überlebensdauer bei fortgeschrittenem (metastasierendem) Brustkrebs um ein Drittel.

Die Diagnose eines HER2-positiven Tumors erfolgt auf Grund eines standardisierten Scoring Algorithmus mit Hilfe der Immunohistochemie (IHC) und in-situ-Hybridisierung (SISH/FISH) zur Genamplifikationsanalyse. Da trotz dieses Tests mehr als 70% der HER2-positiven Frauen nicht auf eine Herceptin Behandlung ansprechen, entstehen unnötige Kosten (mehr als 100 T€ pro Behandlung) und Gesundheitsbeeinträchtigungen (infolge der Nebenwirkungen).

Eine Schwierigkeit dabei ist, dass die morphologische Identifikation des Tumors im H&E-gefärbten Schnitt erfolgt, das Her2-Scoring aber im Her2neu-gefärbten Schnitt, und nicht beide Schnitte gleichzeitig im Mikroskop betrachtet werden können. Hierdurch ergibt sich das Problem des mangelnden Informationsaustausches zwischen den einzelnen Färbungen, der besonders bei schwierig zu beurteilenden Tumoren eine korrekte Diagnose erschwert.

Deshalb ist eine Zielsetzung des Projekts „Histoquant Mamma“ die Verbesserung der Diagnostik durch eine detaillierte automatisierte Bildanalyse der Gewebeproben und eine parallele Visualisierung der unterschiedlichen Schnitte. Der Algorithmus wird nicht nur das HER2-Scoring für eine molekularagnostisch basierte Therapie nutzen, sondern auch andere immunhistologische Biomarker wie KI67, ER (Estrogen Receptor), und PR (Progesteron Receptor) auswerten.

Die automatisierte Quantifizierung und die neuartige Visualisierung wird eine Kombination unterschiedlicher Färbeverfahren wie H&E, IHC, FISH oder SISH erlauben. Der Prototyp wird auf Basis der Definiens Cognition Network Technologie implementiert, die im Gegensatz zu pixel-basierten Verfahren eine robuste Messung von morphometrischen und färbebezogenen Eigenschaften einzelner Zellen und Zellbestandteilen erlaubt. Die aus den automatisch gewonnen Messgrößen abgeleiteten Scores gestatten dem behandelnden Arzt eine umfassende und objektive Grundlage zur Therapie-Empfehlung. Es wird eine grafische Benutzeroberfläche umgesetzt, die dem Pathologen einen möglichst breiten und einfachen Zugang zu allen zur Diagnose relevanten Informationen bietet und eine gleichzeitige Ansicht aller unterschiedlichen Schnitte eines Falles ermöglicht.

Zur Entwicklung und Validierung der Bildanalyse-Verfahren werden Gewebeschnitte von 150 Patientinnen herangezogen.

Gemäß dem Antrag wurden sechs Arbeitspakete definiert, die zusammen mit einer Beschreibung der erzielten Ergebnisse unter II. 1. aufgeführt sind.

2. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde

Die Definiens Cognition Network Technology wurde vom Nobelpreisträger Prof. Gerd Binnig und seinem Team seit 1996 entwickelt und stellt eine robuste Technologie zur Lösung von Bildanalyseaufgaben dar. Die Technologie wird dazu eingesetzt, kognitive Prozesse der menschlichen Wahrnehmung zu modellieren, um Objekte, wie z.B. Gewebe, Zellen oder Zellbestandteile, aus Bilddaten automatisiert zu erkennen und zu quantifizieren. Sie ermöglicht die Erkennung unterschiedlicher Gewebetypen wie beispielsweise Tumor versus Nicht-Tumor Gewebe, die Erkennung und Klassifikation von Zellen und deren Abgrenzung voneinander sowie die Erkennung von Zellkomponenten wie Nukleus, Membran und Zytoplasma. Die Analyse des H&E-Bildes gibt bestmögliche Information zur Gewebe- und Zellmorphologie, IHC/HER2 ermöglicht den Zugang zur Expression des membranständigen HER2-Proteins, IHC/ER, PR und KI67 Zugang zur Expression der entsprechenden nuklearen Proteine und die in-situ-Hybridisierung (FISH/SISH) erlaubt den Zugang zur Genamplifikation.

Für die grafische Oberfläche wird ein neuartiger Viewer entwickelt, der mittels Koregistrierung der einzelnen Schnitte eines Falles eine parallele Ansicht hochaufgelöster Regionen zum unmittelbaren Abgleich erlaubt.

Durch die Zusammenarbeit mit dem Unterauftraggeber Charité Berlin ist der Zugang zu den Gewebeschnitten und zu pathologischem Fachwissen sichergestellt.

3. Planung und Ablauf des Vorhabens

Der geplante Ablauf ist in Abb.1 dargestellt.

Arbeitspaket	Bezeichnung	2010Q4	2011Q1	2011Q2	2011Q3	2011Q4	2012Q1	2012Q2	2012Q3
AP 0	Projektmanagement	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
AP 1	Schnittstellen-Entwicklung	6.0	1.0						
AP 2	Gewebeschnitte und Qualitätssicherung	1.0	1.5						
AP 3	Grafische Benutzerschnittstelle			1.5	3.0	3.0	3.0	2.5	
AP 4	Entwicklung von Bildanalyse-Algorithmen	1.5	6.0	7.0	5.5	3.0	1.0		
AP 5	Entwicklung des Scoring Verfahrens					2.5	4.5	6.0	3.5
AP 6	Validierung und Optimierung								5.0

Abb. 1: Zeitlicher Ablauf der einzelnen Arbeitspakete mit Angabe der erwarteten Aufwände in Personenmonaten

Arbeitspaket 6 (Validierung und Optimierung) enthielt als Unterpunkt die Annotation aller Schnitte der 150 Patientenfälle durch 6 Pathologen. Dieses erforderte mehr Zeit als angenommen, so dass im August 2012 ein Antrag auf Projektverlängerung um zwei Monate gestellt und genehmigt wurde. Das Projekt endete daher mit dem Arbeitspaket 6 am 30. November 2012.

4. Wissenschaftlicher und technischer Stand

Die Definiens Cognition Network Technology stützt sich auf folgende Patente:

1. Method For Processing Data Structures, WO0145033
2. Method For The Processing Of Several Different Data Structures, WO0205198
3. Extracting Information From Input Data Using A Semantic Cognition Network, WO2004036337
4. Cognition Integrator And Language, WO2006106152
5. Automatic Image Analysis And Quantification For Fluorescence In Situ Hybridization US2008137937

5. Zusammenarbeit mit anderen Stellen

Innerhalb des Projektzeitraums fanden unter Beteiligung von Mitarbeitern der Definiens AG folgende projektbezogenen Treffen statt:

- 05.10.2010: Kick Off Meeting, Charité Berlin
- 18.04.2011: Arbeitstreffen, Charité Berlin
- 08.11.2011: Arbeitstreffen, Definiens AG, München
- 25.10.2013: Validierungs-Workshop, Definiens AG, München

Die Zusammenarbeit mit den Projektpartnern vom Institut für Pathologie der Charité gestaltete sich intensiv und produktiv und erlaubte eine pathologisch fachgerecht angeleitete Auswertung der Bilder mittels der automatischen Bildanalysealgorithmen.

II. Eingehende Darstellung

1. Arbeitspaket: Schnittstellen

Das erste Arbeitspaket umfasste die Erstellung von Schnittstellen zu den erforderlichen Scannern für Objektträger mit Gewebeproben sowie die Anpassung des Systems an das Patienteninformationssystem, um die Einbindung in den Alltag des Pathologen zu optimieren:

1. Lesen von Aperio ScanScope Bild- und Metadaten
2. Lesen von 3Dhistech Bild- und Metadaten
3. Lesen von Leica Bild- und Metadaten
4. Lesen und Schreiben von Patientendaten und Scoring Ergebnissen in Roche SWISSLAB (System der Charité)

Beim halbjährlichen Arbeitstreffen bei der Definiens AG in München am 8. November 2011 wurde festgestellt, dass die geplante Anbindung an das Patienteninformationssystem der Charité derzeit noch nicht erfolgen kann, da die Charité ihr System auf Nexus/Pathologie umstellen wird und das neue System noch nicht installiert ist.

Im Änderungsantrag vom 28. November 2011 wurde dieser Punkt daher aus der Vorhabenbeschreibung des Projektes entfernt, im kosten- und aufwandneutralen Austausch gegen die Umsetzung der neuartigen Visualisierung der Koregisterierung der verschiedenen Gewebeschnitte eines Falles. Auf diese wird im Punkt Arbeitspaket 3 weiter eingegangen.

Damit verblieben in diesem Arbeitspaket die Erstellung der genannten Schnittstellen zum Lesen der Bilddaten, die bereits zu Anfang des Projektes, Ende 2010 umgesetzt wurden.

Weiterhin wurde ein Software-Werkzeug zum Konvertieren der 3Dhistech-Annotationen in das Aperio-Format umgesetzt, um die Annotationen der Pathologen möglichst einfach und schnell einlesen zu können.

Eine Schnittstelle zum Lesen von Bilddaten im VMScope-Format wurde ebenfalls entwickelt, da die Daten in diesem Format von der Charité zur Verfügung gestellt wurden und diese zur Entwicklung der Bildanalysealgorithmen erforderlich waren.

2. Arbeitspaket: Gewebeschnitte und Qualitätssicherung

Zu Testzwecken wurden bereits zu Projektbeginn fünf Bilddatensätze zur Verfügung gestellt, die jeweils in sechs Färbungen vorlagen: H&E, Her2, ER, PR sowie die Silber-in-situ-Hybridisierungen (SISH) für Her2 und Chr17. Hier traten folgende Bildartefakte auf, die während der Qualitätskontrolle gefunden wurden:

- Risse und Faltungen im Gewebe
- Wassertropfen und Lufteinschlüsse

- Farbtropfen und unregelmäßige, bzw. übermäßige Färbungen und Farbmarkierungen
- Unschärf eingescannte Regionen
- Artefakte beim Zusammensetzen von Bilddaten des Scanners (Stichingartefakte)
- Kompressionsartefakte durch die JPG-Komprimierung des Bildformats

Bei der Auswahl der Fälle für den endgültigen Satz von Bildern wurde im Folgenden darauf geachtet, dass vermeidbare Artefakte minimiert wurden, so wie dies auch in der Routinediagnostik praktiziert wird. Hierzu zählen insbesondere der Ausschluss von Wasser- oder Farbtropfen, ungleichmäßige Färbungen, Lufteinschlüsse oder große Faltungen im Gewebe.

Unvermeidbare Artefakte durch das Scannen oder Tuschemarkierungen der Tumore müssen aber von der Software behandelt werden und dürfen das Analyseergebnis nicht beeinflussen.

Ebenfalls testweise wurden einige Fälle mit einer Dual-iSH-Färbung angefärbt. Bei diesem seit kurzen von der FDA zugelassenem Verfahren werden sowohl Her2 Gen, als auch Chr17 auf einer einzigen Probe angefärbt. Das Her2 Gen erscheint als roter Punkt, das Chr17 als schwarzer Punkt und der Zellkern durch Hämatoxylin-Gegenfärbung als leicht bläuliche Fläche. Die Dual-iSH ist gut standardisiert und dokumentiert, was eine automatische Analyse vereinfachen würde.

Wie sich jedoch im weiteren Projektverlauf herausstellte, ist die Dual-Färbung noch schwierig im Labor anzuwenden und daher derzeit noch von schwankender Qualität. Es wurde daher beim Arbeitstreffen im November 2011 gemeinsam beschlossen, für das laufende Projekt mit der üblichen und ursprünglich geplanten SISH-Färbung zu beginnen, wobei auf zwei separaten Schnitten die Amplifikation von HER2 und CHR17 bestimmt wird.

Das Scannen der Objektträger hat die Charité an die VMScope GmbH übertragen. Um diesen Unterauftrag transparenter zu gestalten, wurden die Verträge im April 2011 derart umgestaltet, dass auch die VMScope GmbH direkt von der Definiens AG zum Scannen der Objektträger beauftragt wurde.

Im Frühling 2011 begann die Übergabe der Bilder für die tatsächliche Auswertung von der Charité an Definiens. Bis zum Herbst lagen insgesamt 110 Fälle mit 594 Bildern vor. Die restlichen Bilder trafen im Frühjahr 2012 ein.

3. Arbeitspaket: Entwicklung einer Grafischen Benutzerschnittstelle

Der erste Prototyp der grafischen Benutzeroberfläche (GUI) wurde auf dem Arbeitstreffen im November 2011 vorgestellt. Mit dieser kann der User sehr intuitiv das gesamte Slide anschauen und die verschiedenen Färbungen und deren Auswertungsergebnisse auch lokal pro Tumorregion vergleichen.

Verschiedene Regionen des Tumors reagieren unterschiedlich auf die Färbungen, so dass der Pathologe hierdurch Informationen über das Verhalten und Heterogenität des Tumors ableiten kann. Für diese befundungskritische Eigenschaft wurde die Visualisierung der Koregistrierung der Färbungen eines Falles zusätzlich

aufgenommen, im Austausch gegen die im Arbeitspaket 1 erwähnte noch nicht zu realisierende Anbindung an das Patienteninformationssystem (Änderungsantrag vom 28. November 2011).

Bei der Herstellung der verschiedenen Färbungen werden von dem in Paraffin eingebetteten Tumor mit einem Mikrotom ca. 2–5µm dicke Schnitte abgeschnitten, auf dem Objektträger aufgezogen und gefärbt. Im Idealfall erfolgen die Schnitte direkt hintereinander, ohne dass dazwischen Schnitte verworfen werden müssen. Die dann entstandenen Schnitte des Tumors ähneln einander in Lage und Verteilung, so dass ein virtuelles „Übereinanderlegen“ der Schnitte möglich ist.

Durch die Dreidimensionalität des Tumors, sowie durch Verzerrungen, Quetschungen, Spiegelungen oder auch Verdrehungen der Schnitte kann jedoch ein genaues Zusammenpassen durch die Koregistrierung sehr schwierig werden. Die Vorgehensweise besteht darin, dass zuerst „Landmarken“ in den Schnitten gefunden und gesetzt werden. Dies sind auffällige und somit gut identifizierbare Orte im Tumor, die bei allen Schnitten vorhanden sind. Da die Färbungen sich sehr in Farbton und Helligkeit unterscheiden, können ausschliesslich Formeigenschaften des Tumorschnittes hierzu herangezogen werden. Falls mehrere Gewebeteile auf dem Objektträger vorhanden sind, müssen diese einander in den unterschiedlichen Färbungen eindeutig zugeordnet werden. Sind die Landmarken gefunden und gesetzt, wird mit Transformationsformeln das Gewebe zwischen den Marken so verschoben, dass die Landmarken aller Schnitte exakt übereinander liegen.

Die Herausforderung besteht darin, Tumorteile einander zuzuordnen und gute Landmarken zu finden, auch wenn Gewebeteile gefaltet oder verdreht sind oder komplett fehlen.

Gerade bei Stanzbiopsien ist es oft sehr komplex, die jeweils zugehörigen Stenzen der einzelnen Färbungen zu identifizieren. Durch virtuelles „Übereinanderlegen“ einzelner Tumorteile wird anhand eines flächenmäßigen Überlappungsanteils entschieden, welche Teile am besten zueinander passen.

Der Schnitt einer Biopsie kann gegenüber den (Folge-)Schnitten mit anderen Färbungen gedreht (siehe unten) oder auch gespiegelt auf dem Objektträger aufliegen, so dass nach der Detektion der Teile versuchsweise durch Drehung und Spiegelung diejenigen mit dem größten gegenseitigen Überlapp gefunden werden, um eine eindeutige Zuordnung der einzelnen Teile zu gewährleisten. Diese Algorithmen-Entwicklung erfolgte ausserhalb des Projektrahmens durch die Definiens AG.

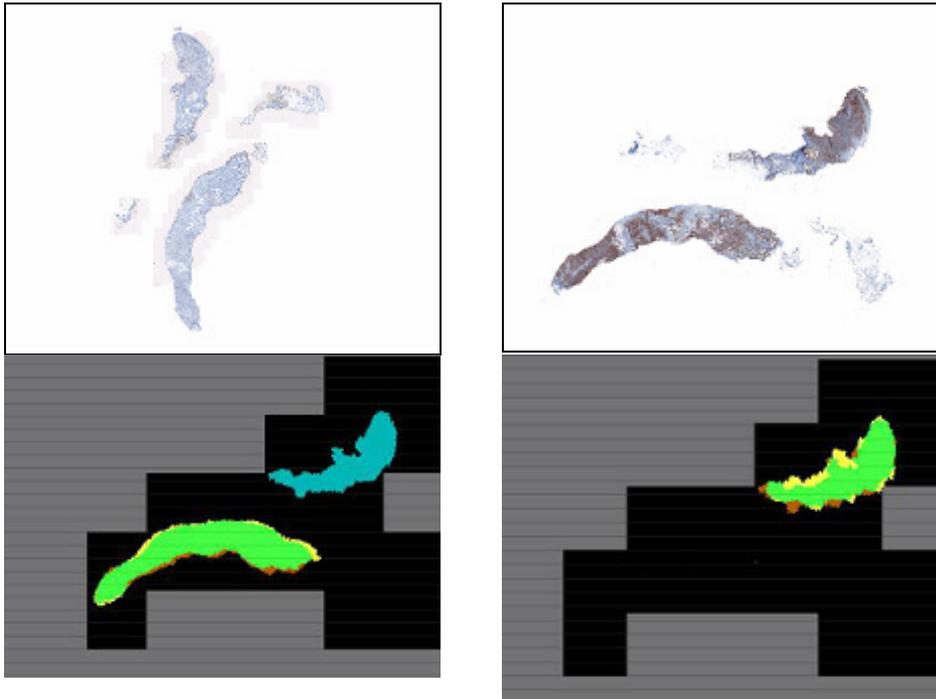


Abb 2: Beispiel einer Koregistrierung einer Stanzbiopsie: Oben die beiden zu registrierenden Bilder: links Ki67, rechts Her2neu; unten jeweils der digitale overlap der beiden Teile der Biopsie (gelb und orange: die Originale; grün: die gemeinsame Fläche)

Im nächsten Halbjahr wurde dann der Prototyp weiterentwickelt und vervollständigt. Eingebaut wurde die Möglichkeit, verschiedene Fälle zu verwalten (s. Abb. 3), dies dient bereits als Platzhalter für die tatsächliche geplante Anbindung an das Patienteninformationssystem, die ja nicht mehr im Zeitrahmen des Projektes fertiggestellt werden kann.

Da die vollautomatische Bildanalyse den gesamten Gewebeschnitt in einem Hintergrundprozess im Anschluss an das Digitalisieren analysiert hat, steht dem Anwender die Möglichkeit zur Verfügung, mit Hilfe der Bildanalyseergebnisse schnell zu "interessanten" Stellen im Gewebeschnitt zu navigieren. Dazu wurde im GUI eine Heatmap-Darstellung gewählt, wo "Hotspots" in Falschfarbendarstellung gezeigt werden. Ein Hotspot kann z.B. durch hohe Proteinexpression (Markerintensität) definiert sein, oder durch hohe Mitosenzahl im H&E-Schnitt.

Im Folgenden einige Screenshots der Oberfläche:

Name	Stains	Received	Ready	Modified	Assigned to	Status	Info
001-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
002-HE (BM3F, Resection)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
003-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
004-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
005-HE (BM3F, Resection)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
005-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
007-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
008-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
009-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
010-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
011-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
012-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	29.04.2013 1...	-	active	invasive duct...
013-HE (BM3F, Resection)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
014-HE (BM3F, Resection)	no information	unknown	22.04.2013 0...	29.04.2013 1...	-	active	no invasive c...
015-HE (BM3F, Resection)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
015-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
017-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
018-HE (BM3F, Resection)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
019-HE (BM3F, Resection)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
020-HE (BM3F, Resection)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
021-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment
022-HE (BM3F, Needle)	no information	unknown	22.04.2013 0...	22.04.2013 0...	-	active	no comment

Abb 3. Fallauswahl: Beispielhafte Übersicht mit Angabe der vorliegenden Färbungen und diversen verwaltungsrelevanten Parametern.

Die Benutzeroberfläche zur Evaluierung einzelner Gewebeschnitte zeigt eine Übersicht aller verfügbaren Färbungen, eine Übersichtsansicht zum schnellen Navigieren im Schnitt und eine große Detailansicht des Bildes (s. Abb. 4).

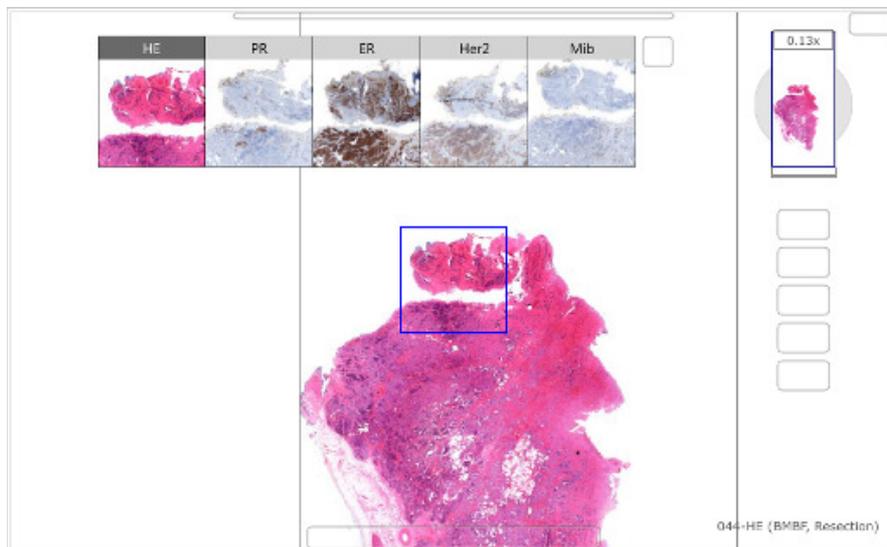


Abb 4. Übersicht eines Falles: oben die sogenannte „Multi-View“-Ansicht, in der der blau markierte Bereich aus der Hauptansicht mit den verschiedenen Färbungen mittels der Koregistrierung angezeigt wird. Der Anwender kann mit je einem Mausklick sehr bequem zu den Ansichten der anderen Färbungen wechseln und diese zur Hauptansicht machen. Rechts oben ist eine Übersicht über das Gesamtbild dargestellt.

Ferner wurde die Koregistrierung der verschiedenen Färbungen eines Falles fertiggestellt, so dass der Anwender Regionen des Tumors in unterschiedlichen Färbungen gleichzeitig ansehen kann (s. Abb. 5). Falls die Registrierung nicht optimal ist (durch die Dreidimensionalität des Tumors oder Schnittartefakte können

die Schnitte erheblich voneinander abweichen), kann der Anwender noch manuell korrigierend eingreifen, um hier eine optimale Registrierung sicherzustellen (s. Abb. 6).

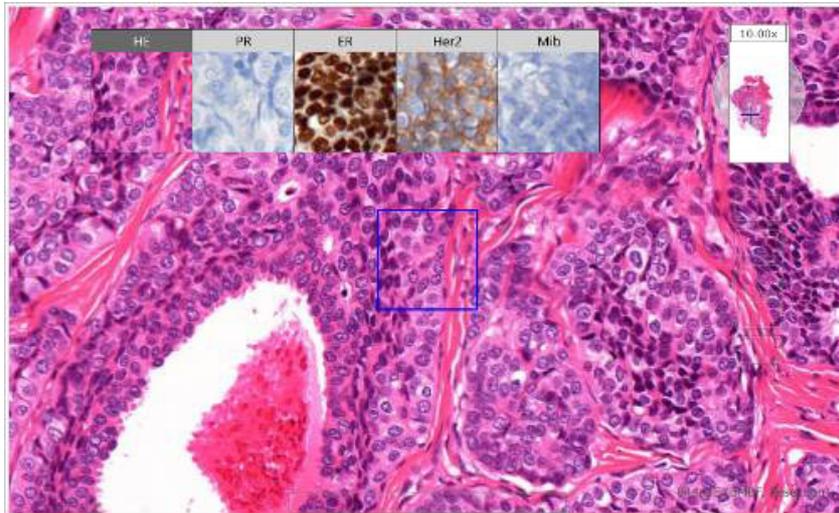


Abb. 5. Parallele Ansicht einer Region eines Tumors in H&E und oben in weiteren Färbungen. Ermöglicht wird die Synchronisation der Ansichten durch die Koregistrierung der Gewebeschnitte.

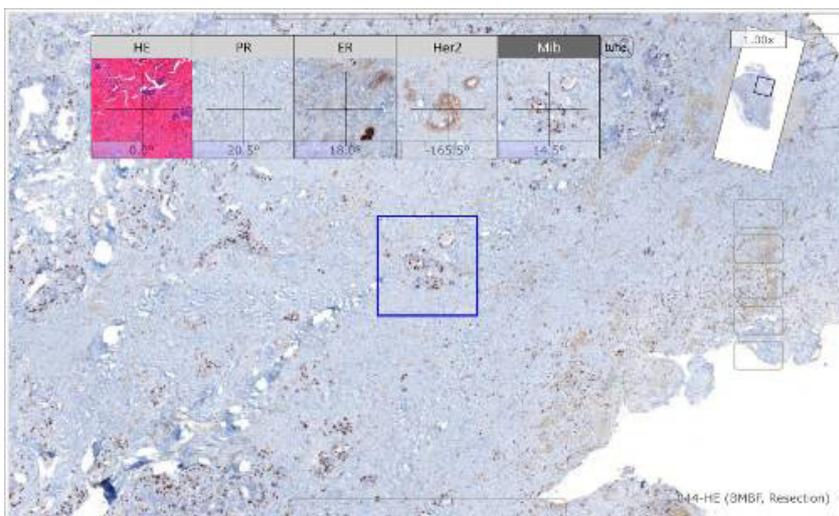


Abb. 6. Es ist möglich, manuell eine nicht optimale Koregistrierung zu editieren.

Der Pathologe kann zu jedem kompletten Fall (s. Abb 7), zu jedem einzelnen Schnitt oder innerhalb eines Schnittes (s. Abb. 8) Kommentare setzen. Mit Annotationen innerhalb eines Schnittes lassen sich Regionen markieren, die zur Diagnose herangezogen wurden oder andere Auffälligkeiten beinhalten. Dies erleichtert auch den Wissensaustausch zwischen mehreren Pathologen.

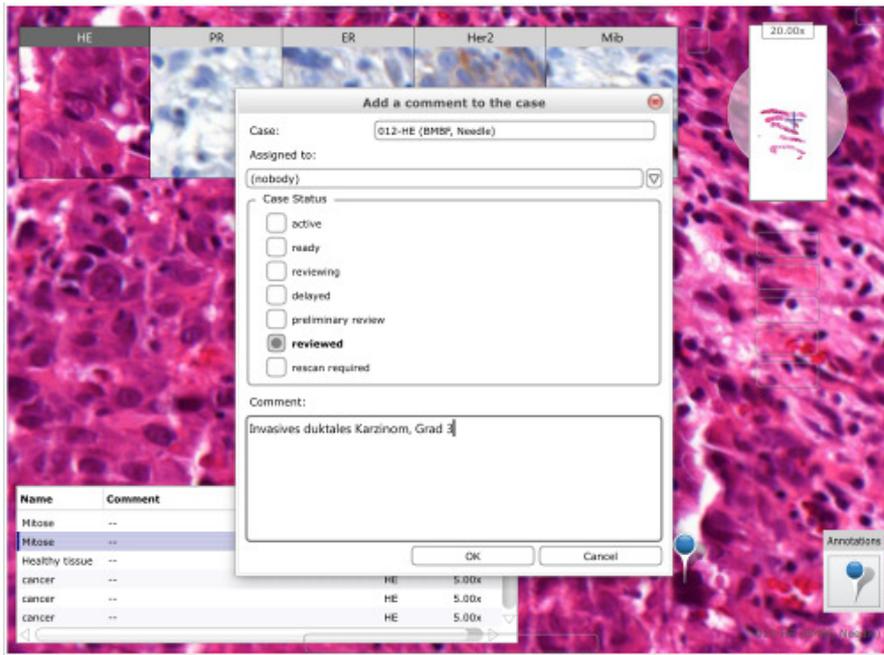


Abb. 7. Zu jeden Fall können Kommentare angefügt werden, die später in einem Bericht zusammengefasst werden.

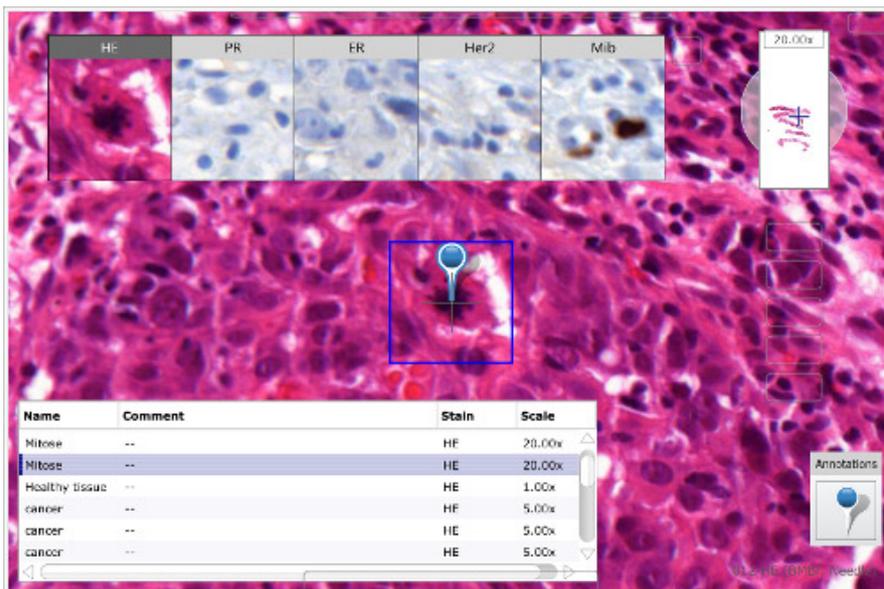


Abb. 8.: Interessante Regionen können per Annotation markiert und mit einem Kommentar versehen werden, so dass später besser nachvollziehbar ist, auf Grund welcher Regionen die Diagnose festgelegt wurde.

4. Arbeitspaket: Entwicklung der Bildanalysealgorithmen

Nachfolgend sind die einzelnen Bildanalysealgorithmen aufgeführt

IHC Nucleimarker: ER, PR, Ki67

Für die Nucleusmarker müssen „negative“ Kerne, d.h. Kerne, die nur mit Hämatoxin blau angefärbt sind und „positive“ Kerne, die mit dem Chromogen über den speziellen Antikörper für das gesuchte Protein braun oder rot angefärbt wurden, aufgefunden werden.

Die Methode für die Detektion der blauen Kerne wird auch bei den Algorithmen für die anderen Färbungen benötigt und war daher einer der ersten Analysen, die fertiggestellt wurden.

Für eine korrekte Auswertung der Nucleusmarkerbilder ist auch eine Selektion der Kerne notwendig, denn nur Tumorkerne dürfen ausgezählt werden. Diese Unterscheidung ist auch für Pathologen in der IHC-Färbung nicht immer sicher durchführbar und wurde hier dementsprechend soweit möglich optimiert. Kerne, die auf Grund ihrer Größe, Form oder ihrer Nachbarschaft nicht einem typischen Epithelkern entsprechen, werden ausgeschlossen.

Die Herangehensweise zur Detektion und Selektion der Kerne in der IHC-Färbung wird in den Abb. 9 und 10 illustriert.

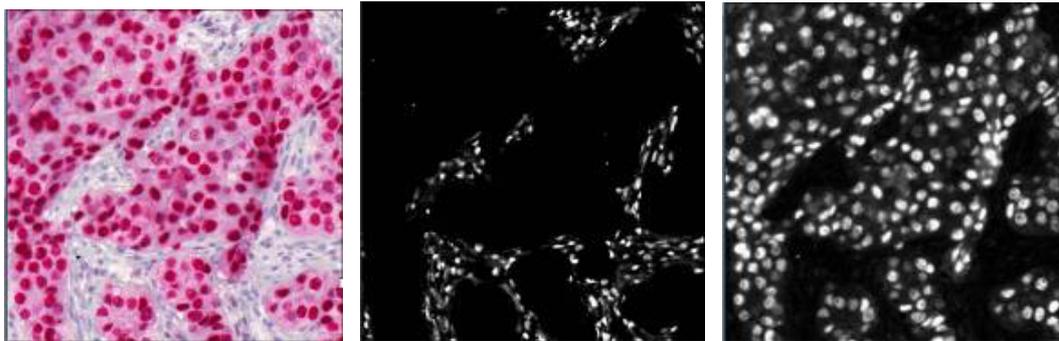


Abb. 9: Trennung des blauen und hier roten Farbkanals, die Kerne werden anschließend auf Grund der Intensität in den getrennten Kanälen detektiert.

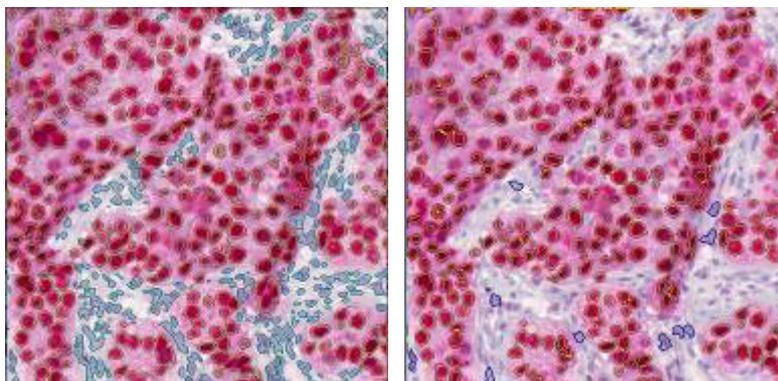


Abb. 10: Links die Detektion aller intensiven Regionen, rechts die Trennung in einzelne Kerne und Selektion auf Tumorkerne.

Der Vergleich der Ergebnisse des Algorithmus mit den in der Validierungsphase erfolgten Annotationen der fünf Pathologen erfolgt beim Arbeitspaket 6.

IHC Her2neu Membranfärbung

Bei diesem Algorithmus werden zuerst die mit Hämatoxylin-angefärbten Kerne zur Kerndetektion aufgefunden. Im Anschluß werden die Zellen identifiziert, die Membranen erstellt und die Zellen auf Grund ihrer Membranfärbung (relativ in Bezug auf die Zytoplasmafärbung) klassifiziert.

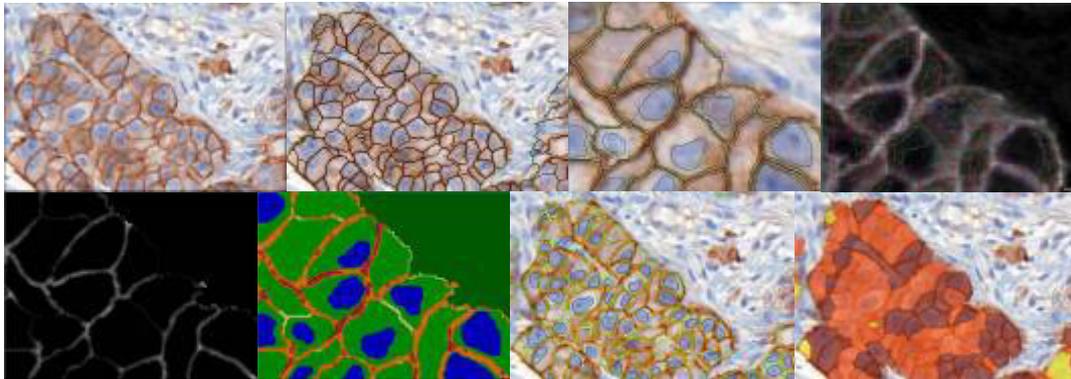


Abb 11: Membrandetektion und Klassifizierung:

Oben von links nach rechts: Originalbild – Detektion der Zellen im Tumorgebiet – Erstellung der gefärbten Membran – Objekte zur Berechnung der lokalen relativen Intensitätsgradienten zwischen Membran und Zytoplasma der Zellen.

Unten von links nach rechts: erstellte Bildebene, die die Intensitätssteigerung in der Membran widerspiegelt – Klassifikation der Membran auf Grund der erstellten Bildebene – Übersicht der Membranklassifikation der Tumorregion – Klassifikation der Zellen auf Grund der Membranklassifikation (gelb: 1+, orange: 2+, rot: 3+)

Die Klassifizierung der einzelnen Zellen wird auf Grund der Membranfärbung (Intensität und Zirkumferenz) nach dem internationalen Standard der ASCO/CAP durchgeführt. Hierdurch können einzelne Regionen ebenso mit einem Score – beruhend auf den selektierten Zellen – versehen werden, wie der ganze Schnitt.

Die Analyse erwies sich als relativ robust gegenüber Artefakten, da die notwendige aufwändige Detektion von Zellen und Membranen eine Fehlinterpretation von Färbungsartefakten minimiert.

Schwieriger war die ursprünglich angedachte Unterscheidung in den invasiven und nicht-invasiven Anteilen des Tumors. Dies konnte so nicht umgesetzt werden, da nach Absprache mit den Pathologen diese Information nicht im Her2neu-gefärbten Schnitt durchgängig vorhanden ist. Hierzu braucht auch ein Pathologe oft die H&E-Färbung oder sogar eine zusätzliche p63-Färbung (IHC). Ausgewertet werden daher von der Software alle gefärbten Tumor-Bereiche. Der Pathologe kann dann die richtigen Bereiche auswählen und sich den entsprechenden Score zur Diagnoseunterstützung ausgeben lassen.

Der Vergleich der Ergebnisse des Algorithmus mit den in der Validierungsphase erfolgten Annotationen der fünf Pathologen erfolgt beim Arbeitspaket 6.

H&E

Der H&E-Algorithmus lag bis zum Validierungsworkshop zum Projektende noch nicht in der endgültigen Fassung vor, da die Schwierigkeit dieser Analyse auf Grund der morphologischen Kriterien erheblich ist.

Dargestellt werden soll der Elston&Ellis-Score, der in der Routinediagnostik ein Grading der Brusttumore erlaubt. Dieser setzt sich aus drei Einzelscores zusammen, die im folgenden getrennt behandelt werden.

Im Algorithmus ist diesen eine Detektion der Zellkerne vorgeschaltet, deren Klassifizierung eine weitgehende Trennung von Stroma- und Tumorbereichen ermöglicht.

Mitosen:

Mitosen (Zellteilungen) werden ausgezählt, um die Proliferationsrate des Tumors zu beurteilen.

Wir haben sehr frühzeitig mit der Entwicklung des Algorithmus begonnen und drei Pathologen um die Annotation kleiner Bildausschnitte gebeten, um eine gute Übersicht über die Variabilität zu bekommen. Das Ergebnis war eher ernüchternd, da es wenig bis kaum Übereinstimmungen in den Annotationen gab:

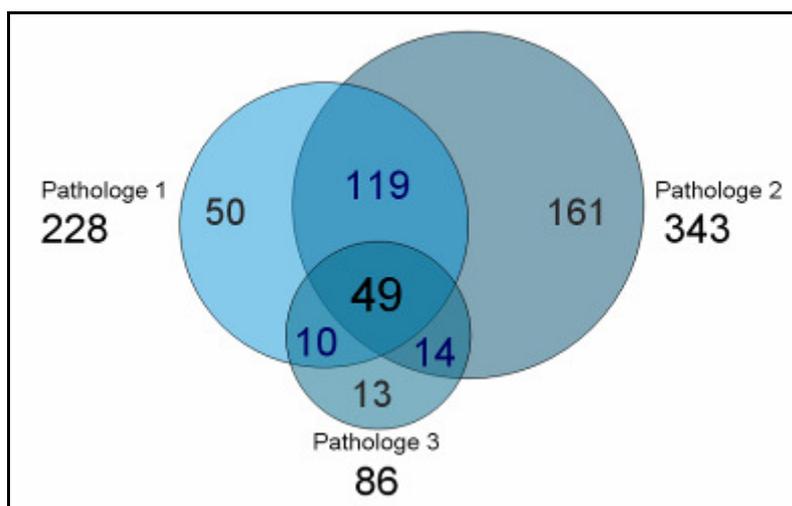


Abb. 12. Gesamtanzahl der von verschiedenen Pathologen festgestellten Mitosen und Anzahl der mit den anderen Pathologen identischen annotierten Mitosen.

In Absprache mit den Pathologen der Charité entwickelten wir dementsprechend einen Algorithmus, der sich auf die eindeutigen Mitosen konzentriert, so wie es in der Leitlinie des Elston&Ellis-Scores empfohlen wird.

Hierzu werden potenzielle Mitosen im Gewebe detektiert und auf Grund ihrer Form und Farbe als Mitosen klassifiziert. Die unterschiedlichen Phänotypen der verschiedenen Mitosephasen werden hierbei berücksichtigt.

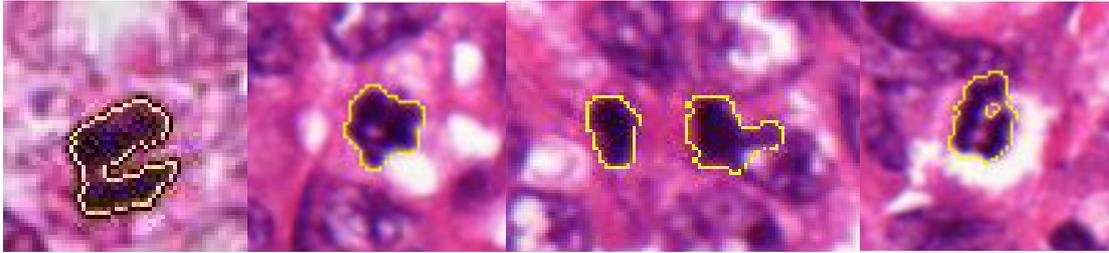


Abb 13: Detektion verschiedener Mitosephasen

Kernmorphologie

Die Kerndetektion erfolgte ja schon zum Zwecke der Erkennung der Tumorregionen. Um die Kernmorphologie zu beurteilen, wurden die Kerne anschließend auf Grund ihrer Größe und Heterogenität der Hämatoxylin-Färbung weiter klassifiziert.

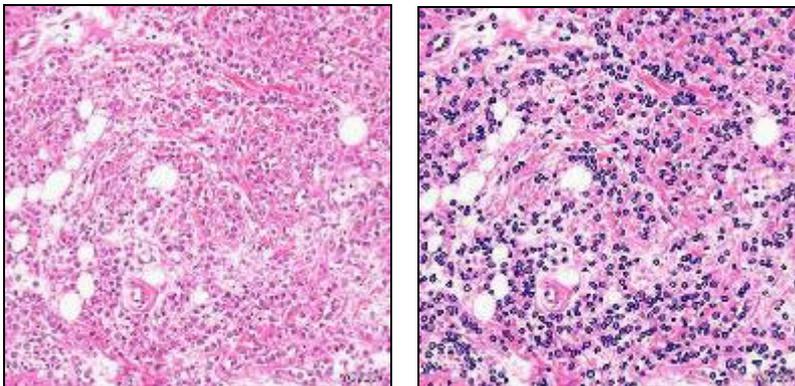


Abb. 14. Detektion von Tumorkernen; auch unregelmäßig gefärbte, nur ringartig sichtbare Kerne werden detektiert

Tubulibildung

Die Tubulibildung ist typisch für das duktales Brustgewebe und nimmt mit zunehmender Entdifferenzierung des Tumors ab. Hierzu werden Lumina („Leerräume“) mit umgebenen Epithelzellsäumen detektiert. Aus dem prozentualen Anteil des Tumors, der noch Tubuli bildet, wird der entsprechende Score ermittelt.

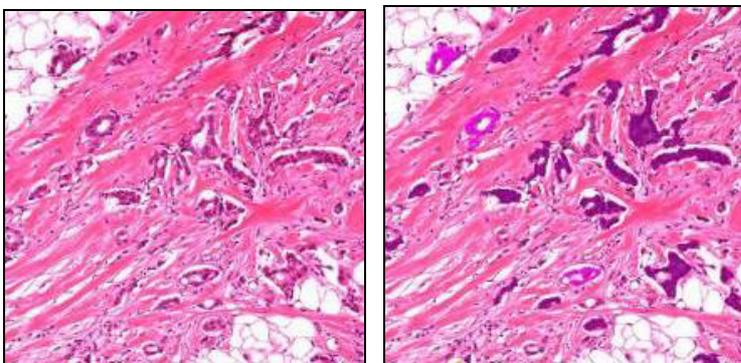


Abb. 15. Links das Original, rechts die Klassifizierung: nicht-tubulibildender Tumor (lila) und tubulibildender Tumor (pink)

Der Vergleich der Ergebnisse des Algorithmus mit den in der Validierungsphase erfolgten Annotationen der fünf Pathologen erfolgt beim Arbeitspaket 6.

SISH

Die Analyse für die SISH-Färbung erfolgt auf zwei Schnitten: Einmal ist das Chromosom 17 mittels Silber angefärbt, einmal das Her2/neu-Gen. Ausgezählt werden die Silber-Signale pro Kern. Typischerweise hat jeder Zellkern (auch nicht-Tumorkerne) jeweils zwei Signale, wobei natürlich durch die Dreidimensionalität des Zellkerns nicht immer alle zwei Signale angeschnitten wurden. In Fällen, bei denen die Her2neu-Auswertung einen Score von 2+ ergeben hat, wird eine in-situ-Hybridisierung (ISH) durchgeführt, um eindeutig eine Amplifikation des Gens nachzuweisen. Für einen positiven Befund muß das Signal bei der Her2/neu-Gen-Färbung mehr als zweimal soviel Signale ergeben, wie für das Chromosom 17. Nur dann ist das Gen überamplifiziert und die gezielte Therapie kann erfolgreich sein.

Für die Auswertung werden nun die dunklen Signale in den bereits detektierten Kernen segmentiert und ausgezählt. Kerne, die zu dicht nebeneinander liegen, sollen ausgeschlossen werden, da die Signale hier nicht eindeutig einem Kern zugeordnet werden können. Eine Schwierigkeit hierbei war die zum Teil nicht ausreichende Qualität der Färbung.

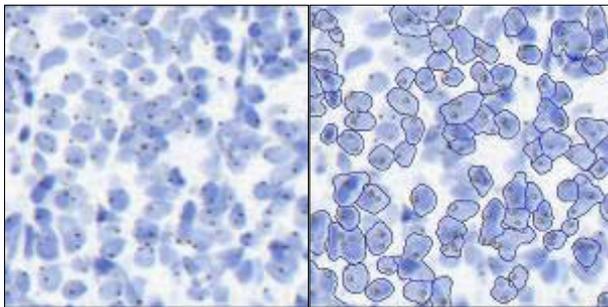


Abb. 16: Links: Originalbild, rechts: Klassifikation der Kerne und Silbersignale.

Der Vergleich der Ergebnisse des Algorithmus mit den in der Validierungsphase erfolgten Annotationen der fünf Pathologen erfolgt beim Arbeitspaket 6.

5. Arbeitspaket: Entwicklung des Scoring-Verfahrens

Die Scores für die verschiedenen Färbungen orientieren sich an der Leitlinie zur Auswertung von Mammakarzinomen, um institutsübergreifende Gültigkeit und hohe Akzeptanz zu erreichen.

Für die Kernfärbungen wird primär der Prozentsatz der positiv angefärbten Kerne ermittelt. Da die exakte Prozentzahl auch vom selbem Pathologen nicht völlig reproduzierbar ist („Intra-Observer Varianz“), werden zum Abgleich mit den Pathologenannotationen die Bilder anhand der erreichten Prozentzahlen der positiven Kerne in Gruppen, sogenannte „Bins“ eingeteilt:

- Bin 0: <1%
- Bin 1: 1-10%
- Bin 2: 10-30%
- Bin 3: 30-100%

Für eine Therapieentscheidung ist oftmals die Unterscheidung des Bin 0 von den anderen Bins am wichtigsten, da z.B. eine Therapie mit Antiöstrogenen schon bei geringster Östrogenreaktion gegeben werden kann. Daher ist die Unterscheidung der niedrigeren Reaktion besonders wichtig.

Für die Her2/neu-Auswertung werden die detektierten Zellen nach Membranfärbung in 0, 1+, 2+ und 3+ klassifiziert. Hierzu wird der international gültige ASCO/CAP Standard verwendet. Aufgrund der prozentualen Anteile der klassifizierten Zellen wird der Gesamtscore für den Schnitt errechnet. Eine Herceptin-Therapie wird nicht verordnet bei den Scores 0 und 1 (therapeutisch equivalent), und immer bei Score 3+. Schnitte, die den Score 2+ erhalten, werden zusätzlich noch mit einer ISH Färbung angefärbt, um den tatsächlichen Amplifizierungsgrad des Her2/neu-Gens festzustellen (siehe SISH).

Bei der H&E-Färbung wird der Elston&Ellis-Score ermittelt. Dieser ergibt sich durch Addition der Scores für die Mitosendichte, die Kernmorphologie und die Tubulibildung des Tumors. Tumore mit einer höheren Summe ergeben einen höheren Grad und gelten als aggressiver als Tumore mit einem niedrigeren Grad.

Die Anzahl der gefundenen Mitosen wird pro „High-Power-Field“ gemessen, d.h. auf eine festgelegte Fläche des Gewebes bezogen und der entsprechende Score für den Elston&Ellis-Score wird nach einer vorgeschriebenen Tabelle abgeleitet.

Bei der Kernmorphologie wird der prozentuale Anteil an veränderten Kernen in einen Score umgesetzt. Ebenso bei der Tubulibildung, hier ist der Score umso höher, je weniger Tubuli gebildet werden.

Bei der Auswertung der SISH-Färbung werden die durchschnittlichen Silbersignale in den Tumorkernen für das Chromosom 17 und das Her2/neu-Gen ausgezählt. Eine tatsächliche Überexprimierung ist erst bei mehr als 2-facher Anzahl der Signale beim Her2/neu-Gen gegeben.

6. Arbeitspaket: Validierung und Optimierung

Zur Validierung der Ergebnisse wurde am 25. Oktober 2012 in München ein Workshop durchgeführt. Hierzu waren Pathologen geladen, die im Vorfeld dieses Treffens alle Bilder annotierten. Durch diese Annotationen wurde eine Vergleichsmöglichkeit geschaffen, anhand derer die automatische Auswertung validiert wurde.

Folgende Mammakarzinom-Experten führten die Annotationen durch:

PD Dr. Doris Mayr
Oberärztin und Privatdozentin am Pathologischen Institut der LMU München
(Insitutsdirektor Prof. Dr. Thomas Kirchner)

Dr. Jörg Nährig
Oberarzt am Institut für Pathologie der TU München
Referenzzentrum für Mammadiagnostik im Rahmen des Dtsch. Mammographie-Screenings
(Insitutsdirektor Prof. Dr. Heinz Höfler)

Dr. Rupert Langer
Oberarzt am Institut für Pathologie der Universität Bern
(Insitutsdirektor Prof. Dr. Aurel Perren)

Dr. Frederick Klauschen
Assistenzarzt am Institut für Pathologie der Charité in Berlin
(Insitutsdirektor Prof. Dr. Manfred Dietel)

Dr. Barbara Ingold-Heppner
Assistenzärztin am Institut für Pathologie der Charité in Berlin
(Insitutsdirektor Prof. Dr. Manfred Dietel)

Zur Validierung wurden für die jeweiligen Färbungen zwei statistische Auswertungen angefertigt: Zum einen die Varianz der Pathologen untereinander, die sogenannte „Inter-Observer-Varianz“. Diese gibt Aufschluß über die Schwankungsbreite der Diagnosen. Erwartung an die Softwareauswertung ist, dass diese nicht mehr von den Pathologen abweicht, als diese untereinander abweichen. Dieser Vergleich der Pathologen und der Software stellt die zweite Auswertung dar. Hierzu wurde ein Konsenswert der Pathologen ermittelt und die Abweichung der Pathologen und der Software zu diesem Wert verglichen.

In den Vergleichen sind die Ergebnisse der Pathologen anonymisiert dargestellt. Im Folgenden sind für die einzelnen Färbungen die Auswertungen aufgeführt.

Kernfärbungen ER, PR, Ki67

Bei den Kernfärbungen werden die ermittelten Prozentzahlen der positiv angefärbten Kerne in vier Bins übersetzt, um die Vergleichbarkeit zu verbessern. Von der Auswertung ausgeschlossen wurden die Fälle, in denen von allen fünf Pathologen keine drei einen gleichen Bin erreichten. Bei den übrigen wurde der gemeinsame Bin der übereinstimmenden mindestens drei Pathologen als Konsens angenommen. Für jeden Pathologen wurde nun ermittelt, bei wieviel Prozent der Fälle er mit diesem Konsens übereinstimmt. Ebenso wurde dies für die Auswertung der Software ermittelt.

Je nach Färbung änderte sich die Verteilung der Bins des Konsens.

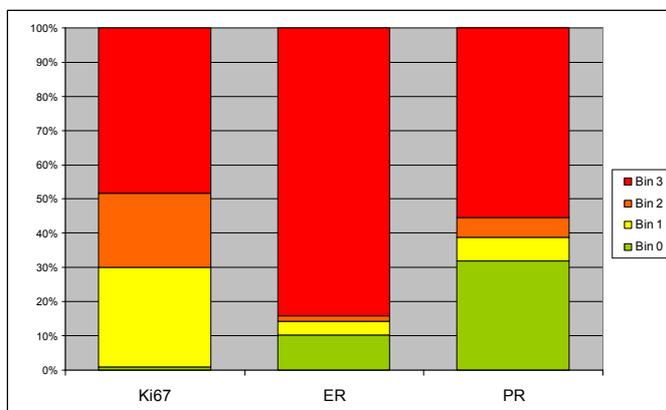


Abb. 17: Vergleich der Bin-Verteilung bei den drei Kernfärbungen

Nachfolgend sind die Ergebnisse dieser Auswertung für die drei Kernfärbungen aufgeführt.

Bei **Ki67** erreicht die Software einen Prozentsatz, der noch über dem eines Pathologen (E) liegt. Insgesamt ist hier die Übereinstimmung auch zwischen den Pathologen nicht sehr hoch. Gründe sind zum einen, dass es bei Ki67 viele Fälle gibt, deren Prozentsatz unter 30% liegt und diese Färbung damit die größte Varianz an Bins aufzeigt (s. Abb. 17). Individuelle Abweichungen um 5% können hier leicht in unterschiedlichen Bins resultieren. Weiterhin reagieren unterschiedliche Regionen des Tumors oft unterschiedlich auf die Färbung, da die Proliferation meist am Tumorrand höher ist als im Tumorzentrum. Je nachdem, wie diese Heterogenität von den Pathologen ausgewertet wurde, kommen hier leicht unterschiedliche Prozentzahlen zustande. Die Software mittelt grundsätzlich über den gesamten Tumor. Es gibt derzeit noch keine Anweisung in den Leitlinien, dies anders zu handhaben, obwohl die Pathologen das oft für sinnvoller halten.

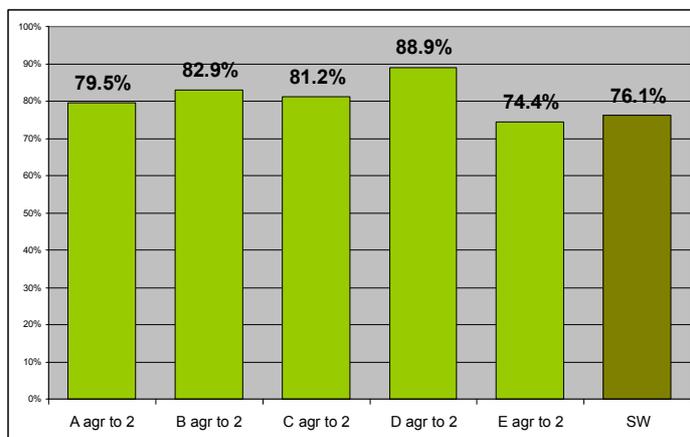


Abb. 18: Ki67: Prozentsatz der Übereinstimmung mit dem Konsens für Pathologen und Software

Für **ER** liegt die Übereinstimmung bei über 90%. Jedoch ist sie bei den Pathologen noch höher. In einem Großteil der uns vorliegenden Fälle ist die ER-Expression entweder null oder knapp 100% (s. Abb. 17), so dass die hohe Übereinstimmung nicht überrascht. Die Fälle, die der Software zur hohen Quote der Pathologen fehlt, sind meist Fälle mit 0% positiven Kernen, die aber Färbeartefakte aufweisen, die zu Fehlinterpretationen führen.

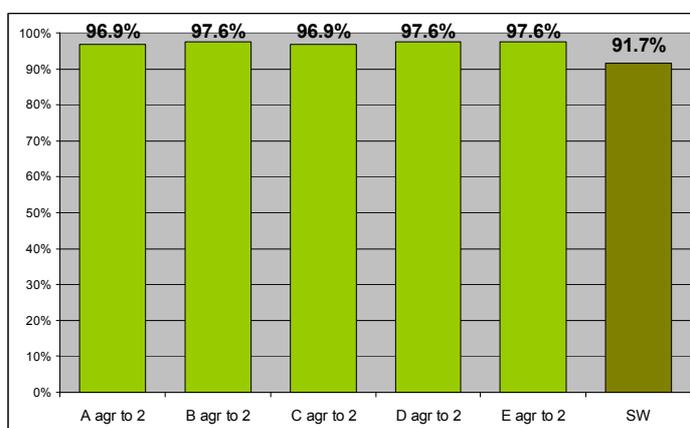


Abb. 19: ER: Prozentsatz der Übereinstimmung mit dem Konsens für Pathologen und Software

Das Ergebnis für **PR** ist noch nicht zufriedenstellend. Die Quote der Software liegt hier unter der niedrigsten Pathologenquote. Sehr viele Fälle wurden hier mit Bin 0 klassifiziert (s. Abb. 17), dieser Bin ist am leichtesten für die Software zu verfehlen: Artefakte oder gesundes Gewebe können hier schnell eine Einteilung in den nächsthöheren Bin verursachen, da schon wenige positive Kerne zum höheren Bin führen.

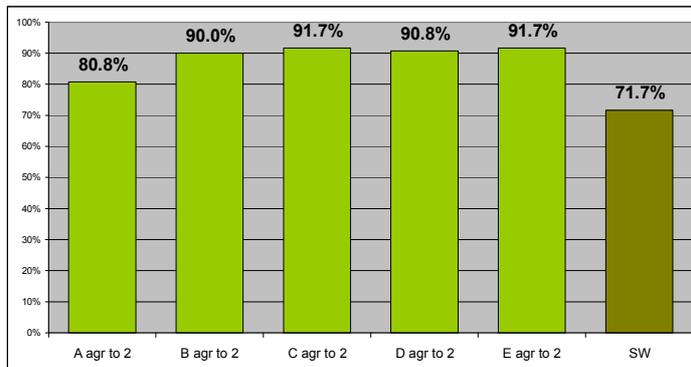


Abb. 20: PR: Prozentsatz der Übereinstimmung mit dem Konsens für Pathologen und Software

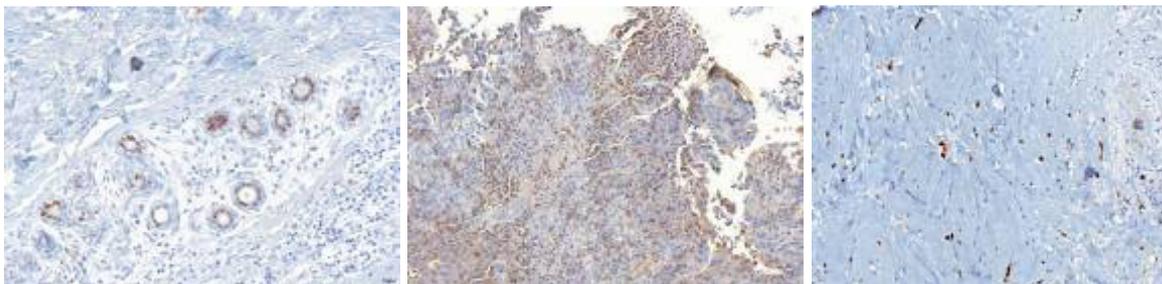


Abb. 21: Ursachen für eine inkorrekte Analyse

Links: Gesundes Brustgewebe ist immer positiv für ER und PR.

Mittel: Artefakte können falsch positive Kerne vortäuschen.

Rechts: Beim lobulären Karzinom ist die Unterscheidung in Tumor-/Nichttumorzelle auch für Pathologen nur schwer zu treffen

Bei der **Her2neu**-Annotation wurde jeder Fall mit einem Score zwischen 0 und 3+ bewertet. Für die Therapie ist aber der Score 0 (keine Färbung) dem Score 1+ (wenig Färbung, Membranen nur teilweise angefärbt) gleichzusetzen, da beide dazu führen, dass eine Therapie mit Herceptin nicht empfohlen wird. Da ein unterschiedlich gesetzter Score hier eine andere – nämlich nur marginale – Bedeutung hat, als bei den höheren Scores, werden für die Auswertung alle Annotation des Scores 0 auf Score 1 gesetzt. Es verbleiben damit noch drei zu vergebene Scores, die jeweils eine unterschiedliche Bedeutung für die Therapie haben. Eine konsistente Vergabe ist demnach von höchster Wichtigkeit.

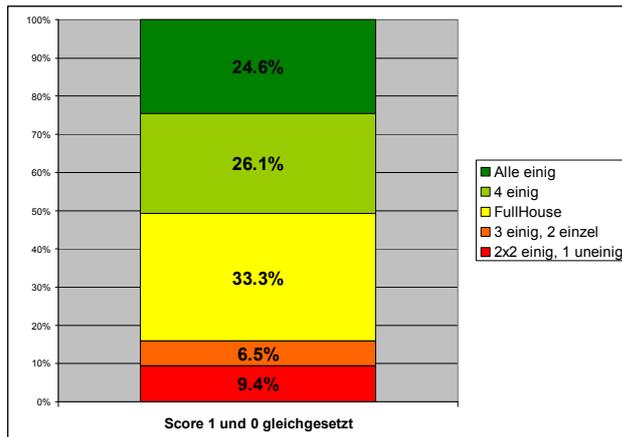


Abb. 22: Unterschiede in der Her2/neu-Annotation der fünf Pathologen: auch nach Gleichsetzung von Score 0 und 1 gibt es in 15,9% der Fälle eine Vergabe von allen drei Scores für einen einzelnen Fall, dabei setzen in 9,4% keine drei Pathologen denselben Score

Auffallend bei der **Her2neu**-Auswertung war der geringe Prozentsatz der Übereinstimmung der Scorevergabe. Dies kann zum einen an unterschiedlichen Färbegewohnheiten liegen. So können Labore unterschiedlich intensiv färben, was die Verteilung der Scores 2+ und 3+, deren Unterschied ausschliesslich auf der Intensität beruht, beeinflusst. Ebenso wird von Pathologen bei Unsicherheit gerne der Score 2+ vergeben, da dieser eine weitere ISH-Analyse nach sich zieht. Einige Tumore zeigten sich auch schwierig bei der Diagnose. Da nur der invasive Anteil des Tumors für die Her2neu-Färbung ausgewertet werden soll, kann es zu Abweichungen kommen, wenn dieser Anteil verschieden beurteilt wird. In dieser Färbung ist es nicht durchweg eindeutig zu klären, welche Anteile schon invasiv oder noch „in situ“ sind (letztere sind oft hoch positiv), sodass verschiedene Pathologen hier zu unterschiedlichen Aussagen kommen können.

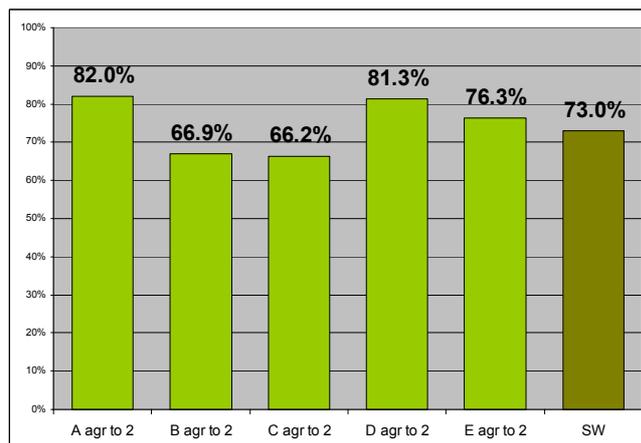


Abb 23: Her2/neu: Prozentsatz der Übereinstimmung mit dem Konsens für Pathologen und Software

Die Auswertung der Software hatte ähnliche Probleme wie die Pathologen. Obwohl wir damit sehr gut im Rahmen der Pathologenabweichung liegen, ist es unser Ziel die Auswertung noch weiter zu verbessern.

Die Auswertung der **SISH**-Bilder war aufgrund der geringeren Bildanzahl erschwert. Auch kam es bei der Annotation durch die Pathologen zu erheblichen Unterschieden in der Bewertung. Ursprünglicher Sinn der Färbung ist es, eindeutig zu klären, ob eine Überexpression des Her2neu-Gens vorliegt und damit ob eine Therapie erfolgreich sein kann oder nicht.

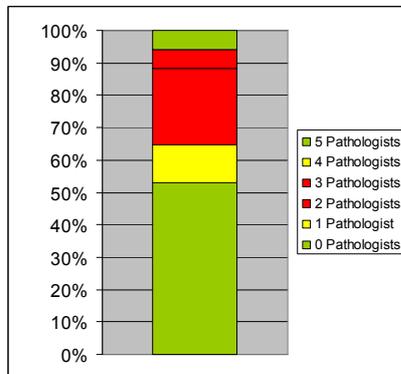


Abb. 24: SISH: Anteil der Fälle, die von mehr oder weniger Pathologen übereinstimmend als positiv bewertet wurden.

In unserem Vergleich der Annotationen gab es aber nur in knapp 60% der Fälle eine Übereinstimmung aller Pathologen, dies vor allem in den negativen – offensichtlicheren – Fällen. Die positiven Fälle wurden von den Pathologen durchweg sehr unterschiedlich bewertet, so dass es auch hier zu erheblichen Therapieunterschieden kommen würde. Dies liegt zum einen daran, dass in den meisten Labors noch die aufwändige FISH(Fluoreszenz-ISH)-Färbung standardmäßig durchgeführt wird, die oft nicht vom diagnostizierenden Pathologen ausgewertet wird. Die Erfahrung mit SISH ist demnach nicht sehr groß. Zum anderen wiesen einige der Bilder eine suboptimale Qualität der Färbung und Scanschärfe auf, so dass Unschärfen zu Abweichungen führen können.

Da auf Grund der geringen Übereinstimmung der Pathologen für die meisten Bilder kein eindeutiger Konsens gesetzt werden konnte, wird der Abgleich mit den Ergebnissen der Software als prozentuale Abweichung in Relation zur Anzahl der positiv bewertenden Pathologen angesetzt (s. Abb. 25).

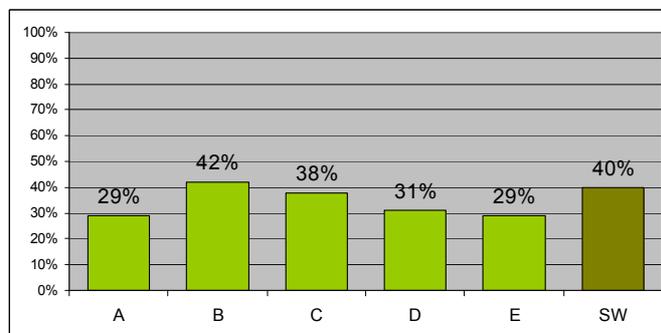


Abb. 25: Prozentuale Abweichung der Bewertung zur unterschiedlichen Abweichung der Pathologen. Ein Wert von 0% wäre eine völlig Übereinstimmung mit dem Durchschnitt, ein Wert von 100% bedeutet keinerlei Übereinstimmung.

Hier zeigt sich, dass die Software sich zwar ähnlich verhält wie die Pathologen, jedoch die Abweichungen untereinander so groß sind (im Schnitt knapp 35%), dass die Auswertung so kaum gewertet werden kann.

Für den Vergleich der **H&E**-Analyse werden die drei Unterpunkte getrennt betrachtet, da der finale Score eine reine Summe dieser Ergebnisse ist. Die durchschnittliche Verteilung der Einzelscores im Vergleich des Pathologenkonsens mit den Ergebnissen der Bildanalysealgorithmen ist bereits sehr vielversprechend.

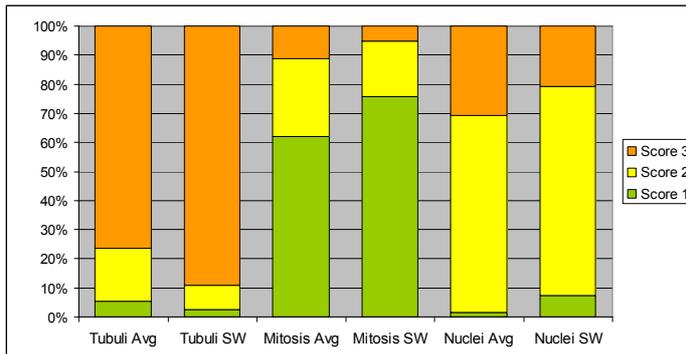


Abb. 26: Durchschnittliche Verteilung der Einzelscores des Pathologenkonsens und Einzelscores der Software im Vergleich

Bei dem Vergleich der Scores für die **Kernmorphologie** fällt die Software gegenüber den Pathologen deutlich ab. Geplante Verbesserungen werden im Bereich der Kerndetektion geschehen, sodass auch gerade die veränderten auffälligen Kerne noch sicherer detektiert werden. Zum anderen wird die Chromatinverteilung in den Zellkernen noch stärker in die Beurteilung einfließen, die sich bisher hauptsächlich auf die Zellkerngröße beschränkte.

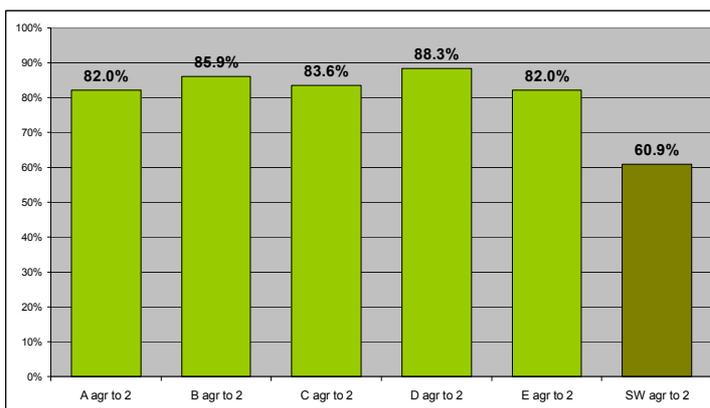


Abb. 27: Vergleich der Übereinstimmung der Scores für die Kernmorphologie

Bei den **Mitosen** wurde ja bereits im AP 4 die äußerst unterschiedliche Auswertung der Pathologen dargelegt. Dieses spiegelt sich auch in der Varianz der Pathologen wieder. Das Ergebnis der Software ist absolut gesehen nicht optimal, erreicht damit aber eine höhere Übereinstimmung zum Konsens als drei der Pathologen. Die unterschiedliche Beurteilung erschwert die Entwicklung eines konsistenten Algorithmus, zeigt aber auch dessen Notwendigkeit.

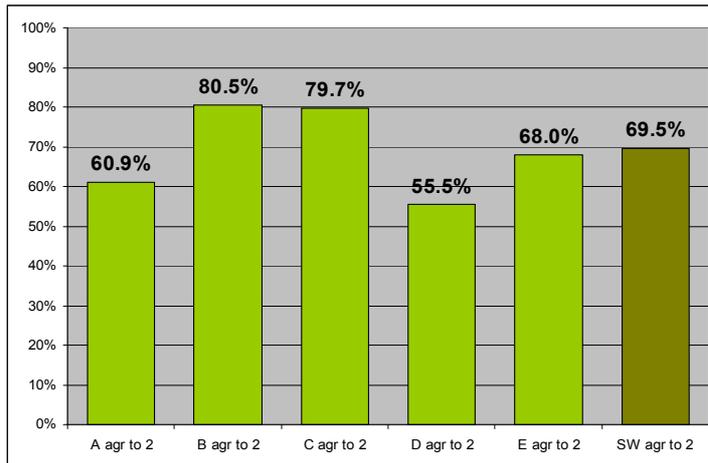


Abb. 28: Vergleich der Übereinstimmung der Scores für die Mitosenzählung

Auch zeigt das Ergebnis, dass die Entscheidung, nur die eindeutigen Mitosen zu werten, soweit korrekt war. Die Pathologen A und D, die nur eine sehr geringe Übereinstimmung mit der Mehrheit haben, haben deutlich öfter Score 2 vergeben als Score 1 (s. Abb 29), finden also mehr Mitosen. Hier werden die unterschiedlichen Bewertungsarten – restriktiv und weniger restriktiv – deutlich. Die Leitlinie fordert daher auch eindeutig den restriktiven Weg, um eine zu große Streuung zu vermeiden.

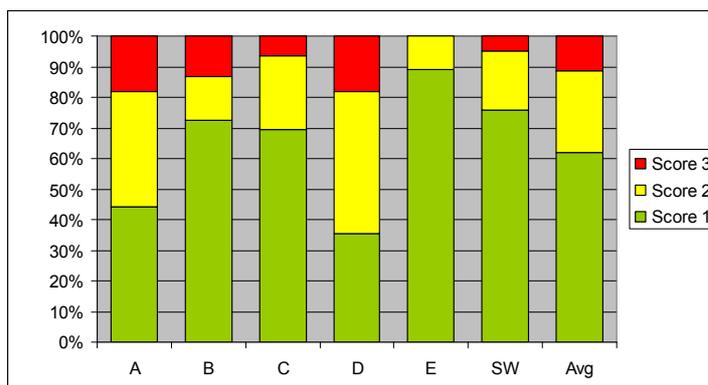


Abb. 29: Verteilung des Scores für die Mitosendetektion der einzelnen Pathologen, der Software und des Durchschnitts der Pathologen.

Bei der **Tubulibildung** zeigte sich, dass nur ein kleinerer Teil der Tumore in unserem Bildersatz überhaupt noch Tubuli ausbildeten (im Durchschnitt der Pathologen: 23,6%), so dass hier die Analyse nur auf einem geringeren Satz von Bildern trainiert werden konnte. Der weitaus größte Teil der Karzinome unseres Satzes ist bereits soweit entdifferenziert, dass keine Tubulibildung mehr festgestellt werden kann. Das Ergebnis der Software ist annähernd so gut wie das der Pathologen untereinander, sollte aber noch anhand zusätzlicher Bilder mit Karzinomen mit Tubulibildung verbessert werden.

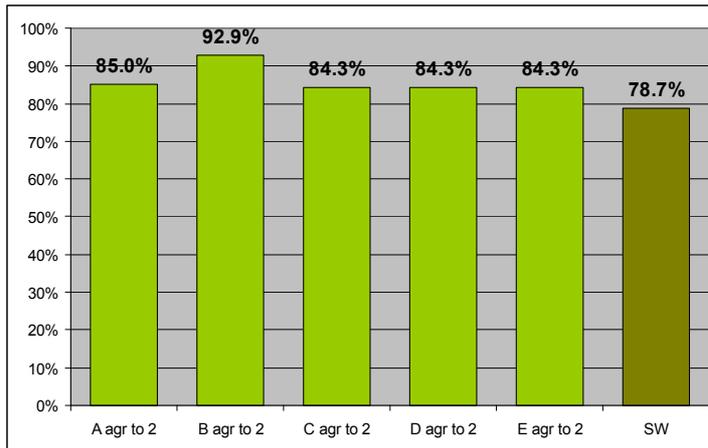


Abb 30: Vergleich der Übereinstimmung der Scores für die Tubulibildung

Fazit:

Die generelle Übereinstimmung mit den Pathologen in den einzelnen Färbungen ist bereits gut bis sehr gut. Jedoch muß die Qualität für ein marktfähiges Produkt noch verbessert werden, um zuverlässig innerhalb der Varianz der Pathologen zu liegen. Insbesondere bei der H&E-Färbung sind noch Erweiterungen notwendig.

Visualisierung der Koregistrierung

Beim Validierungsworkshop in München wurden auch die Bilder besprochen, die eine derartige Schwankung in der diagnostischen Auswertung aufwiesen, dass sie auf Grund des mangelnden Konsens aus dem Vergleich zur Software ausgeschlossen werden mußten. Mit einem Prototyp der neuen Visualisierungs-Oberfläche, die bereits die Koregistrierung zeigte, wurden diese Bilder von den fünf Pathologen und zusätzlich noch von Prof. Dr. Manfred Dietel (Direktor des Instituts für Pathologie der Charité) in drei Teams begutachtet, um vor Ort einen Konsens zu erreichen und die Schwierigkeiten bei den entsprechenden Bildern zu erörtern.

Ein Beispiel ist der Fall 005 unserer Bilddaten in der Färbung ER. Hier liegt der Tumor dicht unter der Haut und das Gewebe ist durch Artefakte derartig unregelmäßig, dass es für die Pathologen sehr schwer war, den Tumor eindeutig zu identifizieren. Jedoch ist es für die Auszählung der positiven Kerne sehr wichtig, die negativ gefärbten – blauen – Kerne soweit zuzuordnen, dass ein korrekter Prozentsatz berechnet werden kann. Die fünf Annotationen umfassten demnach folgende Prozentzahlen: 5% – 20% – 70% – 80% – 90%. Dies ergibt einen Durchschnitt von 53%, jedoch mit einer viel zu hohen Standardabweichung von 34%. Die Software wertete das Bild mit 4% aus, was auf den ersten Blick nicht korrekt erscheint.

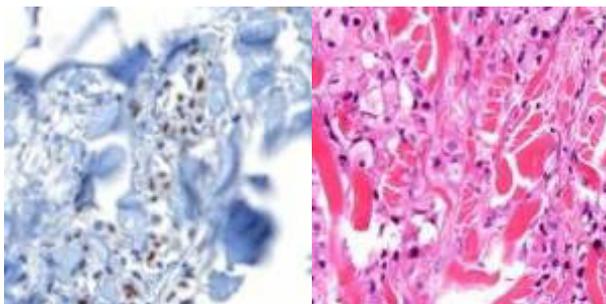


Abb 31: Fall 005, links in ER Färbung mit 34% Standardabweichung in der Pathologenannotation, rechts in der H&E Färbung, deren direkter Vergleich eine konsistente Auswertung erlaubte

Bei der gemeinsamen Besprechung des Bildes offenbarte sich das Problem, dass der Tumor hier nicht identifiziert werden kann. Mit unserem Prototypen konnten die Pathologen jedoch in die relevante Region hineinzoomen und über die „Multi-View“-Ansicht direkt in die entsprechende Region des H&E-Schnittes schalten. Hier konnte die Identifikation des Tumors erfolgen, so dass nun alle drei Teams übereinstimmend den Fall mit 5-10% positiver Kerne auswerteten. Hier liegt nun unsere Software sehr gut im selbem Bin (1-10%) und die Qualität der histopathologischen Diagnose konnte eindrucksvoll gesteigert werden.

2. Positionen des zahlenmäßigen Nachweises

Gemäß dem Finanzierungsplan des Projektantrags wurde das Projekt zu 50% finanziert. Dafür wurde inklusive Reisekosten eine Zuwendung von 471.168€ beantragt.

Diese Zuwendung schlüsselt sich auf in folgende Unterpunkte des Finanzierungsantrags:

I. Punkt 0823: FE Fremdleistungen

Unter diesen Punkt wurden 111.500€ für Fremdleistungen der Charité Berlin veranschlagt. Diese wurden für das Aussuchen, Färben und die Qualitätssicherung der 150 Brustkrebsfälle in 7 Färbungen sowie für die ständige begleitende Beratung in medizinischen Fachfragen während der Entwicklung verwendet. Ein weiterer Punkt ist das Einscannen der Bilder, das von der VMScope GmbH durchgeführt und mit dieser abgerechnet wurde.

Hier wurden letztendlich tatsächlich 111.640,14€ verwendet.

II. Punkt 0837: Personalkosten

Als Personalkosten wurden von der Definiens AG insgesamt 789.337€ berechnet. In diesen Punkt fallen Projektleitung, Entwicklung der GUI und der Schnittstellen und Entwicklung der Bildanalysealgorithmen und des Scoring Verfahrens. Dazu waren bei Definiens ein C++-Softwareentwickler, ein Bildanalyse-Entwickler und ein leitender Akademiker für das Projekt zuständig. Zwischenzeitlich wurde die Arbeit auf mehrere Entwickler verteilt, um in kürzerer Zeit zu dem gewünschten Ergebnis zu gelangen.

Hier wurden letztendlich tatsächlich 789.705,86€ verwendet.

III Punkt 0850: sonstige unmittelbare Vorhabenkosten

Hierunter fällt der im Arbeitspaket 6 angesetzte Validierungsworkshop mit den Annotationen der 5 Pathologen. Ursprünglich wurden hier 41.250€ genehmigt.

Im Antrag auf Entsperrung der Mittel für den Workshop wurde der zu erwartende Betrag nochmal auf 30.800€ gesenkt, da die Workshopdauer deutlich verkürzt wurde. Im Rahmen des Workshops annotierten 5 Pathologen die vorliegenden Fälle unabhängig voneinander. Im Oktober 2012 fand ein gemeinsamer Workshop bei der Definiens AG statt.

Die tatsächliche Abrechnung inklusive der Zeitaufwände der Pathologen, Reisekosten und Spesen betrug 29.079,54€. Hier konnte also gegenüber dem ursprünglichen Plan deutlich gespart werden.

3. Notwendigkeit und Angemessenheit der geleisteten Arbeit

Im Rahmen des Projektes wurde während der Validierungsphase deutlich, dass die Schwankungsbreite zwischen den diagnostizierenden Pathologen zum Teil erheblich ist. Auch wenn unterschiedliche Laborbedingungen teilweise die Unterschiede erklären mögen, so sind die Schwankungen doch so groß, dass die Gefahr, in unterschiedlichen Laboren eine unterschiedliche Therapieempfehlung zu erhalten, durchaus realistisch ist.

Diese Situation zeigt die Notwendigkeit, die histopathologische Diagnose auf konsistente und objektive Grundlage zu stellen und den Pathologen derartig zu unterstützen, dass er seine Diagnose reproduzierbar und nachvollziehbar zum Wohle des Patienten stellen kann.

Der im Projekt entwickelte Prototyp kann diese Unterstützung leisten. Wie gezeigt kann der direkte Vergleich von hochaufgelösten Regionen des Tumors in unterschiedlichen Färbungen, insbesondere zur H&E-Färbung eine signifikante Verbesserung der Auswertung und Diagnose des Falles bewirken. Durch die automatische Analyse der Färbungen wird dem Pathologen zusätzlich eine objektive Bewertung zur Verfügung gestellt, die er bei seiner Diagnosestellung und Quantifizierung zu Rate ziehen kann.

4. Voraussichtlicher Nutzen

Auf Basis des im Rahmen dieses Projektes erstellten Prototyps wird derzeit in der Definiens AG ein Produkt erstellt, dessen Ziel es sein wird, den Pathologen bei der Erstellung einer konsistenten und reproduzierbaren Diagnose zu unterstützen. Dieses Produkt wird eine CE-Zertifizierung erhalten und wird voraussichtlich Ende 2014 auf den Markt kommen.

Die Definiens AG hat bereits langjährige Erfahrung auf dem Gebiet der vollautomatischen Analyse von histopathologischen Schnitten in der Forschung. Mit dem neuen Produkt können in der täglichen Routinediagnostik neue Standards zur Erstellung der Diagnose von Karzinomen gesetzt werden.

5. Während der Durchführung des Vorhabens bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

Im Verlauf des Projekts wurden auch von anderen Unternehmen im Bereich Digitale Pathologie Lösungen zur automatisierten Auswertung von H&E, ER, PR und Her2neu entwickelt. Diese Lösungen beziehen sich allerdings immer auf von Pathologen vorannotierte Bereiche, was dem Aspekt der Standardisierung und Objektivierung der Befundung nicht gerecht wird. Der im Projekt entwickelte Prototyp mit der vollautomatischen Befundung ganzer Gewebeschnitte und der koregistrierten Visualisierung mehrerer Schnitte eines Falls ist das Alleinstellungsmerkmal von Definiens.

6. Veröffentlichungen

Folgende Publikationen von Mitarbeitern der Definiens AG als Hauptautor wurden angefertigt:

- Präsentation des Projektstatus auf der 3. Fachtagung des BMBF zur Fördermaßnahme "KMU-innovativ: IKT" am 17. und 18. September 2012 in Dresden.

Berichtsblatt

1. ISBN oder ISSN geplant	2. Berichtsart (Schlussbericht oder Veröffentlichung) Schlussbericht
3. Titel "KMU innovativ - Verbundprojekt: Histoquant Mamma – Bildanalyse zur Molekulardiagnostik von Brustkrebs	
4. Autor(en) [Name(n), Vorname(n)] Winkler, Bettina Schmidt, Günter Binnig, Gerd	5. Abschlussdatum des Vorhabens 30.11.2012 6. Veröffentlichungsdatum Juni 2013 7. Form der Publikation Technischer Bericht
8. Durchführende Institution(en) (Name, Adresse) Definiens AG Research Bernhard-Wicki-Straße 5 80636 München	9. Ber. Nr. Durchführende Institution -- 10. Förderkennzeichen 01IS10021 11. Seitenzahl 27
12. Fördernde Institution (Name, Adresse) Bundesministerium für Bildung und Forschung (BMBF) 53170 Bonn	13. Literaturangaben -- 14. Tabellen -- 15. Abbildungen 31
16. Zusätzliche Angaben --	
17. Vorgelegt bei (Titel, Ort, Datum) --	
18. Kurzfassung Die primäre Diagnose von Brustkrebs in der Histopathologie erfolgt standardmäßig mit Hilfe von H&E-gefärbten Schnitten einer Biopsie. Eine weitere Einteilung und Untersuchung des hier gefundenen Karzinoms erfolgt mittels weiterer, immunhistochemischen (IHC) Färbungen, auf deren quantitative Auswertung sich eine nachfolgende Therapieempfehlung maßgeblich stützt. Eine objektive und standardisierte histopathologische Auswertung ist entscheidend für den Therapieerfolg, insbesondere die Kombination der IHC-Auswertungen im Rahmen der Personalisierten Medizin. Im Projekt „Histoquant Mamma“ wurden anhand von 150 Brustkrebsfällen mit jeweils 5-7 Färbungen Bildanalyse-Algorithmen entwickelt, die alle Färbungen entsprechend der üblichen Leitlinien vollautomatisch analysieren und quantifizieren. Der Projektvalidierungsworkshop hat gezeigt, dass die Bildanalyse die Abweichungen der individuellen Fallbewertungen verschiedener Pathologen minimiert, was zu konsistenteren Therapieempfehlungen führt. Neben der histopathologischen Standardauswertung ist es mit dem erarbeiteten Softwareprototyp möglich, über eine intuitiv zu bedienende Benutzeroberfläche Aussagen über die Tumoheterogenität zu gewinnen und für einzelne Regionen des Tumors dank Visualisierung der Koregistrierung der einzelnen Schnitte unmittelbar die verschiedenen Färbungen direkt zu vergleichen. Als Methode kommt die Definiens Cognition Network Technology zum Einsatz, die im Zuge des Projekts angepasst und erweitert wurde.	
19. Schlagwörter Bildanalyse, Histopathologie, Brustkrebs, Her2neu, Koregistrierung, IHC, H&E	
20. Verlag --	21. Preis --

Document Control Sheet

1. ISBN or ISSN planned	2. type of document (e.g. report, publication) final report
3. title "KMU innovativ - Verbundprojekt: Histoquant Mamma – Image analysis of molecular diagnostics in carcinoma of the breast"	
4. author(s) (family name, first name(s)) Winkler, Bettina Schmidt, Günter Binnig, Gerd	5. end of project 30.11.2012
	6. publication date June 2013
	7. form of publication Technical report
8. performing organization(s) (name, address) Definiens AG Research Bernhard-Wicki-Straße 5 80636 München	9. originator's report no. --
	10. reference no. 01IS10021
	11. no. of pages 27
12. sponsoring agency (name, address) Bundesministerium für Bildung und Forschung (BMBF) 53170 Bonn	13. no. of references --
	14. no. of tables --
	15. no. of figures 31
16. supplementary notes --	
17. presented at (title, place, date) --	
18. abstract The first diagnose of breast cancer in histopathology is done on the H&E stained tissue section of a biopsy. A more detailed analysis and evaluation of the detected carcinoma is performed using more immunohistochemical (IHC) stains. The quantitative analysis of those stains supports the decision about the suggested therapy. Therefore, an objective and standardized histopathological analysis is important for the success of a therapy, especially the combination of the IHC evaluations in view of the individual, personalized medicine. In this project „Histoquant Mamma” we analyzed 150 cases of breast cancer in 5-7 stains using new image analysis algorithms to quantify all stains fully automatically and on the whole slide. Within a validation workshop with five pathologists it was shown, that the software was able to minimize the significant variation of the individual assessments, improving the resulting therapy suggestion. We developed a prototype, which allows the pathologist using an intuitive GUI to rate the tumor heterogeneity and to compare small regions of a tumor within the different stains easily by visualisation of co-registration of the single stains. The analysis is based on the Definiens Cognition Network Technology, which was adapted and enlarged in course of the project.	
19. keywords image analysis, histopathology, breast carcinoma, Her2, image registration, IHC, H&E	
20. publisher --	21. price --