

# Prototyp eines Mammographie-CAD-Systems auf Basis der Cognition Network Technology

Ralf Schönmeier<sup>1</sup>, Maria Athelougou<sup>1</sup>, Harald Sittek<sup>2</sup>, Peter Ellenberg<sup>1</sup>,  
Owen Feehan<sup>1</sup>, Günter Schmidt<sup>1</sup>, Gerd Binnig<sup>1</sup>

<sup>1</sup>Definiens AG, Trappentreustraße 1, 80339 München

<sup>2</sup>Diagnostisches Mammazentrum München, Feringastrasse 10b, 85774 München  
rschoenmeyer@definiens.com

**Kurzfassung.** In dieser Arbeit wird ein neues Assistenzsystem für die computerunterstützte Detektion und Diagnose von Herdbefunden in der Mammographie vorgestellt, das mit Hilfe der Cognition Network Technology (CNT) eine bildbasierte Ähnlichkeitssuche in einer Referenzdatenbank ermöglicht. Dabei wird vorverarbeitend vollautomatisch eine mehrskalige Segmentierung und Klassifikation auffälliger Objekte durchgeführt, deren Ergebnisse zusammengefasst in der Datenbank zur Verfügung stehen. Die Ähnlichkeitssuche lässt sich interaktiv konfigurieren und der Prototyp ist offen für Erweiterungen und die Einbindung von Daten weiterer Modalitäten.

## 1 Einleitung

Die Früherkennung von Brustkrebs anhand von Mammographiedaten stellt eine wichtige und anspruchsvolle Aufgabe für den Radiologen dar. Hauptanhaltspunkte, ob ggf. eine gut- oder bösartige Brustläsion vorliegt, liefern typischerweise verschiedene Erscheinungsformen von Kalzifikationen und Verdichtungen, den sogenannten Herdbefunden. Da deren morphologische Variabilität sehr groß ist, lassen sich sichere Aussagen über Gut- oder Bösartigkeit oftmals nicht alleine aus der Bildgebung ableiten und erst eine Biopsie liefert im Zweifel eine sichere Diagnose. Die bereits verfügbaren Systeme zur computerunterstützten Detektion und Diagnose (CAD) von Brustläsionen basieren hauptsächlich auf Bildanalysealgorithmen und helfen dem Radiologen auffällige Gebiete in Bildern nicht zu übersehen [1, 2]. Unser Ansatz verfolgt die Strategie, zu einem gegebenen Patientenfall automatisiert in einer Referenzdatenbank relevante ähnliche Fälle zu finden und anhand zugehöriger Referenzbefunde die Qualität von Diagnosen zu steigern. Dazu wurde die CNT dahingehend erweitert, nicht nur wie bisher Bilddaten [3, 4], sondern auch Daten, die in Tabellenform vorliegen und auch textwertige Einträge aufweisen können, zu verarbeiten. Dies ermöglicht die Erstellung eines Datenbanksystems, das die Ergebnisse von bildanalytischen Untersuchungen und anderen Daten, wie z.B. Patienten-Meta-Daten, enthält. Mit CNT werden dann die Beziehungen dieser Einträge heterogenen Ursprungs für eine kontextgetriebene Analyse genutzt.

## 2 Material und Methoden

Der vorliegende CAD-Prototyp wurde auf Basis der in [5] vorgestellten Mammographie-Datenbank mit detaillierten Annotationen und bioptisch abgesicherten Befunden entwickelt. Als Entwicklungsplattform dient die CNT-Software *Definiens XD 1.2.0* [6, 7]. Um der großen Variabilität von Herdbefunden Rechnung zu tragen und für die weitere Verarbeitung adäquate Bildobjekte zu liefern, kommen bei der vorverarbeitenden Bildanalyse der einzelnen Mammogramme unterschiedliche Segmentierungsverfahren zum Einsatz. Daraus gewonnene Merkmalsvektoren werden in einer Datenbank zusammengefasst. Diese direkten Objekteigenschaften (wie z.B. Fläche und Helligkeit) stellen zusammen mit daraus abgeleiteten Kontexteigenschaften (z.B. Distanz zur ebenfalls automatisch detektierten Mamilla) die Entscheidungsgrundlage für einen Klassifikator zur Auffindung von auffälligen Verdichtungen dar. Neben der vollautomatischen Segmentierung und Klassifikation wird ein Konzept umgesetzt, mit dem sich ein Maß für *Ähnlichkeit* konfigurieren lässt, das dem Benutzer zu einem gegebenen Mammographie-Fall auf Knopfdruck *ähnliche* Fälle liefert. Im Folgenden werden die einzelnen Programmmodule etwas genauer skizziert.

### 2.1 Vorverarbeitung

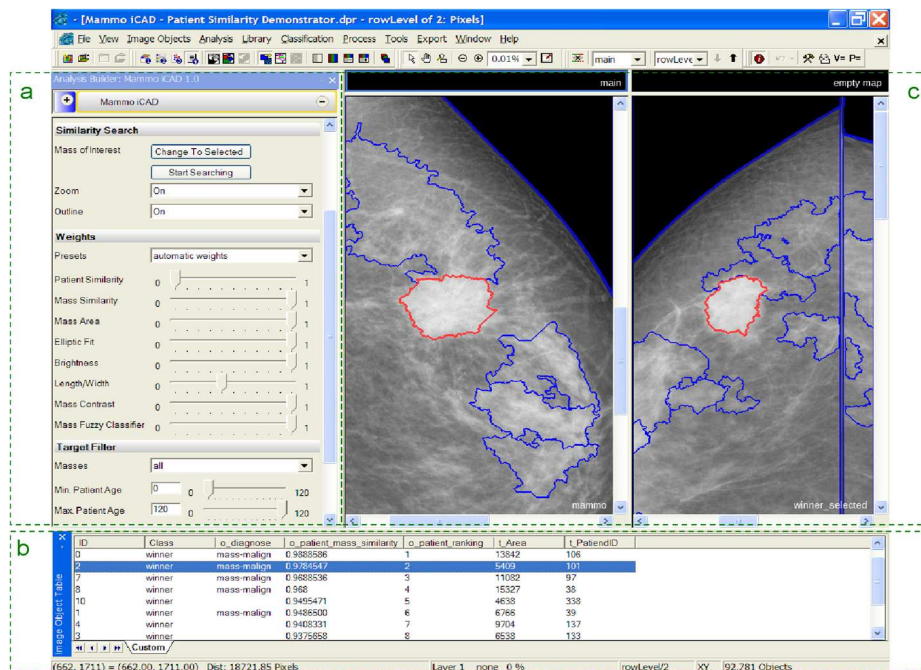
Zunächst erfolgt eine Normalisierung der Grauwerte aller Eingangsbilddaten, um die unterschiedlichen Gegebenheiten verschiedener Gerätehersteller zu berücksichtigen. Es schließt sich eine Segmentierung der Brustkontur, die sich im jeweiligen Mammogramm als Kante darstellt und das Bild in Vorder- und Hintergrund unterteilt, an. Auf ihr erfolgt eine Abschätzung für die Position der Mamilla als dem Punkt mit dem größten Abstand zu einer Basislinie, die in cranio-caudalen (CC) Ansichten der rechten oder linken Brust mit dem rechten bzw. linken Bildrand zusammenfällt. In medio-lateralen (ML) Ansichten ist diese Basislinie um  $-15^\circ$  bzw.  $15^\circ$  gedreht, was sich empirisch für eine ausreichend robuste Erkennung auf allen zur Verfügung stehenden Mammogrammen bewährt hat. Wenn in ML-Ansichten ein Teil des Pektoralmuskels sichtbar ist, so zeichnet sich dieser in der Regel als eine ausgeprägte Helligkeitsstufe aus. Diese kann mit Hilfe eines Kantenfilters auf einem stark geglätteten Bild detektiert werden. Die Abschätzung des Pektoralmuskels wird nach einer Plausibilitätsprüfung durch eine lineare Extrapolation der Hauptrichtung dieser Kante sowie dem Maskieren des Bereichs links (in linken Brustansichten) bzw. rechts (in rechten Brustansichten) davon vervollständigt. Für die Gewinnung von Kandidatenobjekten von potentiellen Herdbefunden werden fünf unterschiedliche Segmentierungen durchgeführt: ein Verfahren, das ausgehend von Startregionen mittels morphologischer Operationen auf relativ kontrastreiche Objekte abzielt, wird für vier verschiedene Größenskalen angewandt. Ein alternatives Verfahren segmentiert vorwiegend Objekte, die sich im Mammogramm recht kontrastarm abzeichnen und in einer relativ homogenen Umgebung befinden.

## 2.2 Tabellenverarbeitung

Die exportierten Daten der Bildanalyse werden in Tabellenform (CSV-Format) zusammen mit Patienten-Metadaten (wie z.B. Alter) als relationales Datenmodell in die Definiens Software geladen, bei dem zugehörige Daten miteinander „verlinkt“ sind. Dadurch stehen sämtliche Informationen in einem einheitlichen Zugriffsmodell zur Verfügung und es lassen sich auch erweiterte Merkmale formulieren, die Beziehungen zwischen Objekten unterschiedlicher Bilder ausnutzen. Diese Merkmale liefern z.B. ein Maß dafür, ob Kandidatenobjekte für eine Verdichtung im selben Abstand zur Mamilla sowohl in der CC- als auch ML-Aufnahme der selben Brust auftreten. Die Klassifikation auffälliger Herdbefunde erfolgt mittels eines hierarchisch aufgebauten Fuzzy-Klassifikators, der sich im Wesentlichen aus zehn Unterklassifikatoren für unterschiedliche Merkmalsprofile zusammensetzt.

## 2.3 Ähnlichkeitssuche

Die Suche nach ähnlichen Fällen zu einem gegebenen Fall basiert auf einem Maß für Patientenähnlichkeit, das sich hierarchisch aus Herdbefund-, Brustansicht- und Patienten-Metadaten-Ähnlichkeit zusammensetzt. Es werden jeweils



**Abb. 1.** Benutzerschnittstelle für die Ähnlichkeitssuche mit der Möglichkeit, die Suchanfrage zu konfigurieren (a), die Ergebnisse auszuwählen (b) und anzuzeigen (c).

fuzzybewertete Differenzen zwischen Merkmalswerten des aktuellen Anfragefalls mit allen Vergleichsfällen berechnet. Bei der Herdbefund-Ähnlichkeit kommt der zuvor beschriebene Klassifikator zum Einsatz, der eine auffällige Verdichtung vorschlägt. Diese wird dann für die Ähnlichkeitssuche mit den Kandidatenobjekten aller anderen Patienten der Datenbank herangezogen. Die Suchanfrage lässt sich interaktiv mittels Gewichtsvariablen konfigurieren, um den Einfluss einzelner Merkmale regeln zu können (Abb. 1 a). Als Ergebnis der Suche werden die für die aktuelle Konfiguration 10 ähnlichsten Fälle aufgelistet (Abb. 1 b) und die zugehörigen Bilddaten per Mausklick angezeigt (Abb. 1 c). Da die Befunde der Referenzfälle in der Datenbank ebenfalls vorliegen, erfolgt mit deren Hilfe ein automatisierter Diagnosevorschlag, der auf einem Mehrheitsvoting beruht.

### 3 Ergebnisse

Die Entwicklungsdatenbank, die auf sämtlichen Bildern aus [5] basiert, enthält ca. 58.000 Herdbefund-Kandidatenobjekte von 303 Patienten. Davon werden 314 Objekte, die mit Annotationen übereinstimmen, vom Klassifikator detektiert. Diese verteilen sich auf 90 unterschiedliche Patienten von den insgesamt 112 mit auffälligen Herdbefunden, was einer Trefferquote von 80% entspricht. Der Klassifikator liefert im Schnitt 0.66 falsch-positive Objekte pro Bild (677 Objekte in 1024 Brustansichten). In einer Testdatenbank, die ca. 34.000 Kandidatenobjekte für Herdbefunde von 120 neuen Patientendatensätzen enthält, detektiert der identische Klassifikator für 7 von den 15 ausschließlich benign annotierten Patienten richtig erkannte Herdbefunde. Bei einer Anzahl von 541 falsch-positiven Objekten aus 480 Bildern entspricht dies durchschnittlich 1.13 falsch-positiven pro Bild. Eine Ähnlichkeitssuche dauert rund 10 Sekunden für den Vergleich mit allen insgesamt 423 Patientenfällen und skaliert linear. Für den Test der Diagnose wurde für jeden dieser Fälle eine Ähnlichkeitssuche mit allen anderen Patienten durchgeführt, um nur anhand deren Referenzbefunde den Diagnosevorschlag zu ermitteln und dann mit dem eigenen Referenzbefund zu vergleichen. Es wurden zwei Szenarien getestet: ein Drei-Klassenproblem, bei dem für die Diagnose die Klassen *malign*, *benign* und *unauffällig* zur Verfügung standen (Tab. 1(a)) und einmal als Zwei-Klassenproblem mit einer Beschränkung des Suchraums auf nur *maligne* und *benigne* Vergleichsobjekte (Tab. 1(b)).

### 4 Diskussion

Die quantitative Leistung der Detektion von Herdbefunden erreicht das Niveau bestehender Verfahren [8], obwohl unser System auf Grundlage einer vergleichsweise großen Datenbasis auch die Erzeugung der ROIs (*regions of interest*) vollautomatisch leistet. Der Benutzer hat die Möglichkeit, diesen Vorschlag für die Ähnlichkeitssuche zu überstimmen und diese Information auch zu speichern. Dadurch „lernt“ die Referenzdatenbank und gewinnt an Wert für künftige Suchanfragen. Die Ähnlichkeitssuche liefert relevante Vergleichsfälle und erlaubt es dem Benutzer somit, die Referenzdatenbank zu durchsuchen, wie es mit manueller

**Tabelle 1.** Konfusionsmatrix für das (links) Drei- und (rechts) Zwei-Klassenproblem mit den relativen Anteilen der Klassifikation basierend auf allen 423 Patientendatensätzen der Datenbank bzw. von den insgesamt 97 (66 malignen und 31 benignen) automatisch detektierten Patienten mit auffälligen Herdbefunden.

Diagnose- vorschlag	Referenzdiagnose			Diagnose- vorschlag	Referenzdiagnose	
	malign	benign	unauffällig		malign	benign
malign	36%	6%	4%	malign	76%	25%
benign	0%	0%	0%	benign	25%	75%
unauffällig	64%	94%	96%			

Durchsicht nicht zu leisten ist. Für eine Verbesserung der Ergebnisse des automatischen Diagnosevorschlags wäre es wünschenswert, wenn eine umfangreichere Referenzdatenbank zur Verfügung stünde, um überhaupt adäquat ähnliche Fälle für die große Variabilität von Herdbefunden zu enthalten. Bereits während der Entwicklungszeit stellte der Prototyp eine leistungsfähige Umgebung zum Entwickeln, Testen und Evaluieren von Bildanalysen auf Basis vieler Patientendatensätze dar. Die Konzepte, die dabei erarbeitet wurden, sind nicht nur auf die Anwendung für die Mammographie beschränkt und weisen durch ihre Erweiterbarkeit den Weg zu CAD-Systemen, die auf multimodaler Bild- und Datenanalyse beruhen und unserer Meinung nach künftig einen großen Stellenwert einnehmen werden.

**Danksagung.** Diese Arbeit wurde von der Bayerischen Forschungsförderung im Rahmen des Projekts Mammo-iCAD gefördert.

## Literaturverzeichnis

1. Destounis S, et al. Special session on breast CAD. *Int J CARS*. 2007;2(Suppl 1):S330–S350.
2. Doi K. Special session on breast CAD. *Int J CARS*. 2009;4(Suppl 1):S171–S176.
3. Schönmeier R, et al. Automated segmentation of lateral ventricles from human and primate magnetic resonance images using cognition network technology. *Magn Reson Imaging*. 2006;24(10):1377–87.
4. Schönmeier R, et al. Automatische Segmentierung des Corpus Callosum aus sagittalen Schichten von kernspintomographischen Datensätzen. In: *Proc BVM*; 2007. p. 389–93.
5. Elter M, et al. Referenzdaten für die computerassistierte Diagnose in der Mammographie. In: *Proc BVM*; 2007. p. 96–100.
6. Schäpe A, et al. Fraktal hierarchische, prozeß- und objektbasierte Bildanalyse. In: *Proc BVM*; 2003. p. 206–10.
7. Schmidt G, et al. Cognition network technology for automated holistic analysis in mammography. In: *Proc BVM*; 2007. p. 282–7.
8. Lladó X, et al. A textural approach for mass false positive reduction in mammography. *Comput Med Imaging Graph*. 2009;33:415–22.