

Texturanalyse zur Detektion gruppierter Mikroverkalkungen bei der Brustkrebsfrüherkennung

T.O. Müller, R. Stotzka, D. Höpfel* und H. Yang*

Hauptabteilung Prozeßdatenverarbeitung und Elektronik
Forschungszentrum Karlsruhe, 76344 Eggenstein-Leopoldshafen

*Institut für Mechatronik und Naturwissenschaften

Fachhochschule Karlsruhe, 76133 Karlsruhe

Email: {Tim.Mueller, Stotzka}@hpe.fzk.de
Dieter.Hoepfel@fh-karlsruhe.de

Zusammenfassung. Bei der Untersuchung von Röntgenbildern der Brust dienen Mikroverkalkungen als erste Hinweise auf einen Tumor. Durch die Unauffälligkeit dieser Merkmale werden 10-30% der Erkrankungen übersehen und daher die Heilungsaussichten drastisch verringert. Ein Zweitgutachter in Form eines Computers kann diese Rate verringern und den Radiologen auf Abweichungen aufmerksam machen. Ein Computer kann für eine Region of Interest (ROI) aus einem Röntgenbild der Brust entscheiden, ob diese Mikroverkalkungen enthält. Für ein solches Klassifikationsproblem müssen aus der ROI Merkmale gewonnen werden. Wir stellen eine neue Methode vor, mittels Texturanalyse automatisch Merkmale für die Klassifikation zu gewinnen.

1 Brustkrebs

Brustkrebs bzw. das Mammakarzinom ist die häufigste Krebserkrankung unter der weiblichen Bevölkerung in der westlichen Welt. Jede zehnte Frau muß damit rechnen, in ihrem Leben mit dieser Erkrankung konfrontiert zu werden. Die wichtigste Voraussetzung bei der Behandlung von Brustkrebs ist eine möglichst frühzeitige Diagnose. Die Heilungschancen reduzieren sich drastisch, sobald das Karzinom Tochtergeschwulste (Metastasen) streut. Ein Tumor mit einem mittleren Durchmesser von 1 cm hat mit einer Wahrscheinlichkeit von ca. 30% Metastasen hervorgebracht, ein Tumor mit mittlerem Durchmesser von 2 cm bereits mit 60% [1]. Die meisten Erkrankungen werden von den Betroffenen selbst durch einen Tastbefund festgestellt und haben zu diesem Zeitpunkt einen mittleren Durchmesser von über 2 cm. Aus diesem Grunde werden große Anstrengungen unternommen, die Erkrankung so früh wie möglich zu erkennen.

Eine anerkannte Technik der Vorsorge ist die Brustkrebsreihenuntersuchung (Mammographie-Screening). Dabei werden Röntgenbilder der weiblichen Brust (Mammogramm) angefertigt. Von jeder Brust werden 2 Aufnahmen von oben (cranio-caudal, cc) und schräg von der Seite (oblique, o) erstellt. Ein Standardmammogramm digitalisiert mit einer Auflösung von 450 dpi hat etwa eine Ausdehnung von 4000 auf 3000 Bildpunkten. Diese Mammogramme werden von einem Radiologen auf winzige Gewebeveränderungen hin untersucht, welche erste Anzeichen auf einen Tumor sein

können. Eine der dabei gesuchten Veränderungen ist das Auftreten gruppierter Mikroverkalkungen. Eine durchschnittliche Verkalkung hat eine Größe von etwa 10 Bildpunkten. Gruppierete Mikroverkalkungen können auf Tumore im Frühstadium mit einer Größe von wenigen Millimetern hinweisen. Zum Beispiel können Gruppen in schlauchähnlicher Form mit 5 bis 20 Verkalkungen erste sichtbare Anzeichen bei Karzinomen in Milchgängen (ductales Karzinom in situ) sein [2]. Solche Gruppen haben eine Länge von 0,5 bis 1 cm. Wird solch eine Gruppe rechtzeitig erkannt, ist das Risiko einer Metastasierung wesentlich geringer als bei einem späteren Tastbefund.

Gruppierete Mikroverkalkungen werden von Radiologen manchmal übersehen, da sie aufgrund geringem Kontrast und überlagertem Gewebe schwer detektierbar sind. Bis zu 30% der Röntgenbilder werden daher falsch beurteilt [3]. Deshalb ist eine zweite Untersuchung durch einen weiteren Experten ratsam. Da ein zweiter Experte die Kosten erheblich steigert, empfiehlt sich der Einsatz von computergestützter Diagnose zur zweiten Untersuchung. Ein Computer unterstützt den Radiologen bei der Detektion der Mikroverkalkungen.

2 Klassifikation

Als Klassifikation bezeichnet man die Zuordnung eines Musters zu einer bestimmten Kategorie. Ein Computersystem, welches eine Klassifikation durchführt, wird als Klassifikationssystem bezeichnet. Im Falle der Mammographieuntersuchung wird „überwachtes Lernen“ für die Klassifikation eingesetzt. Dabei wird dem Klassifikationssystem eine Anzahl befundeter Mammogramme als Trainingsdaten übergeben. Nach der Trainingsphase kann das System selbständig neue, noch unbefundete Mammogramme in Klassen „mit Mikroverkalkungen“ und „ohne Mikroverkalkungen“ einordnen. Der Aufbau eines solchen Systems ist prinzipiell immer gleich und gliedert sich in die Phasen „Vorverarbeitung“, „Merkmalsselektion“, „Klassifikatorauswahl“ und „Qualitätskontrolle“ [4]. In der Vorverarbeitung werden die im allgemeinen als Film vorliegenden Mammogramme zunächst digitalisiert. Anschließend werden sie durch Bildverarbeitungsalgorithmen aufbereitet und verbessert. Dies kann beispielsweise eine lokale Kontrastanpassung und Rauschunterdrückung mit folgendem Schwellwertverfahren sein, um die Verkalkungen vom Hintergrund zu trennen (Segmentierung). Aus den aufbereiteten Bildern werden verdächtige Gebiete (ROIs) ausgeschnitten, woraus wiederum Merkmale extrahiert und als Entscheidungsgrundlage für die Klassifikation verwendet werden. Um eine Einteilung in Klassen vornehmen zu können, muß der Merkmalsraum in Unterräume aufgeteilt werden. Dies ist die Aufgabe des Klassifikators. Während der Trainingsphase wird die Unterteilung vom Klassifikator gelernt. Für eine gute Klassifikation ist entscheidend, die für das konkrete Problem geeigneten Merkmale zu extrahieren. Bei der Untersuchung von Mammogrammen ist beispielsweise die Anzahl ausgeschnittener Bildpunkte ein Merkmal. Ob aber eine Gruppe von Mikroverkalkungen vorliegt, ist noch von vielen weiteren Merkmalen abhängig. Die Merkmalsextraktion erfordert im Allgemeinen tiefgreifende Kenntnisse über das zu klassifizierende Problem. Sie ist im Falle der Detektion von Mikroverkalkungen sehr aufwendig, da umfangreiche, explizit durchzuführende Ver-

arbeitungsschritte notwendig sind, um genügend Merkmale für eine gute Klassifikation zu gewinnen. Durch die Texturanalyse können automatisch weitere wichtige Merkmale extrahiert werden.

2.1 Merkmalsselektion

Zunächst werden alle möglichen Merkmale ohne Vorwissen und Einschränkung aus den Bildausschnitten extrahiert. Eine Zusammenfassung aller Merkmale einer ROI wird als Muster bezeichnet. Da jedoch unnütze Merkmale die Abstraktionsfähigkeit des Systems einschränken und die Gefahr des Auswendiglernens (Overfitting) besteht, werden in dem Schritt der Merkmalsselektion die für die Entscheidung wichtigen Merkmale ausgewählt und die Unbedeutenden verworfen. Für dieses Verfahren existieren eine Reihe von Standardalgorithmen. Da es für eine Bewertung der Texturmerkmale ausreichend ist, eine „gute“ Selektion zu finden, wird ein „Forward-Select-Verfahren“ eingesetzt, bei dem nacheinander das am besten geeignete Merkmal der selektierten Menge von Merkmalen hinzugefügt wird [5]. Die optimale Merkmalsselektion kann nur in Kombination und Rückkopplung mit der Klassifikatorauswahl stattfinden.

2.2 Klassifikatorauswahl

Die Auswahl des für ein spezielles Problem am besten geeigneten Klassifikators erfolgt durch einen interdisziplinären Experten für Klassifikation und das konkrete Problem. Zur Bewertung der Texturmerkmale ist es nicht notwendig, den optimalen Klassifikator zu finden. Daher werden nur zwei Klassifikatoren ausgewählt und mit den Texturmerkmalen trainiert. Dies ist ein Klassifikator mit linearer und ein Klassifikator mit quadratischer Funktion zur Unterteilung des Merkmalsraumes. Die selektierten Texturmerkmale wurden aufgrund dieser Klassifikatoren bewertet.

Das Maß für die Abstraktionsfähigkeit eines Klassifikationssystems ist die Generalisierung, also das Klassifikationsverhalten auf unbekanntem Daten. Idealerweise stehen für Training- und Testphase ausreichend Daten zur Verfügung, um einen Klassifikator mit den statistischen Eigenschaften der Merkmale anzulernen und mit ihm unbekanntem Daten zu testen. In der Praxis ist dies selten der Fall und der Trainingsdatensatz ist zu klein, um die statistischen Eigenschaften der Merkmale korrekt zu erfassen und einen Klassifikator aufzubauen. In diesem Fall benötigt man fast alle vorhandenen Muster des Datensatzes zum Training und schätzt die Generalisierungsfähigkeit mit einer vollständigen Kreuzvalidierung. Dabei wird nacheinander ein Muster aus dem Trainingsdatensatz ausgelassen und der Klassifikator mit dem Rest trainiert. Danach wird mit dem ausgelassenen Muster getestet. Dies geschieht für alle möglichen Kombinationen. Das Verhältnis von richtigen Klassifikationen zu getesteten Mustern ist ein Maß für die Abstraktionsfähigkeit des Klassifikators. Mit diesem Verfahren läßt sich bei kleinen Datensätzen trotzdem eine gute Abschätzung der Generalisierung geben. Da nur 256 Muster zur Verfügung stehen, wird die Generalisierung der Klassifikatoren wie beschrieben ebenfalls mittels Kreuzvalidierung abgeschätzt.

3 Methoden

Da gruppierte Mikroverkalkungen immer in regelmäßiger Struktur auftreten, liegt es nahe, automatisch Merkmale aus dieser Regelmäßigkeit zu gewinnen. Die Regelmäßigkeit wird durch eine gleichmäßige Verteilung von Mikroverkalkungen innerhalb eines Clusters erzeugt. Diese Eigenschaften spiegeln sich in einer für Mikroverkalkungen eigenen Textur wider, welche zwar von Gruppe zu Gruppe unterschiedlich ist, sich aber immer ähnelt.

3.1 Kovarianzanalyse

Eine Methode zur Texturanalyse ist beispielsweise der Einsatz von Cooccurrence-Matrizen. Dabei wird für Abstände zwischen Pixeln eines Bildes jeweils eine Matrix in Form eines zweidimensionalen Grauwertistogrammes angelegt. Dieses gibt die Häufigkeit von Grauwertkombinationen an [4]. Da aber ein bestimmter Abstand untersucht wird, wird dabei nur genau eine „Frequenz“ berücksichtigt. Bei Mikroverkalkungen treten jedoch viele Frequenzen in Form von Kantenanstieg und Abstand voneinander auf, welche die Textur ausmachen. Dem kann beispielsweise durch eine normierte, kompaktere Form der Cooccurrence-Matrizen Rechnung getragen werden. Dabei wird mittels Cooccurrence-Matrizen aller möglichen Abstände ein Summenhistogramm gebildet. Betrachtet man die untersuchte ROI als Pixelmatrix, erhält man diese Information durch eine Kovarianzmatrix, welche auch als Analyse einer Frequenzverteilung interpretiert werden kann.

3.2 Komponentensoftware

Die erstellten Algorithmen wurden in der Komponentensoftware des Forschungszentrums Karlsruhe (FZK) implementiert, wodurch ein schneller prototypischer Aufbau und die Verwendung von bereits vorhandenen Bildverarbeitungskomponenten möglich war. Zudem existieren umfangreiche Bibliotheken für Klassifikation und Merkmalsauswahl. Die Kovarianzanalyse ist als eigenständige Komponente realisiert und kann durch die genormten Schnittstellen in existierende Abläufe zur Detektion von Mikroverkalkungen integriert werden.

4 Ergebnisse

Die verwendeten Mammogramme stammen aus der öffentlichen Bibliothek der Universität von Nijmegen. Die Mammogramme liegen digitalisiert mit einer Quantisierung von 256 Graustufen vor. Hieraus wurde von Hand eine ROI-Datenbank mit 256 ausgeschnittenen Gebieten aus den Mammogrammen erstellt, von denen jeweils 128 aus Brustgewebe mit Mikroverkalkungen und 128 aus Brustgewebe ohne Mikroverkalkungen bestehen. Diese Datenbank dient zur Analyse der entwickelten Algorithmen. Eine ROI hat dabei Ausmaße von 256 mal 256 Bildpunkten. Die erstellten Kovarianzmatrizen werden auf 64 mal 64 Gitterpunkte begrenzt. Zur Bewertung werden jeweils die besten 5, 7, 14, 44 und 275 Merkmale von 4096 durch die Merkmalsselek-

Tabelle 1. Klassifikationsergebnisse mit unterschiedlicher Anzahl selektierter Merkmale.

Merkmale	Linear: Trainingsfehler	Linear: Generalisierungsfehler	Quadratisch: Trainingsfehler	Quadratisch: Generalisierungsfehler
5	38%	42%	-	-
7	32%	33%	28%	37%
14	26%	33%	17%	38%
44	17%	34%	8%	41%
275	3%	38%	2%	43%

tion ausgewählt und anschließend durch den linearen und den quadratischen Klassifikator bewertet (siehe Tabelle 1). Die Anzahl wird willkürlich gewählt, um mit wenig Aufwand den Verlauf einer kontinuierlichen Auswahl abschätzen zu können. Insgesamt kam es durch die Verwendung von mehr Merkmalen wie erwartet zu einer stetigen Verringerung des Trainingsfehlers. Die Generalisierung nahm bei größerer Anzahl von Merkmalen entsprechend ab.

5 Diskussion

Die Ergebnisse mit den vorhandenen Daten entsprechen den Erwartungen. Der Trainingsfehler nimmt bei zusätzlichen Merkmalen ab. Durch die verhältnismäßig kleine Stichprobe und der großen Anzahl von extrahierten Merkmalen ist die statistische Aussagekraft der Muster gering und der Generalisierungsfehler entsprechend hoch. Bei 44 und 275 selektierten Merkmalen ist das Verhältnis von Mustern zu Merkmalen bereits sehr ungünstig. Es wurde jedoch gezeigt, daß sich mittels einer Kovarianzanalyse automatische Merkmale gewinnen lassen, welche bei der Detektion von Mikroverkalkungen wertvoll sein können. Durch den Aufbau in Komponentensoftware können die entworfenen Algorithmen in folgenden Untersuchungen modular eingesetzt und mit verbesserten Methoden verglichen werden. Die entwickelte Methode eignet sich in jedem Fall zur Verbesserung von bestehenden Systemen, da sie neue, wichtige Merkmale bereitstellen kann.

6 Literatur

1. von Fournier D, Anton H-W, Junkermann H und Bastert G: Breast Cancer Screening, Cancer Diagnosis, Springer-Verlag, 1992, ISBN: 3-540-54503-4
2. Stomper PC, Connolly JL: Ductal Carcinoma in Situ of the Breast. American Journal of Radiology, Number 159:483-485, September 1992
3. Giger ML: Computer-Aided Diagnosis of Breast Lesions in Medical Images. IEEE Computer Society, Computing in Medicine, Volume 2, Number 5:39-45, September 2000.
4. Lehmann T, Oberschelp W, Pelikan E, Repges R: Bildverarbeitung für die Medizin. Grundlagen, Modelle, Methoden, Anwendungen. Springer-Verlag, Berlin, 1. Auflage 1997.
5. Witten IH, Frank E: Data Mining. Practical Machine Learning Tools with Java Implementations. Morgan Kaufmann Publishers, 2000, ISBN:1-55860-552-5