

# Merkmale zur Beschreibung der Intensitätsvariation für die Klassifikation von Herdbefunden in Mammogrammen

Florian Wagner, Matthias Elter

Fraunhofer-Institut für Integrierte Schaltungen IIS, Erlangen  
wagnerfn@iis.fraunhofer.de

**Kurzfassung.** Die computerassistierte Diagnose (CADx) für die Unterscheidung von Herdbefunden in Mammogrammen hat das Potential, Radiologen bei der Entscheidungsfindung zu unterstützen. Für ein von uns entwickeltes CADx-System haben wir eine Gruppe von neuen Merkmalen entwickelt, die auf einem Ansatz basieren, bei dem die Läsion in Abhängigkeit vom Abstand zum Rand der Segmentierung in verschiedene Regionen aufgeteilt wird. Für jede Region wird der Mittelwert der Grauwertintensität berechnet. Diese Mittelwerte dienen als Merkmal und als Eingabe für die Berechnung weiterer Merkmale, welche den Übergang zwischen den einzelnen Regionen und das Verhältnis zwischen den am weitesten entfernten Regionen beschreiben. Auf einer Testdatenbank mit 750 verschiedenen Läsionen erreichte die von einem Genetischen Algorithmus ermittelte optimale Merkmalsmenge auf 8 bit Bildern einen  $A_z$ -Wert von 0,72 und auf 12 bit Bildern einen  $A_z$ -Wert von 0,74. Die Ergebnisse zeigen, dass diese einfach zu berechnenden Merkmale gut für die Unterscheidung von benignen und malignen Herdbefunden geeignet sind.

## 1 Einleitung

Die Mammografie ist die aktuell wichtigste, für das Brustkrebs-Screening geeignete Methode und in der Mammadiagnostik unverzichtbar. Leider weist sie bei der Klassifikation von gut- und bösartigen Tumoren eine relative geringe Spezifität auf, sodass von den Patientinnen oftmals als unangenehm empfundene Folgeuntersuchungen durchgeführt werden müssen. Ein Beispiel für eine solche Untersuchung ist die Biopsie, bei der Gewebe zur pathologischen Untersuchung entnommen wird. Studien zeigen, dass nur 20 bis 30 Prozent der Biopsien bösartige Tumore identifizieren [1]. Für die Patientinnen bedeutet dies unnötige und vor allem unangenehme körperliche Eingriffe und für die Krankenkassen unnötige Kosten.

Um zur Reduktion von benignen Biopsien beizutragen, haben wir ein computerassistiertes Diagnose (CADx)-System zur automatischen Klassifikation von benignen und malignen Herdbefunden entwickelt. Das System berechnet für eine gegebene Läsion eine Vielzahl von Merkmalen (statistische Merkmale, Form-

und Texturmerkmale) und ermittelt anhand dieser die Wahrscheinlichkeit für die Bösartigkeit der vorliegenden Läsion. Es werden zudem Merkmale berechnet, die speziell bei Herdbefunden zu beobachtende Eigenschaften beschreiben.

Weisen Herdbefunde eine relativ hohe Dichte auf (d.h. sie erscheinen im Bild relativ hell), ist dies ein Indiz für einen malignen Befund. Transparenz ist hingegen eher ein Indikator für Gutartigkeit. Ein weiteres wichtiges Merkmal zur Charakterisierung ist die Umrandungsschärfe. Benigne Herdbefunde grenzen sich zumeist gut vom umgebenden Gewebe ab. Im Gegensatz dazu ist zu beobachten, dass maligne Läsionen oftmals fließend in das umgebende Gewebe übergehen. In dieser Arbeit wollen wir eine neue Gruppe von Merkmalen vorstellen, welche die Helligkeitsverteilung vom Zentrum des Tumors bis hin zum Rand beschreiben. Die Merkmalsgruppe ist in der Lage, die Dichte innerhalb des Tumors sowie den Übergang in das umgebende Gewebe zu beschreiben.

Varela et al. berechneten in [2] spezielle Merkmale für drei verschiedene Regionen des Herdbefundes. Unter anderem wurden Merkmale implementiert, die den Kontrast und die Umrandungsschärfe des Tumors charakterisieren. Die Klassifikationsleistung wurde anhand eines Datensatzes mit 981 Mammogrammen überprüft. Für jede Region wurde die optimale Merkmalskombination bestimmt. Die ROC-Analyse ergab für die innere Region einen  $A_z$ -Wert von 0.69, für die Randregion einen  $A_z$ -Wert von 0.76 und für die äußere einen  $A_z$ -Wert von 0.75. Die Kombination der Merkmale aller Regionen ergab einen  $A_z$ -Wert von 0.81.

Rangayyan et al. [3] präsentierten Merkmale, welche die Umrandungsschärfe und die Form von Herdbefunden beschreiben. Die Umrandungsschärfe eines Tumors wurde durch paarweise Pixeldifferenzen entlang der Normalen der Konturpixel charakterisiert. Auf einem Datensatz mit 54 Mammogrammen erzielte die beste Merkmalskombination bei der Klassifikation zwischen benignen und malignen Herdbefunden eine Genauigkeit von 94.4 Prozent. Das Merkmal zur Charakterisierung der Umrandungsschärfe zeichnete sich mit einer Genauigkeit von 92.6 Prozent als bestes Einzelmerkmal aus.

In [4] wurden Merkmale vorgestellt, die auf Grundlage eines polygonalen Modells der Kontur des Tumors berechnet werden. Nachdem der Rand des Herdbefundes als ein  $n$ -seitiges Vieleck modelliert wurde, berechnete man darauf Merkmale, welche die Konvexität bzw. Konkavität, die Wahrscheinlichkeit der Spikuliertheit und die Kompaktheit charakterisieren. Die Klassifikationsleistung aller Merkmale wurde auf einem Datensatz mit 54 Mammogrammen getestet. Das Merkmal für die Spikuliertheit eines Tumors wies den besten  $A_z$ -Wert von 0.82 auf. Die Merkmale für die Konvexität und Kompaktheit lieferten  $A_z$ -Werte von 0.75 bzw. 0.76. Die Kombination der drei Merkmale ergab einen  $A_z$ -Wert von 0.79.

Der Rest der Arbeit ist wie folgt gegliedert: In Kap. 2 stellen wir die für die Evaluierung verwendete Mammografie-Datenbank vor und beschreiben den Ablauf der Experimente sowie den Ansatz für die neue Merkmalsgruppe. Die Ergebnisse unserer Experimente werden in Kap. 3 vorgestellt und in Kap. 4 diskutiert.

## 2 Material und Methoden

Für die Auswertungen in dieser Arbeit wurde die öffentlich zugängliche Digital Database for Screening Mammography (DDSM) verwendet [5]. Berücksichtigt wurden alle Mammogramme von Fällen, die durch einen Lumisys Laserscanner digitalisiert wurden und mindestens einen Biopsie-geprüften Herdbefund enthielten. Für jeden Fall existiert eine Datei, welche die Annotation und die Grundwahrheit enthält. Für jedes Mammogramm ist eine manuell markierte Region von Interesse (ROI) gegeben, welche den Herdbefund enthält. Diese ROI wurde von einem Experten definiert, dient zur Lokalisation und als Eingabe für unser semi-automatisches Segmentierungsverfahren. Bilder die Artefakte oder Herdbefunde, die zu nah am Rand liegen enthalten, wurden ausgeschlossen, da unser Segmentierungsverfahren diese Fälle nicht behandeln kann. Letztendlich enthielt unsere Testdatenbank 750 Regionen (387 maligne und 363 benigne) von insgesamt 423 Patienten. Alle Bilder wurden mit einer Farbtiefe von 12 bit und einer Auflösung von  $50 \mu\text{m}$  digitalisiert.

Die Grundidee des neu entwickelten Merkmalsextraktors ist die Beschreibung des Herdbefundes durch eine Vielzahl von konzentrischen Regionen, die in Abhängigkeit vom Abstand zu seinem Rand bzw. zum Rand der Segmentierung definiert werden. Um auch das umgebende Gewebe beschreiben zu können, wird die Segmentierung des Tumors in Abhängigkeit von der kleinsten Achse nach außen vergrößert. Durch die Anwendung der Chamfer-Abstands-Transformation (Details in [6]) auf die Segmentierung erhalten wir von jedem Pixel innerhalb der Segmentierung des Tumors den Abstand zur äußeren Begrenzung. Abhängig vom maximal berechneten Abstand  $dist_{Max}$  teilen wir den Tumor in  $k$  konzentrische Regionen  $R$  auf. Jede Region  $R_i$  enthält alle Pixel  $p(x, y)$  für die gilt

$$i \left( \frac{\max\text{Value}}{k} \right) \leq \text{dist}(p(x, y)) < (i + 1) \left( \frac{\max\text{Value}}{k} \right) + 1 \quad (1)$$

wobei  $dist(p(x, y))$  den Abstand des Pixels  $p(x, y)$  zum Rand des Tumors bezeichnet.

Für die  $k$  ermittelten Regionen werden dann die Grauwertistogramme  $H_R$  bestimmt und daraus der Mittelwert wie folgt berechnet

$$\text{Mean}_i = 1/||H_{R_i}|| \sum_{j=0}^{N-1} H_{R_i}(j)j, \quad i = 0, \dots, k - 1 \quad (2)$$

$||H_{R_i}||$  bezeichnet die Anzahl der Pixel in der Region  $i$ ,  $N$  bezeichnet die Anzahl der Grauwerte und  $H_{R_i}(j)$  steht für die Anzahl der Pixel mit dem Grauwert  $j$  in der Region  $i$ . Die  $k$  Mittelwerte der Regionen werden als Merkmale verwendet. Zusätzlich wird daraus ein Merkmal abgeleitet, welches die durchschnittliche Intensitätsveränderung zwischen den einzelnen Regionen charakterisiert

$$\text{IntensityChange} = 1/(k - 1) \sum_{i=1}^{k-1} |\text{Mean}_i - \text{Mean}_{i-1}| \quad (3)$$

**Tabelle 1.** Klassifikationsleistung ( $A_z$ ) der vorgestellten Merkmale für  $k = 10$ .

Merkmal	8 bit	12 bit	Merkmal	8 bit	12 bit
Mean <sub>0</sub>	0,51	0,54	Mean <sub>8</sub>	0,59	0,63
Mean <sub>1</sub>	0,54	0,54	Mean <sub>9</sub>	0,64	0,66
Mean <sub>2</sub>	0,55	0,59	IntensityChange	<b>0,62</b>	0,63
Mean <sub>3</sub>	0,53	0,54	MaxRatio <sub>1</sub>	<b>0,58</b>	<b>0,55</b>
Mean <sub>4</sub>	<b>0,57</b>	<b>0,59</b>	MaxRatio <sub>2</sub>	<b>0,59</b>	<b>0,56</b>
Mean <sub>5</sub>	0,58	0,61	HistoDiff <sub>1</sub>	0,63	<b>0,63</b>
Mean <sub>6</sub>	0,61	0,65	HistoDiff <sub>2</sub>	<b>0,63</b>	<b>0,62</b>
Mean <sub>7</sub>	0,61	0,64			

Um den Kontrast des Tumors im Vergleich zum umgebenden Gewebe und damit auch die Umrandungsschärfe zu beschreiben, wurden die vier folgenden Merkmale konzipiert

$$\text{MaxRatio}_1 = \text{Mean}_{k-1} / \text{Mean}_0 \quad (4)$$

$$\text{MaxRatio}_2 = (\text{Mean}_{k-1} + \text{Mean}_{k-2}) / (\text{Mean}_0 + \text{Mean}_1) \quad (5)$$

$$\text{HistoDiff}_1 = \sum_{j=0}^{N-1} |H_{R_0}(j) / \|H_{R_0}\| - H_{R_{k-1}}(j) / \|H_{R_{k-1}}\|| \quad (6)$$

$$\text{HistoDiff}_2 = \sum_{j=0}^{N-1} |(H_{R_0}(j) / \|H_{R_0}\| + H_{R_1}(j) / \|H_{R_1}\|) - (H_{R_{k-1}}(j) / \|H_{R_{k-1}}\| + H_{R_{k-2}}(j) / \|H_{R_{k-2}}\|)| \quad (7)$$

Somit werden insgesamt  $k + 5$  Merkmale berechnet, welche die Intensitätsverteilung innerhalb und außerhalb des Tumors sowie den Übergang in das umgebende Gewebe beschreiben.

### 3 Ergebnisse

Die Klassifikationsleistung der einzelnen Merkmale wurde durch die „leave-one-out“-Validation und „receiver operating characteristic“ (ROC)-Analyse ausgewertet. Die Fläche unter der ROC-Kurve  $A_z$  wurde als Referenz verwendet. Ein k-nächste Nachbarn (kNN)-Klassifikator wurde für den Klassifikationvorgang verwendet. Die vorgestellten Merkmale wurden sowohl für die 12 bit als auch für auf 8 bit herunter gerechnete Bilder für  $k = 10$  berechnet. Um die gemeinsame Klassifikationsleistung aller Merkmale zu beschreiben, wurde durch einen genetischen Algorithmus eine möglichst optimale Merkmalsmenge ermittelt (sowohl für die 12 bit als auch für die 8 bit Merkmale).

Tabelle 1 zeigt die Klassifikationsleistungen für jedes einzelne Merkmal. Die  $A_z$ -Werte der jeweils in der optimalen Merkmalsmenge enthaltenen Merkmale

sind fett markiert. Für die auf 8 bit Bildern berechneten Merkmale erhält man eine optimale Gesamtklassifikationsleistung von 0,72 und für die auf 12 bit Bildern berechneten Merkmale einen  $A_z$ -Wert von 0,74.

## 4 Diskussion

Die in Tab. 1 dargestellten Ergebnisse zeigen, dass jedes einzelne Merkmal für die Unterscheidung von benignen und malignen Herdbefunden geeignet ist. Bei den einfachen Mittelwerten der Regionen ( $\text{Mean}_0$  bis  $\text{Mean}_9$ ) fällt auf, dass mit größerer Entfernung zum Rand der Segmentierung die Klassifikationsleistung steigt. Dies zeigt, dass die Grauwertintensität und somit auch die Dichte im Zentrum des Tumors charakteristischer für die Unterscheidung von gut- und bösartigen Läsionen ist als in den äußeren Bereichen. Außerdem ist zu beobachten, dass diese Merkmale, auf 12 bit Bildern berechnet, zu besseren Ergebnissen führen als das jeweilige, auf 8 bit Bildern berechnete Pendant. Somit spiegelt sich hier der Informationsverlust zwischen 12 bit und 8 bit in den Klassifikationsleistungen der einzelnen Merkmale wider.

Wir haben die Läsion in Abhängigkeit vom Abstand zum Rand der Segmentierung in verschiedene Regionen aufgeteilt und für jede Region den Mittelwert der Grauwertintensität berechnet. Diese Mittelwerte dienen als Merkmal und als Eingabe für die Berechnung weiterer Merkmale, welche den Übergang zwischen den einzelnen Regionen und das Verhältnis der am weitesten entfernten Regionen beschreiben. Die von einem Genetischen Algorithmus ermittelte optimale Merkmalsmenge erreichte auf 8 bit Bildern einen  $A_z$ -Wert von 0,72 bzw. einen  $A_z$ -Wert von 0,74 auf 12 bit Bildern berechnet. Die Ergebnisse zeigen, dass diese einfach zu berechnenden Merkmale gut für die Unterscheidung von benignen und malignen Herdbefunden geeignet sind. Es ist denkbar, dass eine Erweiterung der Merkmale, z.B. durch die Varianz der Grauwertverteilung in den einzelnen Regionen, zu einer Verbesserung der Gesamtklassifikationsleistung führen kann.

## Literaturverzeichnis

1. Kopans DB. The positive predictive value of mammography. *AJR Am J Roentgenol.* 1992;158(3):521–6.
2. Varela C, Timp S, Karssemeijer N. Use of border information in the classification of mammographic masses. *Phys Med Biol.* 2006;51(2):425–41.
3. Rangayyan RM, El-Faramawy NM, Desautels JE, et al. Measures of acutance and shape for classification of breast tumors. *IEEE Trans Med Imaging.* 1997;16(6):799–810.
4. Rangayyan RM, Mudigonda NR, Desautels JE. Boundary modelling and shape analysis methods for classification of mammographic masses. *Med Biol Eng Comput.* 2000;38(5):487–96.
5. Heath M, Bowyer K, Kopans D, et al. The digital database for screening mammography. In: *Proc Int Works Digital Mammogr;* 2001. p. 212–8.
6. Butt MA, Maragos P. Optimum design of chamfer distance transforms. *IEEE Trans Image Process.* 1998;7:1477–84.