Automatische Tumorerkennung bei unterschiedlichen Organen mittels Berechnung und Klassifikation von Texturmerkmalen

Thomas Wittenberg, Kurt Neubauer, Christian Küblbeck,
Ignacio Permanyer, Robert Schmidt

Fraunhofer Institut für Integrierte Schaltungen – Angewandte Elektronik
Am Weichselgarten 3, 91058 Erlangen
Email: {wbg, nka, kue, pmr, sch} @iis.fg.de

Zusammenfassung. Die vorliegenden Arbeit beschäftigt sich mit der automatischen Klassifikation von zytologischer Proben des Gebährmutterhalses sowie von Sputum. Die Klassifikation der Stichproben wird jeweils auf einem hochdimensionalen Merkmalsvektor, bestehend aus unterschiedlichen Textur-Merkmalen durchgeführt. Zur Verbesserung der Klassifikationsleistungen und zur idealen Adaption der Merkmale an die Datensätze werden die Merkmalsvektoren mit dem Best-With-All Selektions-Verfahren ausgedünnt

1 Einleitung

Ziel dieser Arbeit ist es, --- langfristig --- ein arztunterstützendes Bild-Klassifikationsverfahren zu erstellen, mit dessen Hilfe zytologisches Bildmaterial automatisch und benutzerunabhängig vorgescreent werden kann. In dem in dieser Arbeit vorgestellten Teilaspekt geht es speziell darum, Bilder mit benignen und malignen Zellen (z.B. Tumorzellen, dysplastische Zellen, etc..) nach Eintrainieren eines Klassifikators einer der beiden Klassen „*ohne Befund*“ und „*mit Befund*“ rechnergestützt fehlerfrei zuzuordnen. Trainieren bedeutet für die vorliegende Problemstellung, mit Hilfe eines Klassifikators (in diesem Fall der *k-nächste Nachbar-Klassifikator*) einen Merkmalsvektor derart zu erstellen, dass dieser möglicht kurz ist, und zum anderen, dass die darin enthaltenen Merkmale die Klassifikation möglichst gut durchführen, und damit das Problem gut beschreiben können.
Für diesen Ansatz wird in dem vorliegenden Beitrag im folgenden ein Verfahren zur automatischen Beurteilung von PAP-gefärbten zytologischen Dünnschichtpräparaten mittels Klassifikation und Berechnung von Texturmerkmalen vorgestellt. In Abschnitt 2 werden die verwendeten Texturverfahren beschrieben sowie die Strategien zur Merkmalsselektion. In Abschnitt 3 werden die beiden verwendeten Stichproben vorgestellt, auf denen die Klassifikation durch geführt wurde. Abschnitt 4 beschreibt die Ergebnisse, die abschließend in Kapitel 5 diskutiert werden.

2 Merkmalsberechnung und -selektion

Von geschultem klinischem Personal lassen sich i.A. benigne von malignen Zellen in PAP-gefärbten Durchlichtbildern u.a. durch die Größe und Form des Zellkerns, die Abgrenzung zum umgebenden Gewebe, die Mitoserate, die Ähnlichkeit zum Muttergewebe und die Kern-/Plasma-Relation unterscheiden. In dem in dieser Arbeit vorgestellten Ansatz fliesen diese morphometrischen Parameter jedoch nur implizit ein, da der verwendete Ansatz ausschließlich auf Texturparametern der Zellbilder arbeitet.

2.1 Merkmalsberechnung

Bei der von uns verwendeten Klassifikationsstrategie werden die zytologischen Durchlichtbilder zunächst auf ihre Grauwert-Komponenten reduziert. Auf jedem Grauwertauszug wird anschließend ein hochdimensionaler Merkmalsvektor mit entsprechenden Texturmerkmalen berechnet. Zur Berechnung dieser Texturmerkmale wurden u.a. folgende Verfahren eingesetzt:

- Grauwertabhängigkeitsmatritzen nach *Haralick* [1],
- Summen- und Differenz-Histogramme nach *Unser* [2],
- Grauwertlauflängen nach *Galloway* [3],
- statistische Kennzahlen nach *Chen* [4]
- Multiresolutionsanalyse durch Wavelet-Transformationen nach *Laine* [5]
- Lokale statistische Maßzahlen nach *Schramm* [6],
- Fraktale Merkmale und Fourier-Spektralenergien nach *Ross* [7] sowie
- Image Convolution nach *Laws* [8]

Eine genaue Beschreibung der einzelnen Verfahren findet sich bei [9].

2.2 Merkmalsselektion

In den Experimenten wurde als Klassifikator der "k-Nächste Nachbar" eingesetzt, der als verteilungsfreier, nichtparametrischer Klassifikator ohne Annahmen über Verteilungen und Trennfunktionen auskommt [6]. Zunächst wurde getestet, wie sich derjenige Merkmalsvektor verhält, der verfügbaren Merkmale vereint. Berücksichtigt man, dass mittels der im letzten Abschnitt genannten Merkmale mehr als 100 Merkmalsdeskriptoren zur Verfügung stehen, ist anzunehmen, dass einige davon redundant bzw. kontraproduktive wirken. Für die konkrete Fragestellung der zytologischen Klassifikation ergibt sich damit die Frage, welche dieser Merkmale sich für die jeweilige Aufgabenstellung eignen, bzw. ob ein reduzierter Merkmalsvektor ein verbessertes Ergebnis liefert.

Ein einfaches Ausprobieren aller Möglichkeiten zur Selektion einer Untergruppe der Merkmale aus einem Pool scheidet prinzipiell aus, da die kombinatorische Komplexität zur Auswahl von beispielsweise nur 10 aus 100 Merkmalen eine Kombinationsvielfalt von ca. $1.7 * 10^{13}$ ergäbe [10]. Eine vollständige Suche in allen diesen Möglichkeiten ist auf Grund des mit jedem Merkmal exponentiell wachsenden Suchaufwandes nicht mehr möglich. Um diese Suche zu reduzieren, können effektive Strategien wie das Bester- Mit- Allen-Verfahren (*Best-With-All, BWA*), das Sequential-Backwards-Verfahren bzw. genetische Algorithmen die verwendet werden. In dieser Arbeit wurde das o.g. BWA-Verfahren verwendet, bei dem sukzessive diejenigen Merkmale zum reduzierten Merkmalsvektor zugefügt werden, die am meisten zur Klassifikationsleistung beitragen.

3 Beschreibung der Stichproben

Das Verfahren der automatischen Klassifikation und Merkmalsselektion wurde an zwei unabhängigen, unterschiedlich präparierten zytologischen Stichproben von verschiedenen menschlichen Organen getestet, einerseits an Bildern von Sputumzellen, andererseits an einer Stichprobe von Zellen des Gebärmutterhalses (Zervix). Beide Stichproben wurden jeweils vollständig durch eine zytologische Fachkraft (eine zytologisch-technische Assistenten bzw. einen Facharzt) befundet.

Die erste Stichprobe besteht 517 Durchlichtbildern mit PAP-gefärbten, mechanisch unvorbehandelten Zellen von Präparaten des menschlichem Bronchialsekrets der Lunge mit benignem (215 Bilder) bzw. maligner (302 Bilder) Befund. Alle Bilder besitzen eine Auflösung von 368x280 Pixel und wurden mit Hilfe einer *Fast Screen-Machine* digitalisiert, und auf 256 Graustufen reduziert, siehe Abb. 1.

Die zweite Stichprobe besteht aus Bildmaterial mit PAP-gefärbten Zellen der Zervix (Gebärmutterhals) mit 20-facher mikroskopischer Vergrößerung, die mittels des sog. „ThinPrep“ Verfahrens mechanisch aufbereitet und mit einem *Zeiss-Axio-Cam* System aufgenommen wurden. Alle Bilder wurden auf 8 bit = 256 Graustufen reduziert. In einem ersten Experiment wurden 267 Bilder mit benignen sowie 98 Bilder mit malignen Zellen mit einer Auflösung von 1000 x 700 Pixel verwendet. In einem zweiten Experiment wurden diese Bilder in kleiner Subbilder mit einer Auflösung von 256x256 Pixel zerlegt, und zwar in 1602 Bilder der Klasse „ohne Befund“ sowie 98 Bilder der Klasse „mit Befund“ .Typische Bilder dieser Stichprobe sind in Abbildung 2 zu sehen.

In der Regel werden für eine Untersuchung die gegebenen Bilder in zwei möglichst gleichgroße Teile zerlegt. Während der erste Teil --- *die Lernstichprobe* --- dazu verwendet wird, einen Klassifikator zu trainieren, verwendet man die zweite Hälfte -- *die Teststichprobe* --- , um zu untersuchen wie gut der Klassifikator trainiert wurde, d.h. festzustellen, wie viele unbekannte Muster mit dem trainierten Klassifikator der richtigen Klasse zugewiesen werden können .

Da jedoch bei dem eingesetzten BWA-Verfahren zur Merkmalsselektion die Ergebnisse der Teststichprobe verwendet werden, um eine möglichst gute Auswahl von Merkmalen zu treffen, besteht die Gefahr, die Testmuster bei der Auswahl "auswendig zu lernen". Durch die Einführung einer dritten Menge von Mustern --- der Verifikationsstichprobe --- kann eine objektive und zuverlässige Aussage über die Güte der Klassifikation gemacht werden.

Die beiden Stichproben (Sputum und Zervix) wurden jeweils in drei möglichst gleichgroße Mustergruppen aufgeteilt. Für die Schritte Klassifikation, Selektion und Verifikation wurden die drei Gruppen jeweils permutiert.

4 Ergebnisse

Die nachfolgenden Tabellen 1 und 2 zeigen die Klassifikationsergebnisse für die Sputum- bzw. die Zervixstichhprobe mit dem 1- bzw. 3-Nächsten-Nachbar-Klassifikator nach der Optimierung des Merkmalsvektoren. Die Permutationen wurden jeweils gemittelt. Die erhöhte Klassifikationsleistung auf der Teststichprobe ist auf das Auswendiglernen der Lernstichprobe während der Merkmalsselektion zurückzuführen.

Klassifikator	Teststichprobe	Verifikationsstichprobe
1-Nächster-Nachbar	82.66	78.94
3-Nächster-Nachbar	83.82	79.53

Tab. 1: Ergebnisse der Sputum Stichprobe

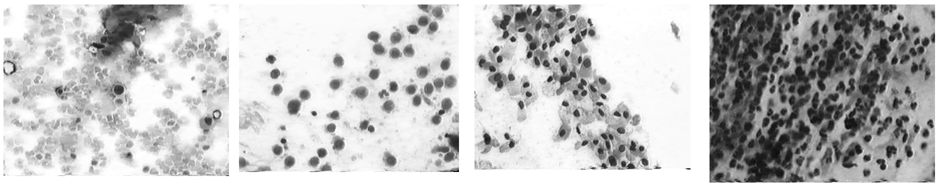
Klassifikator	Teststichprobe	Verifikationsstichprobe
1-Nächster-Nachbar	83.68	63.15
3-Nächster-Nachbar	83,7	66,2*

Tab. 2: Ergebnisse der Zervix Stichprobe (große Bilder). In diesem Fall (*) wurde das Ergebnis der besten Permutation verwendet.

Die Ergebnisse der Zervix-Stichprobe mit kleinen Bildern beträgt 92.53%. Diese stark erhöhte Klassifikationsleistung ist jedoch einzig auf die schlechte Bild-Relation von *"benigne"* -Bilder(534) zu "*maligne*" (33)-Bildern zurückzuführen. Die Klassifikationsmatrizen der Permutationen zeigen, dass die "*benigne*"-Bilder fast ausschließlich richtig, dagegen die "*maligne*"-Bilder überwiegend falsch klassifiziert wurden. Um die Relation der beiden Probenanteile zu verbessern, wurden in einem weiteren Experiment pro Klasse nur noch 34 Muster verwendet. Die Klassifikationsergebnisse auf der zugehörigen Verifikationsstichprobe sind mit 76,1% zwar schlechter als bei der vollständigen Klassifikation der kleinen Bilder, aber um 10% besser als auf den großen Bildern.

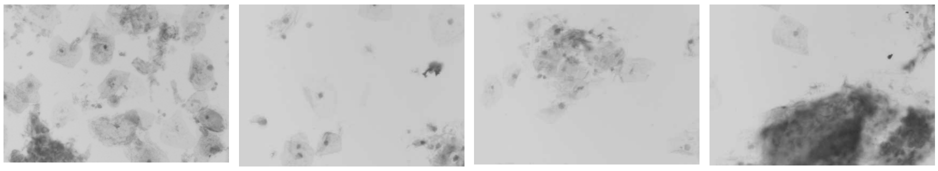
5 Diskussion

Um Verfahren der Texturanalyse für die Klassifikation von zytologischen Proben zu verwenden, wurde eine BWA-Selektion auf dem vollständigen Merkmalsvektor durchgeführt. Die Klassifikationsergebnisse von 79.5 % für die Sputum-Stichprobe bzw. 66.2 % für die Zervix-Stichprobe entsprechen den Erwartungen, die gestellt werden können, wenn ausschließlich Texturmerkmale auf den Grauwertbildern berrücksichtigt werden. Motiviert durch Ansätze der Farb-Textur-Analyse, bei denen nicht nur die einzelnen Farbplanes [11] sonder zudem auch Farbplane-übergreifende Texturmerkmale [12] berechnet werden, ist zu erwarten, das sich die Klassifikationsergebnisse noch verbessern lassen. Ein weitergehender Ansatz zur Verbesserung der Ergebnisse besteht zum einen in der Analyse der h. selektierten Texturmerkmale sowie in einer separaten Voroptimierung der parametrisierbarer Merkmalsfilter [10].



a. b. c. d.

Abb. 1 : Bilder mit Sputumzellen: mit Tumorzellen (a. und b.), ohne Tumorzellen (c. und d.)



a. b. c. d.

Abb. 2 : Bilder mit Zervixzellen: mit Tumorzellen (a. und b.), ohne Tumorzellen (c. und d.)

6 Literatur

1. Haralick M et al: Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics, SMC-3(6):610-621, 1973.
2. Unser M: Sum and difference histograms for texture analysis, *IEEE Trans. Pattern Analysis and Machine Intelligenc*e, 8:118-125, 1986.
3. Galloway MM: Texture analysis using gray level run lengths. *Computer Graphics and Image Processin*g, 4:172-179, 1975.
4. Chen Y, Nixon M, Thomas D: Statistical geometrical features for texture classifcation. *Pattern Recognitio*n, 28(4):537-552, 1995.
5. Laine A, Fun J: Texture classifcation by wavelet packet signatures. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(11):1186-1191, 1993.
6. Schramm U: Automatische Oberfächenprüfung mit neuronalen Netzen. Stuttgart: IRB-Verlag, 1994.
7. Ross T et al: Automatische Klassifkation hochaufgelöster Oberfächenprofle von Hauttumoren mit neuronalen Netzen. In *Mustererkennung 199*5, Sagerer et al (Hrsg) Berlin: Springer-Verlag, 1995.
8. Laws KI: *Textured Image Segmentatio*n. PhD thesis, Faculty of the Graduate School, University of Southern California, 1980.
9. Wagner T: Texture Analysis. In: Jähne, Hausecker, Geissler (Hrsg) Handbook of Computer Vision and Applications, Bd. II, Academic Press, Kapitel12, S. 275-308, 1999.
10. Pannekamp J et al : Adaptex: Adaptive Bewertung texturierter Oberflächen. In *VDI-Berichte 1572,* Bildverarbeitung im industriellen Einsatz, 199-206, 2000.
11. Burmeister A et al: *Tumorerkennung durch Verbindung von Textureigenschaften und Farbinformationen*. In 2. Workshop 'Farbbildverarbeitung' , Ilmenau, 1996.
12. Palm C et al: Color texture analysis of moving vocal cords using aproaches from statistics and Signal theory. In: Advances in Quantitative Laryngoscopy, Voice and Speech Research, Braunschweig et al. (Hrsg.), p.49-56, 2000.