MELDOQ: Ein System zur Unterstützung der Früherkennung des malignen Melanoms durch digitale Bildverarbeitung

R. Pompl[1], W. Bunk[1], A. Horsch[2], W. Stolz[3], W. Abmayr[4], W. Brauer[5], A. Gläßl[3] und G. Morfill[1]

[1]Max-Planck-Institut für extraterrestrische Physik, Garching
[2]Institut für medizinische Statistik und Epidemiologie, TU München
[3]Klinik und Poliklinik für Dermatologie der Universität Regensburg
[4]Fachbereich Informatik/Mathematik, Fachhochschule München
[5]Lehrstuhl für Theoretische Informatik und Grundlagen der KI, TU München
Email: pompl@mpe.mpg.de

Zusammenfassung. Die Dermatoskopie stellt die derzeit leistungsfähigste klinische Methode zur Früherkennung des malignen Melanoms dar. Da deren Anwendung viel Expertenwissen und Erfahrung voraussetzt, wurde im MELDOQ-Projekt ein Bildanalysesystem entwickelt, das Dermatologen bei der Diagnosefindung unterstützen soll. Basierend auf einem Datensatz von 749 standardisiert aufgenommenen dermatoskopischen Aufnahmen wurden 13 Merkmale durch bildanalytische Verfahren quantitativ erfaßt. Zur Klassifikation wird die Methode der linearen Diskriminanzanalyse eingesetzt. Durch exhaustives Testen aller möglichen Merkmalskombinationen und Crossvalidierung wurde ein Klassifikator entwickelt, der bei Verwendung von sechs Merkmalen eine Sensitivität von 91.5% und eine Spezifität von 93.4% erreicht. Ein besonderes Kennzeichen des Systems ist die Unterstützung des Dermatologen durch eine Visualisierung der Bewertungsgründe. In einem angestrebten Feldtest soll überprüft werden, ob diese Ergebnisse im klinischen Alltag bestätigt werden können.

Schlüsselwörter: Quantitative Bildverarbeitung, Diagnoseunterstützung, Dermatoskopie, malignes Melanom

1 Einleitung

Die Inzidenz des malignen Melanoms (*schwarzer Hautkrebs*) hat sich innerhalb der letzten zehn Jahre verdoppelt. Hohe Heilungschancen bestehen nur bei frühzeitiger Diagnose und nachfolgender Exzision. Die leistungsfähigste diagnostische Methode ist die Dermatoskopie [7] (Auflichtmikroskopie bei 10facher Vergrößerung). Die dazugehörige dermatoskopische ABCD-Regel beurteilt semiquantitativ die vier Merkmale Asymmetrie (A), Begrenzung (B), Farbvielfalt (C, *color*) und Differentialstrukturen (D). Jedem der Merkmale wird entsprechend seiner Ausprägung ein Score zugewiesen. Aus diesen Einzelscores wird durch gewichtete Addition $(1.3 \cdot A[0-2] + 0.1 \cdot B[0-8] + 0.5 \cdot C[1-6] + 0.5 \cdot D[1-5],$

die Zahlen in den eckigen Klammern geben den Wertebereich des Merkmals an) der sogenannte dermatoskopische Punktwert berechnet. Anhand dieses empirischen Malignitätsmaßes wird eine Verdachtsdiagnose gestellt. Die korrekte Anwendung dieses diagnostischen Verfahrens erfordert allerdings viel Expertenwissen und Erfahrung. Um keine malignen Melanome zu übersehen, werden viele unnötige Exzisionen gutartiger melanozytärer Hautveränderungen durchgeführt. Im Projekt MELDOQ [3] wurde daher untersucht, ob sich die diagnostische Genauigkeit des Dermatologen durch ein digitales Bildanalysesystem erhöhen läßt. Kernstück des Systems ist die Abbildung der dermatoskopischen Kriterien auf reproduzierbare mathematische Parameterskalen bei hoher Sensitivität und Spezifität. Darüberhinaus wird der Dermatologe durch die objektnahe Visualisierung der Bewertungsgründe unterstützt.

2 Zusammenfassung der Merkmalsquantifizierung

Das zur Verfügung stehende Bildmaterial umfaßt 749 dermatoskopische Aufnahmen melanozytärer Hautveränderungen (189 maligne und 560 benigne), die mit einer 3-Chip-CCD-Kamera [8] aufgenommen wurden. Um eine standardisierte Bildakquisition zu gewährleisten, wurde eine Shadingkorrektur zur Kompensation inhomogener Ausleuchtung und ein Farbabgleich durchgeführt. Die Bilder liegen im RGB-Format vor (Farbtiefe: 24 bit) und haben eine Größe von 512×512 Pixel; dies entspricht einem Bildauschnitt von 11.8×11.8 mm in vivo.

Angelehnt an die dermatoskopische ABCD-Regel wurden 13 Merkmale bildanalytisch modelliert (Details zu einigen der Merkmalsquantifizierungen finden sich in [4] und [5]). Für die Entwicklung der Algorithmen und die Optimierung der dort vorhandenen Parameter wurde ein Trainingsdatensatz von 60 malignen und 60 benignen melanozytären Hautveränderungen verwendet, der dem Gesamtdatensatz entnommen wurde. In Tabelle 1 sind zu jedem berechneten Merkmal drei Charakerisierungen der Separierfähigkeit angegeben, also eine Beurteilung, inwieweit mit diesem Merkmal eine korrekte Zuordnung zur benignen oder malignen Klasse möglich ist. Diese Maße werden hier anhand der *receiver-operating-characteristic*-Kurve (ROC-Kurve) ermittelt, die unter Verwendung aller 749 dermatoskopischen Aufnahmen errechnet wurde. Bei der ROC-Kurve wird für einen Trennparameter systematisch die zu erreichende Sensitivität und Falsch-Positiv-Rate (1-Spezifität) gegeneinander aufgetragen. Mit der *area under curve* (AUC) wird die Fläche unter der ROC-Kurve erfaßt. Weiterhin wird das Wertepaar Sensitivität und Spezifität gewählt, für das der euklidsche Abstand zum Optimum (Sensitivität = 1, Falsch-Positiv-Rate = 0) minimal ist.

Die Separierfähigkeit der Asymmetrie-Merkmale, insbesondere die der Farbasymmetrie (AUC = 0.941), ist am stärksten ausgeprägt. Auch mit den Merkmalen der Gruppen "Farbe" und "Struktur" lassen sich die beiden Klassen relativ gut trennen. Am wenigsten ist dies mit den Berandungsmerkmalen möglich. Diese Ergebnisse korrespondieren mit der Gewichtung in der dermatoskopischen ABCD-Regel, die der Asymmetrie die stärkste und der Berandung die geringste Bedeutung zumißt.

Merkmal	AUC	Sensitivität	Spezifität
Asymmetrie, Form	0.738	62.9%	73.0%
Asymmetrie, Konvextiät	0.701	63.4%	67.1%
Asymmetrie, Farbverteilung	0.941	88.3%	87.5%
Asymmetrie, Strukturverteilung	0.836	81.4%	74.4%
Berandung, strukturelle Beschreibung	0.783	73.5%	73.9%
Berandung, Homogenität	0.797	70.9%	77.8%
Kontur, Regelmäßigkeit	0.808	76.7%	72.5%
Kontur, Homogenität	0.680	61.9%	67.8%
Farbvielfalt	0.923	85.7%	83.0%
Farbhomogenität	0.803	75.1%	72.5%
Strukturvielfalt	0.823	74.6%	76.4%
Strukturhomogenität	0.782	75.6%	68.9%
Größe	0.808	73.0%	76.4%

Tabelle 1. *Zusammenfassung der Separierfähigkeit der einzelnen Merkmale.* Angegeben ist die Fläche unter der ROC-Kurve (AUC) sowie das Wertepaar (Sensitivität, Spezifität) das in der ROC-Kurve dem Optimum am nächsten gelegenen ist.

3 Klassifikationsergebnisse

Als Klassifikationsverfahren wurde die Methode der *linearen Diskriminanzanalyse* gewählt, da diese zum einen bereits in der dermatoskopischen ABCD-Regel Verwendung findet und zum anderen den Klassifikationsvorgang im Gegensatz zu beispielsweise neuronalen Netzen transparent gestaltet. Das Phänomen des *curse of dimensionality* [1] legt eine Reduktion der Dimensionalität der Merkmalsvektoren nahe, wodurch auch die Nachvollziehbarkeit des resultierenden Systems erhöht wird. Da die Menge der möglichen Merkmalskombinationen mit $2^{13} - 1 = 8191$ Elementen eine moderate Größe aufweist, kann die Suche nach der besten Untermenge exhaustiv erfolgen. Dazu wird für jede Kombination eine 10-fache Crossvalidierung durchgeführt. Um die ermittelten Klassifikationsraten (Anteil an richtig klassifizierten Hautveränderungen) korrekt bewerten zu können, werden diejenigen 120 Bilder ausgeschlossen, die zur Optimierung der Quantifikationsalgorithmen verwendet wurden. Die einzelnen Klassifikationsraten der zehn Trainingsmengen werden durch Mittelwertbildung zusammengefaßt. Als optimal wird diejenige Merkmalskombination betrachtet, bei der die gemittelte Klassifikationsrate maximal ist. Als Resultat dieses Verfahrens ergibt sich folgende Merkmalskombination (in Klammern sind die Pendants der dermatoskopischen ABCD-Regel angegeben):

- Asymmetrie, Farbe (dermatoskopisch: A)
- Asymmetrie, Struktur (dermatoskopisch: A)
- Kontur, Regelmäßigkeit (dermatoskopisch: B)

– Farbvielfalt (dermatoskopisch: C)
– Strukturvielfalt (dermatoskopisch: D)
– Größe (dermatoskopisch: -)

Die einzelnen Merkmale sind mit Ausnahme der Größe in der dermatoskopischen ABCD-Regel enthalten.

Anhand der Ergebnisse der 10-fachen Crossvalidierung bezüglich der Testmengen läßt sich die Leistung dieses Klassifikators bei dem System bislang unbekannten melanozytären Hautveränderungen abschätzen. Die Klassifikationsrate beträgt demnach 93.0% bei einer Sensitivität von 91.5% und einer Spezifität von 93.4 %.

Die Parameter des endgültigen Klassifikators werden nun unter Berücksichtigung des gesamten Bildmaterials bestimmt. Die daraus resultierenden Scores sind im linken Teil von Abbildung 1 als relative Häufigkeitsverteilungen und die Klassifikationseigenschaften sind im rechten Teil durch die ROC-Kurve dargestellt.

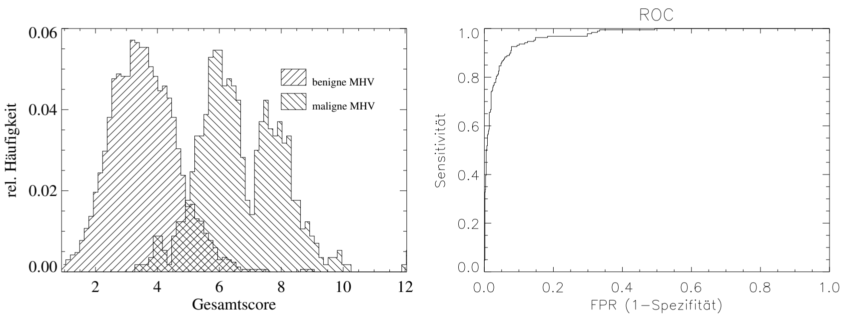

Abb. 1. *Links:* Relative Häufigkeitsverteilung des abgeleiteten Gesamtscores für die malignen und benignen melanozytären Hautveränderungen (MHV). *Rechts:* ROC-Kurve des Gesamtscores.

4 Zusammenfassung und Diskussion

Die Klassifikationsergebnisse zeigen, daß die diagnostische Genauigkeit des entwickelten Systems von über 90% der des Dermatoskopieexperten entspricht. Die Farbeigenschaften der Läsion tragen dabei wesentlich zur Separierung der malignen und benignen Klasse bei. Die hohe Klassifikationsrate konnte jedoch nur in Kombination mit anderen Merkmalen erreicht werden. Die dabei ausgewählten Merkmale finden sich mit Ausnahme der Läsionsgröße in der dermatoskopischen ABCD-Regel wieder.

Vorteilhaft für diese Studie ist das Vorliegen eines umfangreichen Bildmaterials. Während vielen Arbeitsgruppen nur wenig Bildmaterial zur Verfügung stand (beispielsweise werden in [2] 104 Aufnahmen verwendet), konnte hier auf einen vergleichsweise großen Datensatz von 749 standardisierten dermatoskopischen Aufnahmen zurückgegriffen werden. Unterschiedlich zu anderen Ansätzen ist das Bestreben, die Quantifizierung der Merkmale der dermatoskopischen ABCD-Regel zu entlehnen und damit den computergestützten Diagnosevorschlag an der bereits praktizierten Diagnostik zu orientieren. So gehen in den Diagnosevorschlag nur sechs Merkmale ein, wodurch auch eine hohe Transparenz des Systems erreicht wird. Damit hat das System das Potential, in der Dermatoskopie ungeübte Dermatologen mit dieser Diagnosetechnik vertraut zu machen. Im Gegensatz dazu werden beispielsweise in [6] 23 Merkmale verwendet.

In einem angestrebten Feldtest soll überprüft werden, ob diese Ergebnisse in der klinischen Routine bestätigt werden können. Besonderes Augenmerk wird auf die Aussagekraft der Visualisierung und die Nützlichkeit des Diagnosevorschlags gelegt werden.

Literatur

1. Bellman R: Adaptive Control Process: A Guided Tour. Princeton University Press, New Jersey, 1961.
2. Gutkowicz-Krusin D, Elbaum M, Szwaykowski P, Kopf AW: Can early malignant melanoma be differentiated from atypical melanocytic nevus by in vivo techniques? Part II: Automatic machine vision classification. Skin Research and Technology, Vol. 3, pp 15-22, 1997
3. Horsch A, Stolz W, Neiß A, Abmayr W, Pompl R, Bernklau A, Bunk W, Dersch D, Gläßl A, Schiffner R, Morfill G: Improving Early Recognition of Malignant Melanomas by digital Image Analysis in Dermatoscopy. Pappas C, Maglaveras N, Scherrer J (eds): MIE '97, IOS Press 531–535, 1997.
4. Pompl R, Bunk W, Dersch D R, Horsch A, Stolz W, Abmayr W, Brauer W, Gläßl A, Schiffner R, Morfill G: Charakterisierung der Berandungseigenschaften melanozytärer Hautveränderungen zur Unterstützung der Früherkennung des malignen Melanoms. 43. Jahrestagung der gmds, Bremen, 1998. MMV Medien & Medizin Verlag, 309–312
5. Pompl R, Bunk W, Dersch D R, Horsch A, Stolz W, Abmayr W, Brauer W, Gläßl A, Schiffner R, Morfill G: Charakterisierung der Farbeigenschaften melanozytärer Hautveränderungen zur Unterstützung der Früherkennung des malignen Melanoms. Bildverarbeitung für die Medizin 1999, H. Evers, G. Glombitza, T. Lehmann, H.-P. Meinzer (Eds.), Reihe Informatik aktuell, Springer-Verlag, pp 160-164
6. Schindewolf T, Stolz W, Albert R, Abmayr W, Harms H: Classification of Melanocytic Lesions with Color and Texture Analysis Using Digital Image Processing. Analytical and Quantitative Cytology and Histology, Vol. 15, No. 3, pp 1-11, 1993
7. Stolz W, Braun-Falco O, Bilek P, Landthaler M: Farbatlas der Dermatoskopie. Blackwell Wissenschaft, Berlin, 1993.
8. Stolz W, Schiffner R, Pillet L, Vogt T, Harms H, Schindewolf T, Landthaler M, Abmayr W: Improvement of monitoring of melanocytic skin lesions using a computerized acquisition and surveillance unit with a skin surface microscopic television camera. Journal of the American Academy of Dermatology, Vol 35, No 2, 202–207, 1996.