

Automatische Kategorisierung von medizinischem Bildmaterial in einen multiaxialen monohierarchischen Code

M.O. Güld¹, H. Schubert³, M. Leisten³, B. Plodowski¹, B. Fischer¹,
D. Keyzers² und T.M. Lehmann¹

¹Institut für Medizinische Informatik

²Lehrstuhl für Informatik VI

³Klinik für Radiologische Diagnostik

Rheinisch-Westfälische Technische Hochschule (RWTH), Aachen

Email: mgueld@mi.rwth-aachen.de

Zusammenfassung. Diese Arbeit untersucht die automatische Kategorisierung von radiologischem Bildmaterial anhand globaler Merkmale. Zur Bildinhaltsbeschreibung wird ein multiaxialer, monohierarchischer Schlüssel benutzt. Auf einem Datenkorpus mit 2470 Aufnahmen aus der klinischen Routine wird eine Erkennungsrate von 82.2% für eine Einteilung in acht anatomische Kategorien erreicht. Für eine feinere Einteilung des Korpus in 24 anatomische Kategorien wird eine Erkennungsrate von 80.5% erreicht.

1 Problemstellung

Die automatische Identifikation medizinischer Aufnahmen, z.B. die Bestimmung der abgebildeten Körperregion, ist für viele computerunterstützte Anwendungen eine Grundvoraussetzung. Insbesondere die Kategorisierung radiologischen Bildmaterials ist ein wichtiger Abstraktionsschritt in der Architektur [1] von Image Retrieval in Medical Applications (IRMA)¹. Die automatische Kategorisierung erfordert Lösungen auf drei ineinander verzahnten Teilgebieten:

1. Bereitstellung eines geeigneten Kodierungsschemas;
2. Aufbau eines referenzkodierten Datenkorpus für die Evaluierung;
3. Entwicklung und Validierung von automatischen Methoden der Bildverarbeitung und Mustererkennung zur Extraktion, Selektion und Auswertung bzw. zum Vergleich von bildbeschreibenden Merkmalen.

2 Stand der Forschung

Zur einheitlichen, benutzerunabhängigen Kategorisierung medizinischen Bildmaterials wurde in [2] eine Codierung vorgeschlagen, die erstmalig die nötigen

¹ Das IRMA-Projekt wird gefördert von der Deutschen Forschungsgemeinschaft DFG (Le 1108/4)

Anforderungen hinsichtlich Vollständigkeit und Eindeutigkeit erfüllt. Die durch die Modalität bei der Bildgebung bereitgestellte kategorisierungsrelevante Information (z.B. in DICOM) ist jedoch konzeptionell unzureichend und nur eingeschränkt verfügbar [3]. Bisherige Publikationen adressieren häufig ein eingeschränktes Untersuchungsfeld bezüglich der Identifikation pathologischer Fälle, oftmals in Verbindung mit einer speziellen Modalität, z.B. Lungen-CTs [4].

3 Referenzkategorisierung des Datenkorpus

Der zur manuellen Referenzkategorisierung verwendete Schlüssel [2] besitzt für die vier fundamentalen Aspekte bei der Bildaufnahme je eine Achse: (a) Modalität, (b) Aufnahmeparameter (Orientierung, Marker), (c) Anatomie und (d) Biologisches System. Innerhalb jeder Achse ist eine Hierarchie mit einer Tiefe bis zu vier Stufen definiert. In jeder Stufe ist zusätzlich der Wert "nicht näher spezifiziert" wählbar, um eine Kategorisierung bei dieser Tiefe zu beenden.

Der Korpus umfaßt momentan 6514 Aufnahmen verschiedener Modalitäten, größtenteils konventionelles und digitales Röntgen sowie Durchleuchtung mit und ohne Kontrastmittel, konventionelle Mammographie, digitale Subtraktionsangiographie und Computertomographie. Die manuelle Referenzkategorisierung erfolgte durch erfahrene Radiologen. Zur Unterstützung der Radiologen wurde eine webbasierte Benutzerschnittstelle mit kontextsensitiver Eingabehilfe für den Schlüssel implementiert [5]. Bislang wurden 2470 Aufnahmen manuell kategorisiert, wobei der Anatomiecode für 1795 davon fixiert ist. Bei weiteren 675 Aufnahmen existiert eine Kategorisierung bzgl. der obersten Stufe des Anatomiecodes. Tabelle 1 illustriert die Verteilung der Aufnahmen bezüglich der Anatomie-Achse in der Verschlüsselung.

4 Verfahren zur automatischen Kategorisierung

Die Kategorisierung erfolgt anhand globaler Merkmale, d.h. es findet eine starke Komprimierung der Bildinformation auf wenige aussagekräftige Parameter statt. Ein Vorteil ist neben dem geringen Speicherplatzbedarf auch die schnelle Berechnung der Kategorisierungsentscheidung. Drastische Herunterskalierungen der einzelnen Aufnahmen auf eine einheitliche Größe (zunächst unter Verlust des Original-Seitenverhältnisses) haben sich in der Vergangenheit als globale Merkmale gegenüber globalen Texturbeschreibungen (z.B. fraktale Dimension, DCT-basierte Merkmale, Cooccurrence-Matrizen) oder Kanteninformationen als leistungsfähiger erwiesen [6,7]. Wir verwenden in unseren Experimenten einen k-Nearest-Neighbour-Klassifikator mit $k=5$. Als Distanzmaße innerhalb des Klassifikators werden der Euklidische Abstand und die Mahalanobis-Distanz verwendet. Bei Verwendung der skalierten Bilder werden ferner Distanzmaße verwendet, die eine Invarianz bzgl. geringfügiger lokaler Transformationen modellieren: die Tangentendistanz und die Korrelation. Während die bei der Korrelation das Minimum über einem Fenster gegeneinander verschobener Bildpaare gesucht und damit lediglich die Translation variiert wird, modelliert die Tangentendistanz

Tabelle 1. Häufigkeiten der Körperregionen im Referenzkorpus.

Region	Subregion	Code	Anz.	Region	Subregion	Code	Anz.
Ganzkörper	noch nicht gesetzt	1XX	0	Thorax	noch nicht gesetzt	5XX	47
	nicht näher spez.	100	0		nicht näher spez.	500	433
	Rumpf	11.	0		Knochen	51.	0
	Extremitäten	12.	0		Lungen	52.	0
Kopf (Schädel)	noch nicht gesetzt	2XX	164	Hilus	53.	0	
	nicht näher spez.	200	1	Mediastinum	54.	0	
	Gesichtsschädel	21.	5	Herz	55.	0	
	Schädelbasis	22.	18	Zwerchfell	56.	0	
Wirbelsäule	Hirnschädel	23.	6	Brust (Mamma)	noch nicht gesetzt	6XX	0
	noch nicht gesetzt	3XX	129		nicht näher spez.	600	140
	nicht näher spez.	300	0	Bauch (Abdomen)	noch nicht gesetzt	7XX	88
	Halswirbelsäule	31.	51		nicht näher spez.	700	29
	Brustwirbelsäule	32.	23		Oberbauch	71.	40
	Lendenwirbelsäule	33.	57		Mittelbauch	72.	45
Kreuzbein	34.	0	Unterbauch	73.	47		
Steißbein	35.	0	Becken (Pelvis)	noch nicht gesetzt	8XX	29	
Obere Extremität (Arm)	noch nicht gesetzt	4XX	64	nicht näher spez.	800	40	
	nicht näher spez.	400	0	Kreuzbein	81.	0	
	Hand	41.	401	Darmbeinschaukel	82.	2	
	Handgelenk	42.	31	Schambein	83.	0	
	Unterarm	43.	16	Kleines Becken	84.	0	
	Ellenbogen	44.	35	Untere Extremität (Bein)	noch nicht gesetzt	9XX	154
	Oberarm	45.	6		nicht näher spez.	900	2
	Schulter	46.	36		Fuß	91.	15
			Sprunggelenk		92.	16	
			Unterschenkel		93.	15	
			Knie		94.	250	
			Oberschenkel	95.	26		
			Hüfte	96.	9		
							2470

mit Erweiterungen [7] darüber hinaus lokale Transformationen bzgl. Rotation, Scherung, Skalierung und Helligkeit.

5 Untersuchung

Zur Untersuchung, bis zu welcher Schlüsseltiefe eine automatische Erkennung möglich ist, wurden verschiedene Granularitäten bei der Kategorieneinteilung gewählt. Die Anatomiekodierung des verwendeten Schlüssels ist dreistufig, wobei die oberste Stufe acht Kategorien definiert. Der Korpus umfaßt bereits jetzt hinreichend viele Bilder, um die Kategorisierung bis zur zweiten Stufe zu verfolgen. Hierbei werden die nicht näher spezifizierten Bilder des Korpus ausgeklammert. Die Kategorie "Brust (Mamma)" besitzt keine weitere Unterteilung, ferner

Tabelle 2. Einteilung des Referenzkorpus in 24 Kategorien.

Kategorie	Anz.	Kategorie	Anz.	Kategorie	Anz.
Gesichtsschädel	5	Schädelbasis	18	Hirnschädel	6
Halswirbelsäule	51	Brustwirbelsäule	23	Lendenwirbelsäule	57
Hand	401	Handgelenk	31	Unterarm	16
Ellenbogen	35	Oberarm	6	Schulter	36
Thorax	480	Brust (Mamma)	140	Oberbauch	40
Mittelbauch	45	Unterbauch	47	Becken (Pelvis)	71
Fuß	15	Sprungelenk	16	Unterschenkel	15
Knie	250	Oberschenkel	26	Hüfte	9
					1839

wurden die Kategorien “Thorax” und “Becken (Pelvis)” mangels vorliegenden Bildmaterials nicht weiter unterteilt. Insgesamt resultiert eine Untermenge von 1839 Bildern in 24 Kategorien (siehe Tabelle 2). Zwecks Vergleichbarkeit mit früheren Publikationen wurde zusätzlich eine Kategorisierung mit nur sechs Kategorien betrachtet. Bei dieser verschmelzen die Kategorien “Abdomen” und “Becken” sowie “Arm” und “Bein” zu jeweils einer.

6 Ergebnisse

Tabelle 3 zeigt die erzielten Erkennungsraten unter Verwendung von Leaving-One-Out für die unterschiedlichen Auswahlen der Kategorien (6/8 Kategorien: 2470 Aufnahmen, 24 Kategorien: 1839 Aufnahmen). Die Verwendung von 16x16 Pixel großen Skalierungen der Aufnahmen erwies sich in den Experimenten als optimal. Für die oberste Stufe des Anatomiecodes wird eine beste Erkennungsrate von 82.5% unter Verwendung der Tangentendistanz erzielt. Die meisten Erkennungsfehler traten durch die Verwechslung der Kategorien “Bein”, “Arm” und “Wirbelsäule” auf (siehe Tabelle 4). Eine automatische Kategorisierung bis zur zweiten Stufe (24 Kategorien) war mit einer Erkennungsleistung von 80.9% unter Verwendung der Korrelations-Abstandsmaßes möglich. Zum Vergleich ist die Erkennungsleistung eines NN angegeben, der auf globalen Texturmaßen (u.a. der fraktalen Dimension des Bildes) basiert [8].

Tabelle 3. Erzielte Erkennungsraten (Leaving-One-Out).

	Anzahl Kategorien		
	6	8	24
Distanzmaß (5-NN)	6	8	24
16x16, Euklidischer Abstand	83.7%	79.0%	76.5%
16x16, Korrelation (5x5-Fenster)	86.2%	81.5%	80.9%
16x16 Tangentendistanz nach [7], keine Optimierung	86.3%	82.2%	80.3%
7 Texturmerkmale			
Euklidischer Abstand [8]	-	49.3%	-

Tabelle 4. Ergebnis der automatischen Kategorisierung mit 81.5% Erkennung.

	Kopf	Wirbelsäule	Arm	Thorax	Mamma	Abdomen	Becken	Bein
Kopf	156	5	4	2	3	4	1	19
Wirbelsäule	3	190	11	4	2	3	0	47
Arm	2	16	470	21	9	3	0	68
Thorax	4	7	17	435	0	1	0	16
Mamma	1	0	3	0	119	0	0	17
Abdomen	10	10	8	3	1	183	10	24
Becken	0	1	0	0	1	7	48	14
Bein	5	10	30	17	7	4	1	413

7 Diskussion

Eine weitere Verbesserung der Erkennungsleistung ist erforderlich. Hierbei sind hierarchische Ansätze zur gezielten Differenzierung bislang schwer trennbarer Kategorien (hier: Extremitäten und Wirbelsäule) denkbar, etwa durch die Kombination verschiedener Merkmale bzw. Klassifikatoren. Die mittelfristige Erweiterung des Korpus auf 10.000 Aufnahmen ermöglicht eine genauere Untersuchung existierender Verfahren und weitere Experimente mit noch feinerer Granularität bei der Kategorisierung.

Literaturverzeichnis

1. Lehmann TM, Wein BB, Dahmen J, Bredno J, Vogelsang F, Kohnen M: Content-Based Image Retrieval in Medical Applications - A Novel Multi-Step Approach, Procs SPIE 2000; 3972: 312-320
2. Lehmann TM, Wein BB, Keysers D, Kohnen M, Schubert H: A Monohierarchical Multiaxial Classification Code for Medical Images in Content-Based Retrieval, Procs IEEE ISBI 2002; 313-316
3. Güld MO, Kohnen M, Keysers D, Schubert H, Wein BB, Bredno J, Lehmann TM: Quality of DICOM Header Information for Image Categorization, Procs SPIE 2002; 4685: 280-287
4. Shyu CR, Brodley CE, Kak AC, Kosaka A, Aisen A, Broderick LS: ASSERT: a physician-in-the-loop content-based retrieval system for HRCT image databases, Computer Vision and Image Understanding 1995; 75: 111-132
5. Plodowski B, Güld MO, Schubert H, Keysers D, Lehmann TM: Modulares Design von webbasierten Benutzerschnittstellen für inhaltsbasierte Zugriffe auf medizinische Bilddaten, Proc. BVM 2003; in press
6. Paredes R, Keysers D, Lehmann TM, Wein B, Ney H, Vidal E: Classification of Medical Images Using Local Representations, Proc. BVM 2002; 171-174
7. Dahmen J, Keysers D, Motter M, Ney H, Lehmann T, Wein B: An Automatic Approach to Invariant Radiograph Classification, Proc. BVM 2001; 337-341
8. Chaudhuri BB, Sakar N: Texture Segmentation using Fractal Dimension, IEEE PAMI 1995; 17: 72-76