

DISS. ETH NO. 21408

Strategies for False Positive Reduction and Multimodal Lesion Characterization in Computer-aided Diagnosis of Breast Cancer

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by

Jan Martin Lesniak

Diplom-Informatiker, Universität Karlsruhe (TH),
born November 24, 1980
in Frankfurt (Main), Germany

accepted on the recommendation of

Prof. Dr. Gábor Székely, examiner
Prof. Dr. Nico Karssemeijer, co-examiner
Dr. Remi Blanc, co-examiner

2013

Abstract

For 2013, 88886 deaths caused by female breast cancer are estimated in the European Union. One of the key requirements for successful breast cancer treatment is to identify the disease early, which is approached by systematic screening with X-ray mammography. In addition, also multimodal imaging is of clinical interest, including applications such as screening women at high risk of developing breast cancer using combined mammography (MG) and magnetic resonance imaging (MRI). In order to reduce the impact of oversight and misinterpretation errors in the image reading procedure, computer-aided lesion detection and classification systems are developed which aim to provide a second opinion for the radiologist.

Especially the computer-aided detection of breast masses in mammography is challenging and current systems still generate many false positives, hampering their clinical use. Similarly, the proven utility of multimodal imaging in clinical practice opens ground for computer-aided lesion characterization tools, which strive to provide means for automated lesion assessment using multimodal data. This work focuses on machine learning aspects of automated diagnostic support in breast imaging. Approaches for false positive reduction in computer-aided breast mass detection in mammography are investigated and a machine learning framework for multimodal lesion characterization in combined MG-MRI imaging is proposed.

In computer-aided detection of breast masses, first the reduction of false positive marks was approached by improving the interpretation of given image descriptors. A support vector machine (SVM)-based framework was proposed and studied using a large database of film mammograms. A comparative evaluation against several other state-of-the-art classification algorithms suggested its improved false positive reduction capabilities on normals in varying training scenarios. Second, the utility of texture analysis for the purpose of distinguishing false positive against true positive cancer candidates was analyzed. For this purpose, the SVM-based mass detection approach was extended by introducing a rich set of texture descriptors which quantified the properties of region candidates. A systematic analysis of each individual texture descriptor group yielded an improved, yet parsimonious feature set for breast mass detection which provided a statistically significant improvement in system performance.

The third aspect of this work was a machine learning architecture for multimodal MG-MRI lesion characterization, again relying on SVM for classification but also for feature selection. Unimodal respectively multimodal feature selection were subject to investigation in combination with different classification schemes for generation of a multimodal lesion likelihood.

Among the evaluated frameworks, multimodal classifier training in combination with unimodal feature selection provided the best classification performance, while a setup relying on a combination of two independent unimodal lesion assessments provided similar results and showed less dependency on feature selection. The multimodal training approach was adopted for further investigation, in which multimodal classification proved to be superior to an MG- or MRI-only approach and yielded comparable performance to six radiologists in an observer study.

Zusammenfassung

Für das Jahr 2013 werden in der Europäischen Union 88886 Todesfälle durch Brustkrebs erwartet. Eine der Hauptvoraussetzungen für die erfolgreiche Behandlung von Brustkrebs ist dessen frühe Erkennung, für welche Screening durch Röntgenmammographie eingesetzt wird. Zusätzlich wächst das klinische Interesse an multimodaler Bildgebung, bei der verschiedene Bildgebungsverfahren miteinander kombiniert werden, bspw. Röntgenmammographie (MG) und Magnetresonanztomographie (MRT) für Frauen mit erhöhtem Brustkrebsrisiko. Um den Effekt von übersehenen oder falsch interpretierten Tumoren in der radiologischen Inspektion der Bilddaten zu reduzieren, werden computergestützte Verfahren zur Detektion und Klassifikation entwickelt, welche zum Ziel haben eine zweite Meinung für den auswertenden Radiologen bereitzustellen.

Speziell die Identifikation von masseartigen Veränderungen der Brust in der Mammographie stellt hohe Anforderungen an computergestützte Systeme. Aktuelle Methoden generieren viele falsch positive Detektionen, was die klinische Anwendbarkeit dieser Systeme erschwert. Analog erlaubt der sich verbreitende Einsatz multimodaler Bildgebung eine erweiterte Analyse masseartiger Veränderungen in der Brust, welche eine Basis für die Entwicklung von computergestützten Modellen zur Charakterisierung möglicher Tumorregionen bietet. Die vorliegende Arbeit konzentriert sich auf Aspekte des maschinellen Lernens in der computergestützten Erkennung und Diagnose von Brustkrebs. Es werden Ansätze zur Reduktion von falsch positiven Detektionen in der Röntgenmammographie untersucht und ein System zur Charakterisierung von Tumorkandidaten in der multimodalen Mammographie-MRT Bildgebung vorgestellt.

In der computergestützten Erkennung von masseartigen Tumorregionen in der Mammographie wurde zunächst die Interpretation von automatisch extrahierten Bildmerkmalen verbessert. Für diesen Zweck wurde mittels Support Vector Machines (SVM) ein Detektionssystem erstellt und auf einer großen Datenbank mit Filmmammographien ausgewertet. Ein Vergleich mit mehreren aktuellen Klassifikatoren zeigte eine verminderte Anzahl von falsch positiven Detektionen auf normalen Röntgenmammographien unter verschiedenen Trainingsbedingungen. Anschließend wurde das vorhandene System um Texturinformationen aus lokalen Bildausschnitten ergänzt und deren Anwendbarkeit für die Unterscheidung von falschen und korrekten Detektionen untersucht. Zu diesem Zweck wurde eine große Anzahl von Texturdeskriptoren extrahiert und für die Klassifikation von Kandidatenregionen verwendet. Eine systematische Analyse der ver-

schiedenen Texturinformationen erlaubte die Identifikation einer Gruppe von Bildmerkmalen welche eine statistisch signifikant verbesserte Detektionsleistung erbrachte.

Der dritte Schwerpunkt dieser Arbeit war die multimodale MG-MRT-basierte Klassifikation von möglichen Tumorregionen in der Brust, welche wiederum auf SVMs fußte. Zusätzlich zur Klassifikation von Kandidatenregionen wurden SVMs aber auch zur automatischen Merkmalsselektion verwendet. Hier wurden verschiedene Ansätze zur Kombination von multimodalen Deskriptoren und deren automatischen Selektion untersucht. Unter den ausgewerteten Ansätzen lieferte die multimodale Klassifikation zusammen mit unimodaler Merkmalsselektion die besten Ergebnisse, wobei die Kombination von zwei unabhängigen unimodalen Klassifikatoren ähnliche Resultate lieferte und weniger Abhängigkeit von der automatischen Merkmalsselektion zeigte. Weiter wurde der multimodale Trainingsansatz im Rahmen einer Benutzerstudie mit Radiologen ausgewertet, wobei die multimodale Klassifikation bessere Ergebnisse lieferte als ein reiner MG- bzw. MRT-basierter Ansatz. Die Auswertung der Klassifikationsleistungen ergab vergleichbare Resultate für die SVM-basierte multimodale Charakterisierung und die individuelle Diagnose durch sechs Radiologen.