International Journal of Medical Informatics 117 (2018) 44–54

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification

Mugahed A. Al-antari[a,1], Mohammed A. Al-masni[a,1], Mun-Taek Choi[b], Seung-Moo Han[a], Tae-Seong Kim[a,*]

[a] Department of Biomedical Engineering, College of Electronics and Information, Kyung Hee University, Yongin, 17104, Republic of Korea
[b] School of Mechanical Engineering, Sungkyunkwan University, Republic of Korea

ARTICLE INFO

Keywords:
Computer-aided diagnosis (CAD)
Mass detection
You-only-look-once (YOLO)
Mass segmentation
Full resolution convolutional network (FrCN)
Deep learning

ABSTRACT

A computer-aided diagnosis (CAD) system requires detection, segmentation, and classification in one framework to assist radiologists efficiently in an accurate diagnosis. In this paper, a completely integrated CAD system is proposed to screen digital X-ray mammograms involving detection, segmentation, and classification of breast masses via deep learning methodologies.

In this work, to detect breast mass from entire mammograms, You-Only-Look-Once (YOLO), a regional deep learning approach, is used. To segment the mass, *full resolution convolutional network (FrCN)*, a new deep network model, is proposed and utilized. Finally, a deep *convolutional neural network (CNN)* is used to recognize the mass and classify it as either benign or malignant. To evaluate the proposed integrated CAD system in terms of the accuracies of detection, segmentation, and classification, the publicly available and annotated INbreast database was utilized. The evaluation results of the proposed CAD system via four-fold cross-validation tests show that a mass detection accuracy of 98.96%, Matthews correlation coefficient (MCC) of 97.62%, and F1-score of 99.24% are achieved with the INbreast dataset. Moreover, the mass segmentation results via FrCN produced an overall accuracy of 92.97%, MCC of 85.93%, and Dice (F1-score) of 92.69% and Jaccard similarity coefficient metrics of 86.37%, respectively. The detected and segmented masses were classified via CNN and achieved an overall accuracy of 95.64%, AUC of 94.78%, MCC of 89.91%, and F1-score of 96.84%, respectively. Our results demonstrate that the proposed CAD system, through all stages of detection, segmentation, and classification, outperforms the latest conventional deep learning methodologies. Our proposed CAD system could be used to assist radiologists in all stages of detection, segmentation, and classification of breast masses.

1. Introduction

Breast cancer is considered to be one of the most common types of cancer affecting women worldwide. Statistical results published in 2017 categorized breast cancer among the highest levels of all other cancers, accounting for 30% of estimated new cases and 14% of deaths [1]. In 2008, the World Health Organization (WHO) reported that 13.7% of deaths among women worldwide was due to breast cancer [2]. Early detection of breast cancer is a critical requirement for reducing the mortality rate among women [2–5]. At present, digital X-ray mammography is the most reliable screening device for suspicious breast masses and microcalcifications in the early stages [3,4,6,7]. Indeed, women over 40 years old are encouraged by the National Cancer Institute (NCI) to undergo breast screening one or two times per year using both views of mammograms: mediolateral oblique (MLO) and cranio-caudal (CC) [8]. In the diagnosis of breast abnormalities, clinical experts classify suspicious masses as benign or malignant. This task presents a daily challenge for radiologists due to the huge number of mammograms as well as the time and effort to examine each view of a mammogram [4,9,10]. Thus, a tradeoff between sensitivity and specificity has been realized during the diagnosis process. Through the use of a second reading, either by other experts or by a computer-aided diagnosis (CAD) system, the overall accuracy and specificity of mass detection, segmentation, and classification could be improved [3,11] and false positive and negative cases reduced. A reliable and robust CAD system could be of significant assistance in clinical practices [12,13].

* Corresponding author at: Department of Biomedical Engineering, College of Electronics and Information, Kyung Hee University, 1732, Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do, 17104, Republic of Korea.

E-mail addresses: en.mualshz@khu.ac.kr (M.A. Al-antari), m.almasani@khu.ac.kr (M.A. Al-masni), mtchoi@skku.edu (M.-T. Choi), smhan@khu.ac.kr (S.-M. Han), tskim@khu.ac.kr (T.-S. Kim).

[1] Authors contributed equally to this work

https://doi.org/10.1016/j.ijmedinf.2018.06.003
Received 5 April 2018; Received in revised form 22 May 2018; Accepted 6 June 2018

There have been active developments for such a breast CAD system in each specific area of detection, segmentation, and classification. However, there are few studies involving a completely integrated system.

Mass detection from breast images is considered an important preprocessing stage to detect potential regions (i.e., masses) for further analysis by a CAD system. In fact, the variation of the masses within the surrounding tissues in terms of texture, shape, size, as well as the location in mammograms, makes the detection task challenging [8,13,14]. The majority of conventional CAD systems rely on manually detected masses and their hand extracted features to recognize the suspicious masses as benign or malignant via conventional machine learning techniques [11,15–18]. So far, this practice has resulted in a rather significant number of false positives [19–23]. Recently, novel detection approaches based on deep learning were introduced into a CAD system to overcome the challenging tasks of mass detection from mammograms [10,19,24].

Breast mass segmentation also plays a crucial role in accurately extracting discriminative shape features of specific mass regions, while excluding surrounding tissues [19,25]. In fact, improving overall accuracy in addition to reducing false positive and negative rates by mass segmentation is a big challenge due to the strong association between the presence of masses and their irregularities in shape, size, and location with low contrast and ambiguous boundaries [6,26–28]. Many studies involving mass segmentation have utilized region growing, active contour, and Chan-Vese methods [6,26,27]. Unfortunately, these methods still lack performance in handling mass segmentation automatically, because the simple hand-crafted or semi-automatic features based on prior knowledge cannot deal with complex shape variations, as well as the different density distribution of the masses and their surrounding tissues [6]. Recently, a few studies based on deep learning models have offered a good alternative to other conventional segmentation methods, by automatically extracting deep high-level hierarchy features for mass segmentation directly from input raw data to avoid the problems of hand-crafting features [4,19,29,30].

The majority of CAD systems have been developed in the area of mass classification to distinguish breast cancer as either benign or malignant utilizing conventional machine learning classifiers [11,15,16,17,18,31]. To build such systems, a set of hand-crafted or semi-automatic features describing the characteristics of masses are required. These features must have a good discriminative power to distinguish between either benign or malignant mass abnormalities. In fact, the conventional CAD systems based on hand-crafted features suffer due to the high degree of similarity between mass vs. non-mass and benign vs. malignant breast tissues [18,31,32]. Alternatively, some new strategies based on deep learning have recently been proposed to handle the mass classification task [3,31]. These strategies can learn and extract deep high-level features from raw input data directly and achieve much better classification performance in comparison to the traditional approaches [3,4,10,19,33]. In addition, a hybrid CAD system based on a combination of hand-crafted and deep high-level features has presented good results in the mass classification of breast masses [29].

The main contribution of this study is a fully integrated CAD system including three deep learning stages (i.e., detection, segmentation, and classification). Also the newly proposed segmentation method of FrCN is presented for the masses of breast cancer. The main advantage of FrCN is to preserve the high resolution of feature maps. Especially for the edges of the objects, where FrCN learns the full resolution features of each pixel of the original input data to achieve more accurate pixel-to-pixel segmentation. This is achieved by removing the max-pooling and subsampling layers in the networks and enabling the convolutional layers to extract and learn the full resolution spatial features of the input image.

In this paper, a completely integrated CAD system is proposed based on deep learning to automatically detect, segment, and classify breast masses in a single framework. The rest of the paper is organized as follows. First, an automatic deep learning You-Only-Look-Once (YOLO)-based mass detection model is presented. Second, a newly proposed deep learning mass segmentation method, a full resolution convolutional network (FrCN), is proposed and compared against other existing state-of-the-art deep models. Finally, a deep learning convolutional neural network (CNN) classifier is presented to distinguish between benign or malignant detected and segmented masses. We validate the proposed integrated CAD system and compare it to the latest methodologies by utilizing the public INbreast database [34].

2. Literature review

Breast cancer diagnosis via a CAD system can be improved by using the deep high-level features of deep learning, which can represent the characteristics of the masses better [4,10,19]. Mass detection is an important stage in the CAD systems for breast cancer diagnosis [4,35]. It is a challenging problem and has not been fully resolved [19,35]. In general, manual mass detection was utilized in the CAD systems which used deep CNN to classify the masses as either benign or malignant [33,36,37]. In other CAD systems, these manually detected masses are directly fed into CNN to generate the integrated high-level deep features [10,19,37]. In some other CAD systems, the high-level deep features were extracted from the multiple layers of CNN, and then concatenated and fed into the classifier to distinguish between benign or malignant tissues [33]. Most of these CAD systems achieved better classification performance against the traditional machine learning techniques which depend on the hand-crafted features [23,24]. However, the automatic mass detection still remains as a challenge. The need to automatically detect breast abnormalities was addressed in several studies [10,19,24,38]. At present, few deep learning studies present automatic mass detection methods in CAD systems [14,19,24]. The preliminary mass detection results are presented utilizing deep learning YOLO technique using the digital database of mammography (DDSM) [9]. The detection performance via YOLO was better in comparison to other recently published detection methods [14,19]. A CAD system utilizing a deep belief network (DBN) was presented to analyze suspicious regions in mammograms [4]. In this work, for mass detection, adaptive thresholding and morphological operations were utilized achieving an overall detection accuracy of 86% [3,4]. In [35], a new deep model called region-based CNN (R-CNN) was proposed to automatically detect the masses of breast cancer [35]. The entire mammogram was divided into multiple patches to detect the masses locally. Then, R-CNN was trained to classify the detected regions as benign or malignant. In [14], another automatic method using a cascade of deep learning models was proposed for mass detection. This method involved four sequential steps to detect masses in breast abnormalities. First, a multi-scale deep belief network (mDBN) and Gaussian mixture classifier (GMC) was utilized to extract suspicious regions. Second, two-level cascade of R-CNN was used to reduce the false positive rate in these detected regions. All remaining regions were then fed again into a two-level cascade of a conditional random forest (CRF) classifier to enhance the process of false positive reduction. All potential regions that survived in the previous stages were combined using a connected component analysis (CCA) as a post-processing technique [14]. Finally, the refinement algorithm was utilized to improve the precision of mass detection [39]. This refinement algorithm was also implemented with deep R-CNN through two sequential steps [19]. First, Bayesian optimization was used to detect suspicious regions. Then, a deep structure of R-CNN based classifier was utilized to improve the scale and localization of the detected regions. Despite improving the results for automatic mass detection, challenges remained with the high complexities of memory, practical implementation, and long runtime.

Several conventional studies have segmented masses from X-ray mammography images. Growing regions based on gradient filters and simple edge detection have been widely used for mass segmentation [26,40,41]. Other studies have improved the results of mass