Expert Systems With Applications 64 (2016) 490–499

Contents lists available at ScienceDirect

Expert Systems With Applications

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

A computer vision-based algorithm to predict false positive errors in radiology trainees when interpreting digital breast tomosynthesis cases

Mengyu Wang[a,b,*], Meng Wang[c], Lars J. Grimm[a], Maciej A. Mazurowski[a]

[a] Department of Radiology, Duke University School of Medicine, Durham, NC, United States
[b] Schepens Eye Research Institute, Harvard Medical School, Boston, MA, United States
[c] Department of Computer Science, Columbia University, New York, NY, United States

ARTICLE INFO

Article history:
Received 23 April 2016
Revised 13 June 2016
Accepted 3 August 2016
Available online 4 August 2016

Keywords:
Radiology education
False positive error prediction
Training plan optimization
Digital breast tomosynthesis
Image processing
Clustering

ABSTRACT

Objectives: Digital breast tomosynthesis (DBT) is a new imaging modality that improves invasive cancer detection rates compared to mammography. In this work, we aim to advance adaptive computer-based education in DBT by computer algorithm.
Methods: First, a set of potentially difficult locations are identified based on locations marked by other trainees using a regional clustering algorithm. Second, the candidate location is segmented to identify potential abnormal objects. Third, 18 features are extracted from the location from the segmented image. Finally, a classifier uses the 18 features to predict whether the candidate location would result in a false positive error for the trainee. The classifier is personalized for each trainee by using data from the trainee's prior DBT interpretations.
Results: Our algorithm successfully identified locations more likely associated with false positive errors as compared to randomly identified locations. The prevalence of errors among the difficult locations was 20.7% when 1 location per trainee was predicted and 17.2% when 10 locations were predicted. In comparison, the prevalence of errors for random locations generated within a breast region with 1 and 10 identified locations was 0% and 4.8%, respectively.
Conclusions: We developed an algorithm to successfully identify locations on DBT where trainees are more likely to commit false positive errors.
Advances in knowledge: Our user model can be used to select the most challenging cases for each trainee from the perspective of committing false positive errors. Our model improved the status quo of case presentation with random selection to trainee in breast tomosynthesis.

1. Introduction

Digital breast tomosynthesis (DBT) is a relatively new imaging modality for breast cancer detection which reduces breast tissue superposition in comparison to mammography (Baker & Lo, 2011; Sechopoulos, 2013). DBT can reduce screening recall rates and increase invasive cancer detection rates compared to mammography (Ciatto et al., 2013; Rafferty et al., 2013; Skaane et al., 2013). As DBT is being adopted into clinical practice, new educational strategies are needed to teach DBT, as previous work has demonstrated that interpretive skills for mammography may not translate to DBT (Zhang et al., 2015). Ideally, educational approaches to DBT for radiology trainees should be optimized to be as efficient as possible.

Towards this goal, our group has been developing adaptive computer-aided education systems for mammography (Mazurowski et al., 2014; Mazurowski, Baker, Barnhart, & Tourassi, 2010; Mazurowski, Barnhart, Baker, & Tourassi, 2012; Zhang, Silber, & Mazurowski, 2015). Specifically, we have developed algorithms to identify trainee error making patterns, both false positive and false negative, by using computer extracted image features, which eventually can be used to select challenging cases for training (Zhang & Silber, et al., 2015; Zhang et al., 2014). Related areas of investigation explored by other groups include the development of training system ontologies and establishing the relationship between error and eye gaze (Sun, Taylor, Wilkinson, & Khoo, 2008; Voisin, Pinto, Morin-Ducote, Hudson, & Tourassi, 2013).

* Corresponding author.
E-mail addresses: mengyu_wang@meei.harvard.edu (M. Wang), mw2972@columbia.edu (M. Wang), lars.grimm@duke.edu (L.J. Grimm), maciej.mazurowski@duke.edu (M.A. Mazurowski).

http://dx.doi.org/10.1016/j.eswa.2016.08.023
0957-4174/

In this work, we extend our approach to developing a personalized computer-aided education system by a new algorithm that identifies challenging normal locations on DBT that are more likely to cause trainees to make false positive errors. By utilizing computer vision and machine learning, we aim to characterize the error making pattern of each trainee to then identify difficult cases from a large dataset. We expect that radiology education efficiency can be improved by presenting each trainee with cases they are personally more likely to find difficult. Towards this goal, we will focus on modeling false positive errors by trainees in DBT.

2. Methods

Our proposed methodology is described in the following subsections: (1) reader study and false positive error definition, (2) algorithm for false positive error prediction including the clustering algorithm to identify candidate locations, image segmentation, feature extraction and predictive modeling, and (3) evaluation of the predictive model.

2.1. Reader study and the definition of false positive errors

We validated our algorithm for false positive error prediction using data from a reader study in which 3 fellowship-trained breast radiologists and 29 radiology trainees interpreted the same set of 60 DBT studies. All readers were instructed to annotate suspicious locations using a graphical user interface designed to simulate a clinical workstation. Each DBT study consisted of craniocaudal (CC) and medio-lateral oblique (MLO) views of a single breast. The trainees had varying degrees of breast imaging experience, but no formalized training in DBT. The three expert breast radiologists were all certified to read DBT and had all participated in prior DBT reader studies. Institutional Review Board approval was secured for this study.

We used the annotations made by the three expert breast radiologists to determine the true lesion locations. If at least two of the three radiologists put an annotation within a 9 mm distance of another annotation, then we considered this a true lesion and the centroid location of the radiologists' annotations was used as the true lesion location. The distance criteria of 9 mm was used because it was the average radius of breast cancer seen on mammography as previously reported (Timp, Karssemeijer, & Hendriks, 2003). We identified the candidate false positive error locations by taking the locations marked by the trainees and excluding the locations that were within 9 mm from the true lesion locations. The remaining locations identified by the trainees were considered candidate false positive error locations.

2.2. Algorithm for false positive error prediction

To identify locations that are likely to be associated with false positive errors individually for each trainee, we first identified the candidate locations using a clustering algorithm. We then excluded the true lesion locations leaving the candidate locations for false positive errors. Second, we applied image processing algorithms to segment the candidate locations. Then, we extracted the image features from the candidate false positive error locations based on the segmented images. Lastly, we applied a classifier to identify the false positive error locations for each trainee based on the extracted image features. The overall process of the predictive model is shown in Fig. 1.

In addition, we also evaluated our algorithm for all reader annotations including false positive annotations and true positive annotations to see how well our algorithms can predict the trainees' lesion identification behavior in general. Since the main focus of this manuscript is to predict the false positive error from imaging features in radiology trainees, the algorithm is described in the context of false positive error prediction.

Step 1: Clustering algorithm for obtaining candidate locations

The first step is to identify a set of candidate locations that then might be identified as high risk for a false positive error. We developed a regional clustering algorithm to identify the candidate locations. For a specific trainee, the candidate locations were identified based on the locations marked by other trainees.

Fig. 2 shows the overall structure of the proposed clustering algorithm. Starting from $\mathbf{x}_r$ randomly selected in all location set $\mathbf{X}$, we searched and calculated each cluster centroid by implementing Eq. (1) and (2) where $\mathbf{x}_c$ is the average location of members from each cluster. Afterward, we excluded members of each cluster $\mathbf{X_c}$ from $\mathbf{X}$ as shown in Eq. (3). The procedure was repeated until all location set X was empty and same clustering process was performed for next image volume until all image volumes were traversed. After we obtained the candidate locations from clustering, we excluded the cluster locations that were within a 9 mm distance from the true lesion locations to obtain the candidate locations for the false positive error modeling.

$$X_c = \{x_i | \|x_i - x_r\|_{l_2} \leq 9,\ x_i \in X\} \tag{1}$$

$$x_c = \bar{X}_c \tag{2}$$

$$X = X \backslash X_c \tag{3}$$

Step 2: Image segmentation

The proper segmentation step was preceded by image enhancement, including a median filter to smooth the images and a top hat filter to enhance the contrast between the candidate location region and the background tissue. In addition, we constructed a constraint matrix and multiplied it by the original image matrix to suppress the background tissue and enhance the candidate location region. The constraint matrix was generated based on an isotropic radial basis function centered on the candidate location region with a variance σ^2 (σ is 9 mm).

We used a 3D active contour method implemented by a level set algorithm to segment the candidate location region (Chan & Vese, 2001). The level set function φ was defined to be positive inside of a closed surface Γ, negative outside of Γ, and zero on Γ. In the image segmentation, an energy function ν was defined as a function of the current geometry information of Γ (i.e. candidate location region boundary in this context) and the image intensity information inside and outside of Γ. The level set function can be evolved as:

$$\frac{\partial \varphi}{\partial t} = \nu |\nabla \varphi| \tag{4}$$

Where ν is the energy function (also known as velocity) related to the shape information of the current surface Γ (i.e. object boundary) and the image intensity information inside and outside of Γ in the context of the level segmentation as defined below:

$$\nu(\Gamma) = \int |u_0(\vec{x}) - c_1|^2 (d\Gamma_{in}) + \int |u_0(\vec{x}) - c_2|^2 \times (d\Gamma_{out} + \lambda \cdot Area(\Gamma)) \tag{5}$$

Where u_0 is the gray level intensity function of the image, c_1 is the average gray level intensity inside of Γ, c_2 is the average gray level intensity outside of Γ and λ is the area penalty parameter to control the surface smoothness of the segmented 3D object. As such λ is an important parameter in the calculation of energy function. The surface with a zero level set Γ was considered as the segmented candidate location region boundary after