Fine-grained mommagraphy object detection with multiple feature fusion transfer learning

Ri-Gui Zhou[1, 2]**, Shihao Lv**[1, 2]**, Yaochong Li**[1, 2]

[1] College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China.

[2] College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China.

Keywords: CNNs, Random Forest, Learning to Hash, DensNet, Feature Fusion

Abstract

Breast cancer is the second leading cause of cancer deaths among women and screening mammography has been found to reduce mortality. It is necessary to build mammography image recognition systems. Previous solutions for mammography image recognition are usually based on hand-crafted features methods and use but limited in specific situations. Also, some methods utilize simple deep learning network, which can not extract the full feature of mammography, for example, GoogleNet, VGG16, ResNet, DenseNet, etc. In this paper, we propose a deep learning based approach with multiple feature fusion transfer learning strategy. Firstly, we obtain the training data from an open data set called DDSM images. Then we employ data augment methods, and training a deep convolutional neural network to extract image features and conduct the object detection job. A pre-trained model is used to initialize the network and help extract the basic features. Furthermore, we propose a fusion method that makes use of multiple transfer learning models in inference, to improve the accuracy on the test set. Importantly, we take a strategy applied by hash learning in the deep network is cited to enhance the generalization ability of the model and solve the challenge of high-dimensional calculation in deep learning. In the end, regression analysis to analyze the object position. The experimental results prove that our method achieves high accuracy on the mammography image object detection and inspection task.

1 Introduction

The rapid development of machine learning and especially deep learning has attracted the medical imaging community's interest in applying these techniques to improve the accuracy of cancer screening. Breast cancer is the second leading cause of cancer deaths among U.S. women and screening mammography has been proved to reduce mortality **??**. According to a study by the Breast Cancer Surveillance Consortium in 2009, the overall sensitivity of digital screening mammography in the U.S. is 84.4% and the overall specificity is 90.8% **?**. To help radiologists improve the predictive accuracy of screening mammography, computer-assisted detection and diagnosis (CAD) software have been developed and in clinical diagnosis since the 1990s. Unfortunately, data suggests that commercial CAD systems have not led to significant improvement in performance and progress has stagnated in the past decade **?**. With the remarkable success of deep learning in visual object recognition and detection, and many other domains have paid more attention to it **?**. There is much interest in developing deep learning tools to assist radiologists and improve the accuracy of screening mammography.

Early detection of subclinical breast cancer on screening mammography is challenging as an image classification task because the tumors themselves occupy only a small portion of the image of the entire breast. For example, a full-field digital mammography (FFDM) image is typically 4000×3000 pixels while a cancerous region of interest (ROI) can be as small as 100×100 pixels. If ROI annotations were widely available in mammography databases then established object detection and classification methods such as the region-based convolutional neural network (R-CNN) and its variants could be readily applied **???**. However, approaches that require ROI annotations **?** often cannot be transferred to large mammography databases that lack ROI annotations, which are laborious and costly to assemble. Indeed, few public mammography databases are annotated. Yet, deep learning requires large training datasets to be most effective. Thus, it is essential to leverage both the few fully annotated datasets, as well as larger datasets labeled with only the cancer status of each image to improve the accuracy of breast cancer classification algorithms.

Pre-training is a promising method to address the training problem. For example **??**, used layer-wise pre-training to initialize the weight parameters of a deep belief net (DBN) with three hidden layers and then fine-tuned it for classification. They found that pre-training improved the training speed as well as the accuracy of handwritten digit recognition. Another popular training method is to first train a deep learning model on a large database such as the ImageNet **??** and then fine-tune the model for another task. Although the specific task may not be related to the initial training dataset, the model's weight parameters are already initialized for recognizing primitive features, such as edges, corners and textures, which can be readily used for a different task. This often saves training time and improves the model's performance **??**.

To take advantage of feature extraction ability of CNNs, recently researchers have proposed new mammography recognition methods. However, we think the ability to extratct the image features of deep CNNs could be better utilized. In this study, fusing different deep neural network models together to propose a fine-grained mammography image recognition approach via fusing multiple learned features. Besides, utilizing the hash decoder method to simplifying the classification calculation of complex high-dimensional feature vectors, and fusing the results of different classifier in the end. Our work can be summarized as follows:

1. Build a deep convolutional neural network based on well-performed network structures, and design a transfer learning strategy to improve the representation power of the extracted features.

2. To further utilize the feature representation ability of CNNs, we propose a method to fuse extracted features from multiple trained models with similar topological structure to further improve the classification accuracy.

3. The strategy applied by hash learning in the deep network is cited to enhance the generalization ability of the model and solve the challenge of high-dimensional calculation in deep learning.

4. The designed feature fusion model is used in the feature extraction process of Faster RCNN, and a deep hash learning strategy is introduced in the classification stage.

The rest of this paper is organized as follows. Section **??** details of the network structure and the proposed method. Section **??** elaborates the dataset used in our work. Section **??** demonstrates the experiment results and gives discussion. Section **??** makes the conclusion and future work are provided in the end.

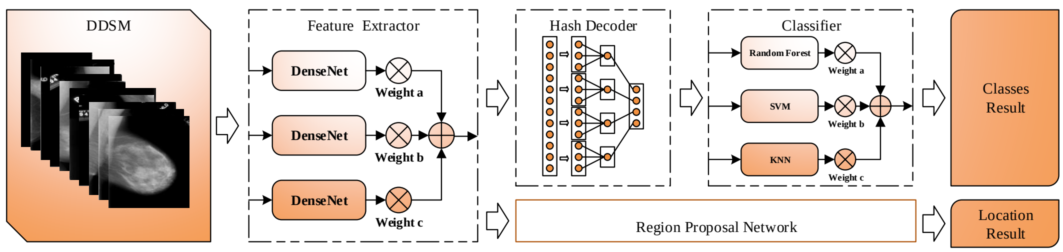

Figure 1: Overview of the proposed approach.

2 Sections and Subsections

This is the section level. The sections are numbered automatically. The $\backslash label$ and $\backslash ref$ can be used to label and refer to particular sections.

2.1 Subsections

This is the subsection level. An example for a reference is given here, see section **??**.

3 Itemize

The simple example for an $\backslash itemize$ is as following

- First item;
- Second item;

The items start with a bullet is as following

- First item;
- Second item;

The items start with a letter is as following

A) First item;

B) Second item;

The items start with a number is as following

1) First item;
2) Second item;

4 Figures

The Figures are included in the $\backslash begin\{figure\}$ environment, the figures in eps format are preferred, but other formats are acceptable as well.

To refer to a figure, we can use the reference, for example, see Figure **??**.

5 Equations

Single line equation is

$$a = b + c. \tag{1}$$

Equation array is

$$x = y + z; \quad (2)$$

$$a = b + c. \quad (3)$$

Equations should be numbered, however, we can generate equations without numbers. Use $\backslash nonumber$ and $\backslash begin\{eqnarray*\}$ to suppress the equation numbers, for example,

$$a = b + c;$$

which is the same as Eq.(**??**).

The equation array without numbers,

$$x = y + z;$$

$$a = b + c.$$

which are same as Eq.(**??**) and Eq.(**??**).

6 Footnote

The footnote command is $\backslash footnote$, for example, footnote [1].

7 Tables

The $\backslash table$ and $\backslash tabular$ environments can be used to generate tables. We give two examples here.

Table with multiple columns.

[1] This generates a footnote.

Table 1: The caption to the table with multiple columns.

Row 1	c1	c2	c3	c4
Row 2	d1	d2	d3	d4
Row 3	e1	e2	e3	e4

Table with single column.

Table 2: The caption to the table with single columns.

Row 1	c1	c2
Row 2	d1	d2
Row 3	e1	e2

References

LastName, A. (2009). Title for the first reference. *Journal of the first reference*, *3*, 18 – 88.

Author1, A., Author2, A., & Author3, A. (2008). Title for the second reference. *Journal for the second reference*, *5*, 188 – 200.

We have illustrated the basic format to the manuscript that you consider to submit to Neural Computation. We hope this is helpful to the authors.

Acknowledgments

The people you want to acknowledge. For this document, we appreciate Jrg Lcke, author of an accepted paper who generously allowed us to use his template.

Appendix

You should put the details that are not required in the main body into this Appendix.

References

LastName, A. (2009). Title for the first reference. *Journal of the first reference*, *3*, 18 – 88.

Author1, A., Author2, A., & Author3, A. (2008). Title for the second reference. *Journal for the second reference*, *5*, 188 – 200.