Automated Classification of Malignant and Benign Breast Cancer Lesions Using Neural Networks on Digitized Mammograms

Mohammed M Abdelsamea[1,2] (iD), Marghny H Mohamed[3] and Mohamed Bamatraf[3]

[1]Department of Mathematics, Assiut University, Assiut, Egypt. [2]School of Computer Science, Nottingham University, Nottingham, UK. [3]Faculty of Information and Computers, Assiut University, Assiut, Egypt.

ABSTRACT: We propose a novel neural network approach for the classification of abnormal mammographic images into benign or malignant based on their texture representations. The proposed framework has the capability of mapping high dimensional feature space into a lower-dimension, in a supervised way. The main contribution of the proposed classifier is to introduce a new neuron structure for map representation and adopt a supervised learning technique for feature classification. This is achieved by making the weight updating procedure dependent on the class reliability of the neuron. We showed high accuracy (95.2%) for our proposed approach in the classification of abnormal real mammographic images when compared to other related methods.

KEYWORDS: mammographic feature classification, neural networks, breast cancer

Introduction

Image feature classification presents a challenge in many computer vision applications, especially in health informatics.[1] With the enormous growth of medical image data that need to be analysed, image classification has become more demanding. A critical task to build a robust image analysis framework is to design an effective classification model that can cope with medical images efficiently, with the least memory requirements and user interaction.

A mammography is a digital image of the breast that is acquired by a low dose X-rays and used by radiologists to detect abnormal regions in the breast. Recently, much attention has been paid to the development of a robust tool for detecting abnormalities in digitized mammograms. A quantum-based approach[2] has been developed based on both quantum signal processing and cellular automata to detect microcalcifications in digitized mammograms. Fisher linear discriminant analysis[3] has been used by integrating features extracted based on neighbourhood structural similarity to distinguish mammographic masses into being benign or malignant. Likewise, several kinds of local features based on local binary pattern have been used to feed linear/nonlinear classifiers,[4] such as support vector, artificial neural network, and random forest, for the characterization of mammographic masses (as benign or malignant). Moreover, a deep learning approach based on convolutional neural network (CNN) has been previously applied to classify abnormalities, benign or malignant, in mammographic images.[5] The Self-Organizing Map (SOM)[6-10] has been extensively used as classifier by projecting the patterns of a higher dimensional input space n into m code-book blocks of size n organized in a two-dimensional grid. SOM provides two fundamental issues: the first is the clustering of patterns and the second is the relationship between their associated clusters. For instance, clustering is a fully unsupervised learning procedure while the relationship between clusters can be visualized in the planar surface by checking the distances between the code-book blocks. Although it is difficult to deduce an exact relationship between those clusters (where the code-book block size is greater than the planar surface size), this gives us an insight about how to bridge the gap between clustering and classification.

In this paper, we describe a simple but effective artificial neural network method, inspired by the network that has been proposed and used in Wang et al.[8] The proposed neural network has been designed to integrate and take advantage of the probability of a particular neuron to be a winner by a voting criteria to classify mammographic images.

Mammographic Image Feature Representation

In this work, image feature representation consists of two main steps: image pre-processing and feature extraction. In the first step, to cope with illumination variations and improve the quality of the images, we equalized the histograms of images and applied a cropping operation to reduce the undesirable background. After this pre-processing step, a feature extraction step was applied to provide a set of robust and reliable features to train the proposed classifier (see next section for details). Among the 14 statistical/texture image features that have been previously proposed in Haralick et al,[11] we first extracted the most practically used features (eg, uniformity, entropy, dissimilarity, and contrast). Those
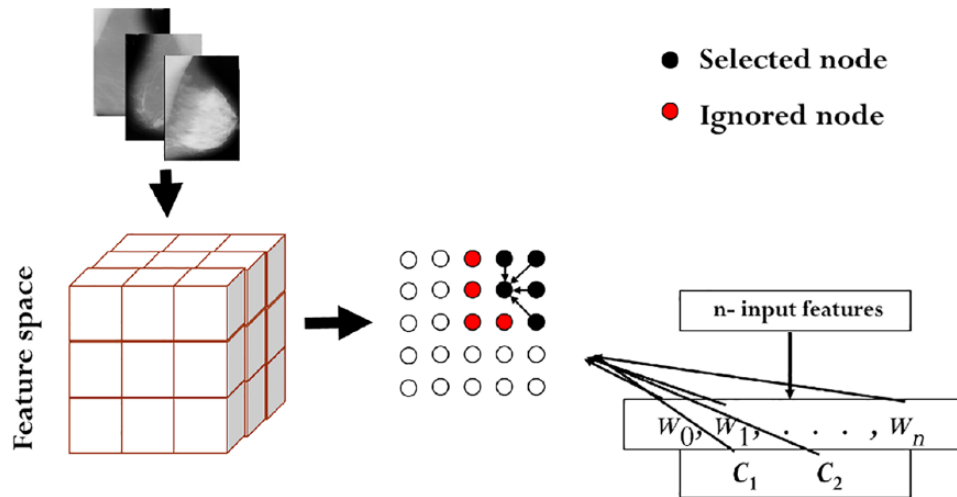
Figure 1. Architecture of proposed framework: the proposed neuron structure and weight updating process.

features are used to build the feature space of the images to teach and test our learning framework. Moreover, to improve the robustness and localization property of the represented features, a block wise partitioning method[12] has been applied. In the following section, we discuss in detail the development of our approach to classify images based on their feature representation.

Proposed Classifier

Here, we discuss the novelty and contribution of the architecture structure of our proposed neural network for supervised learning (classification). Our proposed neural network model was inspired by the network in Wang et al.[8] We introduce a new neuron/node structure and a learning scheme that exploit the class label in the weight update step.

First, every neuron is represented with the set of connection weights $w = \{w_0, w_1, \ldots, w_n\}$, where n is the number of attributes, and a set of winning class counters (WCC^m) $c^m = \{c_1, c_2, \ldots, c_m\}$, where m is the number of classes. This representation provides a robust way to utilize the class label during the training process of the network. This is by introducing WCC to the neuron structure to provide a voting criterion. For instance, neurons with maximum WCC_i are only pulled during the weight-updating process. In this way, the network can automatically shift neurons towards the best matching unit (BMU), which belong to the same class, while leaving out neurons of other class. In each iteration, after computing the distance between the input vector and neurons, the winning neuron is activated. Even though several distance functions could be used, the main idea is to calculate the similarity between patterns in an independent fashion. The next step after identifying the BMU is to increase the WCC_i by one for the ith class, accordingly. This increment gives more confidence that this neuron is targeted by an example of class i; this confidence indicates how similar the input pattern is to the BMU. At the final stage, a 'constrained weight update' is performed.

The term constrained indicates that selecting the neurons that are fitted to the criteria that is clear here, the neurons that are mostly targeted by examples from the same class are supposed to belong to the same cluster, and those neurons not mostly targeted by examples from other classes should come closer to this cluster, so they are ignored. For simplification, when a winning neuron is activated, and before performing the weight update, a vote is conducted: only neurons in the neighbourhood with a maximum class counter will be considered as neighbour neurons. More precisely, E_j belongs to class C_i, $X^T = Max||WCC||$, where Max is a function that returns the set of neurons with maximum WCC_i, the weight update will be performed over X (in a similar way to the self-learning procedure of Kohonen[7]), and, at the same time, the WCC_i will be increased by one.

Figure 1 illustrates the learning process of framework based on the proposed neuron structure. Every neuron is represented with a set of connection weights, and a set of winning class counters, such that with maximum WCC (neurons in black) are pulled during the training. Shifting those neurons (black neurons) towards the BMU of the same class, strengths linkage and connection between such neurons. At the same time, ignoring neurons (neurons in red) from other classes dims the linkage between such neurons.

Experimental Results

To quantitatively demonstrate the accuracy of the proposed framework, a 10-fold cross-validation scheme was used. We evaluated the performance of our method on the classification of cases from the Mammographic Image Analysis Society (MIAS) database (http://peipa.essex.ac.uk/info/mias.html). MIAS database consists of 330 images, which were categorized into 7 classes based on an abnormality criteria. Only 118 images were further categorized based on the severity of abnormality. These images were distributed between classes as follows: benign class (n = 64) and malignant class (n = 54). In this experiment, we demonstrated the performance of our method

in classifying those abnormal cases into benign or malignant when compared to other methods.

The CNN-based model of Li et al[5] has achieved an accuracy of 89.05% in classifying abnormalities, benign or malignant, in MIAS. The best accuracy (92.10%) was obtained by the edge weighted local texture features[4] when used by different classifiers, such as artificial neural network, fisher linear discriminant analysis, support vector machine, and random forest. Likewise, the fisher linear-based model from Rabidas et al[3] has achieved an accuracy of 94.57% when neighbourhood structural similarity was used for the characterization of mammographic masses as benign or malignant. If we compare our model with the mentioned above models, we see that our model outperformed others with an accuracy of 95.2% (when 10-fold cross validation is used) in classifying abnormal cases into benign or malignant. Finally, we showed superior accuracy of our proposed classifier when compared to the dimension reduction and classification algorithim of Wang et al[8] (83.22%) using the same set of features for both methods.

Conclusions

This work presents a novel neural network-based classification model that is able to project high-dimensional input space into low-dimensional space of two, in a supervised fashion. This was achieved by introducing a new neuron structure and adopting the underlying learning procedure based on the class reliability of the neuron. Experimental results confirm the excellent performance of the proposed framework when applied to real mammographic images.

Author Contributions

MA shaped the research, developed the algorithm, and carried out the experiment. MM and MB provided critical feedback and helped shape the research. All authors wrote the manuscript.

ORCID iD

Mohammed M Abdelsamea ⓘD https://orcid.org/0000-0002-2728-1127

REFERENCES

1. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med*. 2019;25:30.
2. Rubio Y, Montiel O, Sepúlveda R. Quantum inspired algorithm for microcalcification detection in mammograms. *Informat Sci*. 2019;480:305–323.
3. Rabidas R, Midya A, Chakraborty J. Neighborhood structural similarity mapping for the classification of masses in mammograms. *IEEE J Biomed Health Inform*. 2018;22:826–834.
4. Midya A, Rabidas R, Sadhu A, Chakraborty J. Edge weighted local texture features for the categorization of mammographic masses. *J Med Biol Eng*. 2018;38:457–468.
5. Li B, Ge Y, Zhao Y, et al. Benign and malignant mammographic image classification based on convolutional neural networks. In: *Proceedings of the 2018 10th International Conference on Machine Learning and Computing*. New York, NY: ACM; 2018:247–251.
6. Kohonen T. The self-organizing map. *Proc IEEE*. 1990;78:1464–1480.
7. Kohonen T. Essentials of the self-organizing map. *Neural Netw*. 2013;37:52–65.
8. Wang M, Li L, Yu C, et al. Classification of mixtures of Chinese herbal medicines based on a self-organizing map (som). *Mol Inform*. 2016;35:109–115.
9. Liu N, Wang J, Gong Y. Deep self-organizing map for visual classification. In: *2015 International Joint Conference on Neural Networks (IJCNN)*. New York, NY: IEEE; 2015:1–6.
10. Dokur Z, Ölmez T. Heart sound classification using wavelet transform and incremental self-organizing map. *Digital Signal Proc*. 2008;18:951–959.
11. Haralick R, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybernet*. 1973;6:610–621.
12. Mohamed M, Abdelsamea MM. An efficient clustering based texture feature extraction for medical image. In: *11th International Conference on Computer and Information Technology*. New York, NY: IEEE; 2008:88–93.