International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 6 Issue 9 ǁ Sep. 2018 ǁ PP.35-46

Malignant classification of mammogram images based on deep learning

Aseem Behera, Sunita Behera, Flora Das, Bhawesh Kumar
Department of Electronics and Communication Engineering, NM Institute of Engineering and Technology,Bhubaneswar , Odisha
Department of Electronics and Communication Engineering, Raajdhani Engineering College,Bhubaneswar,Odisha
Department of Electronics and Communication Engineering,Aryan Institute of Engineering and Technology Bhubnaeswar , Odisha
Department of Electronics and Communication Engineering,Capital Engineering College,Bhubaneswar,Odisha

ABSTRACT
Breast cancer is one of the most common malignant tumors in women, which seriously affect women's physical and mental health and even threat to life. At present, mammography is an important criterion for doctors to diagnose breast cancer. However, due to the complex structure of mammogram images, it is relatively difficult for doctors to identify breast cancer features. At present, deep learning is the most mainstream image classification algorithm. Therefore, this study proposes an improved DenseNet neu-ral network model, also known as the DenseNet-II neural network model, for the effective and accurate classification of benign and malignant mammography images. Firstly, the mammogram images are pre-processed. Image normalization avoids interference from light, while the adoption of data enhancement prevents over-fitting cause by small data set. Secondly, the DenseNet neural network model is improved, and a new DenseNet-II neural network model is invented, which is to replace the first convolutional layer of the DenseNet neural network model with the Inception structure. Finally, the pre-processed mammo-gram datasets are input into AlexNet, VGGNet, GoogLeNet, DenseNet network model and DenseNet-II neural network model, and the experimental results are analyzed and compared. According to the 10-fold cross validation method, the results show that the DenseNet-II neural network model has better clas-sification performance than other network models. The average accuracy of the model reaches 94.55%, which improves the accuracy of the benign and malignant classification of mammogram images. At the same time, it also proves that the model has good generalization and robustness.

Keywords:
Breast cancer
Mammogram images
Deep learning
Inception structure
DenseNet-II neural network model

I. INTRODUCTION

At present, breast cancer has become one of the most common malignant tumors in women [1], and its incidence is rising in both developed and developing countries. Clinically, malignant tumors are usually defined as positive, and benign tumors are defined as negative. Currently, the detection methods applied in the diagnosis of breast cancer include mammography, computed tomography technology, photoacoustic imaging, nuclear magnetic resonance imaging, microwave imaging and other technologies [2,3]. Among them, mammography is a very effective technique for detecting breast cancer [4]. There are two main manifestations of breast can-cer on mammogram images, including masses and calcifications. Typical benign tumors appear on the mass as rounded, smooth and generally clear. The calcification is characterized by a coarser shape, a granular shape, a popcorn shape or a ring shape, and its density is higher and the distribution is more dispersed. Typical malignant tumors have a needle-like shape on the mass, and the edges are irregular and generally fuzzy. The morphology of calcification is mostly small sand-like, linear or branched, with different shapes and sizes, and the distribution is often dense or clustered in a lin-ear shape [5]. Due to the complexity of the mammogram images of early breast cancer, coupled with the low contrast of the mam-mogram images itself, it is difficult for doctors to make a correct diagnosis based on mammogram images. Therefore, it is neces-sary to improve the diagnostic efficiency of doctors by using the computer aided diagnosis system (CAD) of deep learning.

In recent years, with the development of digital image pro-cessing technology and artificial intelligence technology, the use of computer-aided detection and diagnosis system to assist breast radiologists to judge the

classification of benign and malignant breast tumors has become a realistic and significant scientific prob-lem. Many scholars have also used CAD to identify and classify microcalcifications and masses in mammogram images.

On the one hand, the most common method is to use SVM to classify the region of interest (ROI) of mammogram images. Liu et al. used the SVM classifier to classify the masses. The exper-imental results showed that the area under the ROC curve was Az = 0.7 [6]. Anitha et al. used the same idea to conduct a simulation experiment on 44 mammogram images containing masses in the

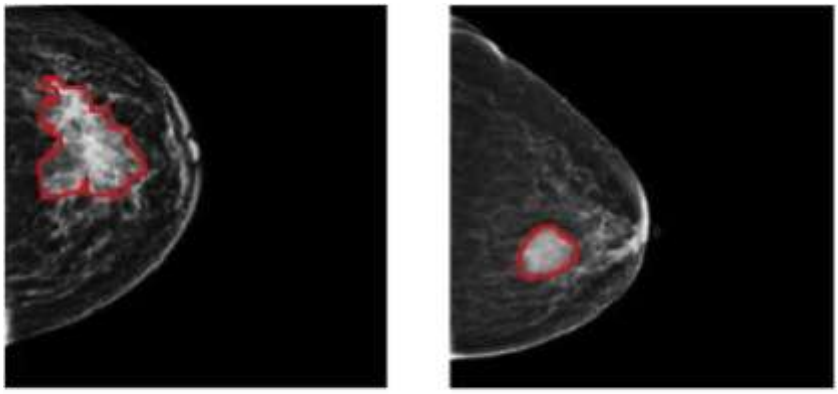

Fig. 1. Benign and malignant breast mass images.

MIAS database, and the author pointed out that the classification accuracy of the masses reached 95% [7]. Combining regression fea-tures to eliminate SVM and normalized mutual information feature selection method, Liu et al. simulated the mammogram images of DDSM library, and the results showed that the classification accu-racy of the method reached 93% [8]. However, these methods had inherent disadvantages, that is, the traditional SVM assumed that all features had the same degree of influence on the classifica-tion target. Since most of the above classification algorithms were carried out on small data sets, they were not suitable for identify-ing and classifying large data sets in hospitals. In addition, there were no uniform comparison between algorithms, and the accu-racy was not comparable. Moreover, these classification algorithms used artificial-based feature extraction methods, which not only required professional domain knowledge, but also required a lot of time and effort to complete. The key was that there were often certain difficulties in extracting features. It seriously restricted the application of traditional machine learning algorithms in mammo-gram images classification.

On the other hand, neural networks based on deep learning classify mammograms. Arbach et al. evaluated the performance of classification of mammograms based on backpropagation neural network (BNN) and compared them with radiologists and resi-dents. The results showed that the area under the ROC curve of BNN was Az = 0.923.The area under the ROC curve of the radiologist was Az = 0.864, and the area under the ROC curve of the resident was Az = 0.648 [9]. The neural network system designed by Shih-Chung et al. detected 91 cases of mammogram images. The test resulted that the area under the ROC curve was Az = 0.78 [10]. Moreover, the hidden layer of the neural network they designed has 125 nodes, which maked the generalization of the neural network not good enough. The common features of these methods were higher sensi-tivity, lower specificity, lower recognition rate, and greater research space.

Therefore, this paper studies and improves a deeper and more complex DenseNet neural network model [11], inventing a new DenseNet-II neural network model, which can maintain the sparse-ness of the network structure. It also prevents the model from overfitting and improves computer performance. At the same time, the data enhancement method is used to prevent the over-fitting caused by the small sample data set, bring about an increase in rate of identification of the mammogram images. Using the DenseNet-II neural network model can not only improve the diag-nostic efficiency, but also provide doctors with more objective and accurate diagnosis results. so it has important clinical application value.

II. DATA COLLECTION AND PREPROCESSING

2.1. Data collection

The datasets used in this paper are taken from mammography images provided by the First Hospital of Shanxi Medical University. Mammary gland lesions are obtained by full-field digital mammog-raphy (FFDM)

examination. A total of 2042 cases of breast disease are usually defined as positive for malignancy and negative for benign [12]. After being confirmed by the experts of the First Hos-pital of Shanxi Medical University, there are 1011 negative cases and 1031 positive cases, all of which are female patients.All of the mammogram images in the database have been marked by experts, as shown in Fig. 1(a) and (b) are images of malignant breast tumors and benign breast tumors, respectively.

2.2. Data preprocessing

Mammography is currently the only widely accepted means of routine breast cancer screening [13] to assist physicians in clinical diagnosis. However, it is a low-radiation, high-resolution image. It is difficult to judge the mammogram images directly with the naked eye. Therefore, it is necessary to preprocess the image datasets. Through image preprocessing, the quality and quantity of images can be improved, making the recognition rate more accurate. Two processing techniques are used for the mammogram images: zero-mean normalization [14] and data enhancement, which can improve the speed and accuracy of mammogram images classifica-tion.

2.2.1. Zero-mean normalization

The main purpose of image normalization [15] is to reduce the interference of medical images due to uneven light. Image nor-malization causes the image to find some invariants, enhances robustness, and speeds up the convergence of the training model. Zero-mean normalization is the normalization of data for the mean and standard deviation of the raw data. The processed data con-forms to the standard normal distribution, that is, the mean is 0, the standard deviation is 1, and the conversion function is as shown in

(1):

$$x^* = \frac{x - }{} \quad (1)$$

Where is the mean of all sample data, is the standard deviation of all sample data.

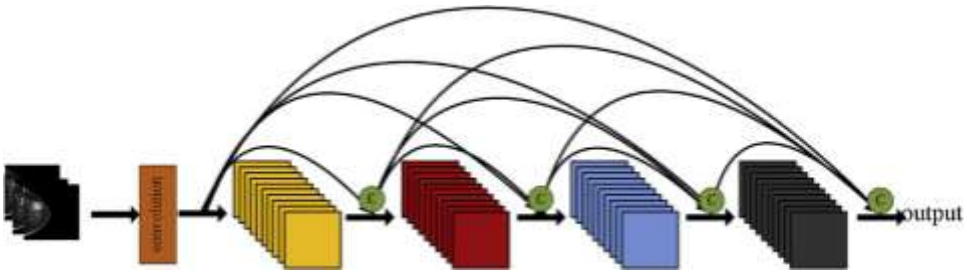

Fig. 2. DenseNet network dense connection mechanism (where c represents the channel level connection operation).

2.2.2. Dataset enhancements

The manually labeled training data set is limited. Due to the limitation of the training samples, CNN is prone to over-fitting during the training process, resulting in low medical image recog-nition rate and unsatisfactory diagnosis results [16]. Therefore, this requires a large increase in the training data set. The solution is to enhance the mammogram images dataset by affine transformation and random cropping.

(1) The mammogram images data set is enhanced by the affine transformation method [17]. It mainly rotates the image by 90 ° / 180 ° / 270 degrees, scales by 0.8, mirrors horizontally and vertically, and combines these operations.
(2) Mammogram images are processed by random cropping to obtain more data sets. The data set can be expanded by a factor of 15 by the above two methods.

III. METHODS

With the continuous improvement of computer computing power and the arrival of the era of big data, deep learning over-comes the limitations of shallow learning, without the need to manually design and extract features. It can extract more expres-sive and abundant essential characteristics from the data set. In recent years,

deep learning has been successfully applied in the fields of computer vision, natural language processing, medical image processing, etc. [18]. Therefore, the use of deep learning technology for the identification of benign and malignant breast tumors has become the focus of clinical diagnosis research, and the accurate identification of benign and malignant breast tumors has important research significance for the diagnosis and treatment of breast cancer [19].

Deep learning can be regarded as a multi-layer artificial neu-ral network [20]. By constructing a neural network model with multiple hidden layers, the low-level features are transformed by layer-by-layer nonlinear features to form a more abstract high-level feature expression to discover the distributed feature representation of the data [21]. As one of the most commonly used deep learning models, convolutional neural networks directly use 2D or 3D images as input to the network, avoiding the com-plex feature extraction process in traditional machine learning algorithms. Compared with a fully connected neural network, con-volutional neural networks use local connections, weight sharing, and downsampling, reducing the number of network parameters and reducing computational complexity. At the same time, the translation, zoom, rotation and other changes of the image are highly invariant.

3.1. DenseNet neural network model

This paper chooses the densely connected DenseNet neu-ral network model. Different from other convolutional neural networks, it has the following two characteristics:(1) Dense con-nection: Each layer is connected to each of the previous layers to achieve feature reuse.(2) Due to feature reuse, each layer uses a small number of convolution kernel extraction features to achieve the purpose of reducing redundancy. In DenseNet, each layer is concated with all previous layers in the channel dimension, ^{and} for an L-layer network, DenseNet contains a total of $^{L(L+1)\,2}$ connections. Fig. 2 shows the dense connection mechanism of DenseNet.

The advantages of DenseNet are mainly reflected in the follow-ing aspects:

(1) Effectively solve the problem of gradient disappearance;
(2) Strengthening feature propagation;
(3) Support feature reuse;
(4) Significantly reduce the number of parameters.

This model can omit the pre-training on the ImageNet [22] dataset and start training directly from the randomly initial-ized model, achieving the goal of saving time and efficiency. In many large-scale practical applications, there are great differences between the actual dataset and the ImageNet dataset. Therefore, it is a good prospect to apply this network model that does not require pre-training to medical image classification [23].

3.1.1. Design of DenseNet neural network structure

DenseNet's network structure consists mainly of the DenseBlock layer and the Transition layer. In order to prevent the network structure from becoming wider, the total number of DenseNet networks is 40 layers, and the growth rate k = 12, including 3 DenseBlock layers and 2 Transition layers. This model does not have too many parameters, while saving computational memory and reducing overfitting. As shown in Fig. 3.

3.1.1.1. DenseBlock layer. In all DenseBlocks, each layer is out-putted with a k-characteristic map after convolution (and the feature maps of each layer are the same size), that is, k convo-lution kernels are used for feature extraction. k is called growth rate in DenseNet, which is a hyperparameter. Since there are many inputs for each layer in the back, DenseBlock can use the bot-tleneck layer internally to reduce the amount of computation, that is, th BN + ReLU+ 1 × 1 Conv + BN + ReLU+ 3 × 3 Conv struc-ture used in the nonlinear combination function in DenseBlock. As shown in Fig. 4. The role of 1 × 1 Conv is to reduce the number of features, which can improve the calculation speed and effi-ciency.

3.1.1.2. Transition layer. It mainly connects two adjacent Dense-Block layers and reduces the feature map size. The Transition layer includes a 1 × 1 convolution and 2 × 2 AvgPooling [24], and the structure is BN + ReLU+ 1 × 1 Conv+ 2 × 2 AvgPooling. Therefore,

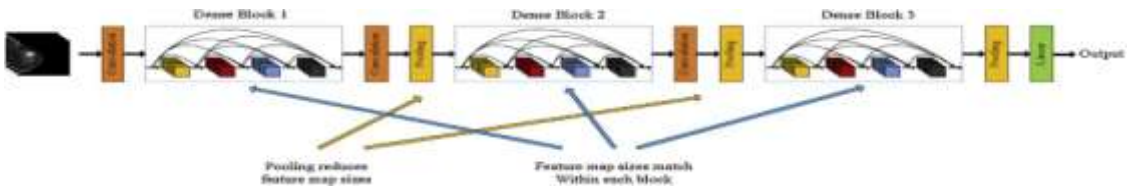

Fig. 3. DenseNet neural network structure.

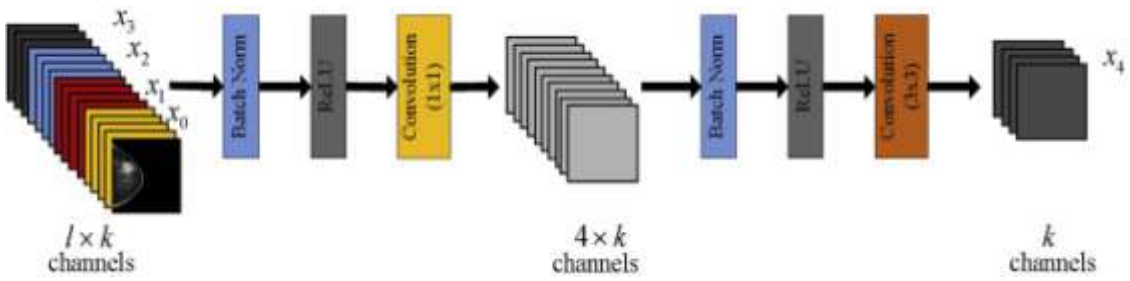

Fig. 4. DenseBlock using the bottleneck layer.

the Transition layer can function as compression model. Each layer will connect all the previous layers as input:

$$x^l = H_l([x_0, x_1, \ldots, x^{l-1}]) \quad (2)$$

ation of multiple features. The calculation formula is as shown in (3)

$$x_j^{(l)} = f\left(x_i^{(l-1)} * k_{ij}^{(l)} + b_j^{(l)}\right) \quad (3)$$

H_1 (•) is a non-linear transformation function, a combination that includes a series of BN(Batch Normalization), ReLU, Pooling, and Conv operations. The final DenseBlock is followed by a global Avg-Pooling layer, which is then sent to a softmax classifier for image recognition classification.

3.2. DenseNet-II neural network structure

Due to the deeper and wider network model, the defects of huge parameters are generated, which leads to over-fitting, increasing the amount of calculation, and consuming more com-puting resources. To avoid this problem, the DenseNet model has been improved and is called the DenseNet-II network model. The Inception structure [25] replaces the first3 × 3 convolution in the DenseNet network before entering the first DenseBlock layer. The improved model is called the DenseNet-II neural network model. Like the DenseNet neural network model, the DenseNet-II neu-ral network model uses 40 layers with a growth rate of k = 12, including 3 DenseBlock layers and 2 Transition layers. As shown in Fig. 5.

The main differences between DenseNet and DenseNet-II neural network structures are:

For the DenseNet neural network, before entering the first DenseBlock, first perform a 3 × 3 convolution (stride = 1). There are a total of 16 convolution kernels. The convolution layer is mainly responsible for extracting features in the image, avoiding the mis-take of manually extracting features. Consisting of a set of feature maps, the same feature map shares a convolution kernel, which is actually a set of weights, also called filters. A learnable convolu-tion kernel is convolved with several feature maps of the previous layer, and the corresponding elements are accumulated and then offset, and then passed to a nonlinear activation function ReLU func-tion to obtain a feature map, that is, the extraction of a feature is achieved. A plurality of different convolution kernels enable extrac-

$i \in M^{l-1}$

Where l represents the number of layers, and k_{ij} represents the convolution kernel of the feature map j connecting the lth layer and the feature map i of the l-1th layer, M^{l-1} represents the input feature maps selected by the l-1layer, and * represents the convolu-tion operation, b denotes the offset, and f (•)denotes the nonlinear activation function.

For the DenseNet-II neural network, the Inception structure is mainly used to replace the first 3×3 convolutional layer of the DenseNet neural network model. Because a deeper and wider net-work structure can obtain better predictive recognition, it will generate huge parameters, which will lead to over-fitting, which greatly increases the amount of calculation and consumes more computing resources. Inception mainly improves the traditional convolutional layer in the network. For deep neural network per-formance and computation, it mainly solves the following problems in deep networks:

(1) too many parameters, easy to over-fit, if the training data set is limited;
(2) The larger the network, the more computational complexity it is, and it is difficult to apply;

(3) The deeper the network, the easier it is for the gradient to disap-pear backwards(gradient dispersion), it is difficult to optimize the model.

The inception structure: not only maintains the sparseness of the network structure, but also utilizes the high computational performance of dense matrices. In literature [24], the Inception module structure is mainly proposed (1 × 1,3 × 3,5 × 5 conv and 3 × 3 pooling combined), as shown in Fig. 6. The basic structure of the Inception module: There are 4 branches, each with a 1 × 1 con-volution. The 1 × 1 convolution is a very useful structure, which can organize information across channels, improve the expressiveness

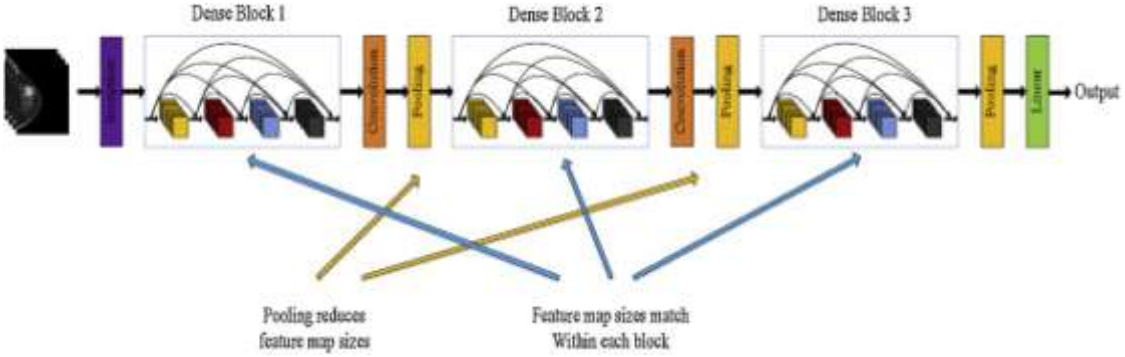

Fig. 5. The structure of DenseNet-II neural network model.

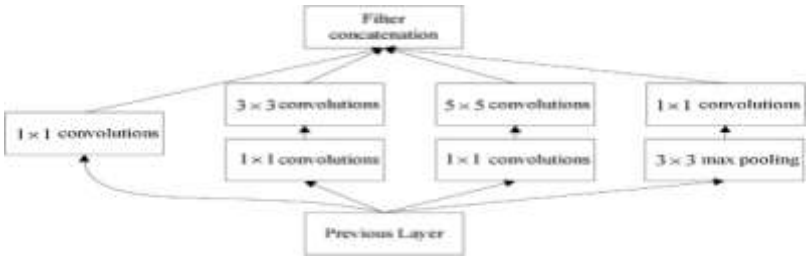

Fig. 6. The inception model.

of the network, and also raising dimension and dimension of the output channel. The module has 3 different sizes of convolution and 1 maximum pooling, which increases the adaptability of the net-work to different scales. It can expand the depth and width of the network with high efficiency, and improve the accuracy without causing over-fitting.

Therefore, the overall structure of the DenseNet-II neural net-work mainly includes Inception, Dense Blocks and transition layers, with fewer parameters and higher precision.

In the DenseNet and the DenseNet-II neural network model, all 3 × 3 3 × 3 convolutions use padding = 1 to ensure that the fea-ture map size remains unchanged. All convolutions use the ReLU function.

IV. ANALYSIS OF THE EXPERIMENT AND RESULTS

4.1. Experimental network parameter settings

This model is based on the Ubantu 18.04 operating system, using two NVIDIA Titan X graphics cards, and experimenting on the pytorch framework. The data enhancement algorithm is imple-mented in python.

In order to test the classification performance of the DenseNet-II model, the AlexNet, VGGNet, GoogLeNet and DenseNet network models are compared with the DenseNet-II model for experimental analysis. The network structure of the selected AlexNet, VGGNet and GoogLeNet models are as follows:

(1) The AlexNet model [26] has eight layers and has a parameter amount of 60 M or more. The first five layers are convolutional layers, the last three layers are fully connected layers, and the output of the last fully connected layer has 1000 outputs. The network structure is shown in Fig. 7.

(2) The VGGNet [27] model established a 19-layer deep network, and achieved the first and second classification results in ILSVRC. The VGGNet network model has many similarities with the AlexNet framework. There are sixteen convolutional lay-ers, two layers of fully connected image feature layers, and one layer of fully connected classification feature layers. The network structure is shown in Fig. 8.

(3) The GoogLeNet model [28] is a 22-layer deep network. Its biggest feature is the introduction of the Inception struc-ture, and solves the problem of huge network parameters and resource consumption. The GoogLeNet model uses the Incep-tion structure, which not only further improves the accuracy of the classification, but also greatly reduces the amount of param-eters. The GoogLeNet network structure is shown in Table 1.

In addition, due to the experimental evaluation of benign and malignant breast tumors, in the AlexNet network and GoogLeNet model, the full-connected layer output of the convolutional neural network was changed from 1000 to 2, which became a two-class problem. The initial learning rates of the five models are all set to 0.01, and the maximum number of iterations is 8000.

4.2. Training strategy

Cross-validation is a method of model selection by estimating the generalization error of the model [29]. It does not have any assumptions, it is an effective model selection method, which has the universality of application and easy operation. In order to accu-rately measure the quality of a model evaluation method, this paper uses K-fold cross-validation [30], K is 10.

In this paper, the data collected from the First Hospital of Shanxi Medical University was expanded 15 times after image preprocess-ing, and 30,630 pairs of mammogram images are obtained. Among them, 15,165 pairs of benign tumor images and 15,645 pairs of malignant tumor images. Using a 10-fold cross-validation method, all data sets were randomly divided into 10 disjoint and identi-

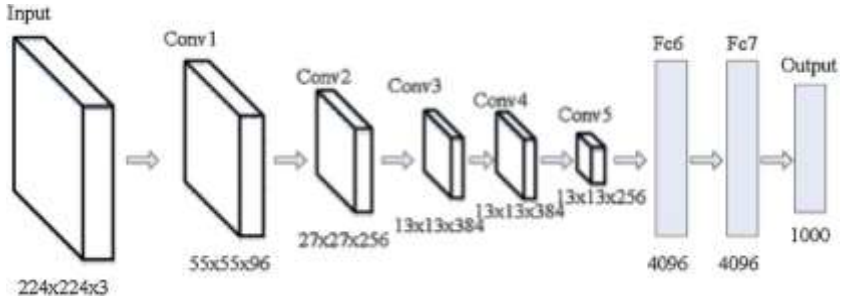

Fig. 7. AlexNet network structure.

Table 1

GoogLeNet network structure.

type	size/stride	output	depth	1×1	3×3	3×3	3×3	3×3	pooling
Conv	7×7/2	112×112×64	1						
max pool	3×3/2	56×56×64	0						
Conv	3×3/1	56×56×192	2		64	192			
max pool	3×3/2	28×28×192	0						
Inception(3a)		28×28×256	2	64	96	128	16	32	32
Inception(3b)		28×28×480	2	128	128	192	32	96	64
max pool	3×3/2	14×14×480	0						
Inception(4a)		14×14×512	2	192	96	208	16	96	64
Inception(4b)		14×14×512	2	160	112	224	24	64	64
Inception(4c)		14×14×512	2	128	128	256	24	64	64
Inception(4d)		14×14×528	2	112	144	288	32	64	64
Inception(4e)		14×14×832	2	256	160	320	32	128	128
max pool	3×3/2	7×7×832	0						
Inception(5a)		7×7×832	2	256	160	320	32	128	128
Inception(5b)		7×7×1024	2	384	192	384	48	128	128
Avg pool	7×7/1	1×1×1024	0						
Dropout40%		1×1×1024	0						
fc		1×1×1000	1						
softmax		1×1×1000	0						

cally sized subsets. Each time a subset is taken as the test data set, and the remaining 9 samples are taken as the training data set, of which 25% of the training set is used as the verification set. Until all 10 subsets

have been tested once, that is, the test set is cycled, the cross-validation process ends. The average of the accuracy of the 10 test results was calculated as the overall evaluation index of the model, and the model with the highest accuracy was obtained. Each data set contains approximately 50% of benign cases and malig-nant cases. Among them, the training set is used for model training and parameter learning; the verification set is used to optimize the model, and the parameters are automatically fine-tuned according to the test results of the model in the training process; the test set is used to test the model's recognition and generalization ability.

Through the K-fold cross-validation principle, the assigned 30,630 mammogram data sets were placed in the AlexNet, VGGNet, GoogLeNet, DenseNet and DenseNet-II models for training, ver-ification and testing according to the unified principle. Different network models use the time.time () function to calculate the time. The difference between the end time and the start time is the time when the network model is classified.

4.3. Evaluation criteria

For the classification of medical images, this paper evaluates the classification performance of the model from sensitivity (Sen), specificity (Spec), and accuracy (Acc) [31]. Specificity, sensitivity and accuracy are defined as follows.

$$\text{Sensitivity(Sen)} = \frac{TP}{TP+FN} \times 100\% \quad (4)$$

$$\text{Specificity(Spec)} = \frac{TN}{TN+FP} \times 100\% \quad (5)$$

$$\text{Accuracy(Acc)} = \frac{TP+TP}{TP+TN+FP+FN} \times 100\% \quad (6)$$

In the formula, TP (True positive) is the number of pictures which diagnostic model can accurately identify as malignant tumors. FP (False positive) refers to the number of pictures in which a diag-nostic model mistakenly identifies a benign tumor as a malignant tumor. TN (True negative) is the number of pictures that the diag-nostic model can accurately identify as a benign tumor. FN (False positive) refers to the number of pictures of a type of benign tumor detected by a diagnostic model malignant tumor.

The receiver operating characteristic (ROC) can be further intro-duced by TP and FP. The ROC curve is a graphical curve on a two-dimensional plane. By using the performance of the classifi-cation model, adjusting the threshold of the classifier, a curve of (0,0), (1,1) can be obtained, which is the ROC curve of the classi-fication model. The AUC value refers to the area enclosed by the ROC curve in the [0,1] interval and the X axis. The overall perfor-mance of the computer-aided diagnostic system is proportional to the AUC value, which means the greater the AUC value, the better the performance [32].

4.4. Experimental results and analysis

To verify the performance of the benign and malignant classifi-cation methods of mammography in this paper, the training results of the five different network models are all K-fold cross-validated, and the K value is 10. The main evaluation classifier performance indicators are sensitivity (Sen), specificity (Spec), accuracy (Acc) and ROC curve area (AUC), etc. The result is a comparative anal-ysis of the 10-fold cross-validation average results. The statistical experiment results are shown in Table 2.

Table 2

The performance table of breast cancer benign and malignant classification(the 10-fold cross-validation average results).

Neural Networks	Performance		
	Acc(%)	Sen(%)	Spec(%)
AlexNet	92.70	93.60	91.78
VGGNet	92.78	93.58	92.42
GooLeNet	93.54	93.90	93.17
DenseNet	93.87	94.59	93.90
DenseNet-II	94.55	95.60	95.36

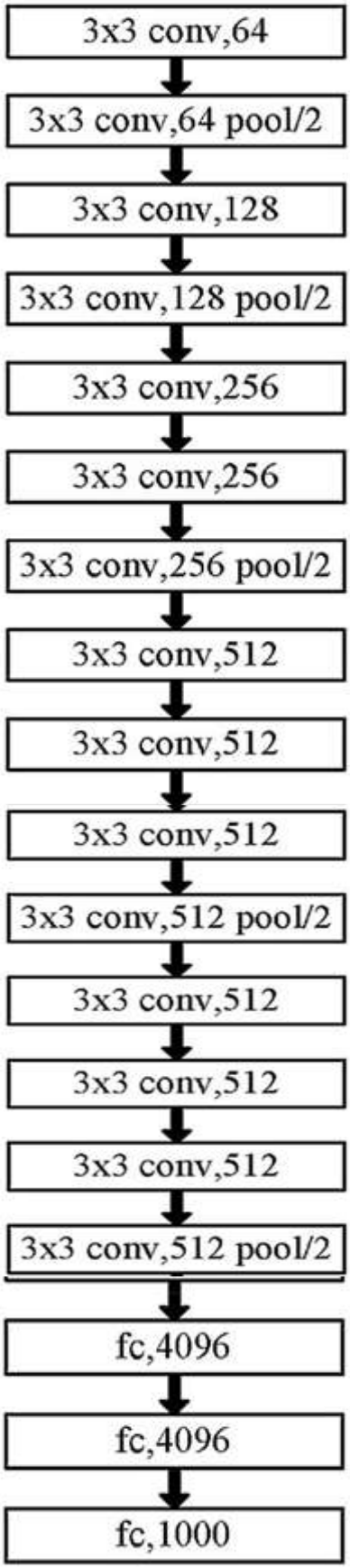

Fig. 8. VGG network structure.

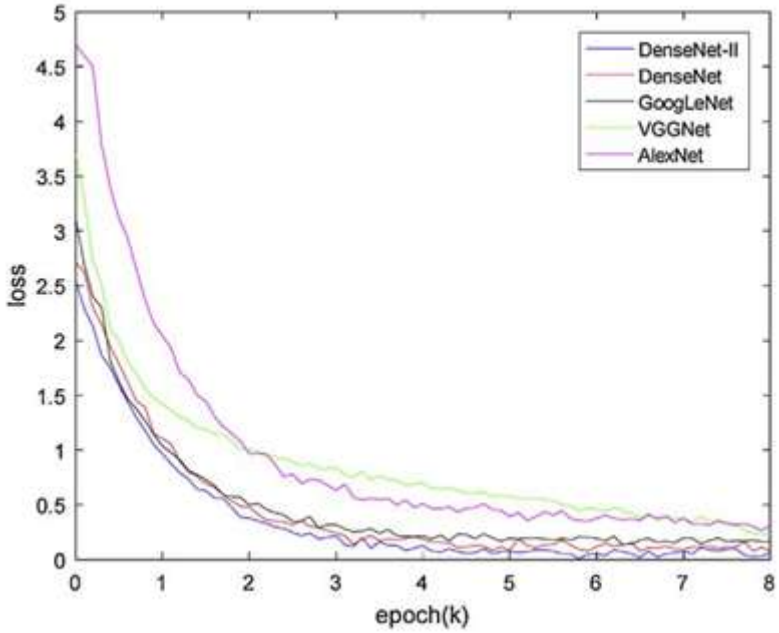

Fig. 9. The training error rate of the five network models (the 10-fold cross-validation average results).

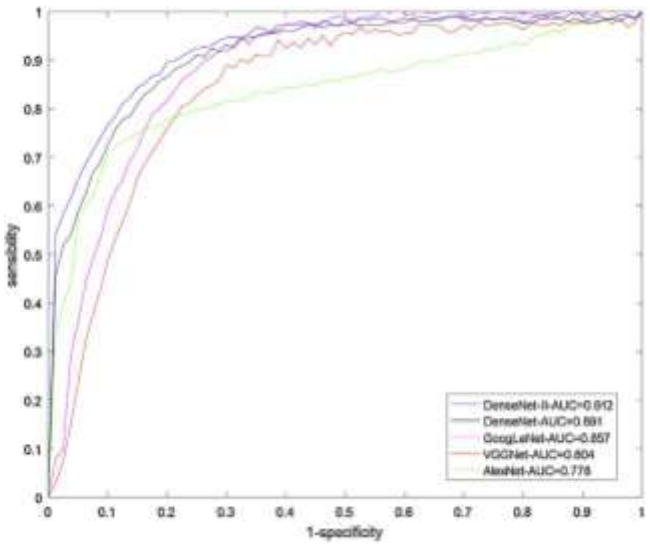

Fig. 10. The ROC curve and AUC value of five network models (the 10-fold cross-validation average results).

It can be seen that when the improved DenseNet-II network model is adopted, the three classification indicators of Spec, Sen and Acc are the highest respectively. Compared with the other four network models, the improved DenseNet-II network model with Inception structure has been added, the average accuracy of classification reaches 94.55%. It shows that the improved model can extract more distinguishing features, so the recognition rate is higher, and it has better robustness and generalization.

The training error rate of the five network models is shown in Fig. 9. The abscissa indicates the number of iterations, and the ordi-nate indicates the training error rate. As the number of training increases, the overall trend is downward. The error rate of the DenseNet-II model has been the lowest, followed by DenseNet, GoogLeNet, and VGGNet, and finally AlexNet.

Fig.10 shows the ROC curve of the five network structure clas-sification performance and the area under the curve AUC. The coordinates are the false positive rate (1-Specificity) and the ordi- nate is the sensitivity (Sensitivity). The DenseNet-II model was initially at a distinct advantage, followed by DenseNet model. Among them, the VGCNet and GoogLeNet models have AUC val-ues of more than 0.8, AlexNet has the worst

classification effect, and the AUC value reaches 0.778. It can also be seen that the clas-sification performance of the DenseNet-II structure is significantly better than other structural models. It can be seen that DenseNet-II has better performance in mammography images classification. With the increase of the number of iterations, the training error rate is steadily reduced, and there is no over-fitting phenomenon.

Due to the advantage of dense connection between DenseNet and DenseNet-II, the gradient disappearance problem is effectively solved, the number of parameters is greatly reduced, and fea-ture propagation and feature reuse are enhanced. Both models have advantages over AlexNet, VGGNet and GoogLeNet in terms of sensitivity, specificity, accuracy, and classification performance. In addition, DenseNet-II has added the Inception structure, which greatly reduces the amount of parameters and solves the problem of gradient descent. The classification performance is better than that of the DenseNet model, which further improves the accuracy of the benign and malignant classification of mammography images.

V. CONCLUSION

This paper studies the use of deep learning methods to achieve the benign and malignant classification of mammograms, and pro-poses an improved neural network model called DenseNet-II neural network model. Firstly, aiming at the insufficiency of sample data set, through the data enhancement method, the over-fitting prob-lems in deep learning due to insufficient data set is effectively avoided. Secondly, Inception Net replaces the first convolutional layer of the DenseNet neural network model. The DenseNet neural network model is improved, and the calculation speed and effi-ciency of the network model are improved. Finally, the 10-fold cross-validation is used to verify the classification results of five network models. The results show that DenseNet-II neural net-work model is superior to other structural models in classification performance, thus achieving a more accurate classification of mam-mography images benign and malignant. The method proposed in this paper not only improves the benign and malignant classifi-cation performance of mammography images, but also provides doctors with more objective and accurate diagnosis results, which has important clinical application value and research significance.

ACKNOWLEDGMENTS

The paper supported by The General Object of National Natural Science Foundation (61772358) Research on the key technology of BDS precision positioning in complex landform; Project supported by International Cooperation Project of Shanxi Province (Grant No. 201603D421014) Three-Dimensional Reconstruction Research of Quantitative Multiple Sclerosis Demyelination; International Coop-eration Project of Shanxi Province (Grant No. 201603D421012): Research on the key technology of GNSS area strengthen informa-tion extraction based on crowd sensing.

REFERENCES

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer Statistics, 2017, CA Cancer J. Clin. 65 (1) (2015) 5.

[2] (a) X. Xiao, L. Xu, B.Y. Liu, Three-dimensional simulation for early breast cancer detection by ultra-wideband, Acta Phys. Sin. 62 (4) (2013) 221–229; (b) J. Ferlay, H.R. Shin, F. Bray, et al., Estimates of worldwide burden of cancer in 2008: globocan 2008, Int. J. Cancer 127 (12) (2010) 2893–2917.

[3] L. Guang-Dong, Three-dimensional microwave-induced thermo-acoustic imaging for breast cancer detection, Acta Phys. Sin. 60 (7) (2011) 074303–074913.

[4] J.M. Timmers, H.J. van Doorne-Nagtegaal, H.M. Zonderland, et al., The Breast Imaging Reporting and Data System (BI-RADS) in the Dutch breast cancer screening programme: its role as an assessment and stratification tool, Int. J. Med. Radiol. 22 (8) (2012) 1717–1723.

[5] N.V.S.S.R. Lakshmi, C. Manoharan, An automated system for classification of micro calcification in mammogr based on Jacobi moments, Int. J. Comput. Theory Eng. 3 (3) (2011) 431–434.

[6] X. Liu, J. Liu, D. Zhou, et al., A benign and malignant mass classification algorithm based on an improved level set segmentation and texture feature analysis, in: International Conference on Bioinformatics & Biomedical Engineering. IEEE, 2010.

[7] J. Anitha, J.D. Peter, A wavelet based morphological mass detection and classification in mammograms, in: International Conference on Machine Vision & Image Processing, 2013.

[8] X. Liu, J. Tang, Mass classification in mammograms using selected geometry and texture features, and a new SVM-Based feature selection method, IEEE Syst. J. 8 (3) (2014) 910–920.

[9] L. Arbach, D.L. Bennett, J.M. Reinhardt, et al., Classification of mammographic masses: comparison between backpropagation neural network (BNN) and human readers, Proc. SPIE 5032 (1) (2003) 1441–1444, vol. 3.

[10] S.C.B. Lo, H. Li, Y. Wang, et al., A multiple circular path convolution neural network system for detection of mammographic masses, IEEE Trans. Med. Imaging 21 (2) (2002) 150–158.

[11] G. Huang, Z. Liu, V.D.M. Laurens, et al., Densely Connected Convolutional Networks, 2016, pp. 2261–2269.
[12] F.A. Spanhol, L.S. Oliveira, C. Petitjean, et al., Breast cancer histopathological image classification using convolutional neural networks, in: International Joint Conference on Neural Networks. IEEE, 2016.
[13] D. Shen, G. Wu, H.I. Suk, Deep learning in medical image analysis, Annu. Rev. Biomed. Eng. 19 (1) (2017) 221–248.
[14] A.N. Nicolaides, S.K. Kakkos, M. Griffin, et al., Effect of image normalization on carotid plaque classification and the risk of ipsilateral hemispheric ischemic events: results from the asymptomatic carotid stenosis and risk of stroke study, Vascular 13 (4) (2005) 211–221.
[15] H. Morgan, M. Druckmüller, Multi-scale Gaussian normalization for solar image processing, Sol. Phys. 289 (8) (2014) 2945–2955.
[16] M. Veta, et al., "Breast cancer histopathology image analysis: a review,", IEEE Trans. Biomed. Eng. 61 (May (5)) (2014) 1400–1411.
[17] X. Song, S. Wang, X. Niu, An integer DCT and affine transformation Based image steganography method, in: Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing. IEEE, 2012, pp. 102–105.
[18] Y. Guo, A. Oerlemans, A. Oerlemans, et al., Deep learning for visual understanding, Neurocomputing 187 (C) (2016) 27–48.
[19] T. Jia, H. Zhang, Y.K. Bai, Benign and malignant lung nodule classification based on deep learning feature, J. Med. Imaging Health Inform. 5 (8) (2015) 1936–1940.
[20] D. Shen, G. Wu, H.I. Suk, Deep learning in medical image analysis, Annu. Rev. Biomed. Eng. 19 (1) (2017).
[21] Y. Bengio, O. Delalleau, On the expressive power of deep architectures, International Conference on Algorithmic Learning Theory (2011) 18–36.
[22] O. Russakovsky, J. Deng, H. Su, et al., ImageNet large scale visual recognition challenge, Int. J. Comput. Vis. 115 (3) (2015) 211–252.
[23] K. He, X. Zhang, S. Ren, et al., Deep residual learning for mage recognition, in: Proceedings of the IEEE Conference on Comp., 2019.
[24] M.D. Zeiler, R. Fergus, Stochastic pooling for regularization of deep convolutional neural networks, Eprint Arxiv (2013).
[25] X. Jin, J. Chi, S. Peng, et al., Deep image aesthetics classification using inception modules and fine-tuning connected layer, in: International Conference on Wireless Communications & Signal Processing. IEEE, 2016, pp. 1–6.
[26] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Commun. ACM 60 (2) (2017) 2012.
[27] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, Comput. Sci. (2014).
[28] C. Szegedy, W. Liu, Y. Jia, et al., Going deeper with convolutions, in: Computer Vision and Pattern Recognition. IEEE, 2015, pp. 1–9.
[29] J. Song, F. Li, K. Takemoto, et al., PREvaIL, an integrative approach for inferring catalytic residues using sequence, structural, and network features in a machine-learning framework, J. Theor. Biol. 443 (2018) 125–137.
[30] A. Suinesiaputra, M.M. Sanghvi, N. Aung, et al., Fully-automated left ventricular mass and volume MRI analysis in the UK Biobank population cohort: evaluation of initial results, Int. J. Cardiovasc. Imaging (2017).
[31] S.Z. Li, K.L. Chan, C. Wang, Performance evaluation of the nearest feature line method in image classification and retrieval, Pattern Anal. Mach. Intell. IEEE Transactions on 22 (11) (2000) 1335–1339.
[32] H. Narasimhan, S. Agarwal, Support vector algorithms for optimizing the partial area under the ROC curve, Neural Comput. (2017) 1–45.