



ARTICLE

Breast Tumor Computer-Aided Detection System Based on Magnetic Resonance Imaging Using Convolutional Neural Network

Jing Lu¹, Yan Wu^{2,*}, Mingyan Hu¹, Yao Xiong¹, Yapeng Zhou¹, Ziliang Zhao¹ and Liutong Shang^{1,†}

¹Department of Radiology, Fourth Medical Center of PLA General Hospital, Beijing, 100048, China

²Qinghe Clinic, Beijing North Medical District of Chinese PLA General Hospital, Beijing, 100085, China

[†]Corresponding Author: Liutong Shang. Email: shangjj79@163.com

^{*}Co-first Author: Yan Wu

Received: 15 June 2021 Accepted: 07 July 2021

ABSTRACT

Background: The main cause of breast cancer is the deterioration of malignant tumor cells in breast tissue. Early diagnosis of tumors has become the most effective way to prevent breast cancer. **Method:** For distinguishing between tumor and non-tumor in MRI, a new type of computer-aided detection CAD system for breast tumors is designed in this paper. The CAD system was constructed using three networks, namely, the VGG16, Inception V3, and ResNet50. Then, the influence of the convolutional neural network second migration on the experimental results was further explored in the VGG16 system. **Result:** CAD system built based on VGG16, Inception V3, and ResNet50 has higher performance than mainstream CAD systems. Among them, the system built based on VGG16 and ResNet50 has outstanding performance. We further explore the impact of the secondary migration on the experimental results in the VGG16 system, and these results show that the migration can improve system performance of the proposed framework. **Conclusion:** The accuracy of CNN represented by VGG16 is as high as 91.25%, which is more accurate than traditional machine learning models. The F1 score of the three basic networks that join the secondary migration is close to 1.0, and the performance of the VGG16-based breast tumor CAD system is higher than Inception V3, and ResNet50.

KEYWORDS

Computer-aided diagnosis; breast cancer; VGG16; convolutional neural network; magnetic resonance imaging

1 Introduction

In current medical imaging and analysis, the methods involved in early tumor diagnosis mainly include: medical imaging examination and genetic diagnosis [1]. Among them, the application of medical imaging examination is far more common than genetic examination [2], and is worthy of investigation. There are currently a variety of diagnostic imaging methods for medical imaging examinations, which include mammography, magnetic resonance imaging, CT imaging, ultrasound imaging and PET imaging. However, mammography is the main diagnostic imaging method commonly used in the pre-diagnosis of breast cancer [3]. The current computer-aided diagnosis



system for breast tumors has played an important role in helping physicians assist in improving the detection accuracy and sensitivity of diagnosis results.

However, the current research methods in the computer diagnosis system are all CAD systems designed based on machine learning methods, and the system improvement of the current machine learning method design encounters a bottleneck. Therefore, it is of great significance for us to explore the use of deep learning methods to improve mammography and magnetic resonance image (MRI) systems. Design based on a unified machine learning method must consider all parts of the design process, including feature extraction, feature selection, dimensionality reduction, and classifier selection. The design of this paper will introduce deep learning [4,5] and apply it to the establishment of our breast tumor diagnosis system. We can then conduct a series of comparative experiments in each link of the system to verify the effectiveness of the system design [6].

There are several methods for breast cancer imaging screening as shown above. Different medical imaging methods are aimed at different diseases and have their own different advantages. Among them, the screening of X-ray images has the most common application, and its detection accuracy and sensitivity for microcalcification points is high [7]. For malignant tumors [8], almost all have the characteristics of microcalcification. Therefore, X-ray imaging has become the main method of early tumor screening, followed by ultrasound imaging, which is more widely used, which is characterized by easy ultrasound detection and low cost. These two image shooting methods are more commonly used in the initial screening. In the later stage of diagnosis, CT and MRI imaging examinations are more commonly used.

The characteristic of MRI imaging is that there is no radiation hazard in the process of imaging and inspection, and it has high resolution for soft tissues and can simultaneously image both breasts. Therefore, MRI performs well in the diagnosis of benign and malignant breast tumors, and it is also widely used in the biological behavior and prognosis of tumors [9]. But at the same time, the imaging of MRI itself has certain limitations. Due to its long imaging time, it is very susceptible to the patient's breathing, heartbeat, and slight changes in the patient's position during the imaging process, making the imaging quality low. However, because the advantages of MRI diagnosis are also obvious, in the general diagnosis process, physicians will combine other imaging methods with MRI as the basis for diagnosis.

This technology is constantly being explored and applied to clinical diagnosis. The diagnosis system based on breast tumors has been commercialized in the United States. The research and development of breast tumor system is also being carried out on a large scale. Large-scale clinical application has helped doctors in improving the efficiency and accuracy of the diagnosis process. However, due to the performance of the current breast tumor diagnosis system, it is difficult to achieve a substantial improvement, and the actual demand forces the research to be further deepened. Regarding the breast tumor diagnosis system, the successful detection and identification of calcification points in the breast needs to be further researched in order to improve the detection of breast tumor.

2 Methodology

2.1 Convolutional Neural Network

The key idea of CNN is multi-layer stacking, local connection, weight sharing and pooling. The layers in CNN are no longer fully connected, but partially connected, which can greatly simplify the complexity of the model and reduce the number of parameters. CNN can learn image features by itself, and does not require professional pathologists to perform complex feature

extraction. It is more suitable for the learning and expression of breast pathological image features. At present, CNN has become the key to deep learning algorithms in breast cancer pathological image research, and is applicable for many medical problems, such as medical image recognition, medical segmentation, etc.

2.1.1 VGGNet

VGGNet has been widely used in medical diagnosis [10–12]. It was developed by one of the Oxford University Computer Vision Group and Deep Mind. VGGNet further deepens the depth of the network. At the same time, the network has good scalability and can be transplanted and applied to various tasks. This paper uses VGG16 for extended applications. In VGGNet, convolution kernels having a size of 3×3 are used, and the maximum pooling operation is still used, and the pooling size of 2×2 is adopted. The structure is simple and it is often used for feature extraction. The deepening of the number of network layers increases the expressiveness of the network and further improves performance. The main parameters in the network are still concentrated in the fully connected layer, and the number of layers in the convolution part is deepened, but the parameter amount has not increased. Due to the use of a small convolution kernel, the parameter amount in the network is reduced, and the convolution layer itself. The number of parameters generated is limited.

Among them, VGG16 and VGG19 are more common in practical applications. Both networks use 1×1 convolutions. The function of the 1×1 convolutional layers plays a linear change in the current network and increases the combination of networks. Since the number of channels before and after the corresponding 1×1 convolution has not changed, there is no dimensionality reduction and dimensionality increase in the network. The advantage of using a small convolution kernel is that the amount of parameters in the network is further reduced, which helps to further deepen the number of network layers, so that the expression performance of the network is improved. The focus is on VGG16 [13].

2.1.2 Inceptionnet

InceptionNet won the championship in the 2014 ILSVRC competition [14], and its top-5 error rate was only 6.67%. InceptionNet includes V1, V2, V3, V4 series, these networks have carried out a series of improvements on the basis of the former. Among them, the InceptionNet that won the championship in the competition is called Inception V1. The network layer is 22 layers deep, making the network deeper, but the amount of calculation is only one-twelfth of AlexNet.

Inception V3 has been improved on the basis of the first two networks. The large convolution kernel is separated into a combination of two smaller convolution kernels. For example, 7×7 is reduced to 1×7 and 7×1 . Reduce the amount of parameters in the network, increase the transformation method, and make the expression of features more abundant, make the speed of calculation faster, and reduce the possibility of network overfitting, and enhance the generalization ability of the network. The size of the convolution kernel is changed in Inception Module. Inception V3 includes 8×8 , 17×17 , and 35×35 blocks, which increases the richness of the network's feature expression, deepens and broadens the network structure.

2.1.3 Resnet

ResNet won the championship in the 2015 ILSVRC competition [15], with a top-5 error rate of 3.57%. It was a network proposed by Kaiming He of Microsoft Asian Research Institute and several other authors. The depth of the network has reached 152 layers, breaking through the depth of all previous designs of convolutional neural networks.

When the network performance drops, it promotes the deepening of the CNN network layer. Therefore, this becomes an important module for the convolutional neural network in the process of deepening the number of network layers. The purpose of the ResNet is to increase the information transmission process; the information of the previous layer of the network is directly transmitted to the design of the next layer. Part of the information will be lost, so the shortcut added by Res Net is directly transmitted to the current layer through the previous layer network signal, which solves the process of network deepening.

2.2 CAD System

2.2.1 Traditional Machine Learning Features

This paper selects Gabor, Gray Level Co-occurrence Matrix (GLCM) and HarrLike, three traditional machine learning features compare with the extracted CNN features. The Gabor's characteristic is that it is sensitive to the boundary of the image, but not sensitive to changes in light. It has good illumination robustness, and has greater selectivity in various directions and different scales. In this paper, the features extracted from 0 to 8 directions are selected, and each direction takes different frequency scales of 0.5, 0.25, 0.125, 0.1, and then filtered by Gabor at different scales, so as to obtain the mean value between the pixels of the image and variance includes a total of 64 dimensions.

GLCM is different from traditional grayscale histograms. It does not only reflect the grayscale value and the grayscale value distribution. It is obtained by statistics of the two pixels in the image that have the same grayscale value at a certain distance. This way can well show the location association between pixels.

2.2.2 CNN Features

Compared with the traditional hand-designed features, CNN extracts the basic features of the picture through the low-level, which has stronger image expression ability. As shown in Fig. 1, the operation extracts the features of the image through convolution kernel, and convolution operation pays more attention to the positional relationship between the features.

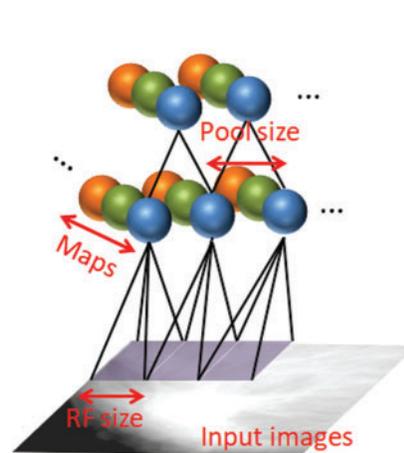


Figure 1: Schematic diagram of feature extraction by convolution operation

As a characterization of the image, it is obvious that the features extracted by CNN are more comprehensive. CNN adopts the form of convolution kernel to realize weight sharing at the same feature level and parallel learning, which greatly reduces the amount of calculation, and improves the efficiency of feature extraction and classification. In actual image processing, we often design specific CNN networks for feature extraction and further image classification.

Use VGG16, Inception V3, and ResNet50 to extract features on the MRI data set. Among them, VGG16 has completed the feature extraction of all layers except for fully connected, and all the layers in Inception V3, ResNet50, including convolutional layer, pooling layer, BN layer, etc., have hundreds of network layers, which is not suitable for each layer. Feature extraction. However, through experiments, it is found that for these two networks: the layer near a certain layer, the performance of feature extraction is similar. The few layers in the network can represent the feature extraction performance of adjacent layers on MRI. Therefore, according to the two network structures of Inception V3 and ResNet50, a small number of network layers are selected in different blocks, and feature extraction is performed on the MRI data set. The performance of the selected few layers can represent the feature extraction performance of adjacent layers on MRI.

2.2.3 CAD System Structure

The computer-aided diagnosis system built under the traditional machine learning method contains the content of these two parts, and in the process of building the system, it covers: image preprocessing, labeling and positioning of the region of interest, segmentation, and feature extraction, feature selection and classification. As shown in Fig. 2.

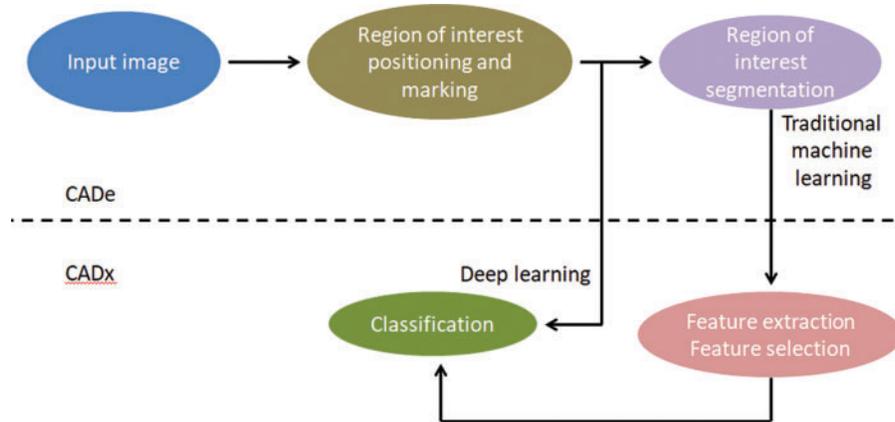


Figure 2: Medical imaging based CAD system framework

The breast tumor diagnosis system studied in this paper belongs to the category of CAD framework. After finding the Region of Interest (ROI), the segmented Region of Interest is processed on the basis of it, including data balance processing and data enhancement. After processing, the final classification of the data set completes the task of the breast tumor diagnosis in this paper. The CAD system design framework of this paper is shown in Fig. 3.

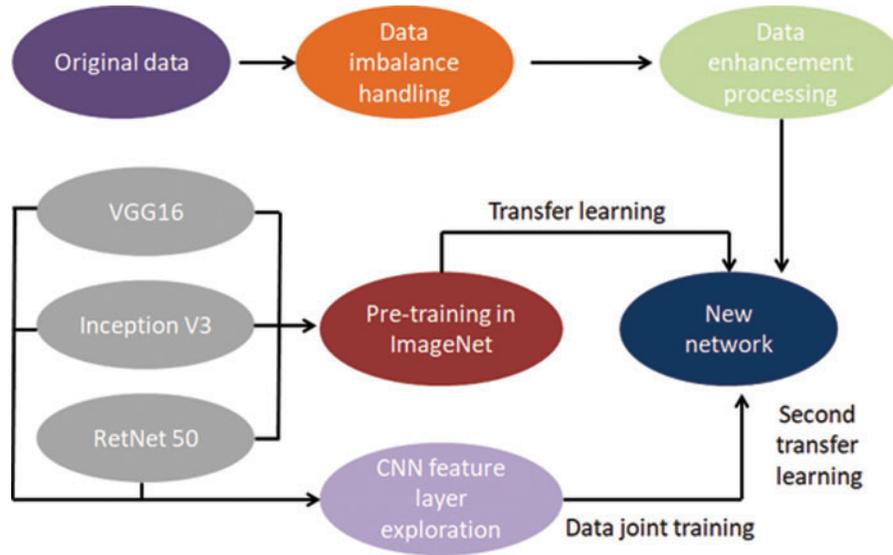


Figure 3: Breast tumor CAD system design framework of this paper

2.3 Convolutional Neural Network-Second Migration

There is a big gap between ImageNet natural images and medical tumor images. When carrying out transfer learning, the mainstream view is that the transfer effect of similar categories is better than the transfer between two relatively different categories. But there is no large-scale data set in the field of medical imaging. Considering the above factors, based on transfer learning, this paper proposes a second transfer method, which consists of two steps and ImageNet is the source domain. We build a new network, introduce the first transfer learning and train the fine-tuning network. The weights after ImageNet pre-training will be migrated to all layers before the L2 layer of the network. On the data set, use network weight migration and fine-tuning. The second migration is based on the network after the first migration. The purpose is to transfer the best model weights obtained in the first training, and perform network fine-tuning on the MRI data set after imbalance processing and data enhancement, which can improve the CAD system performance.

2.4 Performance Evaluation Standards

The evaluation criteria used in this paper are as follows, including: Accuracy, Precision, Recall, and $F1$ value. They are calculated as Eqs. (1)–(5):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Sensitivity = Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Sensitivity = Recall = \frac{TN}{TN + FP} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$F_1 = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (5)$$

Among them, true positive (TP) indicates that the true category of the sample is positive, and the final predicted result is also positive; false positive (FP) indicates that the true category of the sample is negative, but the final predicted result is positive, which is usually called false negative (FN) indicates that the true category of the sample is positive, but the final predicted result is negative, which is commonly referred to as false negative; true negative (TN) indicates that the true category of the sample being negative.

3 Results

3.1 Data Set and Processing

3.1.1 Introduction to Data Set

There is no large-scale annotated data set in the medical field. This paper uses a small public Digital Database for Screening Mammography (DDSM) [16] and a magnetic resonance data set MRI data set constructed by a domestic authority. The MRI data was collected from 58 patients ranging in age from 35 to 60. Each patient received multiple examinations and shots from different angles and positions. The annotations and markings of the region of interest are manually segmented and validated by the radiologist.

The MRI data set contains 623 positive samples and 2457 negative samples, and the data set presents an imbalance ratio of 1:4. The random upsampling algorithm is used to achieve equilibrium in image processing, and data enhancement is performed. This data set is used as a transfer learning and secondary transfer learning method, the final classification task. It is used as the target domain data set of transfer learning and the target domain dataset of the secondary transfer learning method. Examples of samples in the MRI data set are shown in Figs. 4 and 5.

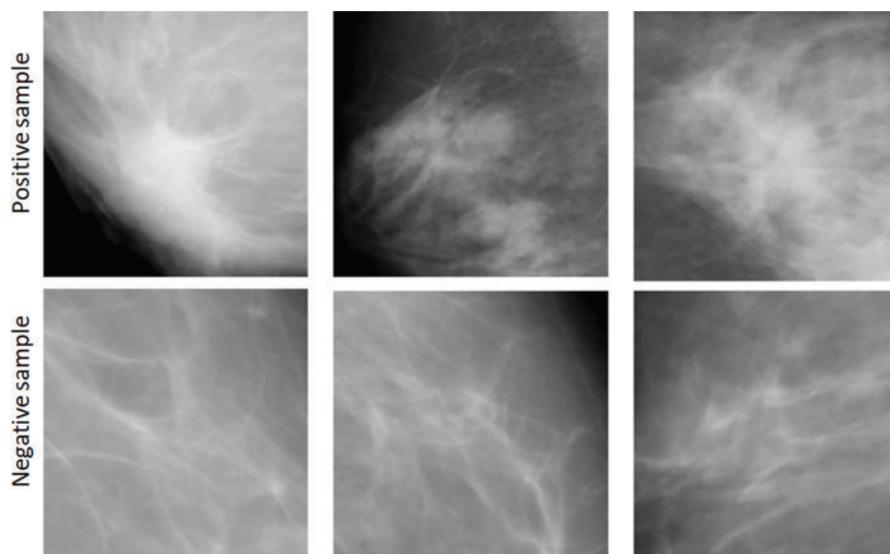


Figure 4: Region of interest in the digital database for screening mammography images

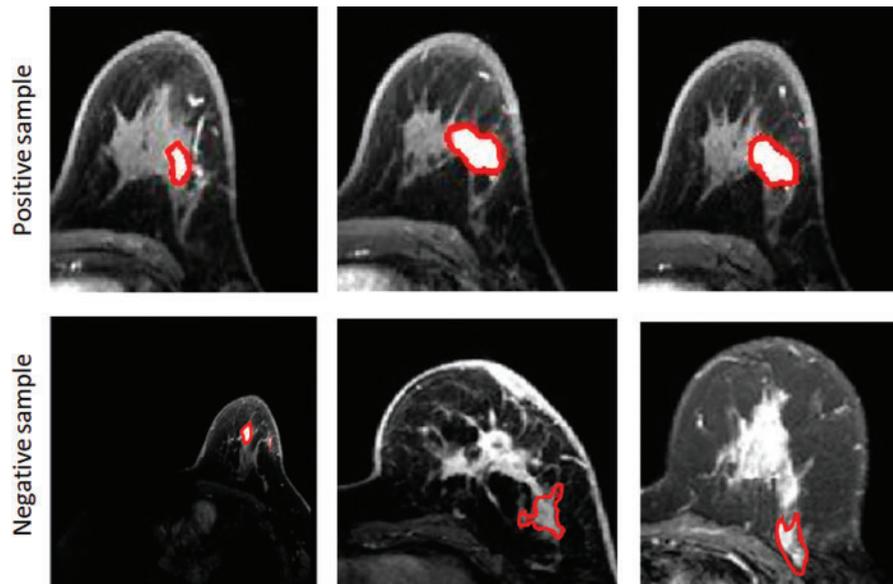


Figure 5: Region of interest marked in magnetic resonance images of digital database for screening mammography

3.1.2 Data Set Imbalance Processing

Data imbalance is a normal state of the data set, and it appears in various forms of data sets. In the medical field, data imbalance is common in image data sets for classification tasks. The imbalance of the data will cause the algorithm and the built network to pay more attention to the samples of most categories in the learning process, and are more inclined to learn more specific details in the majority of samples, while the samples of a few categories will be in the learning process. Among them, the learning of relatively most class samples is ignored. Misdiagnosis of a tumor as a non-tumor condition can endanger the life of the patient. Therefore, the minority samples, that is, the positive samples of the data set in the paper, contain information that is more important for the final classification task, and the misclassification of the positive samples presents a greater cost [17]. The imbalance problem resulted in a poor false positive rate.

The unbalanced data set must be processed before the experiment to reach a balanced state. Algorithms for unbalanced data processing include: data down-sampling, data up-sampling and synthetic data generation, cost-sensitive learning, etc. The random upsampling algorithm [18] is selected to preprocess the unbalanced data set. The random upsampling algorithm will not remove the information in the minority sample set and ensure that the information loss in the minority sample set is zero [19].

3.2 Performance Comparison

3.2.1 Feature Extraction Performance Comparison

Table 1 shows the results of feature extraction. In this paper, Gabor, GLCM and HarrLike traditional machine learning are used. CNN layer features combined with three traditional machine learning methods are abbreviated as V+HarrLike, V+Gabor, V+GLCM, and we compare these with CNN that is represented by VGG16.

Table 1: Performance comparison of different features

Model	Accuracy	Precision	Sensitivity	F ₁ -score
Gabor	0.8642	0.8643	0.8457	0.8854
GLCM	0.8768	0.7564	0.9241	0.8968
HarrLike	0.7563	0.8679	0.7986	0.8147
V+Gabor	0.9002	0.8947	0.8964	0.9148
V+GLCM	0.8948	0.8246	0.9467	0.9166
V+HarrLike	0.8876	0.8712	0.8915	0.8938
VGG16	0.9125	0.8999	0.9145	0.9078

The results show that the feature extracted from CNN, namely VGG16, performs significantly better than traditional machine learning features on F1 parameters, and the fusion of the three features. The features of the three traditional machine learning methods are separately fused with the CNN features to improve the F1-score, making the expression performance of the traditional features close to the feature performance of CNN itself.

Table 2 shows the difference in the performance of features extracted on MRI using the internal VGG16 network, after removing the fully connected layer, the remaining 8 to 17 layers contain different pooling layers and convolutional layers.

Table 2: Comparison of feature extraction performance in VGG16 layer

Layer	Accuracy	Precision	Sensitivity	F ₁ -score
Layer 8	0.6124	0.5249	0.9487	0.7033
Layer 9	0.6579	0.5467	0.9648	0.7024
Layer 10	0.7111	0.5646	0.9622	0.7645
Layer 11	0.8247	0.6998	0.9615	0.8648
Layer 12	0.9045	0.8046	0.9589	0.9286
Layer 13	0.9248	0.9674	0.9876	0.9344
Layer 14	0.8997	0.8623	0.9798	0.9067
Layer 15	0.9478	0.9112	0.9867	0.9463
Layer 16	0.9579	0.9457	0.9742	0.9617
Layer 17	0.9672	0.9489	0.9627	0.9647

Table 2 shows the difference in the feature performance extracted on MRI after removing the fully connected layer from the VGG16 network. The remaining 8 to 17 layers contain different pooling layers and convolutional layers. Among them, the 14th layer has the highest F1 parameter. The results show that within the CNN network, as the network layer deepens, the performance of the features extracted by the network layer shows an upward trend for the final classification task. Some studies have shown that the deeper the CNN network, the extracted features are more high-level, abstract, and relatively the features extracted from the shallower network layer contain more underlying information such as image edge contours and textures.

3.2.2 Network Performance Comparison

We compare the network performance of VGG16, Inception V3, and ResNet50, as shown in Table 3.

Table 3: Comparison of three types of network performance

Model	Accuracy	Precision	Sensitivity	F ₁ -score
Inception V3	0.8241	0.8725	0.8674	0.8836
ResNet50	0.8585	0.8457	0.9005	0.8921
VGG16	0.8978	0.8966	0.9157	0.9033

It can be seen from Table 3 that the Accuracy, Precision, Sensitivity and F1-score of VGG16 are higher than Inception V3 and ResNet50, so our paper uses VGG16 as the basic network.

3.3 New Network Joins Migration Learning

Fig. 6 shows the comparison of F1 scores between three CAD systems built on the basis of VGG16, Inception V3, and ResNet50 networks, with transfer learning added and fine-tuned, and training without transfer learning.

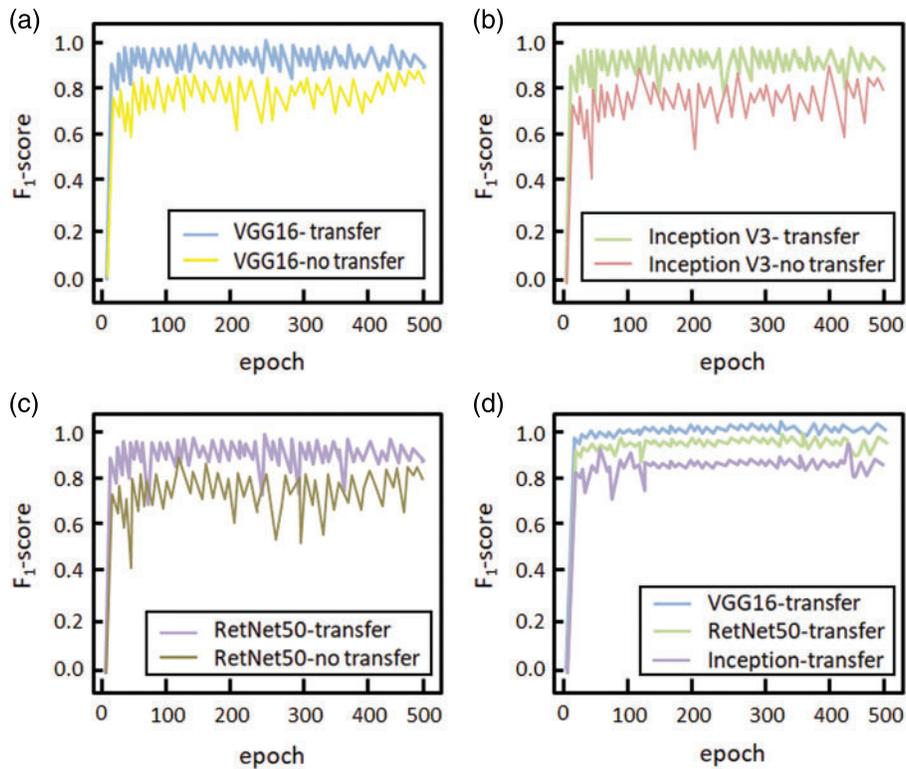


Figure 6: The difference in F1-scores of the three networks based on (a) VGG16 (b) inception V3 (c) ResNet50 (d) Inception-transfer

It can be seen from Fig. 6 that the performance of the network without migration learning is worse than the network with migration learning. As shown in Fig. 6d, the performance of the system based on VGG16 and ResNet50 is better than that based on Inception V3. There is no over-fitting phenomenon in the network that introduces the transfer learning method, and the loss function on the training set and the test set decreases. Therefore, application of a deeper network to a small data set and introducing a migration learning method will help improve the final classification performance of the system.

4 Discussion

The CAD system for breast tumor diagnosis has played an increasingly important role, and has gradually received attention from the research and medical fields. This paper improves the performance of the CAD system, and successfully introduces deep learning into the medical small data set, and has achieved certain results, but the system design details still need to be further improved [20].

Since the learning of the training set by the deep learning network is mainly reflected in the distribution of the learning data [21], the large-scale data volume can more effectively cover the different distribution methods of the data, so the network can learn a richer data distribution situation, and covering more of the data usage. The different forms of the network can improve the final performance of the network on the one hand, and can improve the network further. Therefore, the expansion of the data set can fundamentally improve the performance and generalization ability of the network.

In terms of data set construction, when data is insufficient, transfer learning can be used to make up for the problem of insufficient data. There are also some methods to enhance the amount of data, including data enhancement and synthetic data. Data enhancement increases the sample copy in the original data, but this method does not make the distribution of the data volume more perfect. Therefore, increasing the diversity of samples is also a key link in improving system performance.

In the secondary migration method proposed in this paper, there is still a lot of room for improvement, and more medical image data sets in similar fields can be obtained [22]. Many different forms of data fusion and the application of multiple migration learning may be considered. For example, by obtaining more forms of medical images in similar fields, conducting joint training on data, or using multiple migration methods, the final performance of the network can be improved.

5 Conclusion

The breast tumor CAD system proposed in this paper is clearly ahead of the current similar systems in the classification of breast tumors. The addition of migration learning has enabled deep learning to be successfully introduced into the classification task of medical small data sets, and its performance is significantly higher than that of traditional machine learning methods. To build a CAD system based on the classic CNN network, we use different layers in the network and the feature extraction function of the target data set to guide the construction and fine-tuning of the new network. Based on the transfer learning method, the second transfer method is proposed for the first time, which effectively improves the classification performance of the system, and is also important in improving the CAD diagnostic accuracy, as well as providing reference for other medical small data image classification.

The design method of this paper further improves the performance of breast tumor CAD system. Based on deep learning, this paper introduces transfer learning and proposes secondary transfer learning, which improves the classification performance and alleviates the difficulty of introducing deep learning into medical data sets. Comparative experiments were carried out with traditional machine learning methods, methods without transfer learning, and data joint training methods. The specific introduction method of transfer learning, secondary transfer learning, and data joint training methods provide reference significance for the classification tasks of small data sets and medical imaging data sets.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Yang, Y., Hu, Y., Shen, S., Jiang, X., Gu, R. et al. (2021). A new nomogram for predicting the malignant diagnosis of breast imaging reporting and data system (BI-RADS) ultrasonography category 4A lesions in women with dense breast tissue in the diagnostic setting. *Quantitative Imaging in Medicine and Surgery*, 11(7), 3005–3017. DOI 10.21037/qims.
2. Bicchierai, G., di Naro, F., de Benedetto, D., Cozzi, D., Pradella, S. et al. (2021). A review of breast imaging for timely diagnosis of disease. *International Journal of Environmental Research and Public Health*, 18(11), 5509. DOI 10.3390/ijerph18115509.
3. Fenton, J. J., Taplin, S. H., Carney, P. A., Abraham, L., Sickles, E. A. et al. (2007). Influence of computer-aided detection on performance of screening mammography. *The New England Journal of Medicine*, 356(14), 1399–1409. DOI 10.1056/NEJMoa066099.
4. Oquab, M., Bottou, L., Laptev, I., Sivic, J. (2014). Learning and transferring mid-level imager presentations using convolutional neural networks. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1717–1724. New York.
5. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F. et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42(1), 60–88. DOI 10.1016/j.media.2017.07.005.
6. Pan, S. J., Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. DOI 10.1109/TKDE.2009.191.
7. Brennan, M., Spillane, A., Houssami, N. (2009). The role of breast MRI in clinical practice. *Australian Family Physician*, 38(7), 513–519.
8. Fitzpatrick, J. M., Sonka, M. (2000). Medical image processing and analysis. In: *Handbook of medical imaging*, vol. 2. Washington: The International Society for Optical Engineering Press.
9. Giger, M. L., Chan, H. P., Boone, J. (2008). Anniversary paper: History and status of CAD and quantitative image analysis: The role of medical physics and AAPM. *Medical Physics*, 35(12), 5799–5820. DOI 10.1118/1.3013555.
10. Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint, arXiv:1409.1556.
11. Mudigonda, N. R., Rangayyan, R. M., Desautels, J. E. (2000). Gradient and texture analysis for the classification of mammographic masses. *IEEE Transactions on Medical Imaging*, 19(10), 1032–1043. DOI 10.1109/42.887618.
12. Biglia, N., Bounous, V. E., Martincich, L., Panuccio, E., Liberale, V. et al. (2011). Role of MRI (magnetic resonance imaging) versus conventional imaging for breast cancer presurgical staging in young women or with dense breast. *European Journal of Surgical Oncology*, 37(3), 199–204. DOI 10.1016/j.ejso.2010.12.011.
13. Jain, A. K., Duin, R. W., Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 4–37. DOI 10.1109/34.824819.

14. Szegedy, C., Vanhoucke, V., Ioffe, S. (2016). Rethinking the inception architecture for computer vision. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826. New York.
15. He, K., Zhang, X., Ren, S., Chen, J. (2016). Deep residual learning for image recognition. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778. New York.
16. Heath, M., Bowyer, K., Kopans, D., Robinson, M. (2000). The digital database for screening mammography. *Proceedings of the 5th International Workshop on Digital Mammography*, pp. 212–218. Trabzon.
17. He, H., Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284. DOI 10.1109/TKDE.2008.239.
18. Liu, A., Ghosh, J., Martin, C. E. (2007). Generative oversampling for mining imbalanced datasets. *International Conference on Data Mining*, pp. 66–72. Omaha.
19. Longadge, R., Dongre, S. (2013). Class imbalance problem in data mining review. arXiv preprint, arXiv:1305.1707.
20. Ganesan, K., Acharya, U. R., Chua, C. K., Anand, D. (2013). Computer-aided breast cancer detection using mammograms: A review. *IEEE Reviews in Biomedical Engineering*, 6(1), 77–98. DOI 10.1109/RBME.2012.2232289.
21. Tang, Z., Zhao, G., Ouyang, T. (2021). Two-phase deep learning model for short-term wind direction forecasting. *Renewable Energy*, 173(1), 1005–1016.
22. Wong, K., Fortino, G., Abbott, D. (2019). Deep learning-based cardiovascular image diagnosis: A promising challenge. *Future Generation Computer Systems*, 110(1), 802–811.