https://journal.ziu- مجلة جامعة الزيتونة الدولية – مجلة علمية محكمة تصدر عن جامعة الزيتونة الدولية
university.net
30/08/2024
المجلد الأول
Issue: N25 ISSN: 2958-8537 العدد الخامس و العشرون م.1 : ص.ص 1-10
Al-Zaytoonah University International Journal for Scientific Publishing

A hybrid deep learning model for breast cancer Mammographic Image Classification based on transfer learning and an attention module

Tawfik Ezat Mousa[1], Mohamed S. Geoda[2]

[12] Departement of Computer Technologies,

Higher Institute Of Science and Technology, Tobruk, Libya.

Email: Tawfikezat@yahoo.com

https://orcid.org/0000-0001-8066-738X

Email: Mhomadgg@gmail.com

https://orcid.org/0009-0007-2800-0976

https://journal.ziu- مجلة جامعة الزيتونة الدولية – مجلة علمية محكمة تصدر عن جامعة الزيتونة الدولية
university.net
30/08/2024
المجلد الأول
Issue: N25 ISSN: 2958-8537 العدد الخامس و العشرون 1.م : ص.ص 1-10
Al-Zaytoonah University International Journal for Scientific Publishing

Abstract

Breast cancer is one of the primary causes of death among women. Early detection of breast cancer allows for the receipt of appropriate treatment, thus increasing the possibility of survival. In this paper, we proposed a hybrid deep learning model using a pre-trained VGG16 model with a self-attention mechanism for breast cancer detection. We extract features from the binary class (benign, malignant) dataset of the mammographic image analysis society (MIAS) using pre-trained deep convolutional neural network (CNN) architectures like Xception, MobileNet, DenseNet, and VGG-16. So the results illustrated that the best model is VGG16 with a self-attention module, which achieved an accuracy of 98.77%.

Keywords: Breast cancer, VGG16, MIAS, Mammography, Classification**.**

1. Introduction

Around the world, many women die from breast cancer each year. The existing traditional methods, such as self-breast exams and the doctor examinations, mostly fail to detect the disease earlier. Early detection and treatment are key to improving outcomes for individuals with breast cancer. Therefore, a mammogram uses low-dose x-rays to examine the breast. On the other hand, deep learning has become the state-of-the-art method in image recognition due to its high accuracy. We can use it to predict breast cancer from a mammogram. Therefore, the study will focus on improving classification accuracy in determining whether the tumor is benign or malignant. In the past few years, several researchers have performed classifications of benign and malignant breast cancer using different neural network classifiers. Alruwaili et al. [1] used the ResNet50 model to distinguish between malignant and benign breast cancer. They applied a data augmentation technique to increase the number of training images and prevent the model from overfitting. The MIAS dataset used in the proposed model, achieved an accuracy of 89.5%. Singh et al. [2] proposed an automatic pipeline for the detection and classification of all categories of microcalcification. They applied a transfer learning approach to the pre-trained InceptionResNetV2 model with various optimization methods. Their sysThey assessed their system on the CBIM-DDSM database to detect and classify breast tumors, achieving an accuracy of up to 94%. In other work, Jiang et al. [3] integrated a new dataset of breast mammograms named Film Mammography Dataset number 3 (BCDR-F03). They applied both GoogLeNet and AlexNet models to classify segmented tumors found on mammograms and obtained accuracies of 0.88 and 0.83 for GoogleNet and AlexNet, respectively. Rouhi et al. [4] proposed a new model of primary breast cancer detection

https://journal.ziu-university.net – مجلة جامعة الزيتونة الدولية – مجلة علمية محكمة تصدر عن جامعة الزيتونة الدولية

30/08/2024
المجلد الأول

Issue: N25 ISSN: 2958-8537 العدد الخامس و العشرون م.1 : ص.ص 1-10

Al-Zaytoonah University International Journal for Scientific Publishing

using the region-growing method. Their model is based on the hybridization of cellular neural networks with genetic algorithms, and they achieved an accuracy of 96.47% and 95.13% on the MIAS and DDSM databases, respectively. Pillai et al. [5] evaluated several pre-trained deep learning models, such as EfficientNet, AlexNet, VGG16, and GoogleNet, on the MIAS database. They applied data augmentation to increase the number of training images and prevent the model from overfitting. The VGG16 model yielded the best performance, achieving an average accuracy of 75.46%.

2. Materials and Methods

2.1. Dataset

In our experiments, we used the MIAS dataset, a public mammography dataset that is freely available online via the link: https://www.repository.cam.ac.uk/ handle/1810/
250394.
The Mammography Analysis Society's (MIAS) dataset contains mammograms from 322 images. All images are 1024 x 1024 pixels in Portable Gray Map (PGM) format. The dataset also includes radiologists' actual estimates of the location of abnormalities (benign and malignant), as well as an approximation of the radius surrounding the abnormality's center. In this work, we use the images in the dataset, which include benign and malignant. The following Figures 1 and 2 show the original images of the two categories (benign and malignant).

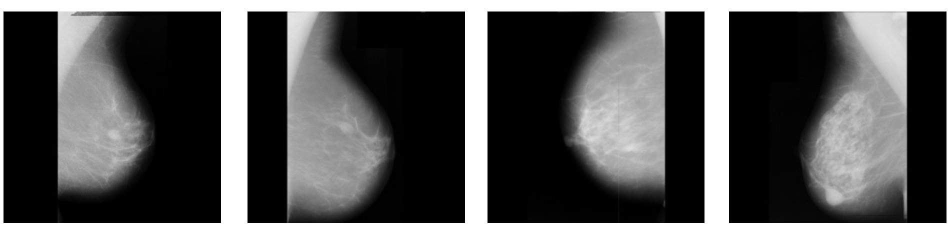

Figure 1. Mammogram (Benign) from the MIAS database (mdb010, mdb012, mdb013, mdb021)

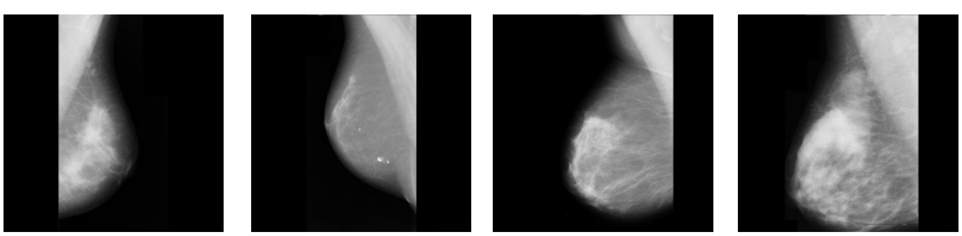

Figure 2. Mammogram (Malignant) from the MIAS database (mdb102, mdb075, mdb095, mdb115)

https://journal.ziu- مجلة جامعة الزيتونة الدولية – مجلة علمية محكمة تصدر عن جامعة الزيتونة الدولية
university.net
30/08/2024
المجلد الأول
Issue: N25 ISSN: 2958-8537 العدد الخامس و العشرون م.1 : ص.ص 1-10
Al-Zaytoonah University International Journal for Scientific Publishing

2.2. Models

The MIAS mammography dataset was applied to four deep CNN models for the classification of breast cancer. The models included Xception, MobileNetV2, VGG16, DenseNet121 networks.

Xception is renowned for its depth-wise separable convolutions, which contribute to a network's low parameter count while preserving excellent performance [6]. On picture classification challenges, this architecture has demonstrated state-of-the-art performance.
MobileNetV2 model is a convolutional neural network architecture that efficiently uses computational resources to maintain high accuracy in image classification tasks. To reduce computational costs, MobileNetV2 uses lightweight, depthwise separable convolutions that separate spatial and depthwise convolutions into separate layers [7]. Various mobile devices have widely adopted this architecture due to its balance between efficiency and accuracy.
Visual Geometry Group (VGG) is a convolutional neural network architecture that consists of multiple blocks of convolutional layers followed by max-pooling layers [8]. The number of filters used in the convolutional layers increases in priority as the network expands. VGG is known for its simplicity and high performance on image classification tasks.
DenseNet is a contemporary architecture with few parameters intended for visual object recognition. The two main components that make up the network's basic structure are the Dense and Transition blocks. DenseNet-121 consists of three transition blocks and four dense blocks. Each layer in the Dense Block is connected to all subsequent layers in a densely manner [9].

3. Proposed methodology

The methodology used proposes applying the VGG16 deep learning model with a self-attention mechanism to improve the accuracy and efficiency of the model's classification.

3.1. VGG16 Architecture

The architecture of VGG-16 [10] is a convolutional neural network model consisting of 16 layers, including 13 convolutional layers and three fully connected layers. The design uses 3x3 filters with a stride of 1 pixel. This design helps the network learn more complex features by stacking multiple convolutional layers on top of each other. A rectified linear unit (ReLU) activation function is applied after each convolutional layer to introduce nonlinearity. To downsample feature maps and reduce spatial dimensions, use the

https://journal.ziu-university.net مجلة جامعة الزيتونة الدولية – مجلة علمية محكمة تصدر عن جامعة الزيتونة الدولية
30/08/2024
المجلد الأول
ISSN: 2958-8537 Issue: N25 العدد الخامس و العشرون م.1 : ص.ص 1-10
Al-Zaytoonah University International Journal for Scientific Publishing

maximum pooling layers. At the end of the architecture, fully connected layers in the network are used for classification tasks.

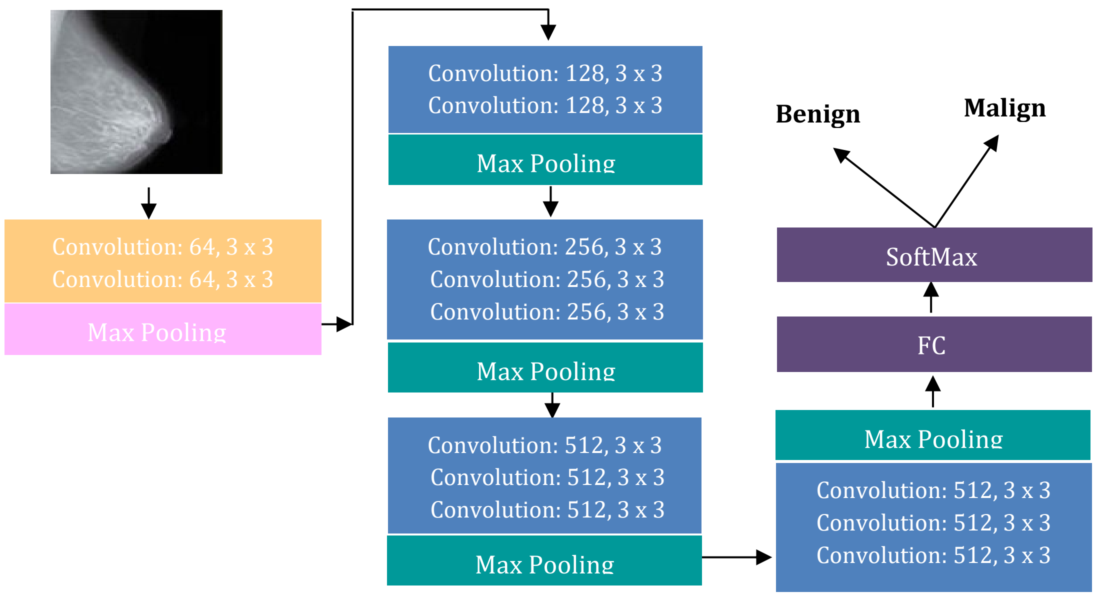

Figure 3. VGG16 architecture

3.2. Self-attention module

We implemented the multi-head attention model to improve model effectiveness, as shown in Figure 4. Multi-head attention serves as a mechanism that enhances the focus on a specific component of the data. It enables the network to concentrate on a few aspects at a time and ignore the rest [11]. Multi-Head Attention consists of several attention layers running in parallel, rather than performing one single attention function. In particular, the input consists of queries and keys of dimension dk (Q and K, respectively), as well as values of dimension dv (V). (Equation 1) computes the scaled dot product of the queries with all keys and applies a SoftMax function to obtain the weights on the values V. We linearly project the attention mechanism h times, using different learned weights (WQ, WK, and WV). (Equation 2) concatenates these different representation subspaces into a single attention head to form the final output result. We implemented a specific version of the attention model known as self-attention, which utilizes identical query, key, and value inputs. Next, we apply a SoftMax function to these scores to obtain attention probabilities.

Finally, we take a linear combination of these distributions with the value input tensors and concatenate them into one channel.

$$\mathbf{Attention}(Q,K,V) = \mathbf{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \times V \qquad (1)$$

$$\begin{cases} \mathbf{MultiHead}(Q,K,V) = \mathbf{Concat}(\mathbf{head}_1,..,\mathbf{head}_h) \\ head_i = Attention(QW^Q, QW^K, QW^V) \end{cases} \qquad (2)$$

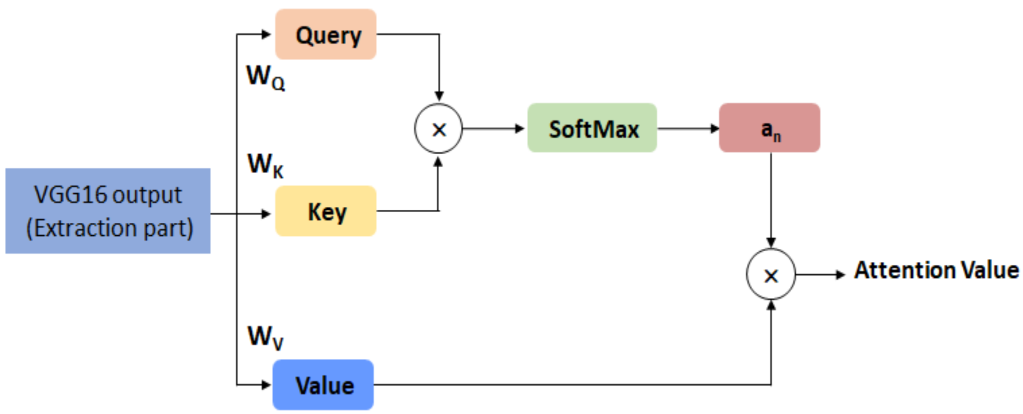

Figure 4. Attention architecture

The suggested methodology involves calculating the dot product of the outputs from the VGG16 model and the self-attention model. Subsequently, we implemented global mean pooling on both the attention model output and the generated dot product tensors. The classification component comprises two densely connected layers with a dropout mechanism to mitigate overfitting of the model. Figure 5 illustrates the distinct components of the proposed breast cancer detection system.

https://journal.ziu- مجلة جامعة الزيتونة الدولية – مجلة علمية محكمة تصدر عن جامعة الزيتونة الدولية
university.net
30/08/2024
المجلد الأول
10-1 ص.ص : م.1 العدد الخامس و العشرون ISSN: 2958-8537 Issue: N25
Al-Zaytoonah University International Journal for Scientific Publishing

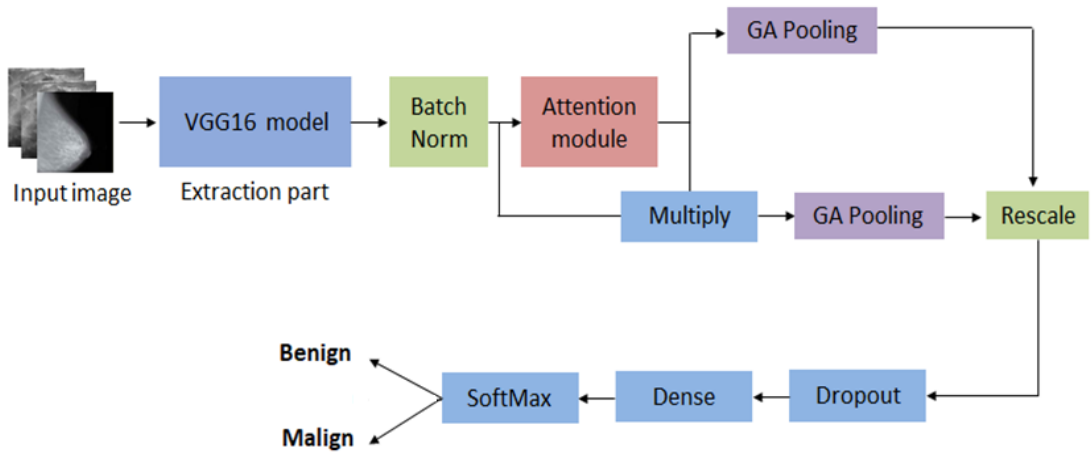

Figure 5. Proposed methodology architecture

4. Experimentation and Results

4.1. Data augmentation

In our experiment on the MIAS dataset, we classify images into two categories (benign and malignant). We divided the data set into an 80% training set and a 20% test set. Due to the limited image set, we applied data augmentation using geometric transformation techniques such as rotate, scale, rotate, and crop to increase the number of training samples and prevent the model from overfitting.

4.2. Experimental setup

During the experiments, we divided the training database into batches of size 32 using the shuffling option to create different min-batch samples in each epoch. Moreover, we computed the loss between the desired and calculated outputs in each iteration using the categorical cross-entropy method. We trained the model with an initial learning rate of 0.001 using the Adam (Adaptive Moment Estimation) optimizer. Moreover, in the multi-head self-attention model, we employed eight parallel attention layers, or heads. For each of these, we used 64 units in the linear projector of both query, key, and value matrices

4.3. Experimental results

In the experiments, the images shape was fixed to (256 × 256 × 3). We also studied several models with varying parameters. We used pre-trained weights for all of these models. First, we evaluated the model's performance without a self-attention mechanism. We achieved the best result with VGG16 (table 1). Applying the multi-head self-attention mechanism

https://journal.ziu- مجلة جامعة الزيتونة الدولية – مجلة علمية محكمة تصدر عن جامعة الزيتونة الدولية
university.net
30/08/2024
المجلد الأول
10-1 ص.ص : 1م العدد الخامس و العشرون ISSN: 2958-8537 Issue: N25
Al-Zaytoonah University International Journal for Scientific Publishing

improved the VGG16 accuracy by 4%, resulting in an accuracy of 0.9877. Figures 6 and 7 represent the confusion matrices for the classification report for 80:20 split ratios and the training and testing accuracy plot, respectively. Using the multi-head self-attention mechanism improved the model's performance.

Table 1. Models accuracies without and with attention

Model	Accuracy (%)	
	without Attention	**with Attention**
MobileNet	94.88	98.50
Xception	92.0	96.21
DenseNet-121	96.43	97.80
VGG-16	95.50	**98.77**

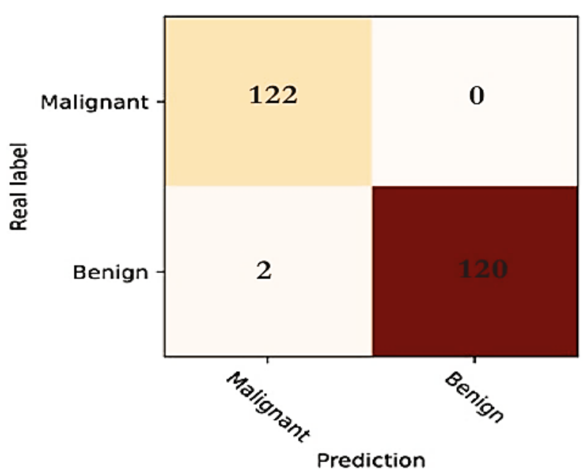

Figure 6. Confusion matrices for VGG16 with self-attention

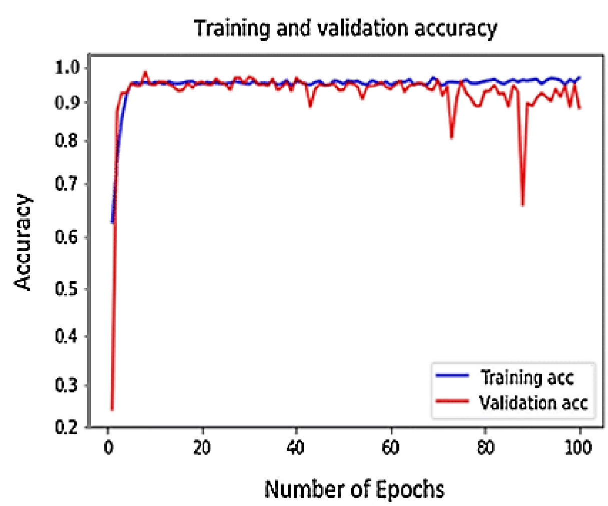

Figure 7. Training and validation accuracy for VGG16 with self-attention.

5. Conclusion

In this work, we proposed a methodology for breast cancer classification based on mammography to assist clinicians in breast cancer detection and diagnosis. This methodology provides a classification of breast images into benign and malignant categories. Our method is advantageous because it combines the VGG16 pre-trained deep convolutional neural network with the self-attention model. We tuned several hyperparameters, including the optimizer and learning rate, during the experiments to enhance the diagnosis efficiency. The proposed methodology outperformed both other pre-trained CNN architectures, achieving a classification accuracy of 98.77% on MIAS data. Moreover, it has good performance in cancer image recognition, which enhances the chances of survival.

References

1. Alruwaili, M., & Gouda, W. (2022). Automated breast cancer detection models based on transfer learning. Sensors, 22(3), 876.
2. Kumar Singh, K., Kumar, S., Antonakakis, M., Moirogiorgou, K., Deep, A., Kashyap, K. L., ..& Zervakis, M. (2022). Deep learning capabilities for the categorization of microcalcification. International Journal of Environmental Research and Public Health, 19(4), 2159.
3. Jiang, F., Liu, H., Yu, S., & Xie, Y. (2017, January). Breast mass lesion classification in mammograms by transfer learning. In Proceedings of the 5th international conference on bioinformatics and computational biology (pp. 59-62).
4. Rouhi, R., Jafari, M., Kasaei, S., & Keshavarzian, P. (2015). Benign and malignant breast tumors classification based on region growing and CNN segmentation. Expert Systems with Applications, 42(3), 990-1002.
5. Pillai, A., Nizam, A., Joshee, M., Pinto, A., & Chavan, S. (2022). Breast cancer detection in mammograms using deep learning. In Applied Information Processing Systems: Proceedings of ICCET 2021 (pp. 121-127). Springer Singapore.
6. Zaalouk, A. M., Ebrahim, G. A., Mohamed, H. K., Hassan, H. M., & Zaalouk, M. M. (2022). A deep learning computer-aided diagnosis approach for breast cancer. Bioengineering, 9(8), 391.
7. Aidossov, N., Zarikas, V., Mashekova, A., Zhao, Y., Ng, E. Y. K., Midlenko, A., & Mukhmetov, O. (2023). Evaluation of integrated CNN, transfer learning, and BN with thermography for breast cancer detection. Applied Sciences, 13(1), 600.
8. Yao, X., Wang, X., Wang, S. H., & Zhang, Y. D. (2022). A comprehensive survey on convolutional neural network in medical image analysis. Multimedia Tools and Applications, 81(29), 41361-41405.
9. Kim, T. H. (2021). Electricity Theft Detection using Fusion DenseNet-RF Model.
10. Sharma, S., Guleria, K., Tiwari, S., & Kumar, S. (2022). A deep learning based convolutional neural network model with VGG16 feature extractor for the detection of Alzheimer Disease using MRI scans. Measurement: Sensors, 24, 100506.
11. Meng, F., & Demeter, D. (2023). Sentiment analysis with adaptive multi-head attention in Transformer. arXiv preprint arXiv:2310.14505.