

An Improved EfficientNet Model and its Applications in Pneumonia Image Classification

Lina Li^{1,*}, Zexuan Tan¹ and Xiangzhu Han²

¹College of Sciences, Jilin Institute of Chemical Technology, Jilin 132022, China

²Department of Statistics, University of Warwick, Coventry CV4 7AL, UK

Received 19 July 2022; Accepted 19 December 2022

Abstract

Since the outburst of COVID-19, the medical system has been facing great challenges due to the explosive growth in detection and treatment needs within a short period. To improve the working efficiency of doctors, an improved EfficientNet model of Convolutional Neural Network (CNN) was proposed and applied for the diagnosis of pneumonia cases and the classification of relevant images in the present study. First, the acquired images of pneumonia cases were divided to determine the zones with target features, and image size was limited to improve the training speed of the network. Meanwhile, reinforcement learning was performed to the input dataset to further improve the training effect of the model. Second, the preprocessed images were inputted into the improved EfficientNet-B₄ model. The depth and width of the model, as well as the resolution of the input images, were determined by optimizing the combination coefficient. On this basis, the model was scaled, and then its ability in extracting the features of deep-layer images was strengthened by introducing the attention mechanism. Third, the learning rate was adjusted by using the Adaptive Momentum (ADAM), and the training efficiency of the model was accelerated. Finally, the test set was inputted into the trained model. Results demonstrate that the improved model could detect 98% of patients with pneumonia and 97% of patients without pneumonia. The accuracy rate, precision rate, and sensitivity of the model were generally improved. Lastly, the training and test results of VGGNet, SqueezeNet-Elus, SqueezeNet-Relu, and the improved EfficientNet-B₄ models were compared and evaluated. The improved EfficientNet-B₄ model achieved the highest comprehensive accuracy rate, reaching 92.95%. The proposed method provides some references to the application of the CNN model in image classification and recognition.

Keywords: Attention Mechanism, Convolutional Neural Network(CNN), Data Augmentation, EfficientNet model

1. Introduction

With the spreading of the Corona Virus Disease 2019 (COVID-19) [1] around the world as a public health emergency with a wide range of infection, rapid transmission and difficult control, the accurate and quick recognition and diagnosis of COVID-19 at a low cost become the key and difficult problem to solve during the medical crisis.

The diagnosis of pneumonia shows that X-ray image is an important diagnosis standard, and the reading efficiency directly determines the detection rate of lesion and working efficiency. Applications of artificial intelligence technology in disease screening can optimize some medical processes under the assistance of a computer [2]-[5]. Abundant information can be obtained from fast analysis as long as data are collected and the trained model is inputted. Later, more accurate auxiliary diagnosis is realized by combining with the background data for comparison, thus enabling to achieve a more accurate diagnosis and prevent danger from happening.

As the first deep neural network that is trained successfully, Convolutional Neural Network (CNN) [6] learns feature expression automatically through large-scaled sample data, which can effectively reduce network

complexity and simulate the complicated hierarchical cognitive laws of the human brain. Recently, CNN has been widely studied in the field of computer vision, and some research progresses have been achieved in image classification.

Especially in 2006, many improved models have been gradually developed after proposing deep learning theory, and they have been applied to the diagnosis and treatment of many diseases in the medical field. However, these models still have some problems, such as they contain abundant parameters to cause excessive training time, low accuracy, and overfitting of the model.

EfficientNet [7] is a new generation of high-performance CNN model that is designed based on Neural Architecture Search (NAS) by Google in 2019. A mainnet was designed through architecture search. Moreover, the depth and width of the model, as well as the resolution of input images, were scaled by a simple yet highly effective compound coefficient, and then the ConvNet model was expanded. Finally, EfficientNet B₀~B₇ models were obtained, thereby improving the operation efficiency of the whole network. It can reach relatively high accuracy at extremely low calculation workloads. This can constrain the target computing resources while assuring the model training efficiency. That is, given the limited computing resources, it can achieve higher image classification accuracy.

Studies on the intelligent intervention of COVID-19 based on image-assisted diagnosis are in the stage of

*E-mail address: linali@jicct.edu.cn

ISSN: 1791-2377 © 2022 School of Science, IHU. All rights reserved.

doi:10.25103/jestr.156.07

continuous trial and there's a promising application prospect. Based on the classification of the pathological images of pneumonia, the improved EfficientNet-B₄ model was applied to X-ray image analysis in this study. The feasibility and effectiveness of the model in the automatic recognition of pathological images of pneumonia were verified by the experiment, aiming to assist doctors in completing diagnosis and relieving pressure of medical workers.

2. State of the art

CNN algorithm attracts wide attention because of its unique superiority in the field of image recognition. The diagnosis method based on Deep Learning (DL) [8] classifies images mainly through the deep network model, which is trained by Backpropagation (BP) [9]. Common basic classification models include VGGNet [10], Inception [11], ResNet [12], DenseNet [13], CapsNet [14], and EfficientNet. With the support of CNN, scholars at home and abroad applied relevant models to the classification and recognition of medical images and gradually improved the performance and accuracy of the network through continuous development and optimization [15].

After the performance evaluation of VGG19, MobileNetV2, Xception, Inception, and Inception-ResNetV2 on the set that involved 1427 X-ray images, Apostolopoulos, et al. [16] detected patients with COVID-19 automatically in a small dataset through Transfer Learning. According to comparison, VGG is superior to other models. Das et al. [17] proposed a deep learning TLCoV model by combining the transfer learning technology. This approach achieved a precision rate of 96.65% and an accuracy rate of 97.67% in pneumonia cases. Wang et al. [18] proposed a COVID-Net model based on ResNet and analyzed how COVID-Net make reasonable prediction through an interpretable method. Narin et al. [19] presented three different diagnosis models of COVID-19 based on ResNet50, ResNet101, ResNet152, InceptionV3, and Inception-ResNetV2, wherein the ResNet50 model achieved the highest classification performance through fivefold cross-validation. Pathak et al. [20] enhanced the results by using the smooth loss function on ResNet50 and overcame the overfitting problem by using Transfer Learning and tenfold cross-validation method. Montalbo et al. [21] proposed the light-weighted model Fused-DenseNet-Tiny by integrating mirror images of the DenseNet model, which solved the problem of extremely large parameters. Zhang et al. [22] used the CAAD model to perform pre-training on ImageNet to realize the classification and anomaly detection of COVID-19 patients and obtained the model sensitivity, specificity, and AUC of 71.7%, 73.8%, and 83.6%, respectively. Nishio et al. [23] trained the model based on VGG16 by using the X-ray images of patients with COVID-19, hence finally achieving an accuracy rate of 83.6% on the test set, and sensitivity for COVID-19 pneumonia was more than 90%. Bhatt et al. [24] enhanced the image classification effects of VGG19 and EfficientNet-B₃ models by using progressive adjustment and transfer learning technology, which achieved a relatively high accuracy rate.

During the continuous improvement and applications of DL and CNN, artificial intelligence (AI) achieved considerable development progresses in the field of computer vision. However, it still has many problems, such as small dataset, excessive long training time of the model, and ignorance of evaluation indexes. Due to the rapid spreading of COVID-19, many researchers are devoted to

the development of new automatic diagnosis systems for COVID-19. Combining with practical needs, this study attempted to improve the comprehensive performance of models in terms of classification, identification and diagnosis of pneumonia images. The improved EfficientNet-B₄ model was used to extract the features of each image in the data-enhanced training set, and a series of model operations and training, such as convolution, pooling, activation, and full connection, were also utilized. Combining with Adaptive Momentum (ADAM) optimization method, the overfitting effect was minimized, and classification effectiveness was strengthened.

The reminder of this study is organized as follows. Section 3 elaborates the experimental design and implementation process of algorithms. Section 4 introduces the classification and evaluation indexes of the model, analyzes the model training and test results, and compares it with the test results of three other classical models. Section 5 summarizes the study and draws the corresponding conclusion.

3. Methodology

3.1 Experimental design

3.1.1 Datasets

All chest X-ray images were selected from the public data of a group of retrospective studies on Kaggle. The dataset contains 5,856 X-ray images (JPEG) and was classified into 3 documents. A total of 5216 images in the training set, 16 images in the validation set, and 624 images in the test set. Tags of images in each document were divided into 2 classes, namely, Pneumonia or Normal.

To analyze the chest X-ray images, all were screened, and quality was controlled by eliminating all low-quality or unreadable scanning. Next, image-based diagnosis was graded by two professional physicians, and the classification tags of images were recorded. To solve possible scoring errors, a third expert was invited to examine the test set [25].

3.1.2 Blocking treatment of data images

The original images in the dataset have tens of millions of pixels. Thus, analyzing the images directly is difficult due to excessive image size. Moreover, the image background accounts for a high proportion, indicating that the information is excessive. This also increases the model training cost significantly. In this study, Mask R-CNN [26] was used to implement object detection and positioning in images, hence realizing the goal of producing high-quality practical segmentation while detecting targets in images effectively and thereby dividing the clipping images into blocks. Small-sized images were obtained through non-repeated sampling on different sizes of case images according to a fixed size. The gained results were processed by binarization. Images were filtered on the basis of the effective information area in images.

3.1.3 Data Augmentation

Due to limited sample size referred by the dataset and imbalanced sample proportions in the training set (pneumonia cases are almost three times of normal cases) (Fig. 1), a series of transformation operations of images was performed through Data Augmentation [27], thus finishing data expansion in the dataset with some randomness. Without changing the original texture structure, organizational morphology, and other symbolic features of

the images, high learning efficiency with fewer samples and decreased overfitting phenomena of the model were effectively achieved, thus avoiding deviations.

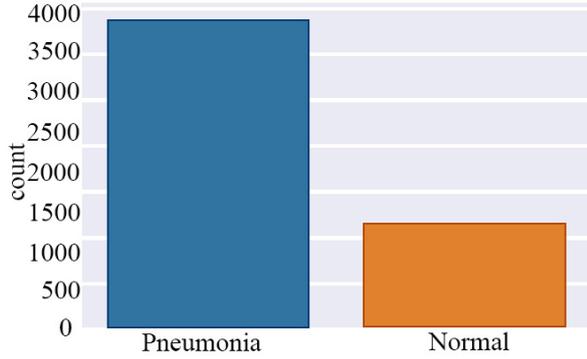


Fig. 1. Histogram of quantities of different image types among samples in the training set

3.1.4 EfficientNet model design

MBCConv is the main component of EfficientNet model [28]. In MBCConv, an optimization block can make models squeezed and activated to realize the effect of adding an attention mechanism into the model. Moreover, a composite coefficient was introduced to control the scaling of model structural parameters. The overall design of the EfficientNet [7] is:

The convolutional layer is expressed as:

$$Y_i = F_i(X_i) \quad (1)$$

where Y_i denotes the output, X_i expresses the tensor of $\langle C, W, H \rangle$ input, and F_i represents various operations of convolution. (C , W , and H are number of image channels, image width, and image height, respectively.) A CNN model can be expressed as a list of a series of convolutional layers.

$$\mathcal{N} = F_k \odot \dots \odot F_3 \odot F_2 \odot F_1(X_1) = \odot_{j=1,2,\dots,k} F_j(X_1) \quad (2)$$

Given that a CNN usually can be expressed as several stages, the convolutional operations in each stage have the same properties. Therefore, a neural network can be expressed as:

$$\mathcal{N} = \odot_{i=1,2,\dots,k} F_i^{L_i}(X \langle H_i, W_i, C_i \rangle) \quad (3)$$

where $F_i^{L_i}$ means that the convolution is repeated by L_i times in Stage i .

A conventional CNN aims to find the optimal network structure F_i , while EfficientNet aims to find the best H_i, W_i, C_i, L_i when F_i is fixed. To improve model accuracy to the maximum extent under limited resources, EfficientNet defined the optimization goal as the following model, which is limited by the condition that

$$\mathcal{N}(d, w, r) = \odot_{i=1,2,\dots,k} \hat{F}_i^{d \times L_i}(X \langle r \times H_i, r \times W_i, w \times C_i \rangle) \quad (4)$$

where d , w , and r are the depth and width of the network model, as well as the resolution of the input images. To solve simultaneous changes in these three parameters, a

composite coefficient (ϕ) was set to control d , w , and r uniformly.

$$\text{depth: } d = \alpha^\phi \quad (5)$$

$$\text{width: } w = \beta^\phi \quad (6)$$

$$\text{resolution: } r = \gamma^\phi \quad (7)$$

The limitations are $\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$ and $\alpha \geq 1, \beta \geq 1, \gamma \geq 1$. The composite coefficient (ϕ) is a user-oriented hyper-parameter, and its value is corresponding to resource consumptions.

3.2 Implementation process of the algorithm

3.2.1 Model preprocessing

The used dataset contains 5216 images in the training set, 16 images in the validation set, and 624 images in the test set. Considering that the images in the validation set are few and develop the reverse control effect of the validation set on model training is difficult, 10% of the images were chosen randomly from the training set and divided into the validation set. The whole dataset was divided into layers. According to the fivefold cross-validation method, the training and validation datasets were divided into 90% training and 10% validation. Finally, 4694 images in the training set, 522 images in the validation set, and 624 images in the test set were obtained.

Images were transformed into images with 64×64 pixels through preprocessing. After color normalization, images were changed images with 30° random rotation, 20% random scaling and 10% random horizontal displacement, 10% random vertical displacement, and random horizontal flip through data enhancement, thus getting the dataset after expansion.

3.2.2 Model construction

Based on the basic framework of EfficientNet-B₄ model in Section 3.1, the model was improved. Specifically, the stem layer was used in preliminary feature extraction. The block layer was used to apply the attention mechanism on the results of depth separable convolution. The global layer was applied to reduce the parameter quantity of the model, which is convenient for training and increases speed. Additionally, two Fully Connected (FC) layers with ReLU activation functions were added in this study, and model overfitting was avoided by setting 30% of loss rate. Finally, a FC layer was applied for binary classification, and the target image classification system was established using the sigmoid activation function. The total parameter number was 17,790,681, which included 125,200 non-training parameters and 17,665,481 training parameters.

3.2.3 Model training and test

The 4,694 images in the training set were inputted in 147 times, 32 non-repeated images at random each time, and a total of 15 epochs were trained. Based on the constructed network model, the cross entropy loss function was applied to train the model during the whole optimization process. Learning rate was updated dynamically by combining with ReduceLROnPlateau to update the learning rate dynamically and use the ADAM optimizer to enhance performance. BP was implemented according to the predicted results and the

loss function of the training set. Parameters were adjusted in the reverse manner layer by layer, and the results were assessed through fivefold cross-validation. The optimal model was recognized and used to verify the test. After finishing model training, the improved and finally optimized EfficientNet-B4 model was applied to test the dataset. The tag prediction results of each image were obtained through the network model to verify its performance.

4. Result Analysis and Discussion

4.1 Model evaluation indexes

To verify the effectiveness of the algorithm, the tags of the output results after the test set was inputted into the model were compared with real tags, thereby obtaining various evaluation indexes of the model in the pathological image classification of pneumonia. The evaluation indexes [29] for binary classification included: true positive (TP), false positive (FP), false negative (FN), and true negative (TN).

Accuracy rate reflects the proportion of accurately classified prediction samples in the total samples.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

Precision rate reflects the probability of actually positive samples in all samples that are predicted positive.

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

Recall rate reflects the probability of predicted positive samples in actual positive samples.

$$recall = \frac{TP}{TP + FN} \tag{10}$$

F_1 value refers to the harmonic mean of precision rate and recall rate.

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{11}$$

The accuracy rate, precision rate, recall rate, and F_1 value are the main parameters that determine the performance of the model.

4.2 Model training results and analysis

Training was performed by using the basic model in Section 3.2 and the improved model. The model training results before and after improvement are listed in Tables 1 and 2, respectively.

Table 1. Model training results before improvement

Epoch	Time	Loss	Accuracy	Recall	Precision	Val Loss	Val Accuracy	Val Recall	Val precision
1	480 s	0.2922	0.8796	0.7481	0.6301	0.7728	0.8314	0.7894	0.7464
2	515 s	0.145	0.947	0.8092	0.7871	0.0456	0.8084	0.8199	0.8166
3	524 s	0.1221	0.9572	0.8271	0.8342	0.1569	0.931	0.8421	0.8485
4	462 s	0.1091	0.9629	0.8521	0.8585	0.101	0.9617	0.8628	0.868
5	497 s	0.1206	0.954	0.8673	0.8712	0.1754	0.9521	0.8734	0.8764
6	521 s	0.0754	0.9717	0.8807	0.8805	0.0022	0.9713	0.8865	0.8863
7	512 s	0.0637	0.9787	0.892	0.892	0.0661	0.9674	0.8961	0.8966
8	536 s	0.0623	0.9806	0.9004	0.901	0.0336	0.9751	0.9044	0.9046
9	684 s	0.0542	0.9815	0.9077	0.908	0.0059	0.9789	0.911	0.9112
10	700 s	0.0526	0.9823	0.9138	0.9142	0.0451	0.9693	0.9161	0.9165
11	733 s	0.0503	0.9813	0.9182	0.9188	0.1943	0.9789	0.9203	0.9209
12	687 s	0.0478	0.983	0.9222	0.9229	0.0116	0.9808	0.9241	0.9246
13	729 s	0.0393	0.9853	0.9262	0.9265	0.0036	0.977	0.9279	0.928
14	689 s	0.0398	0.9851	0.9293	0.9293	0.0626	0.977	0.931	0.9309
15	613 s	0.04	0.9836	0.9323	0.932	0.0082	0.9732	0.9334	0.9332

Table 2. Model training results after improvement

Epoch	Time	Loss	Accuracy	Recall	Precision	Val Loss	Val Accuracy	Val Recall	Val precision
1	437 s	0.2586	0.892	0.7163	0.6646	1.5262	0.8525	0.7739	0.7942
2	427 s	0.1666	0.9382	0.7916	0.8185	0.187	0.9444	0.8177	0.8381
3	419 s	0.136	0.9519	0.8333	0.8532	0.0328	0.9521	0.8477	0.8615
4	392 s	0.1176	0.9582	0.8579	0.8677	0.002	0.9636	0.8677	0.8745
5	358 s	0.1096	0.961	0.8745	0.8798	0.0569	0.9693	0.8802	0.8843
6	422 s	0.0957	0.967	0.8832	0.8898	0.0011	0.9789	0.8878	0.8949
7	331 s	0.09	0.9683	0.8916	0.8981	0.0445	0.8602	0.8956	0.8978
8	611 s	0.0933	0.9687	0.8985	0.898	0.0094	0.977	0.9016	0.9008
9	630 s	0.0898	0.9702	0.9049	0.9033	0.2487	0.9157	0.9069	0.9036
10	677 s	0.0723	0.9704	0.9096	0.9039	0.0295	0.9808	0.9113	0.9062
11	623 s	0.0533	0.9817	0.9139	0.9085	0.0006	0.9866	0.9167	0.9111
12	301 s	0.055	0.9817	0.9189	0.9134	0.0006	0.9789	0.9206	0.9156
13	301 s	0.0515	0.9819	0.9224	0.9176	0.0182	0.9732	0.9237	0.9194
14	301 s	0.0498	0.9842	0.9253	0.9214	0.0254	0.9808	0.9266	0.9232
15	299 s	0.0382	0.9864	0.9283	0.925	0.0003	0.9847	0.9296	0.9265

This model experienced two stages of result verification. One is the fivefold cross-validation, in which the repeated datasets were applied for training and test. The other is to

verify the model performance by using an independent validation set. According to the comparison of data in Tables 1 and 2, the performance of the improved model is enhanced.

4.3 Model test results and analysis

The test set data were inputted into the model trained by the EfficientNet-B₄ network before and after the improvement, thereby getting prediction tags of the model to images in the test set. Through comparative computation with actual tags, the confusion matrices of the prediction results of the model before and after the improvement based on the test set are shown in Figs.2 and 3. The model test results before and after the improvement are listed in Tables 3 and 4.

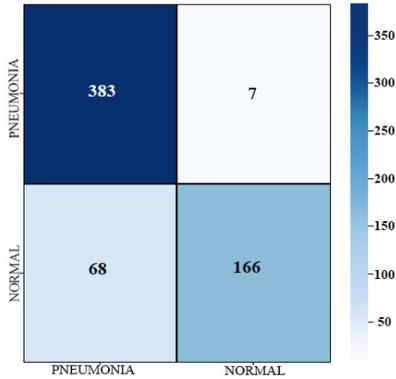


Fig. 2. Confusion matrix of the prediction results of the model test set before improvement

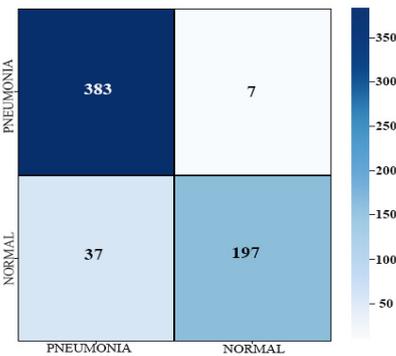


Fig. 3. Confusion matrix of the prediction results of the model test set after improvement

Table 3. Model test results before the improvement

	Precision	Recall	f1-score	Support
Pneumonia	0.85	0.98	0.91	390
normal	0.96	0.71	0.82	234
accuracy	—	—	0.88	624
macro avg	0.9	0.85	0.86	624
weighted avg	0.89	0.88	0.88	624

Table 4. Model test results after the improvement

	Precision	Recall	f1-score	Support
Pneumonia	0.91	0.98	0.95	390
normal	0.97	0.84	0.9	234
accuracy	—	—	0.93	624
macro avg	0.94	0.91	0.92	624
weighted avg	0.93	0.93	0.93	624

According to performance evaluation results and model test results above, the recall rate of the model before improvement for patients with pneumonia reaches 98%, the precision rate for normal people reaches 96%, and the comprehensive accuracy rate is 88%. The recall rate of the model after improvement for patients with pneumonia reaches 98%, the precision rate for normal people reaches 97%, and the comprehensive accuracy rate is 93%. In a word, the improved model has better abilities in classifying and identifying the X-ray images of pneumonia.

4.4 Model calculation results and comparison

To verify the classification and identification effects of the improved model to pneumonia images, the feature extraction network models VGGNet, SqueezeNet-Elus, SqueezeNet-Relu and EfficientNet-B₄ were chosen. The accuracy rate and training time of the training set were compared, thereby obtaining the accuracy rates of the four models on the training set Fig. 4.

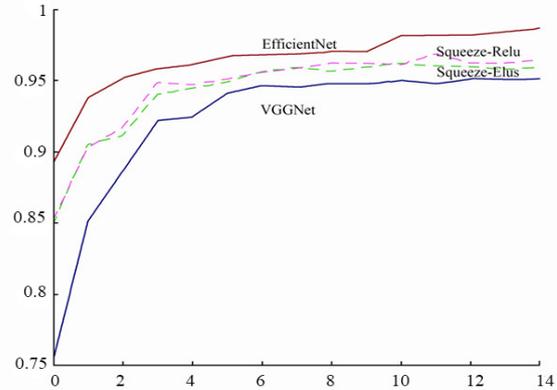


Fig. 4. Accuracy rates of the four models on the training set

The accuracy rates and training time of the four models are listed in Table 5.

Table 5. Comparison of accuracy rates and training time of the four models

Model Name	Accuracy	Training Time
EfficientNetB4	92.95%	1:51:49 s
VGGNet	90.54%	8:37:48 s
SqueezeNet-Relu	88.94%	15:18 s
SqueezeNet-Elus	91.51%	14:22 s

According to the comparison of data, although the improved EfficientNet-B₄ model fails to achieve the fastest network training, the training efficiency and comprehensive performance are better.

5. Conclusions

This study combines CNN with medical image classification and recognition to alleviate the impact of COVID-19 on the medical system, and the EfficientNet-B₄ basic model was improved for the automatic classification of pneumonia images. Some conclusions could be drawn as follows:

(1) The comprehensive performance of the improved EfficientNet-B₄ model is increased by 5% compared with that before the improvement. Moreover, comprehensive evaluation indexes, including accuracy rate, precision rate, recall rate, and F1 value, are improved to some extent. It achieves high-efficiency learning in complicated images with relatively few calculation resources.

(2) The recall rate of the improved EfficientNet-B₄ model for patients with pneumonia reaches 98%, and its precision rate to normal people reaches 97%. The improved EfficientNet-B₄ model has higher sensitivity compared with the model before improvement, indicating that this model has considerable application potentials in the diagnosis and assessment of diseases.

(3) The improved EfficientNet-B₄ model is compared with the training and test results of three other classical models. Its findings show that although the improved

EfficientNet-B₄ model has no absolute advantages in training time, the iteration convergence efficiency and accuracy rate at training are higher than those of the three other models. This result reveals that the improved EfficientNet-B₄ model is feasible in classifying and identifying the pathological images of pneumonia to some extent.

To sum up, the automatic image classification algorithm based on the improved EfficientNet-B₄ model has some references to medical diagnosis, which requires image identification with human eyes. Considering that the used datasets come from only one hospital, it needs more sample data support to increase the universality of the model.

Additionally, performance of the improved EfficientNet-B₄ model is determined on the basis of the weights of the pre-trained models to a very large extent. Hence, with the continuous optimization of the CNN algorithm, the performance of the model will be strengthened during test on dataset which has the more advanced AI technology and great system structure. It can play a greater role in developing a biomedical imaging exploration and analysis system.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License.



References

1. Tang, S., Brady, M., Mildenhall, J., Rolfe, U., Bowles, A., Morgan, Kirsty., "The new coronavirus disease (COVID-19): what do we know so far?". *Journal of Paramedic Practice*, 12(5), 2020, pp.193-195.
2. Zhang, S., Gong, Y. H., Wang, J. J., "The Development of Deep Convolution Neural Network and its applications on computer vision". *Chinese Journal of Computers*, 42(3), 2019, pp.453-482.
3. Zhang, R. R., Pu, L. H., Zhang, W., Qi, X. Y., Yang, F., "Application status and prospects of artificial intelligence in medical imaging education". *Modern Preventive Medicine*, 46(24), 2019, pp.4527-4529.
4. Li, D., Wang, Y. F., Li, Y. X., Huang, W. H., "Application research of artificial intelligence in medical image diagnosis". *Chinese Journal of Clinical Anatomy*, 38(1), 2020, pp.110-113.
5. Qiu, C. H., Huang, C. F., Xia, S. R., Kong, D. X., "Application review of artificial intelligence in medical images aided diagnosis". *Space Medicine and Medical Engineering*, 34(5), 2021, pp.407-414.
6. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., "Gradient-based learning applied to document recognition". *Proceedings of the IEEE*, 86(11), 1998, pp.2278-2324.
7. Tan, M. X., Le, Q. V., "EfficientNet: Rethinking model scaling for Convolutional Neural Networks". *International Conference on Machine Learning*, Long Beach, California: PMLR 97, 2019, pp.6105-6114.
8. LeCun, Y., Bengio, Y., Hinton, G., "Deep learning". *Nature*, 521(14539), 2015, pp.436-444.
9. Ryszard Tadeusiewicz, "Sieci neuronowe". [EB/OL]. Retrieved from http://galaxy.agh.edu.pl/~vlsi/AI/backp_t_en/backprop.html, 2004-06-09/2022-10-20.
10. Simonyan, K., Zisserman, A., "Very Deep Convolutional Networks for Large-Scale image recognition". *International Conference on Learning Representations*, San Diego, USA: ICLR, 2015, pp.1-14.
11. Szegedy, C., Liu, W., Jia, Y. Q., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., "Going deeper with convolutions". *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA: IEEE, 2015, pp.1-9.
12. He, K. M., Zhang, X. Y., Ren, S. Q., Sun, J., "Deep residual learning for image recognition". *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, USA: IEEE, 2016, pp.770-778.
13. Huang, G., Liu, Z., Laurens, VDM., Weinberger, K. Q., "Densely connected convolutional networks". *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, USA: IEEE, 2017, pp.4700-4708.
14. Sabour, S., Frosst, N., Hinton, G. E., "Dynamic routing between capsules". *Conference on Neural Information Processing Systems*, Long Beach, USA: NIPS, 2017, pp.1-11.
15. Ma, J. L., Qiu, S., Ma, Z. P., Chen, Y., "Review of Deep Learning diagnostic methods for COVID-19". *Computer Engineering and Applications*, 58(12), 2022, pp.51-65.
16. Apostolopoulos, I. D., Mpesiana, T. A., "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks". *Physical and Engineering Sciences in Medicine*, 43, 2020, pp.635-640.
17. Das, A. K., Kalam, S., Kumar, C., Sinha, D., "TLCoV-an automated COVID-19 screening model using transfer learning from chest X-ray images". *Chaos Solitons and Fractals*. 144, 2021, 110713.
18. Wang, L. D., Lin, Z. Q., Wong, A., "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images". *Scientific Reports*, 10, 2020, 19549.
19. Narin, A., Kaya, C., Pamuk, Z., "Automatic detection of Coronavirus Disease (COVID-19) using X-ray images and Deep Convolutional Neural Networks". *Pattern Analysis and Applications*, 24(3), 2021, pp.1207-1220.
20. Pathak, Y., Shukla, P. K., Tiwari, A., Stalin, S., Singh, S., "Deep transfer learning based classification model for COVID-19 disease". *Innovation and Research in BioMedical Engineering*, 43(2), 2022, pp.87-92.
21. Montalbo, F. J. P., "Diagnosing COVID-19 chest X-rays with a lightweight truncated DenseNet with partial layer freezing and feature fusion". *Biomedical Signal Processing and Control*, 68, 2021, 102583.
22. Zhang, J.P., Xie, Y.T., Pang, G. S., Liao, Z. B., Verjans, J., Li, W. X., Sun, Z. J., He, J., Li, Y., Shen, C. H., Xia, Y., "Viral Pneumonia screening on chest X-ray images using confidence aware anomaly detection". [EB/OL]. Retrieved from <https://arxiv.org/abs/2003.12338.pdf>, 2020-12-02/2022-10-20.
23. Nishio, M., Noguchi, S., Matsuo, H., Murakami, T., "Automatic classification between COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy on chest X-ray image: combination of data augmentation methods in a small dataset". *Scientific Reports*, 10, 2020, 17532.
24. Bhatt, A., Ganatra, A., Kotecha, K., "COVID-19 pulmonary consolidations detection in chest X-ray using progressive resizing and transfer learning techniques". *Heliyon*, 7, 2021, e07211.
25. Tang, J. P., Zhou, X. F., He, X., Chu, X. W., Li, S. F., Chang, Q. R., Zhang, J. Y., "Survey of studies of COVID-19 diagnosis based on Deep Learning". *Computer Engineering*, 47(05), 2021, pp.1-15.
26. He, K. M., Gkioxari, G., Dollár, P., Girshick R., "Mask R-CNN". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2), 2020, pp.386-397.
27. Dosovitskiy, A., Fischer, P., Springenberg, J. T., Riedmiller, M., Brox, Thomas., "Discriminative unsupervised feature learning with exemplar convolutional neural networks". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(9), 2016, pp.1734-1747.
28. Dai, Z, H., Liu, H, X., Le Q. V., Tan, M.X., "Coatnet: Marrying convolution and attention for all data sizes". *Advances in Neural Information Processing Systems*, 34, 2021, pp.3965-3977.
29. Mujtaba, G., Shuib, L., Idris, N., Hoo, W. L., Raj, R. G., Khowaja, K., Shaikh, K., Nwekea, H. F., "Clinical text classification research trends: Systematic literature review and open issues". *Expert Systems with Applications*, 12(116), 2019, pp.494-520.