SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES
VOL. 4, NO. 1, APRIL 2021
DOI: 10.35377/saucis.04.01.787030

Research Article

Detection of Pneumonia with a Novel CNN-based Approach

Ebru Erdem[1], Tolga Aydın[2]
[1]Corresponding Author; Ataturk University, Department of Computer Engineering, 25240, Erzurum, Turkey; ebruerdem@atauni.edu.tr;
[2]Ataturk University, Department of Computer Engineering, 25240, Erzurum, Turkey; atolga@atauni.edu.tr;

Received 28 August 2020; Revised 25 December 2020; Accepted 28 December 2020; Published online 05 February 2021

Abstract

Pneumonia is a seasonal infectious lung tissue inflammatory disease. According to the World Health Organization (WHO), early diagnosis of the disease reduces the risk of its transmission and death. Various deep learning and machine learning algorithms were used for pneumonia detection. This study aims to analyze the lung images and diagnose pneumonia disease by employing deep learning approaches. We have suggested a novel deep learning framework for the detection of pneumonia in lung. A comparison was made between the proposed new deep learning model and pre-trained deep learning models. 88.62% accuracy rate has been obtained from the proposed deep learning structure. It was observed that by utilizing the new deep neural network developed, the accuracy results of VGG16 (88.78%) and VGG19 (88.30%), which are among the popular deep learning architectures, can be approximated. The test results show that our proposed model has a better recall value (97.43%) (VGG16 (93.33%) and VGG19 (96.92%)), and a better F1-Score (91.45%) (VGG16 (91.22%) and VGG19 (91.19%)).

Keywords: Pneumonia, CNN, VGG16, VGG19

1. Introduction

Lung is a vital organ, and lung abnormalities are highly risky among people. An example of a risk-bearing condition is lung pneumonia. Lung pneumonia (pneumonia) is an inflammation of the lung tissue by various microorganisms. Pneumonia can be detected from chest x-ray images. However, this practice requires highly qualified radiologists. Since there is the risk of confusing pneumonia with other lung diseases, pneumonia detection has turned out to a time consuming process. Computer aided systems (CAD) are being developed to overcome these problems. Thanks to CAD, early detection enables effective treatment and reduces the risk.

With the recently developed deep learning techniques, early diagnosis can be made and the progression of the disease can be prevented. For this reason, the subject to be studied and the technique to be applied is a prominent phenomenon in current medical fields. Automated detection studies of pneumonia with machine learning or deep learning solutions are found in the literature. The detection of the pneumonia was made with AI techniques and results have been obtained as follows;

Among these studies, in the study conducted by Ilyas Sirazitdinov et al, lung pneumonia was detected using images in the "Chest X-ray" database. For this purpose, two convolutional neural networks, Mask R-CNN [1] and RetinaNet [2], were used. The proposed solution was tested on 26,684 image sets from Kaggle Pneumonia Detection Struggle. Good results were obtained for the diagnosis of automatic pneumonia with 79.3% recall [3]. In the study conducted by Enes Ayan et al., CNN [4] [5] networks' performance in pneumonia disease detection was investigated. For this purpose, VGG16 [6] and Xception [7] were compared. VGG16 has been observed to exceed the Xception's accuracy rate of 87%. The Xception network was observed to be more successful than the detection of the VGG16 network. VGG16's success in detecting normal situations was higher than Xception [8]. Abhir Bhandary et al. a deep learning approach that identifies lung abnormalities on chest x-ray images is presented. Performance comparisons were performed by using pre-trained deep learning techniques such as AlexNet [9], VGG16, VGG19, ResNet50 [10]. 96.80% success was achieved with the proposed MAN-SVM method [11]. Gaobo Liang et al. proposed a method of transfer learning for the diagnosis of pediatric pneumonia. The proposed network includes 49 convolutional layers and the ReLU activation function, 1 global mean pooling layer and 2 dense layers. 96.7% recall and 92.7% F1-score rates were achieved in the classification of pneumonia of children. Also, Liang et al. applied CNN and VGG16. They achieved an accuracy rate of, recall rate, precision rate, and F1-score rate of 90.5%, 96.7%, 89.1%, and 92.7%, respectively for CNN model. They achieved an accuracy rate of, recall rate, precision rate, and F1-score rate of 74.2%, 95.1%, 72.3%, and 82.2%, respectively for VGG16 [12]. 96.4% success was achieved with the transfer learning method proposed by Vikash Chouhan et al. on the dataset received from the Guangzhou Women's and Children's Medical Center [13]. Pneumonia was diagnosed using a sequential convolutional neural network customized by Raheel Siddiqi. 93.75% success rate was achieved using an 18-layer neural network [14]. Yadav and Jadhav applied CapsNet. They obtained an accuracy rate of 82.50% [15]. Asnaoui et al. proposed a CNN model and achieved the accuracy, recall, precision, and F1-score rates of 84.18%, 78.33%, 94.05%, and 85.66%, respectively [16]. Mittal et al. applied E3CC and VGG16+CapsNet. They achieved an accuracy rate of 81.54% for E3CC and an accuracy rate of 88.30% for VGG16+CapsNet [17]. Jain et al. proposed a CNN model. They achieved the accuracy, recall, and F1-score rates of 85.26%, 94%, and 89%, respectively [18]. Chakraborty et al. proposed a CNN model. They achieved the accuracy, recall, and precision rates of 95.62%, 95%, and 96%, respectively [19].

CNNs are shown as the most recent technique applied, as can be seen from the studies examined. With the classification made using convolutional networks, the detection of the disease can be done with a high level of success.

The contributions of the paper are as follows:

- We proposed a new CNN model in addition to existing models.
- The classifications were conducted on chest x-ray images with the CNN-based pre-trained models. Comparisons were made between the proposed new model and the pre-trained deep learning models.

2. Dataset and Pre-processing

In this study, evaluations were performed on chest x-ray images. The data set used in the relevant study has open access permission. These images were taken from Guangzhou Women and Children’s Medical Center [20]. Data were selected from past cohort subjects (Figure 1). The dataset contains a total of 5856 chest x-ray images 4273 of which having pneumonia and 1583 normal. The dataset is organized in 3 folders (train, validation, test), as given in Table 1 [21]. Each folder contains x-ray images of all patients.

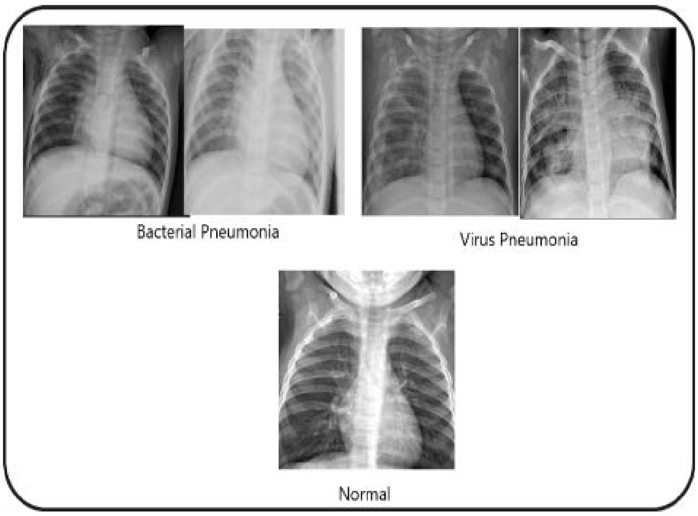

Figure 1 Chest x-ray images

Before the training phase, images with varying width and height values have been reshaped to 224 * 224. Because, pre-trained models accept images at 224x224x3 dimensions. Deep learning algorithms need a huge amount of data to improve performance. Augmentation of data is one of the solutions to get rid of sparse data. Using the train data generator, images were rescaled using 1./255 ratio, zoom which the range for a random zoom that was set to 0.3.

Table 1 Distribution of dataset

	Pneumonia	**Normal**
Training Set	3875	1341
Test Set	390	234
Validation Set	8	8

3. Method

In this study, in addition to the pre-trained deep learning models VGG-16 and VGG-19 applied in the literature, a new deep learning model has been proposed.

3.1. Proposed Model

CNNs contain filters that allow us to collect important information embedded within the image. Since convolutional neural networks do not require pre-treatment or feature extraction on the image, they are run directly on the pixels. The separable convolutional neural networks are a variation of the convolutional neural network. In this network, the filter to be applied on the image is implemented in 2 stages: depth convolution and point convolution. Each color channel is moved on the image during depth convolution phase. Then the resulting images are stacked. Point convolution is applied on the stacked image. At this stage, by increasing the number of channels, the 1x1 filter is passed through every point on the picture. With these two stages, it is aimed to reduce the total number of impacts applied in classical convolutional layers and to save memory and time. Separable convolutional layers, when used instead of convolutional layers in the Inception model, created a new architecture, Xception. This architecture has been shown to perform better in 350 million image classifications than Inception V3 with evolutionary layers [22].

The proposed model consists of 3 convolutional blocks and 3 separable convolutional blocks. 3x3 filters are used in the blocks. Each block is separated by the maximum pooling layer. 2x2 filters are used in the maximum pooling layers. There are 6 convolutional layers in the first 3 blocks and 6 separable convolutional layers in the last 3 blocks. 16, 32, and 64 filters are used in the structure of the first, second, and third convolutional blocks, respectively. 32, 64, and 128 filters are used in the structure of the first, second, and third separable convolutional blocks, respectively (Fig. 2). ReLU function is used as the activation process for hidden layers. To increase the training performance, overfitting has been prevented by using the dropout layer in the model. The first, second, third, fourth and fifth dropout layers have drop rates of 25%, 20%, 90%, 70%, and 50%, respectively. In the last stage, Flatten and 4 fully connected (fc) layers are used. There are 2048 neurons (nodes) in the first fc layer, 1024 neurons in the second fc layer, 512 neurons in the third fc layer, and 1 neuron in the last fc layer. At the last layer of the model, the sigmoid function is implemented. Binary cross-entropy is used as the loss function because of the binary classification. Details of the layers are presented in Table 2.

For the proposed model, images have been reshaped to 150*150 with three channels before the training phase. This is due to obtaining higher accuracy vales when compared to images of 224 * 224 with three channels.

Table 2 Details of layers used in proposed model

Layer	Stride	Filter size	Pool size	Padding	Activation
Input Layer	-	-	-	-	-
Conv1	1	3x3	-	same	relu
Conv2	1	3x3	-	same	relu
MaxPool1	2	-	2x2	-	-
Conv3	1	3x3	-	same	relu
Conv4	1	3x3	-	same	relu
MaxPool2	2	-	2x2	-	-
Conv5	1	3x3	-	same	relu
Conv6	1	3x3	-	same	relu

Table 2 Details of layers used in proposed model (*cont.*)

Layer	Stride	Filter size	Pool size	Padding	Activation
MaxPool3	2	-	2x2	-	-
Dropout1	-	-	-	-	-
SepConv1	1	3x3	-	same	relu
SepConv2	1	3x3	-	same	relu
BatchNorm	-	-	-	-	-
MaxPool4	2	-	2x2	-	-
SepConv3	1	3x3	-	same	relu
SepConv4	1	3x3	-	same	relu
BatchNorm	-	-	-	-	-
MaxPool5	2	-	2x2	-	-
SepConv5	1	3x3	-	same	relu
SepConv6	1	3x3	-	same	relu
BatchNorm	-	-	-	-	-
MaxPool6	2	-	2x2	-	-
Dropout2	-	-	-	-	-
Flatten	-	-	-	-	-
Dense	-	-	-		relu
Dropout3	-	-	-	-	-
Dense	-	-	-		relu
Dropout4	-	-	-	-	-
Dense	-	-	-		relu
Dropout5	-	-	-	-	-
Dense	-	-	-		sigmoid

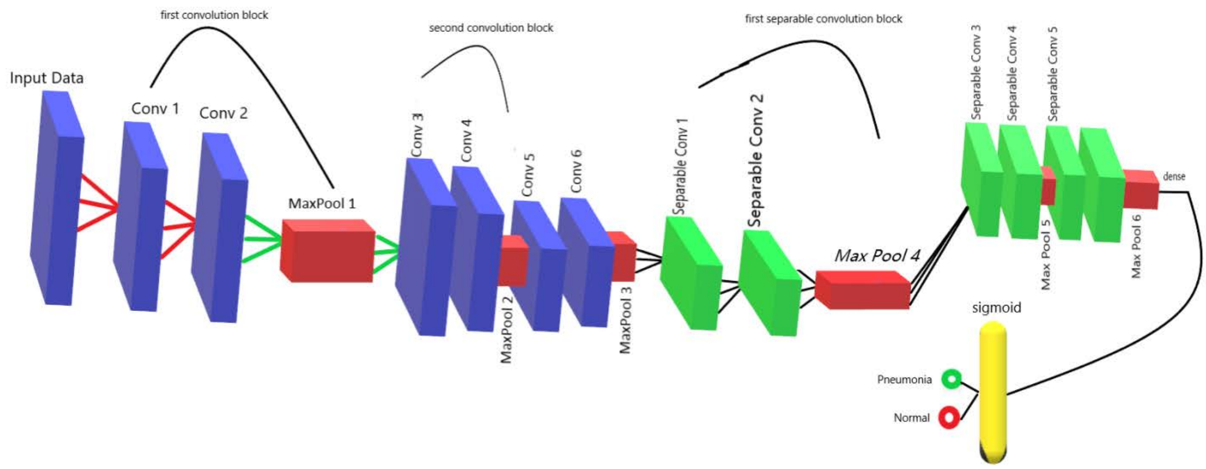

Figure 2 Architecture of the proposed model

3.2. Transfer Learning

With a data set containing a lot of data and classes, the pre-trained models can benefit from the transfer learning method. Using pre-trained models with the transfer learning method can yield successful results, especially in smaller data sets. In this study, deep learning architectures previously trained on larger data sets were used on chest X-ray images with transfer learning.

In 2014, a new model was introduced that improved the success of the AlexNet model. In this model approach, the error rate was decreased by increasing the depth and by reducing the filter size. This model, called VGGNet, has different network structures such as 16-layer VGG-16 and 19-layer VGG-19. The VGG-16 and VGG-19 neural networks have a convolutional layer, a pooling layer, a flatten layer, a dropout layer and a dense layer (Fig. 2.). The images that will enter the models must be 224x224. The weights obtained to recognize the 1000 class problem in the Imagenet competition with these proposed models were used for the problem in this study. The model was customized to solve the 2-class problem by performing fine tuning in the last layer in the study (Fig. 3).

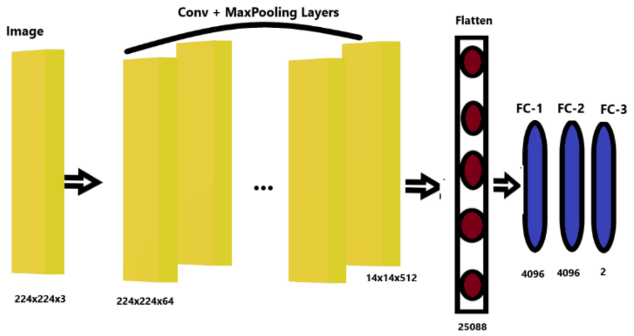

Figure 3 VGG-16 and VGG-19 model architectures

Architectural structures of the proposed and the compared models are given in Table 2. In the new model, additionally, a separable convolution layer is used. The number of convolution layers in VGG-16 and VGG-19 almost corresponds to the number of convolutions and separable convolutions in the new model. Relu, dense and dropout layers are used in all three models. It is remarkable that the number of parameters of the new model is quite low when compared to the number of parameters of the other two models.

Table 2 Details of models

Models	**Convolution Layer Number**	**Separable Convolution Layer Number**	**Relu Layer Number**	**Dense Layer Number**	**Parameter Value**
Proposed Model	6	6	15	4	23,936,513
VGG-16	13	---	15	3	134,264,641
VGG-19	16	---	18	3	139,574,337

4. Results

The deep learning algorithms used in the study were run on the computer with the configurations given in Table 3.

Table 3 Configurations

Memory	245676MB, 24GB
Processor	Intel(R) Core(TM) i7-7700 CPU @3.60GHz (8 CPUs), ~3.6GHz
Graphics Card	Intel(r) Hd Graphics 630, NVIDIA GeForce GT 730
Operating System	Windows 10 Pro 64 Bit (10.0, build 18362)

In the application section, keras, tensorflow and matplotlib libraries were used for all models in Python environment. When the hyper parameters in the model are examined, the number of epoch indicates how many times the data should pass through the model. The model was asked to work in 20 iterations. During the training phase, it was decided to process 32 data (Table 4).

In this study; sensitivity, accuracy, precision and F1-score measurement metrics have been used for performance evaluation. The success rates of the algorithms are given in Table 4. Accuracy is considered as the main performance evaluation metric of the study. Accordingly, while the batch size is 32, the proposed model outperformed VGG-19 and did not exceed VGG-16. VGG-16 gave the best results in terms of precision. On the other hand, for the recall criterion, the proposed model gave the best result. In order to evaluate the success of recall and precision criteria together, the F1-measurement was examined and the proposed model gave the highest result. When the results obtained in Table 4 are analyzed thoroughly, it is observed that the proposed model behaves between the VGG-16 and VGG-19 architectures.

Table 4 Comparative results

Model	**Batch Size**	**Epoch**	**Optimizer**	**Recall (%)**	**Precision (%)**	**F1-Score (%)**	**Test Accuracy (%)**
VGG-19	32	20	ADAM	96.92	86.10	91.19	88.30
Proposed Model	32	20	ADAM	97.43	86.16	91.45	88.62
VGG-16	32	20	ADAM	93.33	89.21	91.22	88.78

Table 5 shows the comparison of the proposed model with the previous studies using the same dataset [20]. Liang and Zheng proposed a CNN-based model to detect the disease. Our model had a higher recall rate (97.43%) when compared to their model (96.7%). Accuracy, precision and F1-Scores were close to the values of their study. Liang and Zheng also studied on VGG16 to detect the disease. In this context, our model has proven to be better in terms of accuracy, precision, recall, and F1-Score. Yadav and Jadhav studied CapsNet to detect the disease and obtained an accuracy of 82.50%. Mittal et al. studied E3CC and VGG16+CapsNet to detect the disease and obtained accuracy rates of 81.54% and 88.30%, respectively. In this context, our model outperformed in terms of accuracy performance criterion when compared to these two methods. Asnaoui et al. proposed a new CNN-based model. Our model was better in terms of accuracy, recall, and F1-Score performance criteria. Ayan and Ünver studied VGG16 and Xception to detect the disease. Our model, again, outperformed in terms of accuracy and recall performance criteria. When compared to another study conducted by Chakraborty et al., we obtained a higher recall value of 97.43% with respect to 95%. Finally, when compared to the study conducted by Jain and et al., our accuracy, recall, and F1-Score performance criteria values were better than their results.

Table 5 Comparison with the previous studies

Article	Year	Method	Accuracy(%)	Recall(%)	Precision(%)	F1-score(%)
Ayan and Ünver [8]	2019	VGG16	87	82	-	-
Ayan and Ünver [8]	2019	Xception	82	85	-	-
Chakraborty et al. [19]	2019	CNN	95.62	95	96	-
Yadav and Jadhav [15]	2019	CapsNet	82.50	-	-	-
Liang and Zheng [12]	2020	CNN	90.5	96.7	89.1	92.7
Liang and Zheng [12]	2020	VGG16	74.2	95.1	72.3	82.2
Asnaoui et al. [16]	2020	CNN	84.18	78.33	94.05	85.66
Mittal et al. [17]	2020	E3CC	81.54	-	-	-
Mittal et al. [17]	2020	VGG16+CapsNet	88.30	-	-	-
Jain et al. [18]	2020	CNN (Model 1)	85.26	94	-	89
Proposed Model	2020	CNN	88.62	97.43	86.16	91.45

"–" denotes that the information is not mentioned in the associated paper.

Considering the processing time as a benchmark criterion, it was observed that the proposed model run in a shorter time. The reason is possibly due to usage of less number of parameters (Table 6).

Table 6 Proposed model, VGG16, and VGG19 running times

	Time to reach result
Proposed Model	2 hours 14 minutes
VGG16	11 hours 7 minutes
VGG19	14 hours 4 minutes

The accuracy and loss graphs of VGG16, VGG19, and the proposed model are shown in Figs. 4, 5, and 6, respectively. When the accuracy and loss graphs are examined, it is observed that the training accuracy values are higher than the test (validation) accuracy values and the test (validation) loss values are higher than the training loss values for the same epoch values.

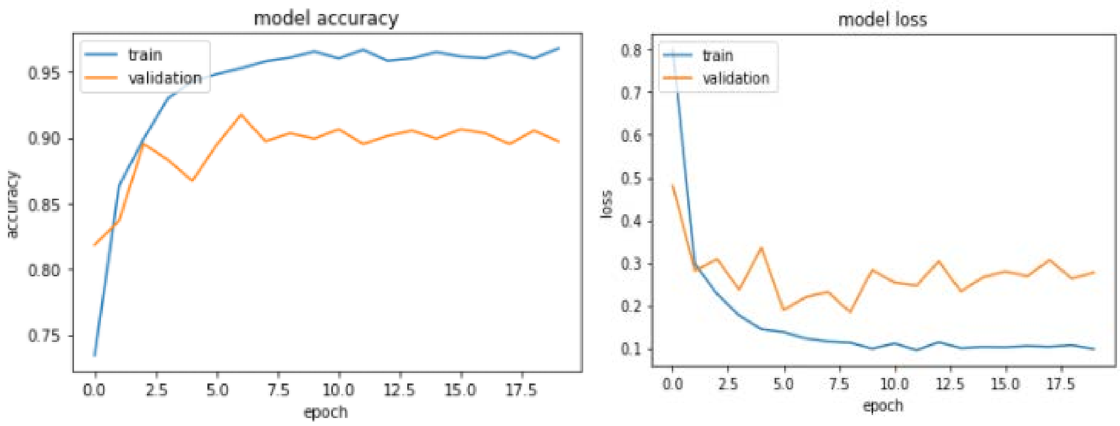

Figure 4 The accuracy and loss graph of VGG16

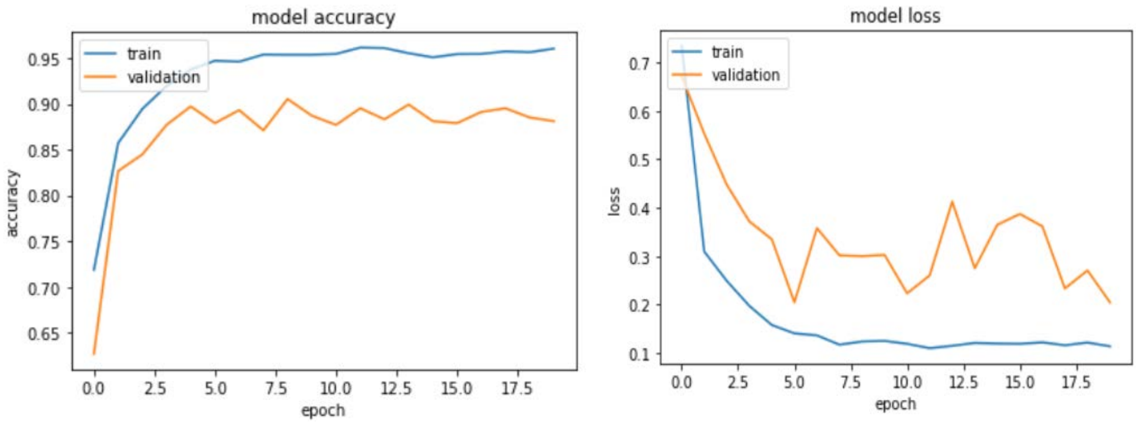

Figure 5 The accuracy and loss graph of VGG19

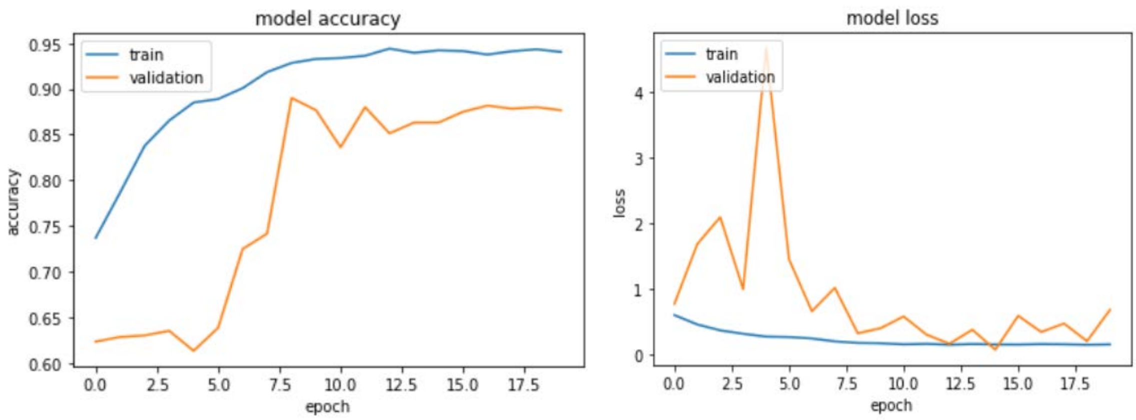

Figure 6 The accuracy and loss graph of the proposed model

5. Conclusions

The study focused on biomedical image interpretation and medical decision in lung pneumonia, which is life-threatening for children and the elderly. As computer aided systems are instrumental in early diagnosis, new models can be proposed in order to reduce risks.

Thanks to their higher performance in solving many related problems, deep learning algorithms have proven to be stronger than artificial intelligence algorithms. Therefore, this study utilized pre-trained deep learning models and a new deep learning model has been developed. Transfer learning has been carried out through pre-trained models. A comparison of these models with the proposed model has been made. It has been observed that the developed model gives better results than VGG-19 and worse results than VGG-16. We think that the proposed model is open to development. Due to this situation, it is concluded that more experiments should be conducted with different hyper parameters of the model. Performance may possibly be improved by making changes in learning ratio, optimizer method and batch size.

In the future studies, it is aimed to detect COVID-19 from lung film images by using the advanced version of this model.

References

[1] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask r-cnn," in Proceedings of the IEEE international conference on computer vision., 2017.

[2] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," in Proceedings of the IEEE international conference on computer vision., 2017.

[3] I. Sirazitdinov, M. Kholiavchenko, T. Mustafaev, Y. Yixuan, R. Kuleev, and B. Ibragimov, "Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database," Computers & Electrical Engineering., vol. 78, pp. 388-399, 2019.

[4] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE., vol. 86, no. 11, pp. 2278-2324, 1998.

[5] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," The handbook of brain theory and neural networks., vol. 3361, no. 10, pp. 1995, 1995.

[6] K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409., vol. 1556, 2014.

[7] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition., 2017.

[8] E. Ayan, and H.M. Ünver, "Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning," in 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)., 2019.

[9] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems., 2012.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition., 2016.

[11] A. Bhandary, G.A. Prabhu, V. Rajinikanth, K.P. Thanaraj, S.C. Satapathy, D.E. Robbins, C. Shasky, Y.-D. Zhang, J.M.R.S. Tavares, and N.S.M. Raja, "Deep-learning framework to detect lung abnormality–A study with chest X-Ray and lung CT scan images," Pattern Recognition Letters., vol. 129, pp. 271-278, 2020.

[12] G. Liang and L. Zheng, " A transfer learning method with deep residual network for pediatric pneumonia diagnosis," Computer methods and programs in biomedicine., vol. 187, pp. 104964, 2020.

[13] V. Chouhan, S.K. Singh, A. Khamparia, D. Gupta, P. Tiwari, C. Moreira, R. Damasevicius, and V.H.C. de Albuquerque, "A Novel Transfer Learning Based Approach for Pneumonia Detection in Chest X-ray Images," Applied Sciences., vol. 10, no. 2, pp. 559, 2020.

[14] R. Siddiqi, "Automated Pneumonia Diagnosis using a Customized Sequential Convolutional Neural Network," in Proceedings of the 2019 3rd International Conference on Deep Learning Technologies., 2019.

[15] S.S. Yadav, S.M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," J Big Data., vol. 6, pp. 113, 2019.

[16] K.E. Asnaoui, Y. Chawki, and A. Idri, "Automated methods for detection and classification pneumonia based on x-ray images using deep learning," arXiv preprint arXiv:2003., pp. 14363, 2020.

[17] A. Mittal, D. Kumar, M. Mittal, T. Saba, I. Abunadi, A. Rehman, and S. Roy, "Detecting Pneumonia Using Convolutions and Dynamic Capsule Routing for Chest X-ray Images," Sensors., vol. 20, no. 4, pp. 1068, 2020.

[18] R. Jain, P. Nagrath, G. Kataria, V.S. Kaushik, and D.J. Hemanth, "Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning," Measurement., vol. 165, pp. 108046, 2020.

[19] S. Chakraborty, S. Aich, J.S. Sim, and H.C. Kim, "Detection of pneumonia from chest x-rays using a convolutional neural network architecture," In International Conference on Future Information & Communication Engineering., vol. 11, no. 1, pp. 98-102, 2019.

[20] D. Kermany, K. Zhang, and M. Goldbaum, "Labeled optical coherence tomography (oct) and chest X-ray images for classification," Mendeley data., vol. 2, 2018.

[21] P. Mooney, "Chest X-Ray Images (Pneumonia)," 2017. [Online]. Available: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia. [Accessed: 27-June-2020].

[22] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, "Rethinking the inception architecture for computer vision," in Proceedings of the IEEE conference on computer vision and pattern recognition., 2016.