–

# Detection of pneumonia from pediatric chest X-Ray images by transfer learning

Yasin DEMİR[*1], Özkan BİNGÖL[2]

## ABSTRACT

When pathogens such as viruses, bacteria and fungi attack the lungs, the alveoli fill with inflamed fluid, causing pneumonia. Early diagnosis of this disease, which has fatal outcomes especially in children under 5 years old, is very important in controlling undesirable situations. Chest X-ray images play an important role in the diagnosis of pneumonia. In addition, the fact that the amount of radiation is lower than imaging devices such as tomography and the possibility of being accessible even from rural areas creates an advantage for X-ray devices. However, X-ray images that are not always clear or human conditions such as fatigue and lack of attention can make it difficult for specialists to detect pneumonia. In this study, a transfer learning-based convolutional neural network (CNN) approach is proposed, which can help specialists in the early and accurate diagnosis of pneumonia in children and, classify healthy and diseased individuals through Chest X-ray images. As a result of the study, an original CNN was proposed by adding additional layers to the AlexNet architecture layers and a test accuracy of 96.31% was obtained.

## INTRODUCTION

Pneumonia is a disease that causes fever, cough, difficulty in breathing and it is a deadly disease, especially for children under 5 years old. According to the World Health Organization (WHO), more than 150 million children under the age of 5 are diagnosed with pneumonia every year, and 20 million of them need to be treated by hospitalization [1]. Approximately 80% of pediatric deaths under the age of 2 in developing countries are due to pneumonia [2]

The first clinical symptom of COVID-19, associated with SARS-CoV-2, which has transformed into a pandemic by affecting the whole world, has been reported as pneumonia [3]. This disease, which was first detected in December 2019, affected the whole world and on June 5, 2021, 48,062 people died in Turkey due to this pandemic [4]. The number of people affected by the COVID-19 disease has doubled each week, meaning that each patient infects at least two people [3]. Therefore, early diagnosis and treatment of pneumonia significantly affect the extent of the disease [5].

As the lungs are invaded by pathogens such as bacteria, viruses, and fungi, the alveoli fill with inflammatory fluid. As a result, pneumonia occurs with symptoms such as shortness of breath, cough, and fever [6]. Accurate results can be obtained with Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) tests for pneumonia detection. However, RT-PCR devices have some disadvantages (such as not being applicable in many hospitals, time-consuming test results, RT-PCR kits not being common, and supply shortages). False negatives from the traditional RT-PCR test kit cause the disease to spread rapidly and disrupt the treatment process. Samples taken at the beginning or end of the disease, failure to obtain the sample at the desired level, and delayed delivery to the laboratories may cause the result to be negative.

Chest radiographs play a key role in the diagnosis and treatment of pneumonia [7]. This imaging method, which gives less radiation than computed tomography, can be easily reached in rural areas [8]. Other imaging methods are used in cases where the treatment with X-ray images is insufficient and the disease is still critical [9].

The ambiguity of Chest X-ray images and the similarity with benign diseases can cause conflicts even among radiologists [10]. In pandemic situations such as COVID-19, the insufficient number of radiologists, human defects such as distraction and fatigue that occur with long shifts make it even more difficult to read X-ray images. Rajpukar

et al. (2017) compared the performances of the CNN, which they recommend for pneumonia detection, with the performances of four expert radiologists and reported that the performance of the proposed network was higher [10].

Ayan et al. (2021), the same data set was used, and the data set was balanced by increasing the number of normally labeled images in order to prevent uneven distribution in the data set, and real-time data augmentation methods were used. Comparing the ensemble method developed with ResNet-50, Xception, MobileNet and PNet, it reached AUC of 93.67%, 94.23%, 94.19%, 93.08% and 95.21, respectively [11].

In the study of Bozkurt and Bayram (2021), local binary pattern (LBP) and textural features were extracted using lung x-ray images, k-nearest neighbor (kNN), Naive Bayes (NB), Artificial Neural Network (ANN) and Support Vector Machine ( Different classifiers such as DVM) have been used. The highest accuracy rate of 99.28 was achieved with DVM [12].

In Dey et al. (2021), a customized VGG19 network was proposed by combining conventional feature extraction with deep features obtained using transfer learning method. Compared with different classifiers, they obtained higher accuracy of 95.70% of VGG19 with Random Forest classifier [13].

Kundu et al. (2021) combined the three CNN model with the precision, recall, f1-score and AUC metrics scores to form the weight vector, using the transfer learning method due to the small size of the data set. The proposed ensemble model reached 98.81% and 86.85% accuracy rates and 98.80% and 87.02% accuracy rates in Kermany and RSNA datasets, respectively. [14]

Han et al. (2021) proposed a model combining radiomic features and comparative learning methods to detect pneumonia in their study. It has been observed that the proposed model achieves high accuracy in the RSNA dataset. [15]

In this study, a transfer learning-based CNN was designed to help experts for early and accurate detection of pneumonia. The proposed method provides classification of healthy and pneumonia individuals from Chest X-ray images. AlexNet, ResNet, Inceptionv3, and GoogleNet, which are well known in the literature, were first used for transfer learning. Then, new layers were added on AlexNet, where the best values were obtained, and a high success rate was achieved.

## EXPERIMENTAL STUDY

### Dataset

In the study, a dataset of chest X-ray images taken by Paul Mooney (2018) from Guangzhou Women and Children Medical Center was used [16]. This dataset consists of 5840 images of children aged 1-5 years, labeled as normal and pneumonia by expert radiologists. The training data in the images was divided into two subsets, 1341 healthy images and 3875 diseased images. Likewise, the test data in the images were divided into two subsets as 234 healthy images and 390 diseased images. 80% of the training data was used for training and 20% for validation. The test data is already reserved in the used dataset. In Figure 1, an example image of individuals labeled as healthy and with pneumonia disease in the data set is given. This data set has been used in many academic studies and classification processes have been carried out with different methods [17,18]

**Figure 1.** Chest X-ray images  (a) Healthy (b) Pneumonia

Children's X-ray images are more problematic than adults. In particular, posture disorders and the presence of limbs such as arms and necks make it difficult to train the net and detect the pneumonia that belongs to the problem. Two sample images containing the mentioned problems are shown in Figure 2.
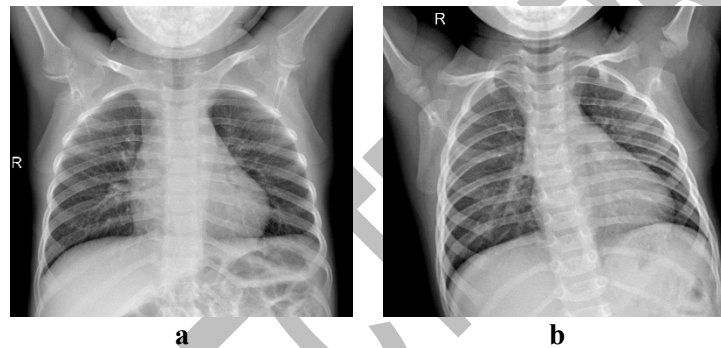


**Figure 2.** Chest X-ray images  (a) Arm and neck limbs (b) Posture disorder

## Transfer Learning and Data Augmentation

The transfer learning approach was preferred due to the relatively low number of images in the data set and the use of hardware resources. Commonly used and pre-trained networks (AlexNet, ResNet derivatives, Inceptionv3, GoogleNet and SqueezeNet) for this learning method have been tested and, the training and test accuracies obtained from these networks are shown in Table 1. Based on the results obtained, AlexNet, which has the highest test accuracy with 82.69% and short training time, was preferred for the following parts of the study. It has been tried to produce a unique network structure by making improvements on this network.

**Table 1.** Performance values of pre-trained networks

| Model | Training Accuracy (%) | Test Accuracy (%) |
|---|---|---|
| Alexnet | 98.47 | 82.69 |
| Googlenet | 98.18 | 82.53 |
| Inceptionv3 | 99.14 | 79.65 |
| Resnet18 | 99.04 | 79.65 |
| Resnet50 | 98.18 | 81.25 |
| Resnet101 | 98.85 | 79.49 |
| Squeezenet | 99.33 | 74.20 |

It can be seen that the test accuracy reaches lower values compared to the training accuracy in table 1. This situation arises when the network encounters the problem of overfitting or when the training data is insufficient.

Methods such as data augmentation, adding drop layers have been applied to reduce the difference between training and test accuracy and also to increase accuracy rates.

Increasing the problem-specific dataset size in deep learning networks generally affects the network positively. The first step is to improve the performance of the network with data augmentation methods. Horizontal scaling and shifting in the x-axis were preferred as data augmentation methods. Horizontal scaling is the operation of enlarging the image relative to the x-axis in the [1.0 1.9] scale range. Shifting in the x-axis is the shifting of pixels in the x-axis in the range of [-25 -15] pixels. As a result of shifting, the remaining empty pixels are filled with black color. Figure 2 shows the normal, horizontally scaled and shifted results of a sample image in the data set.
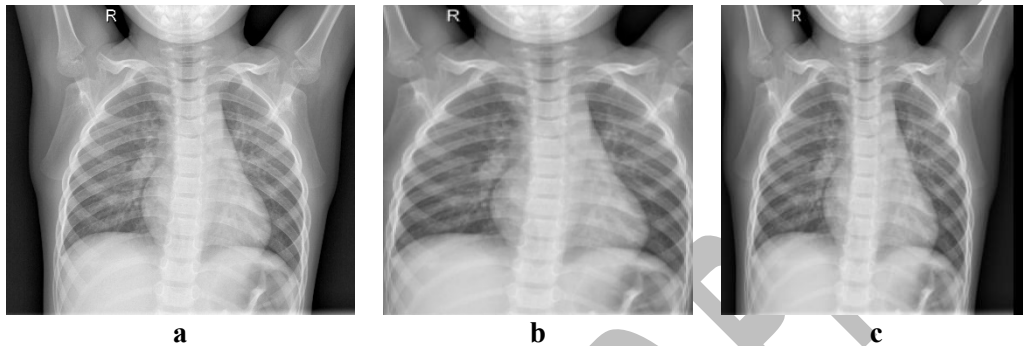


**a**  **b**  **c**

**Figure 3.** Data Augmentation (a) Normal, (b) Horizontal Scaling, (c) X Scroll

**Modified AlexNet**

Although the Alexnet architecture has a good feature extraction capability, it cannot achieve the desired classification success in child X-ray images. Therefore, in our study, the weights of the feature extracting layers of AlexNet were taken with transfer learning, and a classifier arranged in a pyramid structure was added behind the network. Instead of Fully Connected Layers with 1000 Neurons in AlexNet, three Fully Connected Layers with 2048, 1024, 2 neurons have been added, respectively. Here, the number of neurons is the numbers determined as a result of different experiments. In addition, a ReLU activation layer and a Batch Normalization layer have been added for the first two layers. By following the train and test accuracy values in the experiments, the final form of the network was obtained by adding the Dropout layers after the Batch Normalizaton layers. Detailed information about the layers in the proposed Network are shown in Table 2 and Figure 4.

**Table 2.** Modified AlexNet Layers and Parameters

| Layers | Parameters |
|---|---|
| Input Layer | 224x224x3 |
| Convolution Layer | 11x11x96 |
| Max Pooling | 3x3 |
| Convolution Layer | 5x5x256 |
| Max Pooling | 3x3 |
| Convolution Layer | 3x3x384 |
| Convolution Layer | 3x3x384 |
| Convolution Layer | 3x3x384 |
| Max Pooling | 3x3 |
| Fully Connected Layer | 4096 |
| Fully Connected Layer | 4096 |
| *Fully Connected Layer* | *2048* |
| *ReLU* | |
| *Batch Normalization* | |
| *Dropout Layer* | *0.5* |
| *Fully Connected Layer* | *1024* |
| *ReLU* | |
| *Batch Normalization* | |
| *Dropout Layer* | *0.5* |
| *Fully Connected Layer* | *2* |
| *Softmax* | |
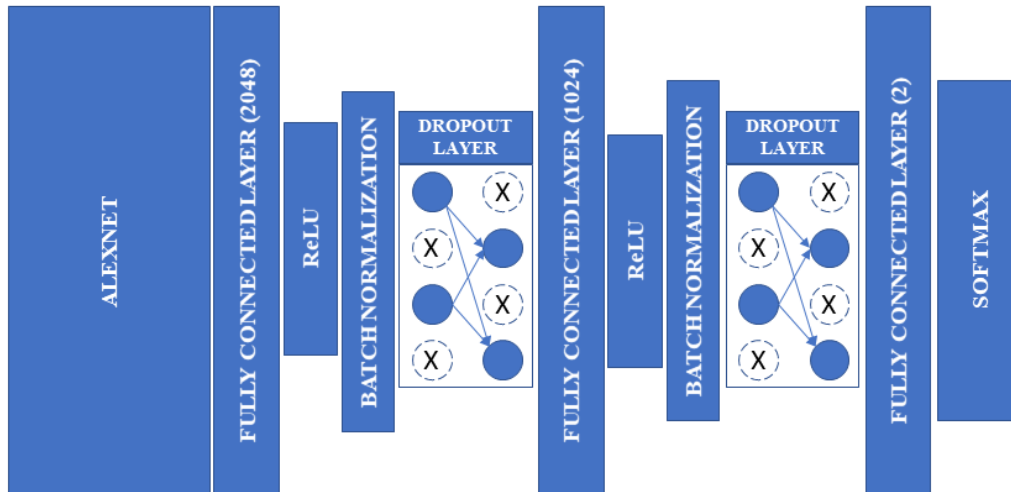


**Figure 4.** The architecture of the proposed network

The network architecture described above has been implemented for the hyperparameters given in Table 3. The mini-batch size was chosen as 128 for efficient use of existing hardware resources. The optimization algorithm for the training of the network was chosen as RMSprop. It was observed that the performance of the network increased

with L2 normalization. The test result obtained after applying data augmentation, drop layer operations and additional layers is given in Table 4.

**Table 3.** Hyperparameters of the proposed network

| Hyperparameters | Values |
|---|---|
| Optimization algorithm | RMSprop |
| Mini batch size | 128 |
| L2 normalization | 0.0004 |
| Epoch | 30 |
| Learning Rate | 0.0005 |
| Validation frequency | 50 |

**Table 4.** The performance of the proposed network

| Additional Layers | Training Accuracy (%) | Test Accuracy (%) |
|---|---|---|
| Alexnet+(F2048+R+B+D.5+F1024+R+B+D.5+F2) | 95.69 | 96.31 |

The proposed network architecture is trained with the training images in the dataset and, the accuracy and loss graphs shown in Figure 5 are obtained. The blue lines in the figure show the accuracy rate, the red lines imply the loss rate, and the black dots show the validation. As can be seen from the figure, the training success of the network starts with 50% accuracy at the beginning, reaches 94% levels in the 5th Epoch and 95% after the 10th Epoch. The training of the network has been tested at different Epoch levels, and the best test accuracy has been achieved at the 30th Epoch. In the network with overfitting problem, the difference between training and test accuracies is significantly improved. When the loss graph is examined, starting from 1% levels, the error rate decreases to 0.15% from the 1st Epoch to the 30th Epoch. Although rare, it is seen that instantaneous losses increase.
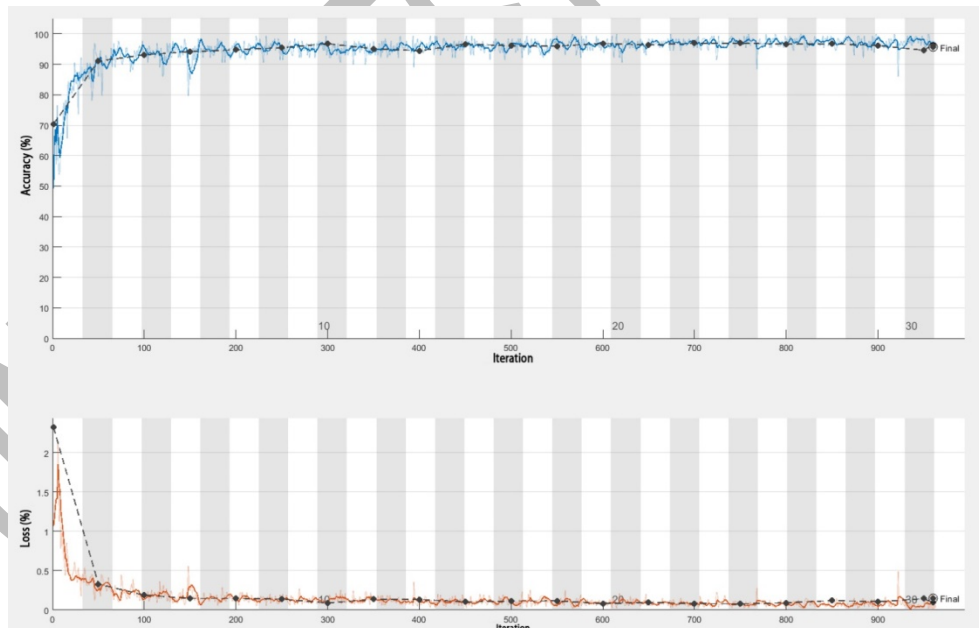


**Figure 5.** Accuracy-Epoch and Loss-Epoch graph of the trained network

To measure the validity of this study, true positive (TP), true negative (TN), false positive (FP), false negative (FN) metrics are used in relation to sensitivity, precision, accuracy and F1 score. Table 5 shows the complexity matrix of the proposed network, with "0" denoting normal and "1" pneumonia. The 218 and 383 in Table 5 show the number of correct classifications by the trained network. Of the 624 test images, 218 were classified as normal, making up 34.9% of all images. Pneumonia was detected in 383 out of 624 test images, which is 61.4% of the entire test set. 7 of the pneumonia images were classified as normal, which appears to be 1.1% of all test data. Of the normal test set, 16 were classified as pneumonia, representing 2.6% of all test data. Of the 225 normal predictions, 96.9% were classified as correct, 3.1% as incorrect, and of the 399 pneumonia predictions, 96.0% were classified as correct and 4.0% as incorrect. Of the 234 normal chest X-ray images, 93.2% were estimated to be normal and 6.8% to be pneumonia. Of 390 pneumonia chest X-ray images, 98.2% were classified as pneumonia and 1.8% as normal. In general, 96.3% of the predictions were classified as correct and 3.7% as incorrect.

**Table 5.** Confusion matrix

| | | Actually | | |
| --- | --- | --- | --- | --- |
| | | Normal | Pneumonia | Total |
| **Predicted** | **Normal** | 218 %34.9 | 7 %1.1 | 225 %36.0 |
| | **Pneumonia** | 16 %2.6 | 383 %61.4 | 399 %64 |
| | **Total** | 234 %37.5 | 390 %62.5 | 624 |

Accuracy shows the closeness of the measured value to the true value and this is expressed in equation 1. The precision shows how many of the class predicted as pneumonia actually belong to pneumonia, and this is shown in equation 2. Sensitivity indicates how many of the class that should be predicted as pneumonia are detected as pneumonia, and this is formulated in equation 3. The F1 score metric in equation 4 is the harmonic mean of precision and sensitivity values. Table 6 shows the performance metrics of the proposed model.

TP, Pneumonia X-ray images are classified as pneumonia.
TN, Normal X-ray images are classified as normal.
FP, Normal chest X-ray images are classified as pneumonia.
FN, Pneumonia X-ray images were classified as normal.

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

$$\text{Precision} = \frac{TP}{TP + FP} \qquad (2)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \qquad (3)$$

$$F1 - \text{Score} = \frac{2 \times \text{precision} \times \text{sensitivity}}{\text{Precision} + \text{Sensitivity}} \qquad (4)$$

**Table 6.** Performance metrics for the recommended network

| Model | Accuracy | Precision | Sensitivity | F1 – Score |
|---|---|---|---|---|
| Proposed Model | 0.9631 | 0.9821 | 0.9599 | 0.9708 |

**CONCLUSION**

Deep learning networks are widely used in the literature for pneumonia detection. Computed tomography, chest x-ray images play a key role in pneumonia. Due to the fact that computed tomography is not common in rural areas and the amount of radiation, chest x-ray images come to the fore. Lung x-ray images detect the presence of pneumonia together with the clinical process early and with very high accuracy. Child lung x-ray images are a challenging task for deep learning networks compared to adult lung x-ray images. Child X-ray images are louder due to shoulder, neck and posture disorders. In the literature, convolutional neural networks are a very successful deep learning model on image recognition. In this study, the ESA model was used to detect pneumonia, which can be fatal for children.

Developing and training the network from scratch takes a lot of time. Transfer learning method can be a solution to small data set problems by shortening the training time compared to the network developed from scratch. The high test accuracy even in a specific area such as lung x-ray images makes the transfer learning method stand out compared to other methods. Another advantage of the transfer learning method is that it uses hardware resources efficiently. Even in very deep layer networks, training can be done quickly.

When comparing the performance of known networks with transfer learning applied, the test accuracy was found to be 82.69% with the highest AlexNet. It has been seen that changes in layers made in known networks can significantly improve the performance of the network. In this study, AlexNet architecture was used to increase the performance of the network and classification was carried out by adding new layers. As a result of the study, an accuracy rate of 96.31% was achieved. The high accuracy of the proposed network shows that it can be recommended as an auxiliary system for early diagnosis of pneumonia and specialists.

The performance comparison of the other approaches in the literature using the data set discussed in this study and the proposed CNN model is presented in Table 7.

**Table 7.** Performance values of the networks in the literature with the same data set

| Author | Method | Accuracy |
|---|---|---|
| Gülgün & Erol [17] | Data Augmented + CNN | %93.40 |
| Toğaçar vd. [7] | AlexNet+SVM | %95.80 |
| Demir & Bingöl, 2021 [19] | Data Augmented + CNN | %94.55 |
| **Proposed** | **Modified AlexNet** | **%96.31** |

**DISCUSSION AND SUGGESTION**

It is predicted that early diagnosis and treatment of pneumonia, which can progress to death in children, will significantly reduce mortality rates. In the detection of pneumonia, X-ray images of the lungs have an important place in the detection and treatment process. However, interpreting chest X-ray images can be a difficult process for radiologists in some cases. Misdiagnosis also increases the risk of transmission of the disease, especially in epidemic pneumonia.

Insufficient data while obtaining medical images and uneven distribution for pairwise comparison is an important problem in the field of health. It has been observed that when deep learning networks are trained, a small data set generally causes the network to memorize, and when the data is increased, this problem disappears and the performance of the network increases. The increase in data size also affects the training time of the network. Applying the data augmentation method to randomly selected data significantly shortened the training time of the network. Updating hyperparameters and network layers requires multiple testing of the network model. This increases the training time

of the network considerably. It was seen that the applied transfer learning method gave successful results in the small data set.

In the field of health, lung x-ray images are used in the diagnosis of many diseases, except pneumonia. With high accuracy achieved; It has been shown that with the help of transfer learning-based ESA, it is possible to study other diseases from X-ray images.

Clinical evaluation and imaging methods are used together in pneumonia disease. In the detection of pneumonia, the absence of other test results such as patient history and blood test, pulse oximetry, sputum test, immune system, apart from lung x-ray images are limited for detection alone. Lung x-ray images are included in the clinical evaluation. In future studies, it is planned to develop a system for meaningful interpretation of other parameters affecting pneumonia disease with deep learning methods.

## REFERENCES

[1] Rudan, I. (2008). Epidemiology and Etiology of Childhood Pneumonia. Bulletin of the World Health Organization, 86(5), 408-416. https://doi.org/10.2471/BLT.07.048769

[2] URL-1. (2020, Nov 21). https://www.ncbi.nlm.nih.gov/books/NBK536940/

[3] Velavan, T. P., & Meyer, C. G. (2020). The COVID-19 epidemic. Tropical Medicine & International Health, 25(3), 278.

[4] URL-2. (2021, Feb 6). https://www.who.int/news-room/fact-sheets/detail/pneumonia

[5] Aydogdu, M., Ozyilmaz, E., Aksoy, H., Gursel, G., & Ekim, N. (2010). Mortality prediction in community-acquired pneumonia requiring mechanical ventilation; values of pneumonia and intensive care unit severity scores. Tuberk Toraks, 58(1), 25-34.

[6] Toikka, P., Virkki, R., Mertsola, J., Ashorn, P., Eskola, J., & Ruuskanen, O. (1999). Bacteremic pneumococcal pneumonia in children. Clinical infectious diseases, 29(3), 568-572.

[7] Toğaçar, M., Ergen, B., & Sertkaya, M. E. (2019). Zatürre Hastalığının Derin Öğrenme Modeli ile Tespiti. Firat University Journal of Engineering, 31(1). World Health Organization.

[8] Gaál, G., Maga, B., & Lukács, A. (2020). Attention u-net based adversarial architectures for chest x-ray lung segmentation. arXiv preprint arXiv:2003.10304.

[9] Rand, T. (1997). A Radiologic approach to Diseases of the Chest. European Journal of Radiology, 3(25), 251-252.

[10] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... & Ng, A. Y. (2017). Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225.

[11] Ayan, E., Karabulut, B., & Ünver, H. M. (2021). Diagnosis of Pediatric Pneumonia with Ensemble of Deep Convolutional Neural Networks in Chest X-ray Images. Arabian Journal for Science and Engineering, 1-17.

[12] Bozkurt, F., & Bayram, E. (2021). Local Binary Pattern Based COVID-19 Detection Method Using Chest X-Ray Images. 2021 29th Signal Processing and Communications Applications Conference (SIU), 1-4.

[13] Dey, N., Zhang, Y.-D, Rajinikanth, V., Pugalenthi, R., & Raja, N. S. M. (2021). Customized VGG19 architecture for pneumonia detection in chest X-rays. Pattern Recognition Letters, 143, 67-74.

[14] Kundu, R., Das, R., Geem, Z. W., Han, G.-T., & Sarkar, R. (2021). Pneumonia detection in chest X-ray images using an ensemble of deep learning models. PloS One, 16(9).

[15] Han, Y., Chen, C., Tewfik, A., Ding, Y., & Peng, Y. (2021). Pneumonia detection on chest x-ray using radiomic features and contrastive learning. 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI 2021), 247-251.

[16] Mooney P. (2018, Mar. 24). Chest X-Ray Images (Pneumonia). https://www.kaggle.com/ paultimothymooney/chest-xray-pneumonia

[17] Gülgün, O. D., & Erol, Prof. Dr. H. (2019). Classification Performance Comparisons of Deep Learning Models In Pneumonia Diagnosis Using Chest X-Ray Images. Turkish Journal of Engineering. https://doi.org/10.31127/tuje.652358

[18] Hidayatullah, R. C., & Violina, S. (2020). Convolutional Neural Network Architecture and Data Augmentation for Pneumonia Classification from Chest X-Rays Images. 5(2), 7.

[19] Demir, Y., & Bingöl, Ö. (2021, June). Evrişimsel Sinir Ağı Yöntemi ile Pediatrik Akciğer Röntgen Görüntülerinden Pnömoni Tespiti. In 2021 29th Signal Processing and Communications Applications Conference (SIU 2021) (pp. 1-4). IEEE.