Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2021) 29: 1615 – 1627

doi:10.3906/elk-2009-1

Determination of Pneumonia in X-ray Chest Images by Using Convolutional Neural Network

Özlem POLAT[1,*], **Zümray DOKUR**[2], **Tamer ÖLMEZ**[2]
[1]Department of Mechatronics Engineering, Faculty of Technology, Sivas Cumhuriyet University, Sivas, Turkey
[2]Department of Electronics and Communication Engineering,
Faculty of Electrical and Electronics Engineering, İstanbul Technical University, İstanbul, Turkey

Received: 01.09.2020 • **Accepted/Published Online:** 19.11.2020 • **Final Version:** 31.05.2021

Abstract: Pneumonia is one of the major diseases that cause a lot of deaths all over the world. Determining pneumonia from chest X-ray (CXR) images is an extremely difficult and important image processing problem. The discrimination of whether pneumonia is of bacterium or virus origin has also become more important during the pandemic. Automatic determination of the presence and origin of pneumonia is crucial for speeding up the treatment process and increasing the patient's survival rate. In this study, a convolutional neural network (CNN) framework is proposed for detection of pneumonia from CXR images. Two different binary CNNs and a triple CNN are used for determining: (*i*) normal or pneumonia, (*ii*) pneumonia of bacterium or virus origin, and (*iii*) normal or bacterial pneumonia or viral pneumonia. In this approach, CNNs are trained with Walsh functions to extract the features from CXR images, and minimum distance classifier instead of a fully connected neural network is employed for classification purpose. Training with Walsh functions maintains the within-class scattering to be low, and between-class scattering to be high. Preferring the minimum distance classifier reduces the number of nodes used and also allows the network to be controlled with fewer hyperparameters. These approaches bring some advantages to the system designed for the classification process: (*i*) easy determination of hyperparameters, (*ii*) achieving higher classification performance, and (*iii*) use of fewer neurons. The proposed small-size CNN model was applied to CXR images from 1- to 5-year-old children provided by the Guangzhou Women's and Children's Medical Center (GWCMC). Three experiments have been conducted to improve the classification performance: (*i*) the effect of different sizes of input images on the performance of the network was investigated, (*ii*) training set was augmented by randomly flipping left to right or down to up, by adding Gaussian noise to the images, by creating negative images randomly, and by changing image brightness randomly (*iii*) instead of RGB CXR images, gray component of the original image and four 2D wavelet images were given as input to the network. In these experiments, no major changes were observed in the classification results obtained by using the proposed CNNs. The proposed method has achieved 100% accuracy for normal or pneumonia, 92% for pneumonia of bacterium or virus origin, and 90% for normal or bacterial pneumonia or viral pneumonia. It is observed that higher classification performances were obtained with approximately five times less parameters compared to the networks that gave the best results in the literature. Thus, the applied CNN model is promising in medicine and can help experts make quick and accurate decisions.

Key words: Deep neural networks, X-ray chest image classification, determination of pneumonia

*Correspondence: ozlem.polat@cumhuriyet.edu.tr

1. Introduction

Pneumonia is infection in one or both lungs. Bacterial, viral, and fungal infections can cause pneumonia. Inflammation in the air sacs of lungs may occur because of these infections. Symptoms of pneumonia may vary among the patients, and may include sputum, fever, chills, shortness of breath, and cough[1].

Chest X-ray (CXR) images are used for diagnosing pneumonia. However, diagnosis of pneumonia from CXR images is not an easy task, even for specialist radiologists. Pneumonia's appearance on CXR images is often indistinguishable, it can be confused with other diseases and it behaves like many other benign abnormalities. These inconsistencies have caused important subjective decisions and variations among radiologists in the diagnosis of pneumonia [1]. Therefore, computerized support systems are needed to help radiologists diagnose pneumonia from CXR images. The discrimination of whether pneumonia is of bacterium or virus origin has also become more important during the COVID-19 pandemic. In this context, there is a need to quickly determine whether there is a viral disease in the CXR images when it is known to have a virus pandemic.

Determining pneumonia from CXR images is an extremely difficult image processing problem. Both textural and anatomical features need to be used together for the classification process. The disease would be tried to be determined by considering textural (tissue) changes on a known template lung image. However, it is quite difficult to determine the boundaries of the lung in the CXR image, as the images are obtained from children in the age group of one to five. Moreover, pneumonia disease affects the boundaries of the lungs. Textural properties would depend on the resolution of the image. Since there are differences in the resolution of the images, textural properties are different even within images of the same class. It is very difficult to design algorithms that automatically overcome these difficulties and classify the images with high performance by using conventional methods. It is observed that large numbers of preprocessing stages have been used to increase the classification performance [2, 3].

Deep neural networks (DNN), popular in recent years, are widely used in CXR image classification to detect pneumonia, and high performance is obtained [4–18]. There are several databases of CXR images. The studies below will be discussed in two categories: studies [4–9] using other databases, and studies [1, 10–18] using the same database with this study. All of these studies have classified CXR images into two broad categories: normal and pneumonia. Only Kermany et al. [18] classified it as bacterial or viral pneumonia, as in this study, in addition to the classification of normal or pneumonia. The following studies in the literature were reviewed according to classification performance, the number of classes, the chest database used, preprocessing stage, and structure of the DNN.

Li et al. [4] combined segmented lung images and original images as training dataset to form a pneumonia detection model. They modified the structure of SE-ResNet to classify images and used the dataset obtained from the Radiological Society of North America (RSNA) pneumonia detection challenge. They classified the CXR images as normal or pneumonia with a success rate of 83.5%. Varshni et al. [5] used DenseNet-169 for feature extraction and support vector machine as classifier. They applied their model on a dataset obtained from ChestX-ray14 which is also publicly available on the Kaggle. Preprocessing stages were not applied, and CXR images were classified into two categories with a success rate of 80.02%. In the study by Li et al. [6], a CXR images dataset (PneuX-rays) was collected at Shenzhen No.2 People's Hospital in China. Preprocessing stages were not used. CNN and fully connected neural network (FCNN) with two hidden layers were applied on the dataset, and the number of network parameters was optimized. A classification performance of 92.79%

[1]NIH: Health Topics Pneumonia. Website: www.nhlbi.nih.gov/health-topics/pneumonia [accessed 17 Jan 2021].

was achieved in two categories. Transfer learning including fine-tuned VGG16, Inception V3, and ResNet-50 networks was used by Aledhari et al. [7], and the results of experiments were compared in terms of classification performance. They classified the CXR images in the ChestX-ray14 database with a succes rate of 75%. Tilve et al. [8] focused on the survey and comparison of lung disease detection by using different methods, and suggested a model for the detection of pneumonia. In the study, first of all, histogram equalization as preprocessing stage and then lung segmentation stage were applied. Improved VGG16 model was tested with ChestX-ray14 image dataset, and CXR images were classified into normal and pneumonia classes with a success rate of 96.2%. O'Quinn et al. [9] developed an algorithm using CNNs to detect visual signals for pneumonia. To classify CXR images into two categories, they used deep learning network AlexNet. They applied the model on the images obtained from the RSNA pneumonia detection challenge. CXR images were classified with 72% success without applying preprocessing.

In the studies [1, 10–18], CXR images were selected among pneumonia patients at the Guangzhou Women and Children's Medical Centre (GWCMC) and CXR images were classified into two categories: normal and pneumonia. Ayan et al. [1] used VGG16 and Xception deep learning models for the diagnosis of pneumonia. They did not use a preprocessing stage, and reported 87% success rate with VGG16. Vijendran et al. [10] used FCNN which is named online sequential extreme learning machines, and classified CXR images with an accuracy of 92.5% without applying any preprocessing stage. Islam et al. [11] suggested a compressed sensing-based deep learning method to detect pneumonia on CXR images. In the study, the discrete cosine transform was used as a preprocessing stage. They obtained 97.34% classification performance using CNN and FCNN with two hidden layers. Bhagat et al. [12] preprocessed the data with the histogram equalization, image sharpening, and contrast enhancing techniques. Moreover, they augmented the data using generative adversarial networks. They applied a variant of AlexNet as the DNN on the extended dataset, and they classified CXR images into two classes with an accuray of 79.5%. In the study by Mahajan et al. [13], preprocessing stages were not applied. Inception V3 and DenseNet121 deep learning models were compared in terms of classification performance. Densenet121 network was trained on a finely tuned CheXNet model, and 88.78% test accuracy was achieved for the classification of CXR images. Sharma et al. [14] proposed different DNN models to extract features from CXR images and classify these images to determine the pneumonia. They enriched GWCMC CXR image database by some operations such as rescaling, rotation, zoom, width shift, height shift, shear, and horizontal flip; and they classified CXR images into normal and pneumonia with a success rate of 90.68% by using CNN and FCNN. Labhane et al. [15] developed deep learning models to detect pneumonia from CXR images. They augmented the data and then applied CNN, Inception V3, VGG16, and VGG19 models without using preprocessing stages. Then, they compared the models according to classification performances. As a result of the tests, they classified the CXR images into two categories with 98% success rate. In the study by Talo [16], after the augmentation process, ResNet-152, Inception V3, and VGG16 models were used for classification. ResNet-152 was customized to recognize pneumonia, and 97.4% performance was achieved without applying preprocessing stages. In the study by Stephen et al. [17], preprocessing stages were not applied. GWCMC CXR image dataset was enriched by some operations such as rescaling, rotation, zoom, width shift, height shift, shear, and horizontal flip; and 95.3% success rate was achieved with the use of CNN and FCNN. Kermany et al. [18] also used GWCMC CXR image dataset. They applied two different binary DNNs for classification processes. The Inception V3 model pretrained on the ImageNet dataset was adapted for classification, and CXR images were classified as normal/pneumonia, and bacterial pneumonia/viral pneumonia with success rates of 90.7%

and 92.8%, respectively, without preprocessing stages.

It is observed that the augmentation process for CXR images in [5, 9–11, 13, 16, 18] and preprocessing stages in [1, 5–7, 9, 10, 13–18] were not preferred. Transfer learning has been used to strengthen the training of DNNs, especially when the dataset is of small size. However, it is observed that DNNs using transfer learning contain an excessively large number of weights and also do not give higher classification performance than DNNs trained with random weights. Moreover, the size of datasets used for CXR images is not small. Therefore, in some studies [6, 10, 11, 14, 17], CNN plus FCNNs are preferred to classify the CXR images.

In recent years, the use of algorithms inspired by nature has become widespread for increasing the classification performance. In the study by Erkaymaz et al. [19], classification performance was further improved by modifying the structures or training algorithms of neural networks. In [21], the structure of a conventional feed-forward neural network was modified by including new connections into the network during training, and high classification performance was achieved. Moreover, it is observed that deep neural networks were applied to interdisciplinary subjects such as learning physical properties of liquid crystals [20]. The CNN was employed to determine the parameters of liquid crystals with high accuracy by analyzing the input liquid crystal images. The study has demonstrated the usefulness of deep CNNs for predicting physical properties of liquid crystals directly from their optical textures.

It can be expected that high classification performance is obtained after a strong training process, even if DNNs with small size are used. In this context, determining the structure of DNN correctly and applying the strong training process on this structure are extremely important for increasing the success rates of the DNN. We will depart from the literature on the basis of (*i*) the design of the DNN structure, and (*ii*) the training of this structure.

DNNs generally consist of two parts; a CNN and an FCNN. While the CNN is generally used to extract features from images, the FCNN is used for classification purpose. The performance of the network depends on the determination of the hyperparameters (number of layers, number of features in the layers, size of the filter, learning rate etc.) correctly. The smaller the number of hyperparameters, the easier it is to determine the correct values for these parameters. The proposed DNN contains only the CNN structure. Minimum distance classifier (MDC) instead of FCNN is preferred as the classifier in this study. Since there is no hyperparameter of the MDC, and FCNN is not used in the proposed DNN, the number of hyperparameters decreases automatically. At the same time, the disadvantages of the FCNN are avoided with this preference.

The proposed DNN is trained by taking into account the divergence measure in the feature space, and thereby, it is aimed to achieve low intraclass and high interclass distributions. In the proposed methods, it is possible to train each part of the DNN (CNN+FCNN) individually. Because the proposed DNN, which is composed of only the feature extractor part (CNN), has fewer weights and does not contain the weights of the FCNN, the training algorithm is totally reserved to focus on the features (reducing the scattering within the classes,(*1). In the study, since Walsh vectors (functions) are selected as the centers of classes, the classes automatically locate apart from each other in the feature space as a result of the properties of the Walsh vectors. Thus, the distances between the classes become automatically maximized, (*2). These two approaches (*1, *2) enforce the training strategy to decrease the within-class scatterings, and to increase the distances between the classes, which in turn leads to achieving the best divergence measure for the features. By means of strong training strategy, the classes can be distinguished using linear classifiers.

The remainder of this paper is as follows. In Section 2, the structure and training of the proposed DNN are explained. In Section 3, the dataset is described, and computer simulations and results are given. Section

4 discusses and concludes the paper.

2. Methods

2.1. Structure of the proposed DNN

DNNs are generally composed of two structures: CNN and FCNN. While CNNs are used for feature extraction, FCNNs are used for classification of images. The proposed DNN contains only the CNN structure. The MDC, instead of FCNN, is preferred as the classifier in this study and has no hyperparameters. At the same time, the disadvantages of the FCNN are avoided, such as: (i) failing to achieve optimal solution, (ii) high memory requirements because of a lot of neurons in FCNN, (iii) incompatibility for real-time applications due to high computing time, and (iv) spending a long time because of the determination of extra hyperparameters (such as the learning rate, the number of hidden layers, and the number of nodes in these hidden layers).

Figure 1 shows the proposed DNN structure which contains feature extractor (FE) and MDC. The feature extractor has six convolution layers and one single dense layer. The single dense layer is used here to equalize the size of the network output to the size of the Walsh vectors. In this layer, filter size is equal to the size of the image. Therefore, this layer is also considered as the convolution layer. Input data is represented as an $(N \times N)$ dimensional image ${}^{i}X$ (the class label is represented by i), and presented to the input of the proposed FE.

${}^{i}FeO = [FeO_1, FeO_2, ..., FeO_M]$ is the proposed FE's output. During the training, the network weights are updated in such a way that FE's output converges to a Walsh vector. Therefore, the proposed FE provides a mapping from images to Walsh vectors after the training. MDC determines the classes for the input images according to the following equations:

$$D_i = \sum_{j=1}^{M} (FeO_j - {}^{i}Wf_j)^2 \quad , \quad D_{cl} = min_i(D_i) \quad , \quad class \quad label = cl \tag{1}$$

where M represents the Walsh matrix's dimension (M is equal to 4 in this study), FeO_j is the jth output of the proposed FE, ${}^{i}Wf_j$ is the jth component of Walsh vector representing the ith class. The reason for the use of Walsh vector, ${}^{i}Wf = [{}^{i}Wf_1, {}^{i}Wf_2, {}^{i}Wf_3, {}^{i}Wf_4]$ will be mentioned in the following section.

The proposed DNN contains convolution layers and minimum distance classifier. There exist ReLU layers, max pooling, and batch normalization process at the output of each convolution layer. Input image size is decreased to 384×384 pixels. The size of filter is 3×3 for each convolution layer, and the number of feature maps depends on the type of the classification processes. These values will be shown in the Computer Simulations section. CXR images of 384×384 size are applied to the input of the proposed FE. There are four nodes at the output of the proposed FE. An MDC is used to determine the class of the feature vectors. The nodes of the MDC consist of Walsh vectors. Training is applied only to the feature extractor layers. Detailed information on the usage of the Walsh vectors will be mentioned in the following Training of the Proposed DNN section.

2.2. Training of the Proposed DNN

A measure for determining how well the feature vectors scatter in the feature space while maintaining the class separability at its maximum is the divergence value which is calculated over the samples in the training set. It

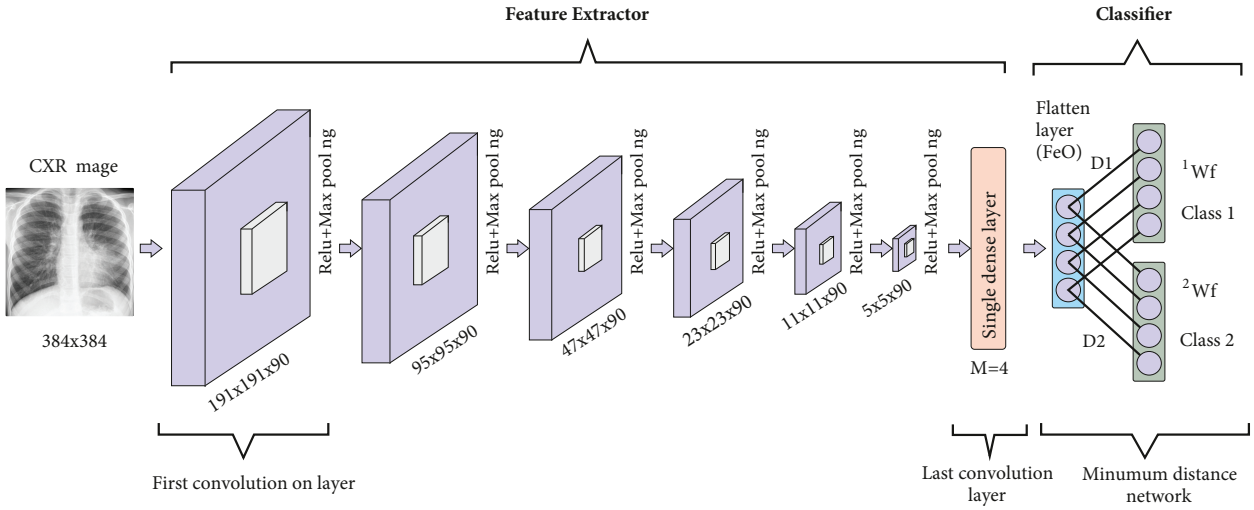


Figure 1. The structure of the proposed binary DNNs for CXR image classification. The network comprises convolutional layers, a single dense layer, and an MDC. After applying max pooling five times, image size decreases to 5×5. Single dense layer contains four nodes and each node has $5 \times 5 \times 90$ parameters. MDC consists of two Walsh nodes. ${}^{i}Wf$ represents the Walsh vectors being the weights of the nodes of MDC.

calculates the interclass/intraclass ratio according to the following equations:

$$divergence = det(C^{-1}B_c) \quad , \quad C = C_1 + C_2 + + C_K. \tag{2}$$

Here B_c is the between-class (interclass) scatter matrix obtained by calculating the covariance of the classes' mean vectors. While K represents the number of classes, C_i is covariance matrix of the features of the ith class. In this case, C, the within-class (intraclass) scatter matrix of the whole training set, is computed by summing the covariance matrices of all the classes.

Since ReLU is used as the activation function, the minus 1 values in the original Walsh matrix are replaced with zeros in the modified matrix. Walsh matrices for 2-, 4-, and 8-dimensional feature spaces are defined in equation (3):

$$W_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \qquad W_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \qquad W_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \tag{3}$$

Walsh vectors have the following properties: vectors are equidistant from each other, and the distance between vectors is half of the rank of Walsh matrix (rank is equal to the output vectors' dimension). Therefore, the centers of classes are selected among the columns (rows) of Walsh matrix. The rank was selected as four for both the two- and three-class problems in this study. Any column (row) in the matrix is randomly chosen to be the center of one class. When the Walsh matrix rank increases, class centers become more distant from one another. However, over-increasing the rank will cause the computational load and training duration to increase.

It is desired that class centers locate at the furthest distances from one another in the feature space; and Walsh vectors ensure this situation [21].

The proposed DNN's training essentially considers the maximization of the divergence measure; hence, it is aimed to achieve low within-class and high between-class scatterings in the feature space. Training is carried out only in the feature extraction process, and MDC is not used in the training process. CXR image is given as input to the feature extractor. The feature extractor generates an output by using its own weights. There is a unique Walsh vector representing each class. The desired output at this stage is the Walsh vector, which represents the class of the input image. The weights of the network are updated taking into account the difference between the desired output (Walsh vector) and network output. After the training process is completed, the nodes of MDC consist of Walsh vectors. Figure 2 shows the training processes of the proposed DNNs.

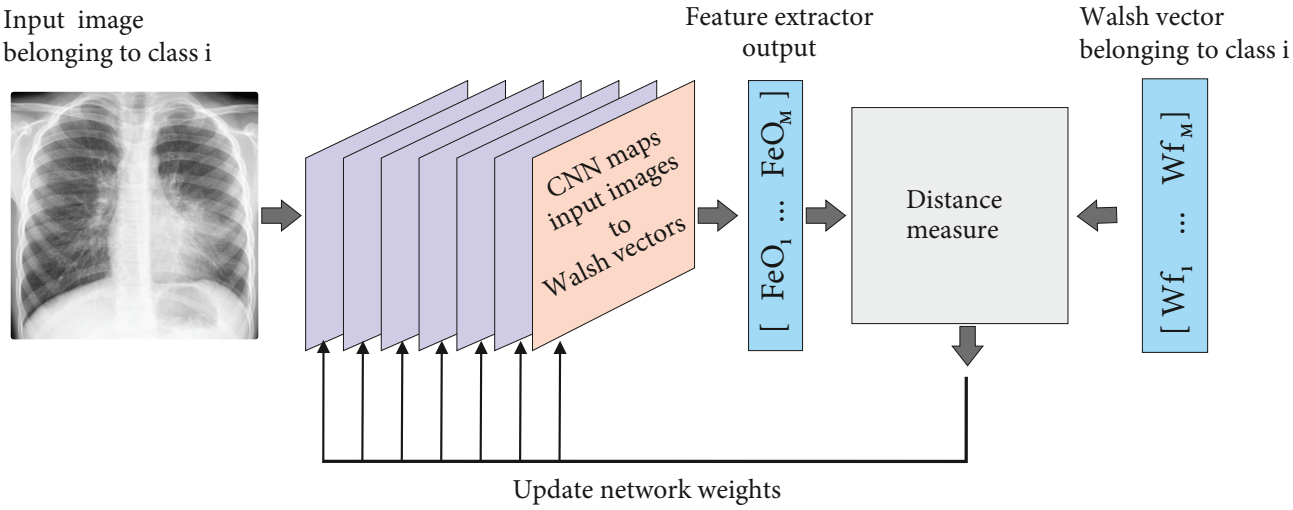


Figure 2. Training process of the proposed DNNs. FeO represents the output of the proposed FE for the input image belonging to the ith class (and also the outputs of the flatten layer in Figure 1), and Walsh vector belonging to class i is the desired output. In the training process, all weights of the filters of CNN are modified according to the difference between the output (FeO) of CNN and desired output (WF). M is equal to 4 in this study. ${}^{i}WF_{m}$ represents the mth element of Walsh vector belonging to the ith class. As an example, ${}^{i}WF = [{}^{i}WF_1, {}^{i}WF_2, {}^{i}WF_3, {}^{i}WF_4]$ for two classes: ${}^{1}WF = [1, 1, 0, 0]$ and ${}^{2}WF = [1, 0, 1, 0]$.

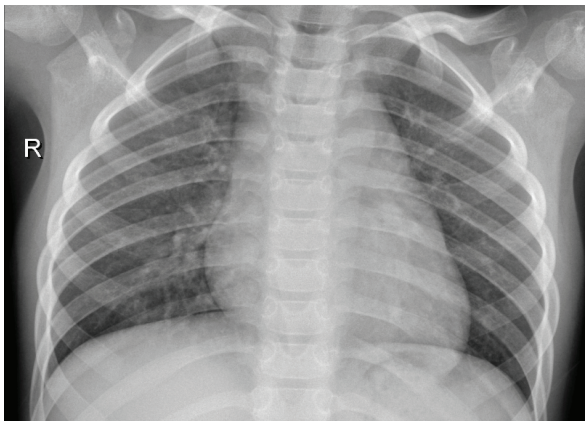
3. Computer simulations

The algorithms are coded in Python using Tensorflow library, and running on a workstation which has 32 core - 2.7 GHz CPUs and GeForce GTX2080 Ti Graphics card.

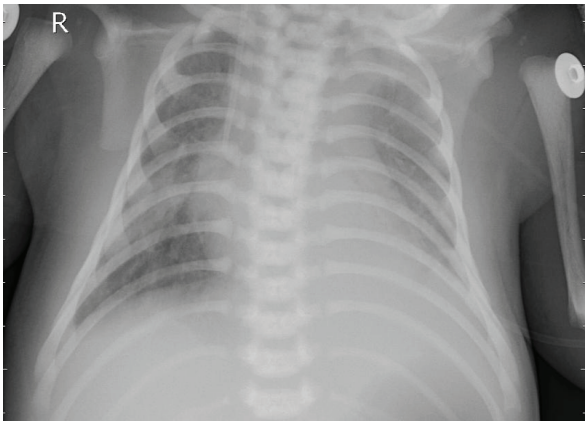
3.1. Dataset

In this study, CXR images of 1- to 5-year-old children provided by the GWCMC were used. In this dataset, there are a total of 5840 images, each in different sizes. Of these images, low-quality or illegible ones have been eliminated and the remaining images have been graded by a total of three experts for use in artificial intelligence studies. Thus, to train the proposed CNN network, a total of 5232 CXR images from 5856 patients were used. While 1349 of 5232 images belong to the normal class, 3883 of them belong to the pneumonia class. In addition, of the 3883 images, 1345 were labeled as bacterial pneumonia and 2538 as viral pneumonia. Sample images

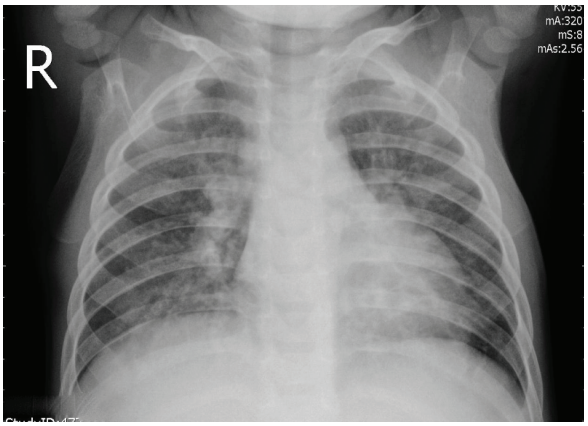
related to normal, bacterial pneumonia, and viral pneumonia classes are presented in Figure 3. The test set contains 390 pneumonia (242 bacterial and 148 viral) and 234 normal images from 624 subjects.



(a) Normal



(b) Bacterial pneumonia



(c) Viral pneumonia

Figure 3. CXR image samples for each class (a) normal, (b) bacterial pneumonia, and (c) viral pneumonia. While 1349 of 5232 images belong to the normal class, 3883 of them belong to the pneumonia class. In addition, of the 3883 images, 1345 were labeled as bacterial pneumonia and 2538 as viral pneumonia. The test set contains 390 pneumonia (242 bacterial and 148 viral) and 234 normal images from 624 subjects. The size of all images is set to $384 \times 384 \times 1$.

3.2. Classification by using the proposed DNN

In this study, three different DNNs are developed to classify the CXR images as: (1) normal or pneumonia (DNN-NP), (2) viral pneumonia or bacterial pneumonia (DNN-VB), and (3) normal, viral pneumonia, or bacterial pneumonia (DNN-NVB). These DNNs are tested on GWCMC CXR image dataset.

Firstly, the effect of different sizes of input images on the performance of the network was tested. The changes of the classification results for 256×256, 512×512, and 1024×1024 input image sizes were examined. No major changes were observed in the classification results obtained by using the proposed DNNs. In order to adjust the filter size in the layers as desired, the image size was selected as 384×384.

Each image was randomly flipped left to right or down to up, and Gaussian noise was added to the

images. Negative images were created randomly. Image brightness was changed randomly. Thus, the number of images in the training set has been increased five times by using the above augmentation processes. The proposed DNNs were tested with the extended dataset, and no major changes were observed in the classification performances.

Instead of RGB CXR images, the gray component of the original image and four 2D wavelet images were given as input to the network. In this condition, too, there was no change in the classification performance of the proposed network.

In all experiments, CXR images are given as input to DNNs without enriching the training set by using augmentation and without applying any preprocessing step to the input images.

In the structure of the DNN-NP, feature extractor has six convolution layers and a single dense layer. The filter size is 3×3 in each convolution layer and there are 90, 90, 90, 90, 90, and 90 feature maps, respectively. The single dense layer has four output nodes due to using four dimensional Walsh vectors for training, so it has $90\times5\times5\times4$ weights. Like DNN-NP structure, the feature extractor of DNN-VB also has six convolution layers and a single dense layer. The filter size is 3×3 in each convolution layer and there are 90, 90, 90, 90, 90, and 90 feature maps, respectively. The single dense layer has four output nodes and $90\times5\times5\times4$ weights.

In this study, GWCMC CXR image dataset was used, but there are studies in the literature using different datasets. Here we will first examine these studies. Table 1 shows the classification results obtained from the studies [4–9] using other datasets. According to Table 1, detection of pneumonia from CXR images is best performed using the transfer learning method with 96.2% success. To evaluate DNN-NP and DNN-VB networks, 10 experiments are performed by initializing the weights randomly. To calculate the classification performance, accuracy (4) is used as the performance metric. Classification performance of the proposed DNNs are given as the average of these experiments. Table 2 shows the classification results related to DNN-NP, DNN-VB, and the study by Kermany et al. [18]. As seen in Table 2, the proposed study perfectly detects pneumonia from CXR images. In determining the origin of pneumonia, a result close to that of Kermany et al. [18] is obtained. Looking at the average achievements, it is seen that the proposed study has better classification performance. In the literature, there are also studies [10–18] detecting pneumonia with the dataset[2] used in this study, and Table 3 represents the classification performances of these studies. In the tables, classification results, the need for augmentation processes, and the need for preprocessing stages are shown. As can be seen in Table 3, the proposed study gives better classification results than other studies using the same dataset, even without preprocessing and without enlarging the dataset.

$$\text{Average Accuracy} = \frac{1}{10}\sum_{j=1}^{10}\frac{TP+TN}{TP+FN+TN+FP}\times100 \tag{4}$$

After binary classification with DNN-NP and DNN-VB, the DNN-NVB network was tested for three classes: normal, viral pneumonia, and bacterial pneumonia. As in DNN-NP and DNN-VB, feature extractor of DNN-NVB network has six convolution layers and a single dense layer. The filter size is 3×3 for all six layers; and the number of feature maps is 60, 60, 60, 60, 60, and 60, respectively. The single dense layer has four output nodes and $60\times5\times5\times4$ weights. After the experiments 90% performance is achieved for three classes by using DNN-NVB. The classification performance is slightly lower than the results of the DNN-NP and DNN-VB

[2]Mooney P (2018). Chest X-Ray Images (Pneumonia). Website: www.kaggle.com/paultimothymooney/chest-xray-pneumonia [accessed 17 Jan 2021].

Table 1. The classification results obtained from the studies using other databases.

Studies	Binary classification results: normal/ pneumonia (%)	Usage of transfer learning	DNNs	The need for augmentation	The need for preprocessing
Li et al. [4]	83.5	Yes	SE-Resnet	Yes	Yes
Varshni et al. [5]	80.02	Yes	DenseNet-169	No	No
Li et al. [6]	92.79	No	CNN+FCNN	Yes	No
Aledhari et al. [7]	75	Yes	VGG16	Yes	No
Tilve et al. [8]	96.2	Yes	VGG16	Yes	Yes
O'Quinn et al. [9]	72	Yes	AlexNet	No	No

Table 2. The classification results obtained by using the DNN-NP and DNN-VB.

DNNs	Binary classification results: normal/ pneumonia (%)	Binary classification results: bacterial pneumonia/ viral pneumonia (%)	Average classification results: (%)	The need for augmentation	The need for preprocessing
DNN-NP	100			No	No
DNN-VB		92	96	No	No
Kermany et al. [18]	90.7	92.8	91.75	No	No

binary classifiers, but it was expected that the performance would decrease slightly as the number of classes increased.

In this study three classifiers (DNN-N, DNN-VB, and DNN-NVB) were used. The presence of pneumonia and whether it was bacterial or viral were determined by DNN-NP and DNN-VB networks, respectively. Three class problems (normal, bacterial pneumonia, and viral pneumonia) were tried to be solved with DNN-NVB. To the best of our knowledge, there is no study that classifies the CXR images into three classes using the GWCMC CXR dataset.

4. Discussion and conclusions

In this study, the presence and origin of pneumonia from CXR images are determined using DNNs. In the proposed DNNs, CNN trained with Walsh functions is used to extract features from CXR images and MDC is used for classification purpose. DNN-NP and DNN-VB binary classifiers are used to classify normal vs pneumonia and bacterial pneumonia vs viral pneumonia, respectively, while DNN-NVB is used to solve three-class problems: normal, bacterial pneumonia, and viral pneumonia.

In the augmentation process, training and test sets should first be separated, then the training set should be enriched by using some functions. Otherwise, if the entire dataset is firstly enriched and then divided into training and test clusters, some samples will be included both in the training set and the test set. In this case,

Table 3. The classification results obtained from the studies using the same database with this study.

Studies	Binary classification: normal, pneumonia (%)	Usage of transformation	DNN	The need for augmentation	The need for preprcessing
Vijendran et al. [10]	92.5	No	FCNN	No	No
Ayan et al. [1]	87	Yes	VGG16	Yes	No
Islam et al. [11]	97.34	No	CNN+FCNN	No	Yes
Bhagat et al. [12]	79.5	Yes	AlexNet	Yes	Yes
Mahajan et al. [13]	88.78	Yes	DenseNet121	No	No
Sharman et al. [14]	90.68	No	CNN+FCNN	Yes	No
Labhane et al. [15]	98	Yes	VGG16	Yes	No
Talo et al. [16]	97.4	Yes	ResNet-152	No	No
Stephen et al. [17]	95.3	No	CNN+FCNN	Yes	No
Kermany et al. [18]	90.7	Yes	Inception V3	No	No
Our study	100	No	CNN+MDC	No	No

a very high performance will be obtained for the test set. There are several CXR images of the same person in the dataset used. Therefore, the training and test sets had been prepared by ensuring that there were no images of the same person in the training set and the test set. In this study, training and test sets are constituted in the same way as in the literature.

It is observed from the tables that preprocessing stages are not needed. In some studies [6, 10, 15–17], high performance was obtained without using preprocessing stages. In systems that do not use preprocessing stages, the classification response is produced faster. This situation is an advantage for real-time applications. Therefore, preprocessing stages were not used in this study.

In networks using transfer learning, the number of nodes becomes excessively high. If the database is small, transfer learning is generally preferred to increase classification performance. However, the dataset used in this study is not small. Therefore, there is no need to use large network structures. In the literature, there are many studies [6, 10, 11, 14, 17] in which the networks are trained with random weights.

In Li et al.'s study [6], the number of weights used in the AlexNet (42,725,889 weights), VGG16 (27,560,769 weights), and their network (1,695,777 weights) are compared with each other. The network, which is called PNet, developed in [6] showed higher performance with less weights. PNet contains CNN and FCNN. The DNN-NP proposed in our study gives 100% classification performance and contains $1\times3\times3\times90 + 90\times3\times3\times90 + 90\times3\times3\times90 + 90\times3\times3\times90 + 90\times3\times3\times90 + 90\times3\times3\times90 + 90\times5\times5\times4$ weights (374,310 weights in total). Generally, the block which contains most of the weights in a conventional DNN structure is FCNN. The DNN-NP proposed in our study does not contain the FCNN; hence, this preference causes the number of weights of DNN to be less than that of the structures in the literature.

As observed in Tables 2 and 3, the highest classification performance in the literature is achieved by using the proposed DNN. The DNN-VB proposed in our study gives 92% classification success and contains $1\times3\times3\times90 + 90\times3\times3\times90 + 90\times3\times3\times90 + 90\times3\times3\times90 + 90\times3\times3\times90 + 90\times3\times3\times90 + 90\times5\times5\times4$ weights (374,310 weights in total). The DNN-NVB proposed in our study gives 90% classification success

and contains $1\times3\times3\times60$ + $60\times3\times3\times60$ + $60\times3\times3\times60$ + $60\times3\times3\times60$ + $60\times3\times3\times60$ + $60\times3\times3\times60$ + $60\times5\times5\times4$ weights (168,540 weights in total). It is observed that the proposed DNNs achieve high performances by using fewer nodes.

Acknowledgments

All simulations are realized on the GeForce GTX2080 graphics card supported by the Scientific Research Project Unit of the İstanbul Technical University (project no. MYL-2018-41621). The authors thank Muazzez Buket Darıcı who prepared the Chest X-Ray image database for us.

References

[1] Ayan E, Unver HM. Diagnosis of pneumonia from chest x-ray images using deep learning. In: 2019 Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science (EBBT); Istanbul, Turkey; 2019. pp. 1-5. doi: 10.1109/EBBT.2019.8741582

[2] Sharma A, Raju D, Ranjan S. Detection of pneumonia clouds in chest x-ray using image processing approach. In: Nirma University International Conference on Engineering (NuiCONE); Ahmedabad, India; 2017. pp. 1-4. doi: 10.1109/NUICONE.2017.8325607

[3] Singh N, Sharma R, Kukker A. Wavelet transform based pneumonia classification of chest x-ray images. In: 2019 International Conference on Computing, Power and Communication Technologies (GUCON); Delhi, India; 2019. pp. 540-545.

[4] Li B, Kang G, Cheng K, Zhang N. Attention-guided convolutional neural network for detecting pneumonia on chest x-rays. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Berlin, Germany; 2019. pp. 4851-4854. doi: 10.1109/EMBC.2019.8857277

[5] Varshni D, Thakral K, Agarwal L, Nijhawan R, Mittal A. Pneumonia detection using CNN based feature extraction. In: 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT); Coimbatore, India; 2019. pp. 1-7. doi: 10.1109/ICECCT.2019.8869364

[6] Li Z, Yu J, Li X, Li Y, Dai W et al. PNet: An efficient network for pneumonia detection. In: 2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI); Suzhou, China; 2019. pp. 1-5. doi: 10.1109/CISP-BMEI48845.2019.8965660

[7] Aledhari M, Joji S, Hefeida M, Saeed F. Optimized CNN-based diagnosis system to detect the pneumonia from chest radiographs. In: 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); San Diego, CA, USA; 2019. pp. 2405-2412. doi: 10.1109/BIBM47256.2019.8983114

[8] Tilve A, Nayak S, Vernekar S, Turi D, Shetgaonkar PR et al. Pneumonia detection using deep learning approaches. In: 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE); Vellore, India; 2020. pp. 1-8. doi: 10.1109/ic-ETITE47903.2020.152

[9] O'Quinn W, Haddad RJ, Moore DL. Pneumonia radiograph diagnosis utilizing deep learning network. In: 2019 IEEE 2nd International Conference on Electronic Information and Communication Technology (ICEICT); Harbin, China; 2019. pp. 763-767. doi: 10.1109/ICEICT.2019.8846438

[10] Vijendran S, Dubey R. Deep online sequential extreme learning machines and its application in pneumonia detection. In: 2019 8th International Conference on Industrial Technology and Management (ICITM); Cambridge, United Kingdom; 2019. pp. 311-316. doi: 10.1109/ICITM.2019.8710700

[11] Islam SR, Maity SP, Ray AK, Mandal M. Automatic detection of pneumonia on compressed sensing images using deep learning. In: 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE); Edmonton, AB, Canada; 2019. pp. 1-4. doi: 10.1109/CCECE.2019.8861969

[12] Bhagat V, Bhaumik S. Data augmentation using generative adversarial networks for pneumonia classification in chest x-rays. In: 2019 Fifth International Conference on Image Information Processing (ICIIP); Shimla, India; 2019. pp. 574-579. doi: 10.1109/ICIIP47207.2019.8985892

[13] Mahajan S, Shah U, Tambe R, Agrawal M, Garware B. Towards evaluating performance of domain specific transfer learning for pneumonia detection from x-ray images. In: 2019 IEEE 5th International Conference for Convergence in Technology (I2CT); Bombay, India; 2019. pp. 1-6. doi: 10.1109/I2CT45611.2019.9033555

[14] Sharma H, Jain JS, Bansal P, Gupta S. Feature extraction and classification of chest x-ray images using CNN to detect pneumonia. In: 2020 10th International Conference on Cloud Computing, Data Science and Engineering (Confluence); Noida, India; 2020. pp. 227-231. doi: 10.1109/Confluence47617.2020.9057809

[15] Labhane G, Pansare R, Maheshwari S, Tiwari R, Shukla A. Detection of pediatric pneumonia from chest x-ray images using CNN and transfer learning. In: 2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE); Jaipur, India; 2020. pp. 85-92. doi: 10.1109/ICETCE48199.2020.9091755

[16] Talo M. Pneumonia detection from radiography images using convolutional neural networks. In: 2019 27th Signal Processing and Communications Applications Conference (SIU); Sivas, Turkey; 2019. pp. 1-4. doi: 10.1109/SIU.2019.8806614

[17] Stephen O, Sain M, Maduh UJ, Jeong DU. An efficient deep learning approach to pneumonia classification in healthcare. Hindawi J Healthc Eng 2019. doi:10.1155/2019/4180949

[18] Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 2018; 172(5): 1122-1131. doi: 10.1016/j.cell.2018.02.010

[19] Erkaymaz O, Ozer M, Perc M. Performance of small-world feedforward neural networks for the diagnosis of diabetes. Appl Math Comput 2017; 311: 22-28. doi: 10.1016/j.amc.2017.05.010

[20] Sigaki HYD, Lenzi EK, Zola RS, Perc M, Ribeiro HV. Learning physical properties of liquid crystals with deep convolutional neural networks. Sci Rep 2020; 10(7664). doi: 10.1038/s41598-020-63662-9

[21] Dokur Z, Olmez T. Heartbeat classification by using a convolutional neural network trained with Walsh functions. Neural Comput & Applic 2020; 32: 12515-12534. doi: 10.1007/s00521-020-04709-w