Pneumonia Detection Using Convolutional Neural Networks (CNNs)

Harsh Shukla, Divyam Tyagi, Rihan Sadiq, Shweta Shekhawat

Under the supervision of: Asst. Prof. Gaurav Singhania
Information Technology,
Noida Institute of Engineering and Technolohy,
Greater Noida, India

***Abstract*—The main cause of death in children under five is pneumonia, an interstitial lung disease. According to a UNICEF study, it was responsible for about 16% of all deaths among children under the age of five in 2016, killing about 880,000 of them. Most of the affected children were under two years old. Timely identification of pneumonia in children can help to fast-track the process healing. In order to accurately identify pneumonic lungs from chest X-rays, this paper presents convolutional neural network models. These models can be used by medical professionals to treat pneumonia in the real world. The Chest X-Ray Images (Pneumonia) dataset from Kaggle was used for the experimentation.**

There are one, two, three, and four convolutional layers in the first, second, third, and fourth models, respectively. The accuracy of the first model is 89.74%, the accuracy of the second is 85.26%, the accuracy of the third model is 92.31%, and the accuracy of the fourth model is 91.67%. In the second, third, and fourth models, dropout regularization is used to reduce overfitting in the fully connected layers. For better evaluation, recall and F1 scores are also computed from each model's confusion matrix.

Keywords— Convolutional neural networks(CNNs), Pneumonia detection, ReLU, Max-pooling, Forward and backward propagation

1. INTRODUCTION

Indoor air pollution is one of the main causes of pneumonia in children. In addition to this, undernourishment, a lack of clean water, sanitation, and basic health facilities are significant contributing factors. The interstitial lung disease known as pneumonia is brought on by bacteria, fungi, or viruses. It killed about 880,000 kids in 2016, which was about 16% of the 5.6 million under-five fatalities [1]. Most of the casualties were under two years old. Early pneumonia diagnosis can reduce the risk of child fatalities. In order to reliably identify pneumonic lungs from chest X-rays, convolutional neural network models are presented in this research. These models can be used by medical professionals worldwide to treat pneumonia [2]. In order to aid in the early detection of pneumonia, these models have been trained to quickly distinguish between normal and pneumonia in chest X-ray pictures. Although convolutional neural network-based transfer learning models with pre-trained weights like AlexNet, ResNet50, InceptionV3, VGG16, and VGG19 are among the most successful ImageNet dataset models, they were not trained on this dataset because it is not as large as those that typically use transfer learning [3]. For the purpose of preventing pneumonia in children and other age groups, four classification models were created using CNN to identify the disease from chest X-ray images. The size of the dataset is directly correlated with the model's accuracy, meaning that using large datasets enhances model accuracy. However, there is no direct relationship between the model's accuracy and the number of convolutional layers.

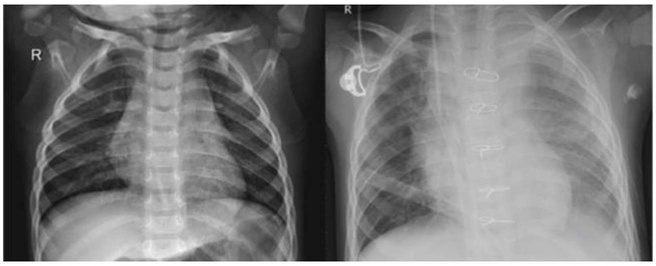

Fig. 1 Left image depicts normal lungs and Right image depicts pneumonic lungs

The models must be evaluated after each execution in order to train a specific amount of combinations of convolution layers, dense layers, dropouts, and learning rates in order to achieve the best outcomes. A model that not only attained the specified accuracies but also excelled other models in terms of recall and F1 scores was obtained by first training simple models with one convolution layer on the dataset. The goal of the study is to create from scratch CNN models that can categorize and subsequently identify pneumonic patients from their chest X-rays with high validation accuracy, recall, and F1 scores. Recall is frequently preferred over other performance assessing factors in medical imaging scenarios because it provides a measure of false negatives in the data. The number of false negatives in the outcome is extremely important in figuring out how well models work in the actual world [4]. A model is considered to be underperforming, ineffective, and even risky if it has high accuracy but poor recall values because this means that there are more occasions when the model incorrectly predicts a patient as healthy when they are actually ill. Therefore, the patient's life would be in danger. Only models with high recall values, respectable accuracies, and high F1 scores would be the emphasis in order to avoid this [5].

2. RELATED WORK

The challenge of accurately categorizing images has been taken up by numerous scholars.

In order to identify common thorax disease from frontal and lateral chest X-ray pictures, Rubin et al. [6] built a CNN model. These photos were subjected to extensive automatic recognition using the MIMIC-CXR dataset. The dataset was divided into training, testing, and validation sets with respective weights of 70%, 20%, and 10%. Performance was enhanced via pixel normalization and data augmentation. For PA and AP, their DualNet CNN model obtained average AUCs of 0.72 and 0.688, respectively. Lakhani et al. [7] created a deep convolutional neural network to categorize pulmonary TB. In order to categorize chest X-ray pictures, transfer learning models like AlexNet and GoogleNet were also used. The dataset was divided as follows: training set: 68%; testing set: 14.9%; and validation set: 17.1%. The best performing model, with an AUC of 0.99, was obtained by using data augmentation and pre-processing approaches. The model's accuracy and recall were 100 and 97.3%, respectively. Guan et al. [8] created an AG-CNN model to identify thoracic illness. Thorax illness was identified from chest X-ray pictures using the ChestX-ray14 dataset. Attention-guided global and local branch CNN was utilized for categorization. With an AUC of 0.868, their model outperformed the other models stated in their study report. To categorize chest X-ray images into pneumonia and other 14 diseases, Rajpurkar et al. [9] created a deep convolutional neural network model. A dataset called ChestX-ray14 was utilized to train the model. They compared academic radiologists in practice to their ChXNet model (a 121-layered model). In comparison to radiologists, whose F1 score (95% CI) was 0.387, their ChXNet model outperformed them with an F1 score (95% CI) of 0.435.

Krizhevsky et al. [10] developed a deep convolutional neural network model with five convolutional layers, some of which were followed by max-pooling layers, and three fully connected layers. There were 60 million different parameters in this network. This model earned a top-five error rate of 17% by using dropout. To achieve top-five test accuracy of 92.7%, Simonyan et al. [11] created a highly accurate model using several small kernel-sized filters. The ImageNet dataset was used to train this model, which was then submitted to the ILSVRC 2014 competition. Xu et al. [12] created a convolution neural network for the classification and segmentation of brain tumor MRIs. This model used a number of strategies, including feature selection, data augmentation, and pooling. Anthimopoulos et al. [13] created a convolutional neural network with five convolution layers, leaky ReLU, average pooling, and three fully connected layers to identify interstitial lung disease patterns in a dataset of 14,696 pictures divided into seven classes. The classification accuracy of this model was 85.5%. A residual neural network (RNN) was created by Heet al. [14] to classify the images in the ImageNet dataset. To address the issue of vanishing gradients, RNN developed the idea of shortcut connections. This model reached cutting-edge classification accuracy and was submitted to ILSVRC 2015. Glozman et al.[15] created a transfer learning model as an extension of AlexNet utilizing data augmentation approaches. The ADNI database was used to train this model. Hemanth et al. [16] presented the MCPN and MKNN neural network models, which are two neural network models. These models dealt with large convergence times for artificial neural networks and accurately identified MRIs.

3. MOTIVATION AND CONTRIBUTIONS

As was already mentioned, pneumonia affects many people, especially children, and is most common in developing and underdeveloped nations, which are known for risk factors like crowding, unsanitary living conditions, and malnutrition as well as a lack of access to the necessary medical facilities. To completely cure pneumonia, an early diagnosis is essential. The most frequent method of diagnosis is the examination of X-ray scans, but this method depends on the radiologist's ability to interpret the images and is frequently not agreed upon by other radiologists. In order to diagnose the illness, an automatic CAD system with generalizing capabilities is needed. To the best of our knowledge, the ensemble learning paradigm has not been investigated in this classification job, and the majority of prior techniques in the literature concentrated on creating a single CNN model for the categorization of pneumonia cases. However, the ensemble learning model was used in this work because it integrates the discriminative data from all of the component base learners, enabling it to generate better predictions. Transfer learning models were utilized as base learners, and the decision scores of these models were ensembled, to deal with the limited quantity of biological data that was accessible.

The main contributions of this study are as follows.

1. For improving the performance of the base CNN learners in the classification of pneumonia, an ensemble framework was developed. A weighted average ensemble technique was used for this.

2. Precision, recall, f1-score, and AUC, four assessment measures, were used to generate the classifier weights. We employed a hyperbolic tangent function instead of choosing the weights purely based on the performance of classifiers or in accordance with the findings of tests.

3. The Kermany dataset [4] and the RSNA Pneumonia Detection Challenge [33] dataset, two publicly accessible chest X-ray datasets, were used to test the proposed model using the five-fold cross-validation setting. The method is viable for use in the real world because the results outperform those of state-of-the-art techniques.

4. METHODOLOGY

Using the Chest X-Ray Images (Pneumonia) dataset on Kaggle, CNN models that were built from scratch were trained. The models have been implemented using the Keras neural network package with a TensorFlow backend. 624 testing photos, 16 validation images, and 5216 training images make up the data set. To get better results from the dataset, data augmentation has been used. The four models, each with a different number of convolutional layers, were trained using the training dataset. With training and testing batch sizes of 32 and 1, respectively, each model was trained for 20 iterations. The stages above are further detailed in the sub-headings that follow.

3.1 CNN ARCHITECTURE

CNN models are feed-forward networks with fully connected layers, flattening layers, pooling layers, and convolutional layers using the appropriate activation functions.

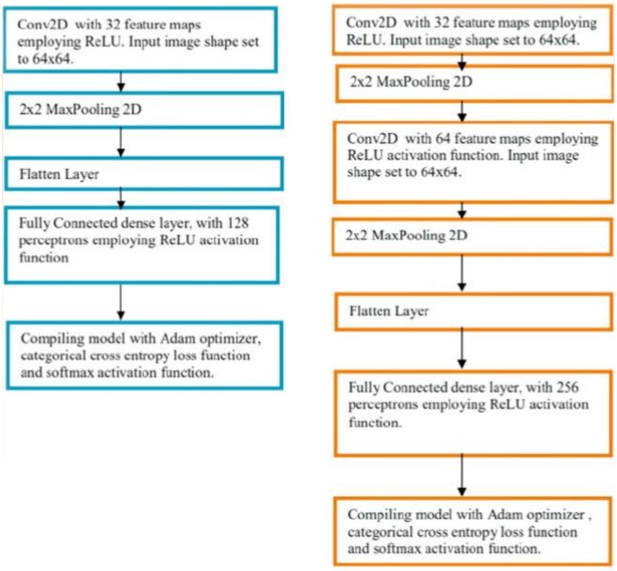

Fig. 2 Algorithms of CNN classifier model 1(left) and model 2 (right)

Convolutional layer. It serves as the foundation for CNNs. In mathematics, the convolution process is used to combine two functions [17]. The input image is initially transformed into a matrix for the CNN models. The input matrix is given a convolution filter, which glides across it and performs element-wise multiplication while storing the result. Thus, a feature map is produced. When black and white photos are present, the 3 3 filter is typically used to generate 2D feature maps. When the input image is represented as a 3D matrix, where the RGB color represents the third dimension, convolutions are carried out in three dimensions. The input matrix is used to operate a number of feature detectors, creating a layer of feature maps that makes up the convolutional layer.

Activation function The two separate activation functions—ReLU activation function and softmax activation function—are used by each of the four models in this research. Rectified linear function is the name of the ReLU activation function [18]. It is a nonlinear function that, when the input is negative, produces zero, and when the input is positive, produces one. The following formula provides the ReLU function: Because it solves the issue of vanishing gradients and helps to increase the nonlinearity of layers, this kind of activation function is frequently used in CNNs. Numerous variations of the ReLU activation function exist, including Noisy ReLUs, Leaky ReLUs, and Parametric ReLUs. ReLU has computational ease and representational sparsity advantages over other activation functions. The four models in this paper's presentation all make use of the softmax activation function. This broadly use activation function is employed in the last dense layer of all the four models. This activation function normalizes inputs into probability distribution. Categorical cross-entropy cost function is mostly used with this type of activation function.

Pooling Layer Pooling layers come after convolutional layers. All four models use max-pooling layers as the type of pooling layer. The maximum pixel intensity values from the window of the image that the kernel is currently covering are chosen by the max-pooling layer, which has a dimension of 2 2. Max-pooling is used to downsample images, hence reducing the dimensionality and complexity of the image [20]. The usage of universal pooling and overlapping pooling are two more forms of pooling layers. The models in this research make use of the max-pooling approach since it makes it easier to identify important characteristics in an image.

Flattening layer and fully connected layer The input image is fed into the flattening layer after passing through the convolutional layer and the pooling layer. This layer reduces the computational complexity of the input image by flattening it into a column. The thick layer or completely linked layer is then fed this. Every node in the first layer of the fully linked layer [21] is connected to every node in the second layer. Every layer in the fully connected layer collects features, and the network then produces a prediction based on these features [22,23]. Forward propagation is the name of this procedure. A costfunction is computed following forward propagation. It is a gauge of how well a neural network model performs. All four models employ categorical cross-entropy as their cost function. After calculating the costfunction, back propagation occurs. Until the network operates at its best, this procedure is repeated. In all four models, the Adam optimization technique has been applied. This process is repeated until the network achieves optimum performance. Adam optimization algorithm has been used in all four models.

Reducing overfitting As a result of the first model's significant overfitting, the dropout technique was used in the later models [24]. The dropout technique addresses the issue of vanishing gradients and reduces overfitting. Each neuron is urged by the dropout technique to create its own unique representation of the input data. During the training process, this method randomly breaks connections between neurons in successive layers [25]. The rate of model learning was also changed to lessen overfitting. It is also possible to use data augmentation techniques to lessen overfitting.

Algorithm of CNN classifier After training and testing several CNN models during the course of the study, the number of epochs for all the classifier models reported in this work was set at 20. Classifier models with longer training times have displayed overfitting. Additionally, a number of optimizers were trained and investigated. After providing the best results, the Adam optimizer function was decided upon and used for all classifiers. A straightforward classifier model using ReLU activation function and a convolutional layer with an image size of 64 * 64 and 32 feature mappings was first developed. It used a fully linked dense layer with 128 perceptions. The second classifier model was trained using an additional convolutional layer of 64 feature maps for enhanced feature extraction in order to improve the results. In order to improve learning, the dense layer's perceptron were increased from 64 to 256. For more precise feature extraction, the third model was trained for three convolutional layers using 128 feature mappings in the third convolutional layer. The dense layer was left alone.

Dataset The 1.16 GB-sized Chest X-Ray Images (Pneumonia) dataset was imported from Kaggle [26], and it has 5856 jpeg images altogether, separated into Train, Test, and Valid folders, with each folder including categories for Pneumonia and Normal. From pediatrics patients aged one to five at the Guangzhou Women and Children's Medical Center, chest X-ray photos (front and rear) were chosen.

5. CONCLUSION AND FUTURE WORK

For the disease to be properly treated and to avoid endangering the patient's life, early diagnosis of pneumonia is essential. The most popular method for diagnosing pneumonia is a chest radiograph, but there is inter-class variability, and the diagnosis depends on the clinicians' skill at spotting early pneumonia traces. To assist medical practitioners, an automated CAD system was developed in this study, which uses deep transfer learning-based classification to classify chest X-ray images into two classes "Pneumonia" and "Normal." An ensemble framework was developed that considers the decision scores obtained from three CNN models, GoogLeNet, ResNet-18, and DenseNet-121, to form a weighted average ensemble. Four assessment metrics—precision, recall, f1-score, and AUC—were combined using the hyperbolic tangent function in a unique way to determine the classifier weights. Using a five-fold cross-validation scheme, the framework was tested on two publicly available datasets of pneumonia chest X-rays and achieved accuracy rates of 98.81%, sensitivity rates of 98.80%, precision rates of 98.82%, and a f1-score of 98.79% on the Kermany dataset and accuracy rates of 86.86%, sensitivity rates of 87.02%, precision rates of 86.89%, and a f1- In these two datasets, it fared better than cutting-edge techniques. The strategy is viable, according to statistical studies of the suggested model utilising McNemar's and ANOVA tests. Furthermore, the proposed ensemble model is domain-independent and thus can be applied to a large variety of computer vision tasks.

As was already mentioned, the ensemble framework occasionally failed to make the right predictions. We may look into methods like picture contrast enhancement or other pre-processing processes in the future to increase the image quality. Before classifying the lung image, we might also think about segmenting it to help the CNN models extract features more effectively. Additionally, the computation cost is higher than that of the CNN baselines developed in studies in the literature because three CNN models are needed to train the proposed ensemble. In the future, we might try to use techniques like snapshot ensemble to reduce the computational requirements.

REFERENCES

[1] https://data.unicef.org/topic/child-health/pneumonia/. Accessed on 15 July 2019.

[2] 2. Jaiswal, A.K., Tiwari, P., Kumar, S., Gupta, D., Khanna, A., Rodrigues, J.J.: Identifying pneumonia in chest x-rays: a deep learning approach. Measurement 145, 511–518 (2019)

[3] Kim, D.H., MacKinnon, T.: Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. Clin. Radiol. 73(5), 439–445 (2018).

[4] Bernal, J., Kushibar, K., Asfaw, D.S., Valverde, S., Oliver, A., Martí, R., Lladó, X.: Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review. Artif. Intell. Med. 95, 64–81 (2019).

[5] Arthur, F., Hossein, K.R.: Deep learning in medical image analysis: a third eye for doctors. J. Stomatology Oral Maxillofac. Surg.

[6] Rubin, J., Sanghavi, D., Zhao, C., Lee, K., Qadir, A., Xu-Wilson, M.: Large Scale Automated Reading of Frontal and Lateral Chest X-Rays Using Dual Convolutional Neural Networks (2018). arXiv preprint arXiv:1804.07839

[7] Lakhani, P., Sundaram, B.: Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology 284(2), 574–582 (2017).

[8] Guan, Q., Huang, Y., Zhong, Z., Zheng, Z., Zheng, L., Yang, Y.: Diagnose Like a Radiologist: Attention Guided Convolutional Neural Network for Thorax Disease Classification (2018). arXiv preprint arXiv:1801.09927

[9] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M.P.: Chexnet: Radiologist-Level Pneumonia Detection on Chest X-rays with Deep Learning (2017). arXiv preprint arXiv:1711.05225

[10] Krizhevsky, A., Sutskever, I., Hinton, G.E:. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105

[11]

[12] Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition (2014). arXiv preprint arXiv:1409.1556

[13] Xu, Y., Jia, Z., Ai, Y., Zhang, F., Lai, M., Eric, I., Chang, C.: Deep convolutional activation features for large scale brain tumor histopathology image classification and segmentation. In: 2015 international conference on acoustics, speech and signal processing (ICASSP), pp. 947– 951 (2015)

[14] Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., Mougiakakou, S.: Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans. Med. Imaging 35(5), 1207–1216 (2016)

[15] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)

[16] Glozman, T., Liba, O.: Hidden Cues: Deep Learning for Alzheimer's Disease Classification CS331B project final report (2016)

[17] Hemanth, D.J., Vijila, C.K.S., Selvakumar, A.I., Anitha, J.: Performance improved iterationfree artificial neural networks for abnormal magnetic resonance brain image classification. Neurocomputing 130, 98–107 (2014)

[18] Bi, X., Li, S., Xiao, B., Li, Y., Wang, G., Ma, X.: Computer aided Alzheimer's disease diagnosis by an unsupervised deep learning technology. Neurocomputing (2019)

[19] Eckle, K., Schmidt-Hieber, J.: A comparison of deep networks with ReLU activation function and linear spline-type methods. Neural Netw. 110, 232–242 (2019)

[20] Ren, S., Jain, D.K., Guo, K., Xu, T., Chi, T.: Towards efficient medical lesion image superresolution based on deep residual networks. Sig. Process. Image Commun. 75, 1–10 (2019)

[21] Zheng, Y., Iwana, B.K., Uchida, S.: Mining the displacement of max-pooling for text recognition. Pattern Recogn. 93, 558–569 (2019)

[22] Bhumika, P.S.S.S., Nayyar, P.A.: A review paper on algorithms used for text classification. Int. J. Appl. Innov. Eng. Manage. 3(2), 90–99 (2013)

[23] Kumar, A., Sangwan, S.R., Arora, A., Nayyar, A., Abdel-Basset, M.: Sarcasm detection using soft attention-based bidirectional long short-term memory model with convolution network. IEEE Access 7, 23319–23328 (2019)

[24] Saeed, F., Paul, A., Karthigaikumar, P., Nayyar, A.: Convolutional neural network based early fire detection. In: Multimedia Tools and Applications, pp. 1–17 (2019)

[25] Kukkar, A., Mohana, R., Nayyar, A., Kim, J., Kang, B.G., Chilamkurti, N.: A novel deeplearning-based bug severity classification technique using convolutional neural networks and random forest with boosting. Sensors 19(13), 2964 (2019)

[26] Khan, S.H., Hayat,M., Porikli, F.: Regularization of deep neural networks with spectral dropout. Neural Netw. 110, 82–90 (2019)

[27] https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia.

[28] ALzubi, J.A., Bharathikannan, B., Tanwar, S., Manikandan, R., Khanna, A., Thaventhiran, C.: Boosted neural network ensemble classification for lung cancer disease diagnosis. Appl. Soft Comput. 80, 579–591 (2019)

[29] Vora, J., Tanwar, S., Polkowski, Z., Tyagi, S., Singh, P.K., Singh, Y.: Machine learning-based software effort estimation: an analysis. In: 11th International Conference on Electronics, computers and Artificial Intelligence (ECAI 2019), pp. 1–6, University of Pitesti, Pitesti, Romania, 27–29 June 2019

[30] Chakraborty, S., Aich, S., Sim, J.S., Kim, H.C.: 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)

[31] Liang, G., Zheng, L.: A transfer learning method with deep residual network for pediatric pneumonia diagnosis. In: Computer Methods and Programs in Biomedicine (2019)

[32] . Du, S. S., Zhai, X., Poczos, B., Singh, A.: Gradient Descent Provably Optimizes OverParameterized Neural Networks (2018). arXiv preprint arXiv:1810.02054

[33] Radford, A., Metz, L., Chintala, S.: Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks (2015). arXiv preprint arXiv:1511.06434

[34] Krizhevsky, A., Sutskever, I., Hinton, G.E:. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105