Detecting Pneumonia With TensorFlow and Convolutional Neural Networks

Dejan Babic
University of Donja Gorica
Podgorica, Montenegro
dejan.babic@udg.edu.me

Ivan Jovovic
University of Donja Gorica
Podgorica, Montenegro
ivan.jovovic@udg.edu.me

Tomo Popovic
University of Donja Gorica
Podgorica, Montenegro
tomo.popovic@udg.edu.me

Stevan Cakic
University of Donja Gorica
Podgorica, Montenegro
stevan.cakic@udg.edu.me

Luka Filipovic
University of Donja Gorica
Podgorica, Montenegro
luka.filipovic@udg.edu.me

2022 IEEE International Conference on Omni-layer Intelligent Systems (COINS) | 978-1-6654-8356-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/COINS54846.2022.9854948

Abstract—**Artificial intelligence is getting more and more involved in our everyday life as a result of enormous amounts of data available for feeding the machine and deep learning algorithms. Deep learning introduced new dimensions and possibilities of applications in medical science. With COVID-19 outbreak in 2020 at global level, the health systems of many countries were overwhelmed. With many patients infected, health system is pressured to correctly diagnose patient's state of illness. In a lot of occasions, it was almost impossible to correctly diagnose many COVID-19 positive patients that have pneumonia due to many outbreaks in many areas. The intelligent system that could detect pneumonia with certainty could help in easing the pressure on the health system and make doctors focus on more severely ill patients. This paper describes development of pneumonia detection model using TensorFlow to processes the chest X-ray images to determine whether the patient has pneumonia. The model is based on deep learning algorithm supported through convolutional neural network. The model presented in this paper has achieved rather high accuracy (over 95%) in analyzing X-Ray images and could be used to speed up decision process in healthcare.**

Keywords—Convolutional Neural Networks, COVID-19, Deep Learning, Machine Learning, Pneumonia Detection

I. INTRODUCTION

The philosophy of information technologies (IT) is such that it always leaned on improving every area of human activity. IT aims to support the decision processes and provide faster information flow. This showed to be true in many areas of everyday life in last years. Especially during the COVID-19 pandemic full swing throughout the world which impacted every single point of human activity.

With IT's role in supporting medicine through creating healthcare information systems, remote treatments, computer assisted diagnosis and telemedicine came a new era of exploring use of artificial intelligence (AI) in assisting of decision-making process in medicine [1]. Big Data concept, that emerged in the last years, gave AI a full boost in exploring its possibilities. The birth of AI concept happened in the 1950s but there were many limitations to its acceptance. Mainly, they were computing and algorithmic limitations. In the early 2000s, the finding of deep learning and first neural networks and algorithms made a good starting point for its wide acceptance, and namely its acceptance in medicine. With Deep Learning concept and its understanding of the human brain, there is a promising era of very potent use of machine learning and its technologies in improving diagnostic accuracy as well as health system efficacy and therapeutic monitoring [2-3].

According to World Health Organization more than 470 million of people were sick from COVID-19 [21]. The pandemic changed the way of life for many people, and since its global ascending in 2020 the way that businesses, governments, school and health systems work has drastically changed. There is no doubt it took toll on many people, but mostly health professionals and health systems, where many countries faced overloaded public health and lack of health workers [4].

In severe cases COVID-19 sometimes causes pneumonia. With uncertain development of COVID-19 illness a lot of covid-positive patients have to take chest X-rays or CT scans through which health workers can diagnose illness severity and therefore treat the patient accordingly [5]. Main problem occurs in cases of major outbreaks in many different areas in which health system is unable to go through all the data on time that is in many COVID-19 cases crucial.

Having assistance in medical sector, when there is such alarming situation could be of more help and it could go in a way of saving peoples' lives. In that fashion there are many offered solutions, such as detecting COVID-19 without PCR tests, through symptomatic analysis and chest X-rays with models that come from AI world. These proposals offer an automatic analysis of patient's health state [6-7]. Some of the research works could be used in situations where there are problems of identifying pneumonia in covid-positive tested patients, such as pneumonia detection models on chest X-rays, based on deep learning [8-9].

In last years, a sub-discipline in biomedical engineering, called medical imaging provided a potent field for research of AI use in medicine, mainly through Deep Learning method. As for the nature of biomedical engineering that involves around image analysis, Convolutional Neural Networks (CNN) found their field of use, as neural networks that are known for best image processing and classifying [10].

In this paper, main focus is on developing the highly accurate pneumonia recognition model through patient's chest X-ray images. This could be of much help in assisting of diagnosing pneumonia when there are big outbreaks of COVID-19 and it could increase the speed of diagnosing process in these cases. Paper describes a standard problem of classifying images, with the help of Convolutional Neural Networks methods. CNNs are one of the most known artificial neural networks and are designed to automatically and adaptively learn spatial hierarchies through different methods [11]. One of the main goals of this research was to obtain high accuracy of the model while not having an overfitted or underfitted model. The final model had 95,45% accuracy and

This research was supported in part by the EUROCC project, European High-Performance Computing Joint Undertaking (JU) under grant agreement No. 951732. The JU receives support from the EU's Horizon 2020 research and innovation programme and EUROCC project participating institutions.

more than satisfactory results in the rest of the model performance metrics.

II. Materials and Methods

A. Dataset, images of Chest X-rays

The dataset used for this research was found through Mendeley Data website and is provided with the help of Indian Institute of Science, PES University and M S Ramaiah Institute of Technology, Concordia University [12]. The authors of this dataset integrated all of the online available datasets that are related to pneumonia and chest X-Rays, and eliminated duplicates. With the use of Inception V3 architecture and unsupervised learning algorithm, all defective images were removed [12].

This refined dataset consists of 12111 images which are distributed in four folders. Three folders contain images of COVID-19 pneumonia, 3179 images; viral pneumonia, 1807 images; and bacterial pneumonia, 3198 images. The last folder contains healthy lungs X-Ray images, 3927 images. All the images, originally in high quality, were resized to 400x300 resolution and as such made available in the dataset. There was no split made for training or folder classification by normal/pneumonia relation when the dataset was downloaded.

For model training purposes all the X-Ray images that represent pneumonia were put in one folder, so the dataset that would be used for training consisted of the folder named “Pneumonia” and folder named “Normal”, which would then be used for binary classification problem. The example of X-Ray images with their labels is shown in figure below.

A)

B)

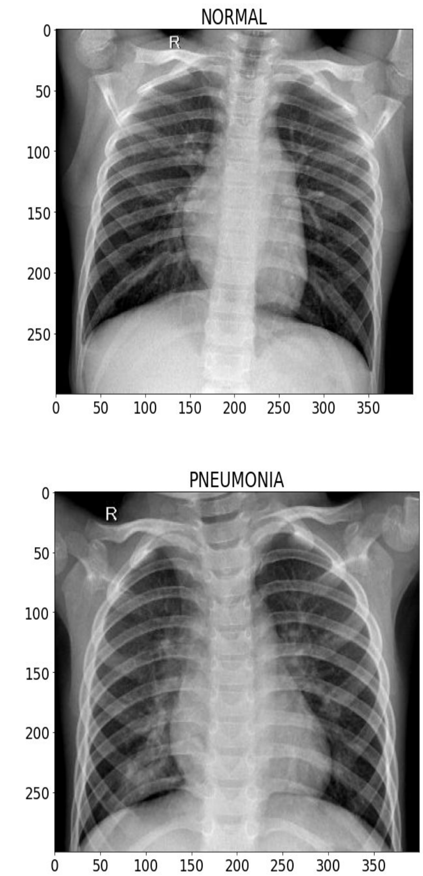

Fig. 1. The X-ray images labeled: A) Chest X-ray shows normal lungs, B) Chest X-ray shows pneumonia lungs

B. Dataset augmentation and pre-processing

The dataset used for training the model had fair amount of data (images) that could be used for training the model without any augmentation. However, as providing more different representations of data to the neural network improves diversibility [22], the dataset was augmented. Augmentation increased the volume to 22110 images. Augmentation was done over Roboflow [23] and the number of images is increased through the random image rotation. All possible rotations were used, horizontal and vertical.

Images were also preprocessed. Images were stretched to 1024x1024. Also, the images were in RGB scale, so the preprocessing step of grayscaling was added. These steps of preprocessing were also done in Roboflow. Next, the pixel values were divided by 255 so that the pixels would be represented with floating points between 0 and 1. It is known that it enhances the performance of CNN [8, 14]. Another preprocessing step was performing the split, which used the usual proportion of 80/10/10, and it was done through code.

C. Tools, Methods and CNN building

The model was built with Python programming languages and TensorFlow framework. Pre-processing and post-processing is performed through NumPy, Pandas and Seaborn libraries. All the training and various experiments with different configurations of the CNN while developing this model were done remotely, on an HPC cluster. Training was finished in 3.5 hours using HPC’s Tesla M60 GPU with 1.1775 GHz of computing power.

As the dataset had inconsistency in the proportion of “Normal” versus “Pneumonia” images. Proportionally bigger weight was assigned to the “Normal” class as it was underrepresented compared to “Pneumonia” class. Weights were used because the training of the model could go in a way of showing bias to the imbalance in the dataset. Therefore, weights were used to force the model to learn from errors it makes during the training process and prevent early overfitting.

With the problem of classifying X-rays by labeling them as “Pneumonia” or “Normal”, CNN deep learning algorithm was chosen. It is one of the best algorithms when it comes to analyzing image and their classification. CNN algorithm does not recognize patterns based on their spatial dependency, instead, it aims to detect any feature regardless of their position in the images [13].

CNN is the type of the Artificial Neural Network (ANN). ANN algorithm consists of multi layers, where the CNN emerged from with adding the convolution mathematics operation between the matrices [13]. Its distinction from other classifying neural networks comes from not fully connected layers, as well as weights sharing concept and less parameters used for training. With those characteristics, CNN improves generalization, its training has less overfitting and runs more smoothly [15]. As for the basic architecture of the CNN, it consists of input and output layers. Between those layers are hidden layers which hold convolutional layers which are the main reason why CNNs are so good at learning the spatial features, as they do the most of calculations. Between the input and output layers, there can be pooling and fully connected layers [15,16]. The way that CNN works is through backpropagation algorithm which is similar to the way the human brain learns through response. With the multi-layered hierarchical structure, CNN manages to extract low, mid, and high-level spatial features. These features represent the level of abstractness. High-level features are a combination of low and mid-level features, and are more abstract. Extraction of these features happens between convolutional layers where neurons of first hidden layer are connected to smaller number of neurons in next hidden layer and so on. This process allows transforming the low-level feature to a higher-level feature in next layers, leading to the final output [16, 17].

When it comes to architecting the CNN used for this research, main tool used was TensorFlow framework. TensorFlow allows sequential model building, and through that method the CNN can be built by following necessary steps. It is important to note that there were many training experiments included through this research. Training experiments conducted, went through many different configurations of the CNN. Different input shapes, different number of hidden layers, number of filters in hidden layers, different number of units for dense layers, batch normalization and dropout techniques combined, as well as different activation functions were all used for these experiments. There were many far more complex CNN models trained, who followed different standards for calculating the preferred output size, preferred layer structure with specific input shapes etc. All these model metrics weren't satisfactory in our case. The model that showed the best overall performance metric is the one described.

In Figures 2 and 3, the summary and graphic representation of Convolutional Neural Network that is presented in this paper are shown.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 62, 62, 16)	160
max_pooling2d (MaxPooling2D)	(None, 31, 31, 16)	0
conv2d_1 (Conv2D)	(None, 29, 29, 16)	2320
max_pooling2d_1 (MaxPooling2	(None, 14, 14, 16)	0
conv2d_2 (Conv2D)	(None, 12, 12, 16)	2320
max_pooling2d_2 (MaxPooling2	(None, 6, 6, 16)	0
flatten (Flatten)	(None, 576)	0
dense (Dense)	(None, 128)	73856
dense_1 (Dense)	(None, 1)	129

Total params: 78,785
Trainable params: 78,785
Non-trainable params: 0

Fig. 2. The model summary of the CNN used in the experiment

Firstly, the input layer took image of 64x64 shape. Then, the image goes through 3 hidden convolutional layers with every layer consisting of 16 filters. After every convolutional layer, there is a pooling layer that downsamples the image. Pooling layers are used in situations when the input data size needs to be reduced, as it could produce a lot of network parameters that could also lead to longer training time and overfitting. Main characteristic of pooling technique is that it downsamples the input image, without losing its important information. There are many pooling methods, but the one used was Max Pooling as the most popular one used for CNN [18]. The activation function that was used is ReLu as it is the most preferred choice when working with hidden layers [19]. The number of perceptrons in fully connected dense layer is 128. This number was determined through training experiments. Models trained with less neurons in dense layer had cases of underfitting or overfitting. Adam optimizer function was used for finalization as it performed the best in some previous works related to this field of research [20]. Binary cross entropy loss function was used to further penalize bad predictions during training [24]. The architecture of this model has some similarities in configuration of the neural network as some of the models presented in work [20].

III. Results and Discussion

As stated previously, the model's training has been early stopped by showcase of overfitting. It resulted in a training that lasted for 105 epochs. Aim was to train the model for 150 epochs. Until 40th epoch, the model showed signs of slight underfitting and after that it adjusted to the learning curve. Loss and validation loss metric were dropping proportionally until 105th epoch when the model started showing signs of overfitting. The figure below shows the loss and validation loss of the model training for all 105 epochs.

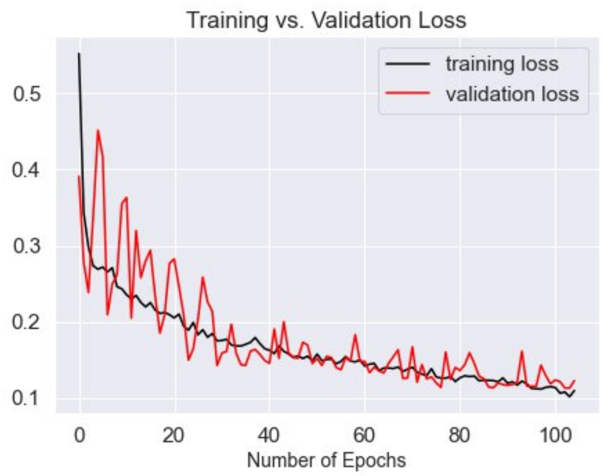

Fig. 4. Training and validation loss during model training process

When it comes to other performance metrics of the model, such as precision, specificity and recall, results are in respectful order: 95,21%, 90,28% and 98,07%. As the classes were imbalanced, there is notable difference in specificity metric compared to other metrics. To summarize the model's classification performance, confusion matrix was used.

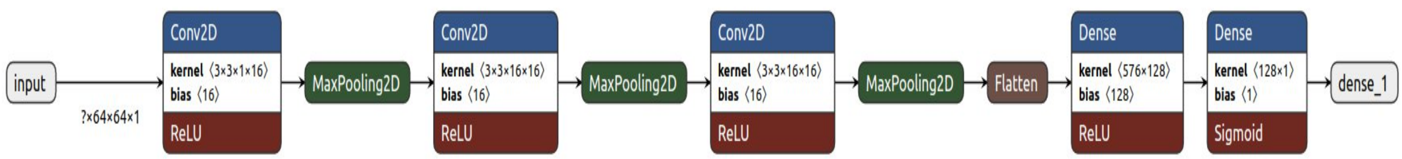

Fig. 3. Architecture of the CNN that is described in this paper

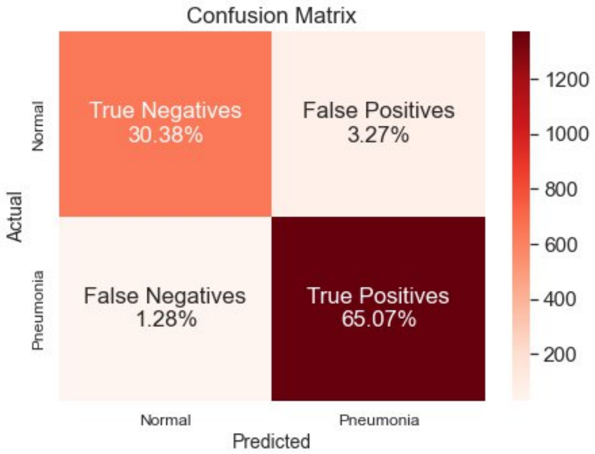

Fig. 5. Confusion matrix representing the performance of the model on the test dataset

As for the interpretation of the confusion matrix, it can be seen that the model is really good in terms of recognizing the pneumonia X-Rays, as it falsely predicted negatives for 1.28% of whole test dataset, achieving more than satisfying metric of recall that is 98,07%. For predicting the negatives, it performed worse, achieving 90,28% specificity. These metrics could be slightly less if the test dataset was more evenly spread across the 2 classes, as we can see that 66,35% of whole test dataset were pneumonia X-rays. For evaluating the all-around strength of this model, ROC curve is used. It shows the performance of the model on all classification thresholds [16]. Area under the ROC curve for this model was 0.9913, which was satisfactory.

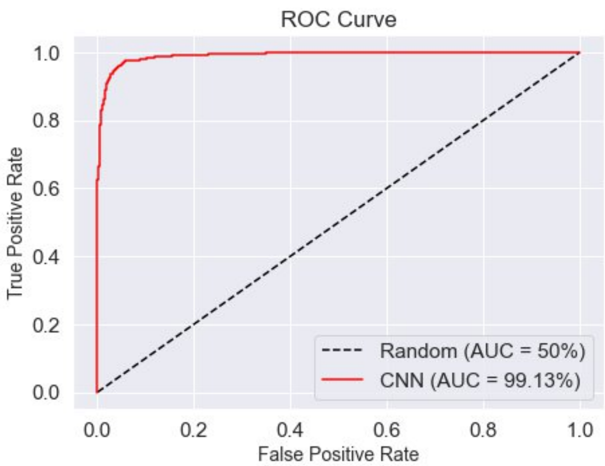

Fig. 6. Visual representation of the strength of the model with the ROC curve on the test dataset

This model's performance was better than the performances of the models in cited works [8,9,20]. But what should be pointed out is that the works cited, used different datasets and different CNN structures [8,9,20], than the ones presented in this research paper.

IV. Conclusion

This paper described use of CNN in classifying images of X-rays related to pneumonia illness. It described a real-life problem related to the COVID-19 pandemic where the pneumonia recognizing models could be of big help in faster diagnosing patients with pneumonia. The process of the development the model, preparing dataset and augmentation was described. With 95% accuracy on test dataset, the models best feature is its ability to recognize pneumonia with 98,07%% accuracy. Even though it has high specificity of 90,28%%, it is its biggest flaw.

The next steps in the research would be to further explore dataset and develop the model which detects different types of pneumonia. Using different classification algorithms and making an comparative analysis with other works are also the motives for further research topics.

References

1. A. Amisha, P. Malik, M. Pathania, V. Rathaur, "Overview of artificial intelligence in medicine", J Family Med Prim Care, 2019, pp. 2328-31.
2. M. Haenlein, A. Kaplan, "A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence", California Management Review, 2019.
3. V. Kaul, S. Enslin, S. Gross, "History of Artificial intelligence in medicine", Review Article, 2020.
4. X. Ouyang et al. "Dual-Sampling Attention Network for Diagnosis of COVID-19 from Community", Acquired Pneumonia, 2020.
5. G. Fagherazzi, C. Goetzinger, M. Rashid, G. Aguayo, L. Huiart,. "Digital Health Strategies to Fight COVID-19 Worldwide: Challenges, Recommendations, and a Call for Papers", J. Med. Internet Res., 2020.
6. U. Helal, M. Hossain, M. Islam, M.A. Zubaer, S. Yang, "Detecting COVID-19 Status Using Chest X-ray Images and Symptoms Analysis by Own Developed Mathematical Model", A Model Development and Analysis Approach. COVID 2022
7. Z. Kahran, F. Akal, "Covid-19 Classification Using Deep Learning in Chest X-Ray Images", 2020.
8. L. Racic, T. Popovic, S. Sandi, S. Cakic, "Pneumonia Detection Using Deep Learning Based on Convolutional Neural Network", 2021 25th International Conference on Information Technology (IT), 2021
9. R. Kundu, R. Das, Z. Geem Woo, G. Han, R. Sarkar, "Pneumonia detection in chest X-ray images using an ensemble of deep learning models", PLoS ONE 16(9):e0256630, 2021.
10. S. Wang et al, "Review and Prospect: Artificial Intelligence in Advanced Medical Imaging", Frontiers in Radiology, 2021.
11. R. Yamashita, M. Nishio, R. Gian Do, K. Togashi, "Convolutional neural networks: an overview and application in radiology", Springner Open, 2018.
12. U. Sait et al., "Curated Dataset for COVID-19 Posterior-Anterior Chest Radiography Images (X-Rays)", Mendeley Data, V3, doi: 10.17632/9xkhgts2s6.3.
13. A. Saad, A. Tareq, "Understanding of a Convolutional Neural Network", 2017 International Conference on Engineering and Technology (ICET), 2017.
14. R. González, R. Woods. "Digital Image Processing", Prentice Hall, 2007, pp. 85.
15. I. Sakshi, G. Anil, S. Mishra, A. Pooja," Conceptual Understanding of Convolutional Neural Network – A Deep Learning Approach", Procedia Compuer Science, vol.132, pp. 679-688, 2018.
16. A. Geron, "Hands-On Machine Learning with Scikit-Learn, Keras,and TensorFlow", O'Reilly Media, Inc., Canada, 2019.
17. K. Asifullah, S. Anabia, Z. Umme, S. Aqsa, "A Survey of the Recent Architectures of Deep Convolutiona Neural Networks", Artificial Intelligence Review, Springer Nation, 2020.
18. G. Hossein, K. Hossein, "Pooling Methods in Deep Neural Networks – a review", ResearchGate (Preprint), 2020
19. S. Sharma, S. Sharma, A. Athaiya, "Activation Functions in Neural Networks", International Journal of Engineering Applied Sciences and Technology, 2020 Vol. 4, Issue 12, ISSN No. 2455-2143, pp. 310-316, 2020.
20. N. Anand, J. Rachna, "Pneumonia Detection Using Convolutional Neural Networks", Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019) , pp 471–483, 2020.
21. WHO Coronavirus (COVID-19) dashboard, Available online: https://covid19.who.int/ , Last accessed: March 28, 2022.
22. C. Shorten, T.M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning", J Big Data, doi: 10.1186/s40537-019-0197-0, 2019.
23. Roboflow online platform, available: https://roboflow.com/, Last accessed: June 10, 2022.
24. A.U. Ruby et al., "Binary cross entropy with deep learning technique for Image classification", doi: 10.30534/ijatcse/2020/175942020, 2020.