Automatic detection for COVID-19 and Pneumonia from chest X-Ray images using convolutional neural networks

Andrew Muhsen*, Doaa Salah**, Habiba Ahmed ***, Zainab Emam ****
System and Biomedical Engineering Department, Faculty of Engineering, Cairo University, Giza, Egypt
*Andrew.naaem99@eng-st.cu.edu.eg** doaa.ibrahim00@eng-st.cu.edu.eg *** habiba.abdelaal00@eng-st.cu.edu.eg **** zainab.gad00@eng-st.cu.edu.eg

Abstract—An automatic disease detection framework assists doctors in the diagnosis of disease, provides fast results, and reduces the death rate. X-Ray machines are available in most hospitals. Therefore, an automated detection system is the fastest diagnostic option. We introduce a deep learning technique to diagnose COVID-19 and Pneumonia automatically from chest X-ray images. two deep learning models are built, one based on the combination of a convolutional neural network (CNN) and long short-term memory (LSTM) and the other based on (CNN) only. A collection of 4575 chest X-ray images, including 1525 images for each class. The results show that CNN model achieved better detection than CNN with LSTM.

Keywords— *COVID-19, Pneumonia, Deep learning, Chest X-ray, Convolutional neural network, long short-term memory*

I. INTRODUCTION

A new coronavirus called COVID-19 was identified as a disease that can cause catastrophic heart infections, dangerous respiratory issues, and even death. As a result, on January 30, 2020, The World Health Organization labeled this epidemic a Public Health Emergency of Worldwide Concern [1]. Despite a shortage of infrastructure, poor healthcare systems, and subpar diagnostic tools, all countries are involved in the fight against this tragic epidemic.

COVID-19 symptoms might vary from a cold and cough to a fever, shortness of breath, and acute respiratory syndrome [2]. One of the most often utilized assays for COVID-19 detection is real-time reverse transcription polymerase chain reaction (rRT-PCR), which extracts DNA via reverse transcription and then utilizes PCR to amplify DNA for analysis. While COVID-19 exclusively contains RNA, it can detect it [3]. Nevertheless, the RT-PCR technique takes 4-6 hours to complete and the test kit is not commonly available and quite costly. It is also untrustworthy. CT scans and chest radiography image analysis may be useful in this scenario. [4] The sensitivity of CT scans for the COVID-19 infection rate was around 98%, compared to 71% for RT-PCR, [5] but chest X-ray radiography is more often utilized in clinical practice owing to its benefits, including cheap cost, low radiation dosage, easy-to-operate, and broad accessibility in general or community hospitals [6].

This paper compares two deep learning-based system models: CNN and a fully connected layer neural network, while the other combines CNN, LSTM, and fully connected layer neural networks to automatically identify COVID-19 from X-ray pictures.

II. LITERATURE REVIEW

In recent years, the use of deep learning techniques, particularly convolutional neural networks (CNNs), has shown promising results in automated COVID-19 detection from chest X-ray images. Several studies have been conducted to investigate the effectiveness of CNNs in detecting COVID-19 from chest X-ray images. In a research [7], the authors introduced an automated detection scheme named EMCNet was proposed to identify COVID-19 patients by evaluating chest X-ray images. They used an ensemble of machine learning classifiers which ensures better results for the dataset of various sizes and resolutions. EMCNet has achieved 98.91% accuracy, 100% precision, 97.82% recall, and 98.89% F1 score.

In another study [8], the authors used fine-tuned deep-learning architectures which have been made to speed up the detection and classification of COVID-19 patients from other pneumonia ones. The models used are MobileNetV2, ResNet50, InceptionV3, DenseNet121, InceptionResNetV2, NASNetMobile, VGG16, Xception. DenseNet121 has obtained best accuracy of 97%.

In a study [9], authors proposed a system where nine different pre-trained models extracted features from chest X-ray images. The system used SVM to classify COVID-19 using extracted features. Among all the models, ResNet50 was considered best for feature extraction. The system obtained an accuracy of 95.33% and an F1-score of 95.34%.

In the study [10], the authors introduced a transfer learning strategy with CNN that makes an automatic diagnosis of COVID-19 cases by feature extraction from chest X-rays. The system used the five CNN variants VGG19, Inception, MobileNet, Xception, and Inception-ResNetV2 to classify COVID-19 images. Among the pre-trained models, MobileNetV2 achieved 96.78% accuracy, 98.66% sensitivity, and 96.46% specificity.

In another study [11], the authors have introduced a CNN-based model namely CoroNet, which uses Xception architecture for Covid-19 detection. The model has achieved 89.6 % accuracy, 93% precision, and 98.2% recall. Another method of Covid-19 detection has been proposed in research [12], the authors proposed a neural network that is a concatenation of Xception and ResNet50V2 networks. This network achieved the overall average accuracy for all classes which is 91.4%.

In another study [13], the authors proposed a new deep CNN model DeTraC. The system simplified the local structure of the dataset by applying the class decomposition layer. The training of the system was performed using the pre-trained ResNet model. The model achieved 95.12% accuracy, 97.91% sensitivity, and 91.87% specificity. In a study [14], the authors proposed a deep learning model called COVID-Net, which is designed to detect COVID-19 cases from chest X-ray images. The model is based on a deep CNN architecture and uses a tailored design to optimize its performance. The study achieved 92.4% accuracy in COVID-19 detection.

III. DATASET DESCRIPTION

Deep learning models are data-driven. The training of a deep learning model requires a large amount of data. The dataset used in this analysis contains chest X-ray images of patients with reported COVID-19 disease, Pneumonia, and Normal cases. The dataset that is used in these two deep learning models is taken from the dataset of a paper [15,16]. The dataset contains 4575 chest X-ray images. These images are 1525 COVID-19 cases, 1525 Pneumonia cases, and 1525 Normal cases, according to the paper[15] covid samples were rare in which some of the images were obtained using data augmentation. The Input shape of our models is the universally known ImageNet shape 224 x 224.

IV. METHODOLOGY

1. Combined CNN with LSTM model

CNN-LSTM is a hybrid deep learning architecture that combines convolutional neural networks (CNNs)[14] which are learning techniques for computer vision applications and long short-term memory (LSTM) which is a specialized RNN used for processing sequential data networks. This combination is specifically designed to handle spatiotemporal data, that can selectively store and retrieve information over extended periods, the CNNs are used to extract spatial features from the data, while LSTMs are used to model temporal dependencies. The CNN-LSTM architecture has been applied to various applications, such as video analysis and action recognition, and has shown improved performance over traditional methods, and we're going to utilize this decent architecture for Covid-19 detection from a chest x-ray image Having around 17.8 million parameters our intended model is in Fig. 1.a consists of 5 convolutional blocks, representing the feature extractor, the output of the feature extractor is noticed to be 7 x 7 x 512, which needs to be reshaped to be an input for the LSTM layer. A reshape layer was used to convert these 7 x 7 into 1D vectors each with a length of 49, which is fed to the LSTM layer as sequences to return another sequence to be fed to a classifier layer with 3 neurons to predict a class from all the classes in the dataset (COVID-19, pneumonia, and normal)

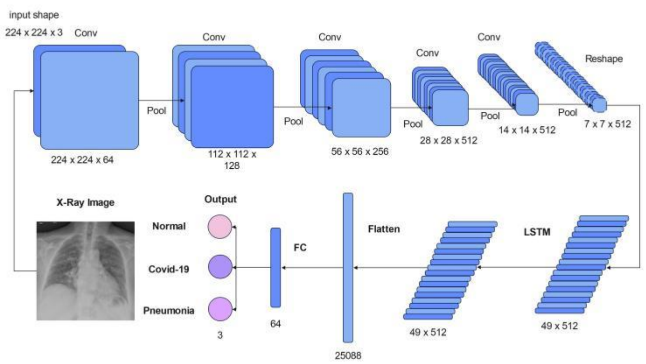

Fig 1. a CNN-LSTM architecture

The architecture used Relu as the activation function in all layers that required activation except for the output layer, using the Sigmoid function. A dropout layer was included after the FC layer with a drop ratio of 0.15, and batch normalization layers were inserted after each Max-pooling layer to improve the model's generalization. The LSTM layer was trained with a reduced risk of vanishing or exploding gradients. Additionally, the Convolutional layers were initialized with the L2 kernel regularizer for improved performance.

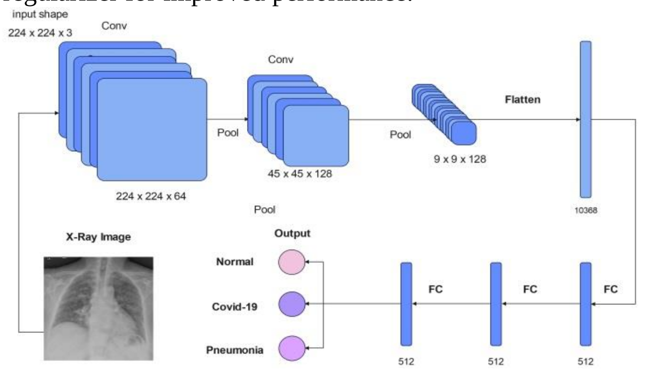

Fig 1. b CNN architecture

2. CNN model

The CNN Model depicted in Fig 1. b having around 7.8 million parameters comprises a feature extractor consisting of 2 convolutional blocks and a Multi-layer Perceptron (MLP) as a classifier with 3 FC layers. Each block is made up of 4 2d convolutional layers, max-pooling, batch normalization, and a dropout layer with a drop ratio of 0.3.

Table 1: The full summary of the CNN-LSTM network.

Layers	output shape	filters/neurons	Kernel size
Convolutional 2D	224 x 224 x 64	64	3 x 3
Convolutional 2D	224 x 224 x 64	64	3 x 3
Max pooling 2D	112 x 112 x 64	-	2 x 2
Convolutional 2D	112 x 112 x 128	128	3 x 3
Convolutional 2D	112 x 112 x 128	128	3 x 3
Max pooling 2D	56 x 56 x 128	-	2 x 2
Convolutional 2D	56 x 56 x 256	256	3 x 3
Convolutional 2D	56 x 56 x 256	256	3 x 3
Max pooling 2D	28 x 28 x 256	-	2 x 2
Convolutional 2D	28 x 28 x 512	512	3 x 3
Convolutional 2D	28 x 28 x 512	512	3 x 3
Convolutional 2D	28 x 28 x 512	512	3 x 3
Max Pooling 2D	14 x 14 x 512	-	2 x 2
Convolutional 2D	14 x 14 x 512	512	3 x 3
Convolutional 2D	14 x 14 x 512	512	3 x 3
Convolutional 2D	14 x 14 x 512	512	3 x 3
Max Pooling 2D	7 x 7 x 512	-	2 x 2
reshape	49 x 512	-	-
LSTM	49 x 512	512	-
Flatten	25088	-	-
Dense	64	64	-
Dense - output	3	3	-

Table 2: The full summary of the CNN network.

Layers	output shape	filters/neurons	Kernel size
Convolutional 2D	224 x 224 x 64	64	5 x 5
Convolutional 2D	224 x 224 x 64	64	5 x 5
Convolutional 2D	224 x 224 x 64	64	5 x 5
Convolutional 2D	224 x 224 x 64	64	5 x 5
Max pooling 2D	45 x 45 x 64	-	5 x 5
Convolutional 2D	45 x 45 x 128	128	5 x 5
Convolutional 2D	45 x 45 x 128	128	5 x 5
Convolutional 2D	45 x 45 x 128	128	5 x 5
Convolutional 2D	45 x 45 x 128	128	5 x 5
Max Pooling 2D	9 x 9 x128	-	5 x 5
flatten	10368	512	-
Dense	512	512	-
Dense	512	512	-
Dense	512	512	-
Dense - output	3	3	-

The CNN model convolutional layers were initialized with an L2 kernel regularizer with a factor of 0.0001. LeakyRelu activation functions were used for all layers that required them, with SoftMax activation used for the output layer. The extracted features are flattened and passed to an MLP classifier with 3 hidden and an output layer with 3 neurons representing the classes.

V. RESULTS

A. Metrics

To evaluate the performance of our models, we utilized several metrics, including precision, recall, and F1 score To provide a more comprehensive view of model performance than just loss and accuracy metrics. By using a combination of metrics, we were able to better assess the accuracy of our models over the long term.

$$precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN}, F1\ score = \frac{2 * TP}{2 * TP + FP + FN}$$

B. Experiment Setup

The dataset was divided into 80% for training and 20% for testing. SGD optimizer was used with a learning rate of 0.0001 and a decay rate of 1e-6. The maximum number of epochs was set to 125 with a batch size of 24 images, but an early stopping criterion was employed in the experiments with a tolerance of 20 epochs. Both models stopped before the 50^{th} epoch with remarkable performance, The experiments were conducted using Kaggle's Free GPU P100 with 16 GPU RAM, taking the CNN model 82 seconds per epoch, and 172 seconds per epoch for the CNN-LSTM model

C. analysis and comparison

The chosen loss metric for both models was the Categorical Cross Entropy (CE) loss. The CNN showed better performance in terms of CE loss with 0.14 on training and 0.3 on validation, while the CNN-LSTM had a higher loss of 0.41 on training and 0.56 on validation, Also performing much better the CNN model had a training accuracy of 97.6% and validation accuracy of 95.41% while the CNN-LSTM had a training accuracy of 98.88% and validation accuracy of 94.31% which indicated the effect of larger convolutional kernels in the CNN model.

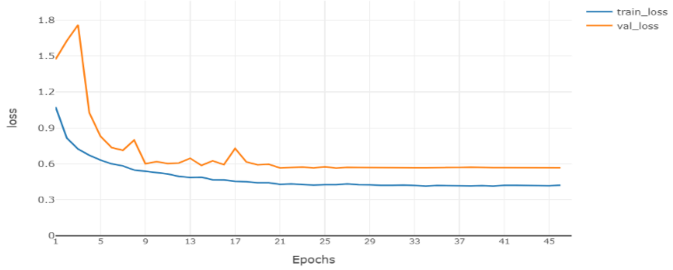

Fig 2. a – CNN-LSTM model loss

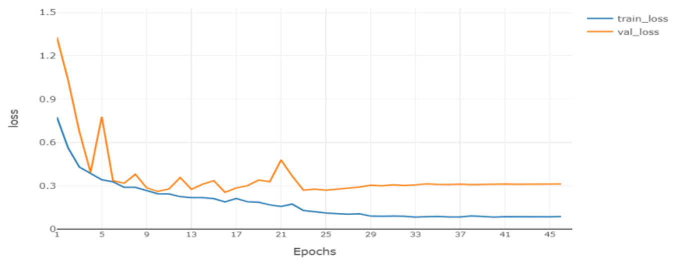

Fig 2. b – CNN model loss

As we previously mentioned the duo to limitations of using only loss and accuracy as metrics for evaluating models. In Fig 4. a and Fig 4. b, we presented other metrics to evaluate the performance of the CNN-LSTM and CNN models, respectively.

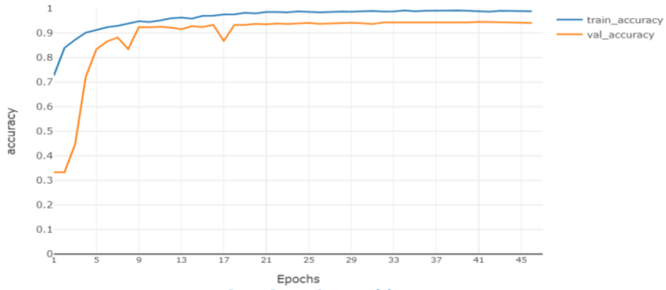

Fig 3. a – CNN-LSTM model accuracy

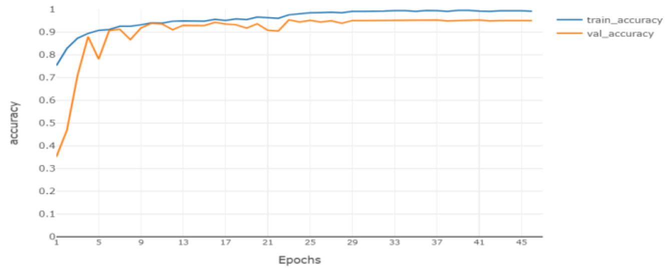

Fig 3. b – CNN model accuracy

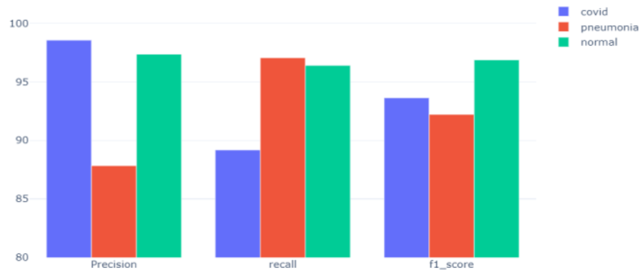

Fig 4. a – CNN-LSTM model precision, recall, and F1 score

Table 3: Performance of the CNN- LSTM model.

	precision	recall	f1-score
COVID19	98.55 %	89.18 %	93.63 %
PNEUMONIA	87.83 %	97.05 %	92.21 %
NORMAL	97.35 %	96.39 %	96.87 %

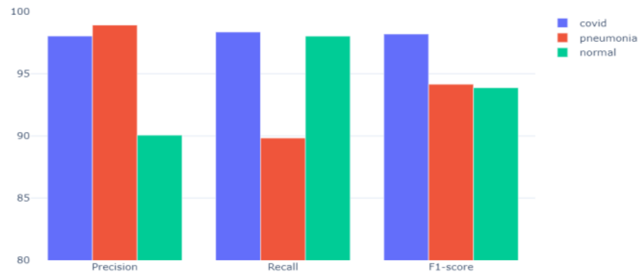

Fig 4. b – CNN model precision, recall, and F1 score

Table 4: Performance of the CNN model.

	precision	recall	f1-score
COVID19	98.04 %	98.36 %	98.20 %
PNEUMONIA	98.91 %	89.83 %	94.15 %
NORMAL	90.06 %	98.03 %	93.87 %

The CNN-LSTM model had a slightly higher precision of 98.55%, but a much lower recall of 89.18% and an F1 score of 93.63%, while the CNN model outperformed with a high precision of 98.03% with a higher recall of

98.36% and an F1 score of 98.19%. To support our results we also plotted a confusion matrix for each model performance, shown in fig 5. a for CNN-LSTM and fig 5. b for CNN, both models showed great performance in differentiating between covid and normal cases which is the main of the experiment, on the hand both quietly failed to differentiate well between pneumonia cases and normal cases, which demonstrates the major common features between both classes.

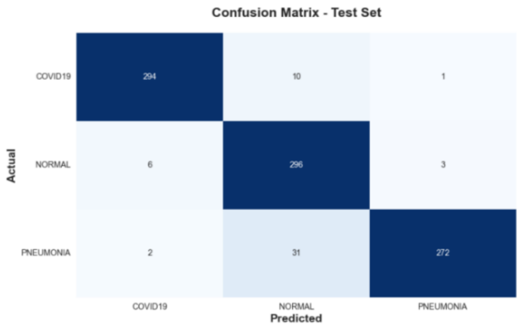

Fig 5. a – CNN-LSTM model Confusion matrix

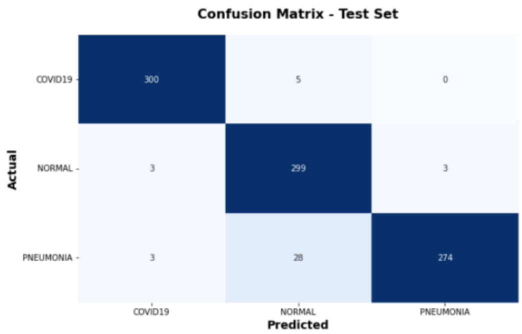

Fig 5. b – CNN model Confusion matrix

We also added the receiver operating characteristic (ROC) curve which is a great metric for multi-classification problems, to measure the performance of each model by the area under the ROC curve in fig 6. a for the CNN-LSTM model and fig 6. b for the CNN model.

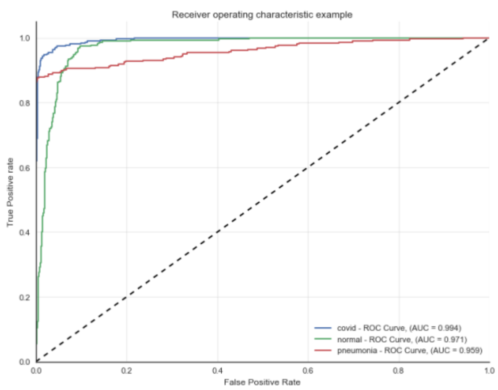

Fig 6. a – CNN-LSTM model ROC

VI. CONCLUSION AND FUTURE WORK

From the previously shown results, it's safe to assume that the CNN model performed slightly better, having larger convolutional kernels, it was able to distinguish between covid and non-covid effectively, on the other side, the CNN-LSTM showed to be to distinguish pneumonia with slightly higher accuracy than the CNN model, Although the High accuracy of both models, on such a heavily augmented dataset, it was hard for both not to misclassify when it came to pneumonia vs normal cases, in which both models detected pneumonia cases as healthy.

The future of the project is to push models' performance through parameter hyper tuning while applying the same analogy to CT scan images as they are more effective than X-ray images and finalize our work with a graphical user interface for testing and follow-ups.

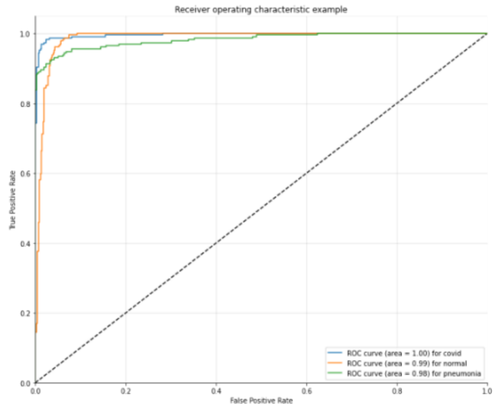

Fig 6. b – CNN model ROC

REFERENCES

[1] Culp WC. Coronavirus disease 2019. A & A Pract 2020;14:e01218.

[2] Zu ZY and M.D. Coronavirus disease 2019 (COVID-19): a perspective from China. Radiology 2020;296(2):E15–25..

[3] Corman VM, O.L, M. K and R.M. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill 2020;25(3).

[4] Xia W, Shao J, Guo Y, Peng X, Li Z, Hu D. Clinical and CT features in pediatric patients with COVID-19 infection: different points from adults. Pediatr Pulmonol 2020;55(5):1169–74.

[5] Fang Y, Zhang H, Xie J, Lin M, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR.

[6] A. Narin, C. Kaya, Z. Pamuk, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, arXiv preprint arXiv: 2003.10849 (2020).

[7] Prottoy Saha , Muhammad Sheikh, Md. Milon . EMCNet: Automated COVID-19 diagnosis from X-ray images using convolutional neural network and ensemble of machine learning classifiers.

[8] Sonam Aggarwal, Sheifali Gupta, Adi Alhudhaif, Deepika Koundal, Rupesh Gupta, Kemal Polat. Automated COVID-19 detection in chest X-ray images using fine-tuned deep learning architectures.

[9] Sethy PK, Behera SK, Ratha PK, Biswas P. Detection of coronavirus disease (COVID- 19) based on deep features and support vector machine.

[10] Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks.

[11] Asif Iqbal Khan, Junaid Latief Shah , Mohammad Mudasir Bhat. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images.

[12] Rahimzadeh M, Attar A. A new modified deep convolutional neural network for detecting COVID-19 from X-ray images.

[13] Abbas A, Abdelsamea MM, Gaber MM. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network.

[14] Wang, L., Lin, Z. Q., & Wong, A. (2020). COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images.

[15] Md. Zabirul Islam, Md. Milon Islam, Amanullah Asraf. A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray image.
https://doi.org/10.1016/j.imu.2020.100412

[16] Chest X-ray images. https://data.mendeley.com/datasets/jctsfj2sfn