A Review of CNN on Medical Imaging to Diagnose COVID-19 Infections[*]

Samuel Clark[1, 2], Ehsan Kamalinejad[2], Christian Magpantay[1],
Suritaneil Sahota[1, 3], Jiaofei Zhong[1], and Yanke Hu[4]

[1]Department of Computer Science, California State University East Bay, Hayward, CA
[2]Department of Mathematics, California State University East Bay, Hayward, CA
[3]Department of Biological Science, California State University East Bay, Hayward, CA
[4]Evomics Medical LTD., Shanghai, China
sclark40@horizon.csueastbay.edu, ehsan.kamalinejad@csueastbay.edu,
cmagpantay2@horizon.csueastbay.edu, ssahota7@horizon.csueastbay.edu,
jiaofei.zhong@csueastbay.edu, yanke.hu@evomics.net

Abstract

In this paper, we study the *Convolutional Neural Network* (CNN) applications in medical image processing during the battle against *Coronavirus Disease 2019* (COVID-19). Specifically, three CNN implementations are examined: CNN-LSTM, COVID-Net, and DeTraC. These three methods have been shown to offer promising implications for the future of CNN technology in the medical field. This survey explores how these technologies have improved upon their predecessors. Qualitative and quantitative analyses have strongly suggested that these methods perform significantly better than the commensurate technologies. After analyzing these CNN implementations, it is reasonable to conclude that this technology has a place in the future of the medical field, which can be used by professionals to gain insight into new diseases and to help in diagnosing infections using medical imaging.

Keywords: *CNN, COVID-19, image processing, X-ray*

1 Introduction

Prior to the pandemic of *Coronavirus Disease 2019* (COVID-19), there have been two epidemic level coronavirus diseases in the world, i.e., *Severe Acute Respiratory Syndrome* (SARS), and *Middle East Respiratory Syndrome* (MERS) [1]. Since the first case of COVID-19 was identified about one

---

and a half years ago, there have been over 200 million confirmed cases of COVID-19, including more than 4 million deaths worldwide, according to the World Health Organization.

It has been possible to confirm or diagnose the COVID-19 infections through the *Reverse Transcription-Polymerase Chain Reaction* (RT-PCR) process, in which it would take 4 to 6 hours to get the results [1]. This may be a concern as it only takes a significantly less amount of time such as 15 minutes to infect another with COVID-19. It is important to get the test results as fast as possible so that the patient may proceed to further treatments or be able to quarantine. Previous studies [1, 2, 3, 4, 5, 6, 7] have shown that radiological images such as X-rays, *computed tomography* (CT), and *magnetic resonance imaging* (MRI), could be used to detect abnormalities in patients and diagnose COVID-19 infections. We further investigate the deep learning and CNN methods for image classification to detect the abnormalities in the chest X-ray images, so that it may be possible to diagnose COVID-19 faster than several hours to almost instantaneously after the patient has been X-rayed.

This survey explores several CNN technologies with unique implementations that enhance their performance when compared to conventional CNN architectures. In this study, we try to find out answers to the following questions: (1). What CNN architectures are best suited for diagnosing COVID-19 from X-rays? (2). What possible improvements can be made to the existing CNN technologies for tackling COVID-19? (3). How effective can we expect CNN to process medical imaging such as chest X-rays? (4). Will there be better types of neural networks to process images? Therefore, we focus on three of the state-of-the-art CNN methods for medical image processing in diagnosing COVID-19 infections in this paper.

2   Related Work

Slowing down the spread of the COVID-19 is one major way to combat this pandemic. The analysis of medical imaging, more specifically, chest X-ray images and CT scan images, can be utilized as an exceedingly important tool in the diagnosis of COVID-19 infections. Machine learning approaches have been employed for the medical image processing applications, such as identifying cancerous and non-cancerous from digital images [8]. Deep learning approaches such as CNN are not only rivaling but also surpassing the typical machine learning approach.

The CNN architecture has proven itself to be highly applicable in the medical image applications, specifically in making diagnoses using digitized medical images, as it is the effective and efficient deep learning approach, due to the success shown in previous medical imaging studies [9]. With the help of CNN and the related approaches, it is faster to detect the abnormalities in patients based on medical images to diagnose COVID-19 infections, which will aid in preventing the spread of the COVID-19 at a much more accelerated rate through these tools.

One method that is of interest to introduce is using CNN and the *Long Short-Term Memory* (LSTM) in combination to achieve better and refined results, which is named deep CNN-LSTM network [1]. In this CNN-LSTM architecture, LSTM is a modified and improved version of the *recurrent neural networks* (RNNs). The proposed CNN-LSTM network applies CNN to extract features, and then utilizes LSTM network as a classifier to detect COVID-19.

Among the motivations for the neural networks is the necessity for a greater degree of auditing than is required for neural network applications in other industries. With the ability to audit decisions made by CNN, more accurate networks which are influenced less from extraneous data can be developed. COVID-Net [2] is a response to this demand for thorough auditing of CNN technologies in medicine. As a deep CNN tailored for the detection of COVID-19 cases from chest X-ray images, COVID-Net is open source and available to the general public, along with COVIDx, an open access benchmark dataset comprising of 13,975 chest X-ray images across 13,870 patient cases.

Another research study adapts a deep CNN, called *Decompose, Transfer, and Compose* (DeTraC) model [3], to improve the performance of the pre-trained models on the classification of COVID-19 chest X-ray images. DeTraC is able to detect irregularities in the image dataset through inspecting its class boundaries by a class decomposition method. DeTraC shows effective solutions for the classification of COVID-19 cases and its ability to cope with irregular data and limited number of training images.

3 Method

In this section, we focus on three state-of-the-art methods of medical image processing based on CNN in classifying chest X-ray images of COVID-19 infections.

3.1 CNN-LSTM

The CNN-LSTM method is proposed to diagnose COVID-19 cases based on X-ray images, where LSTM stands for long short-term memory. The reason for using LSTM is because it provides memory cells as opposed to the normal RNN which helps resolve any issues pertaining to the vanishing and exploding gradient problem. The CNN process works in the following way that kernels receive the input and convolve them using stride but no padding, because it is needed to maintain the dimensions of the input volume with ordinary features. The convolution layer operation is as follows:

$$F(i, j) = (I * K) (i, j) = \Sigma \Sigma I (i + m, j + n) K (m, n) \quad (1)$$

where I is the input matrix, K is the filter of size $m \times n$, and F is the output of the feature map. The convolution layer is $(I * K)$ [1]. Then, max pooling is used for the pooling layers and the ReLU activation function is used to increase nonlinearity in the mapping [1]. This is how the CNN process follows until it reaches into the LSTM network.

LSTM, a modified version of RNN, is combined with an update gate, a forget gate, and an output gate [1]. Its main difference is that the LSTM model has a memory cell in which it can save long-term. Thus, An LSTM network can save and link the information obtained previously to the data obtained in the present [10]. With the combined structure, CNN-LSTM is used to detect COVID-19. It is composed of 12 convolution layers, 5 pooling layers, 1 fully connected layer, one LSTM layer, and one output layer with the softmax function. Figure 1 shows the system architecture of the CNN-LSTM [1].

The system goes into different phases in which the combinatorial neural network may detect COVID-19 within the X-rays. First, raw X-ray images entered a preprocessing pipeline which involves data resizing, shuffling, and normalization. Next, the preprocessed data is then divided into the training set and testing set, when the training set is used to train the CNN, for comparison, and CNN-LSTM.
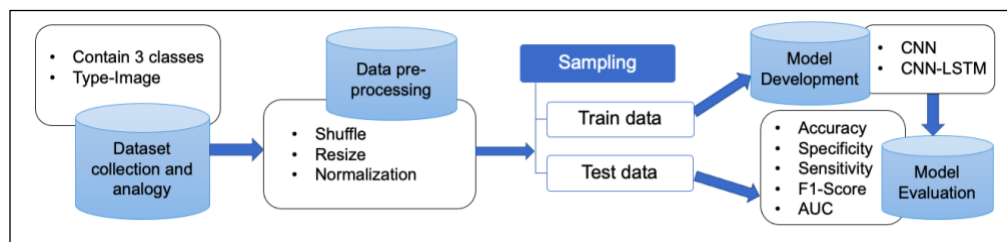

Figure 1: System architecture of the CNN-LSTM

3.2   COVID-Net and Enhanced Auditing of CNNs

The nature of medicine and medical imaging demand more extensive auditing than what would be suitable for other CNN applications. It is crucial for these CNNs to make decisions based on relevant information. The goal of "explainability"-focused auditing is to ascertain that the CNN's decisions are not influenced by extraneous data in the medical images, such as markup editing and image distortions. To accomplish this, the developers of COVID-Net [2] used an analysis tool known as GSInquire [11], which enables COVID-Net developers to explore different architectures and determine the optimal design through a process known as "generative synthesis" [12].

GSInquire uses a generator-inquisitor pair of functions that work together to form an interpretation of a decision. The generator function creates new neural networks based on a set of possible seeds. Then, the inquisitor function takes these new networks and determines which parameters influence decisions. This allows for CNNs to be audited to ensures that the decisions are based on relevant information. Figure 2 shows the critical factors identified by GSInquire in sample chest X-ray image of COVID-19 positive case [2]. According to the interpretation by GSInquire, the COVID-Net mainly focuses on the highlighted areas of the lungs in the chest X-ray images, which are considered as the main critical factors in determining whether the image is of a patient with COVID-19 infections [2].

Figure 2: Highlighted chest X-ray of COVID-19 case

The architecture chosen for COVID-Net utilizes a *projection-expansion-projection-extension* (PEPX) design pattern. This architecture features a series of steps which project an input feature into a smaller dimension, expand it into a higher dimension, followed by a depth-wise representation step for learning spatial characteristics, and an extension step, which extends channel dimensionality, for producing final features. The resulting neural network named COVID-Net has significant architectural diversity and strong representational capacity. This is partly due to the fine level of granularity that generative synthesis permits architectural designs to be explored.

Among other possibilities, COVID-Net has the potential to assist medical professionals in identifying key factors of COVID-19 infection. This could lead to novel discoveries and more efficient triaging. Furthermore, the considerable auditing of this CNN could serve to increase our medical community's trust in using COVID-Net and other CNNs as a tool. This auditing methodology not only served to increase the integrity of COVID-Net, but also resulted in an efficient and powerful architecture.

3.3   DeTraC

A major contributor to the success of CNN in the field of medical image processing is due to its capability of transfer learning in this architecture. Considering a network that was pre-trained, with a large dataset that was annotated with given metadata for a specific task, transfer learning means

transferring the knowledge learned from the pre-trained network into a new task, without the need to train the new task with a massive amount of annotated data values used as part of the training, which is especially useful and efficient for medical image processing. There are three different scenarios of transfer learning, i.e., shallow tuning, deep tuning, and fine tuning, where the third one provided excellent results for classifying X-ray images.

DeTraC [3] utilizes the fine-tuning scenario that seeks to teach more layers, in a step-by-step manner, adjusting the learning parameters in order to increase the efficiency and effectiveness. Class decomposition is a major component of the DeTraC method. Its ability to deal with inconsistencies in data is an invaluable and critical asset when dealing with medical images.

The DeTraC architecture is composed of three integral phases [3]. In phase 1, it is important to identify the deep local features from each image, which is achieved by further training an ImageNet pre-trained CNN model. Following that step is the class decomposition component that is employed for its ability to break down and simplify the local structure. After that, phase 2 of the DeTraC architecture uses an advanced gradient descent optimization algorithm in order to complete the training of the network. Finally in phase 3, the class composition layer is used to refine the classification of the medical images and generate the final prediction. In short, the DeTraC method functions by first taking a functional CNN architecture that has already been trained, then applying class decomposition, followed by the transfer learning via a gradient descent optimization method, finally a class composition layer to form the fine-tuned qualifiers. The class decomposition and class composition components can be thought of in a recursive manner. That is to say, the decomposition will break the problem into smaller independent subclasses, and the composition component will construct all those subclasses back to the final stage.

4   Performance Evaluation

In this section, we compare the system performance of the aforementioned three CNN-based methods in classifying medical images to diagnose COVID-19 infections.

4.1   Results for CNN-LSTM

The system performance of CNN-LSTM is measured with these metrics: confusion matrix, accuracy, AUC using ROC, specificity, sensitivity, and F1-score [1]. In the experiment setup, the data was divided into 80% for training and 20% for testing, using the 5-fold cross-validation technique for the results to be attained. The learning rate is 0.0001 and the maximum epoch number is 125. Python and the Keras package were used with Tensorflow to implement the CNN and CNN-LSTM networks on an Intel(R) Core (TM) i7-2.2 GHz processor and a (GPU) NVIDIA GTX 1050 Ti with 4 GB and 16 GB RAM [1].

The images of X-rays were tested first with CNN and then CNN-LSTM after, so that a comparison can be made on the difference of which network would perform better. According to the study in the previous work [1], with the CNN architecture, among the 915 images used to classify COVID-19, 14 images were not properly classified with three of them for COVID-19. On the other hand, using the CNN-LSTM architecture, only 8 were misclassified including two images that were given with COVID-19. This shows that the CNN-LSTM does in fact, performs better than the standalone CNN. The comparisons of the performance of the CNN and CNN-LSTM architectures are shown in Table 1 and Table 2 [1].

According to the experimental results, CNN-LSTM architecture is able to distinguish the COVID-19 images from pneumonia and normal images. When compared with other systems through 4575 images, including 1525 images of COVID-19 infected patients, the experiments showed that CNN-

LSTM achieves outstanding performance with 99.4% accuracy as in Table 3 [1], with 99.9% AUC, 99.2% specificity, 99.3% sensitivity, and 98.9% F1-score.

Class	Accuracy (%)	Specificity (%)	Sensitivity (%)	F1-Score (%)
COVID-19	98.5	98.2	99.0	97.7
Pneumonia	98.6	99.7	96.4	97.8
Normal	99.9	99.8	100.0	99.8

Table 1: Performance of the CNN network

Class	Accuracy (%)	Specificity (%)	Sensitivity (%)	F1-Score (%)
COVID-19	99.2	99.2	99.3	98.9
Pneumonia	99.2	99.8	98.0	98.8
Normal	99.8	99.7	100.0	99.7

Table 2: Performance of the CNN-LSTM network

Architecture	Accuracy (%)	Accuracy (COVID-19) (%)
COVIDiagnosis-Net [16]	98.3	100.0
Sgdm-SqueezeNet [17]	98.3	96.7
ResNet-50 [18]	96.2	100.0
CNN-LSTM [1]	99.4	99.2

Table 3: Comparison of CNN-LSTM with other systems

4.2 Results for COVID-Net

The performance of COVID-Net was evaluated and analyzed both quantitatively and qualitatively. In the pre-processing step, the chest X-ray images were cropped before training to cut the embedded textual information in the X-ray images. Data augmentation was employed including translation, rotation, horizontal flip, zoom, and intensity shift. Two more deep neural network architectures were chosen to compare against COVID-Net. VGG-19 [13] was one of them, with low architectural diversity but without residual design principles and lightweight design patterns. ResNet-50 [14] is another architecture, with moderate architectural diversity, residual design principles and lightweight design patterns, but without lightweight PEPX design patterns and selective long-range connectivity.

Architecture	Accuracy (%)	Sensitivity (COVID-19) (%)	Positive predictive value (COVID-19) (%)
VGG-19	83.0	58.7	98.4
ResNet-50	90.6	83.0	98.8
COVID-Net	93.3	91.0	98.9

Table 4: Performance of COVID-Net compared with other systems

COVID-Net has promising performance compared to other architectures. When tested with the same dataset of COVIDx, COVID-Net achieved 93.3% accuracy, whereas ResNet-50 obtained 90.6% accuracy, and VGG-19 obtained 83.0% accuracy, as shown in Table 4 [2]. COVID-Net not only outperforms these other two networks, but also has less complexity compared to ResNet-50 and VGG-19. Thus, it requires significantly less computations than the other two architectures.

4.3   Results for DeTraC

For the DeTraC method, the results were based on the assessment of three main criteria: accuracy, specificity, and sensitivity. Accuracy is defined as the total sum of true positive cases of COVID-19 and true negative cases of non-COVID-19 (normal or other diseases), all divided by the number of cases. Specificity is defined as the true negative cases of non-COVID-19 divided by the total sum of true negative cases of non-COVID-19 and the incorrect predictions for COVID-19 or other cases. Sensitivity is defined as the true positive cases of COVID-19 divided by the total sum of true positive cases of COVID-19 and the incorrect predictions for COVID-19 or other cases.

The dataset includes 80 normal, non-disease chest X-rays, 105 COVID-19 positive X-rays, and 11 SARS X-rays. Data augmentation of the provided image samples allowed for a total of 1764 images. ResNet [15] was used as a pre-trained network in the transfer learning component, which achieved an effective performance with 92.5%, accuracy, 65.01% sensitivity, and 94.3% specificity, compared to the DeTraC architecture accomplishing 95.12% accuracy, 97.91% sensitivity, and 91.87% specificity [3], in the classification of X-ray images.

5   Conclusion and Discussion

Utilizing CNNs for the classification of medical images has made major progress. The CNN-LSTM architecture achieved promising performance and accuracy in the classification of medical images to diagnose COVID-19 infections. This is significant in hopes of reducing the false negatives that leads to false diagnostics of the COVID-19 cases in patients, which is very serious if a positive COVID-19 case is not identified and ignored.

COVID-Net exemplifies the benefits of human-machine collaborative design and demonstrates a significant innovation of neural networks in medical applications. The principles of transparency and integrity behind COVID-Net give it a promising outlook in the future of medicine, and the impactful innovations of COVID-Net will ripple throughout the future of CNN technology for medical imaging applications.

The DeTraC method shows promising potential for diagnosing COVID-19 from chest X-rays in an efficient manner with a decreased error rate. Transfer learning was utilized with the DeTraC architecture, as a method to combat issues where there is a limited number of training images. The DeTraC architecture has illustrated its success in the diagnosis of COVID-19 from chest X-rays by efficiently increasing the accuracy and sensitivity by significant margins, also with limited datasets available for training and the inconsistent nature of medical imaging.

References

[1]   M.Z. Islam, M.M. Islam, and A. Asraf, "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," *Informatics in Medicine Unlocked, 20,* p. 100412, 2020.

[2]   L. Wang, Z.Q. Lin, and A. Wong, "Covid-net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Scientific Reports, 10(1), pp.1-12, 2020.,* vol. 10, no. 1, pp. 1-12, 2020.

[3]   A. Abbas, M.M. Abdelsamea, and M.M. Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network," *Applied Intelligence,* vol. 51, no. 2,

pp. 854-864, 2021.

[4] S. Hassantabar, M. Ahmadi, and A. Sharifi, "Diagnosis and detection of infected tissue of COVID-19 patients based on lung X-ray image using convolutional neural network approaches," *Chaos, Solitons & Fractals,* vol. 140, p. 110170, 2020.

[5] M. Heidari, S. Mirniaharikandehei, A.Z. Khuzani, G. Danala, Y. Qiu, and B. Zheng, "Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms," *International journal of medical informatics,* vol. 144, p. 104284, 2020.

[6] S.S. Yadav, S.M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *Journal of Big Data,* vol. 6, no. 1, pp. 1-18, 2019.

[7] S. Kremer, et al., "Brain MRI findings in severe COVID-19: a retrospective observational study," *Radiology,* vol. 297, no. 2, pp. E242-E251, 2020.

[8] M.I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," *Classification in BioApps,* pp. 323-350, 2018.

[9] R. Yamashita, M. Nishio, R.K.G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into imaging,* vol. 9, no. 4, pp. 611-629, 2018.

[10] G. Chen, "A gentle tutorial of recurrent neural network with error backpropagation," arXiv preprint arXiv:1610.02583, 2016.

[11] Z.Q. Lin, M.J. Shafiee, S. Bochkarev, M.S. Jules, X.Y. Wang, and A. Wong, "Do explanations reflect decisions? A machine-centric strategy to quantify the performance of explainability algorithms," arXiv preprint arXiv:1910.07387, 2019.

[12] A. Wong, M.J. Shafiee, B. Chwyl, and F. Li, "Ferminets: Learning generative machines to generate efficient neural networks via generative synthesis," arXiv preprint arXiv:1809.05989, 2018.

[13] K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE conference on computer vision and pattern recognition*, 2016.

[15] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Thirty-first AAAI conference on artificial intelligence*, 2017.

[16] F. Ucar, and D. Korkmaz, "COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images," *Medical hypotheses,* vol. 140, p. 109761, 2020.

[17] M.E.H. Chowdhury, et al., "Can AI help in screening viral and COVID-19 pneumonia?," *IEEE Access,* vol. 8, pp. 132665-132676, 2020.

[18] M. Farooq, and A. Hafeez, "Covid-resnet: A deep learning framework for screening of covid19 from radiographs," arXiv preprint arXiv:2003.14395, 2020.