Automated COVID-19 screening framework via Deep Convolutional Neural Network with Chest X-ray Medical Images

Waris Damkham
Faculty of Information and Communication Technology, Mahidol University, Thailand
waris.dam@student.mahidol.ac.th

Jia-Ching Wang
Department of Computer Science and Information Engineering National Central University, Taiwan
jcw@dlmsl.csie.ncu.edu.tw

Tipajin Thaipisutikul
Department of Computer Science and Information Engineering National Central University, Taiwan
tipajin.tha@mahidol.ac.th

Akara Supratak
Faculty of Information and Communication Technology, Mahidol University, Thailand
akara.sup@mahidol.ac.th

Pattanasak Mongkolwat
Faculty of Information and Communication Technology, Mahidol University, Thailand
pattanasak.mon@mahidol.ac.th

***Abstract*— COVID-19 screening using chest X-rays plays a significant role in the early diagnosis of COVID-19 illness during the ongoing pandemic. Manually identifying this infection from chest X-ray films is a challenging and time-consuming technique due to time restrictions and the competence of radiologists. Also, the manual Covid-19 identification technique is made much more difficult and opaquer by the feature similarity between positive and negative chest X-ray images. Therefore, we propose an Automated COVID-19 screening framework that utilizes Artificial intelligence techniques with a transfer learning approach for COVID-19 diagnosis using chest X-ray images. Specifically, we employ the transfer learning concept for feature extraction before further processing with modified deep neural networks. Also, Grad-CAM visualization is used for our case study to support the predicted diagnosis. The results of the experiments on the publicly accessible dataset show that the convolutional neural network model, which is simple yet effective, performs significantly better than other deep learning techniques across all metrics, including accuracy, precision, recall, and F-measure.**

Keywords*— *COVID-19, Medical imaging, Deep neural networks, Transfer learning, Classification

I. INTRODUCTION

Due to the limited resources and the amount of data accessible for research, early diagnosis of COVID-19 continues to be a complex problem despite the global research efforts over the past several months. In particular, the accuracy of the Chest X-ray (CXR) diagnosis of COVID-19 infection depends heavily on radiographic competence in identifying patterns of COVID-19 infection chest X-rays of lung involvement, which can fluctuate in size and appearance over time.

Artificial intelligence has been used as a developing discipline to tackle various complicated issues, such as image categorization, object recognition, and medical imaging analysis [1]. In particular, Convolutional neural networks (CNN) have recently outperformed humans in a variety of computer vision tasks. It has already been used to identify and classify pneumonia [2,3] and other diseases on radiography. As a result, it has emerged as the obvious choice for CXR image analysis to handle the automated COVID-19 screening. Also, some new transfer learning techniques described in [4,5] have shown encouraging results in diagnosing COVID-19.

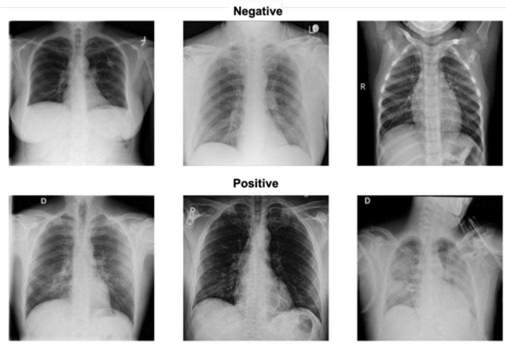

Fig. 1. Examples of negative and positive COVID-19 cases in Chest X-ray Medical Images.

Nevertheless, the recent studies still have limitations in accurately classifying COVID-19 with typical chest X-ray images. In particular, as shown in Figure 1, we can observe a high resemblance between each image, which raises the possibility of incorrectly classifying some chest X-rays. Also, traditional and manual COVID-19 screening is pruned to extracting unnecessary and duplicative features that could diminish classification performance.

Therefore, in this work, we present an end-to-end framework based on deep learning for categorizing COVID-19 into either positive or negative groups automatically with chest X-ray images. Our framework is made up of three primary parts: data management, model learning, and model classification.

In particular, given the chest X-rays, our proposed framework performs data pre-processing to get the proper representation, including resizing, normalizing, and augmentation. Various supervised deep learning models are extensively re-trained for categorizing chest X-rays as positive or negative. Specifically, we conduct experiments on

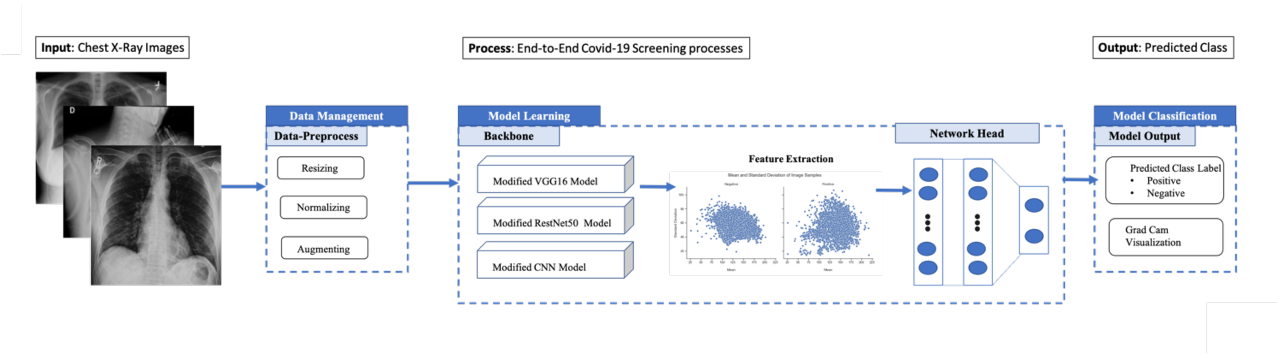

Fig. 2. Overall architecture of our proposed automated Covid-19 screening framework.

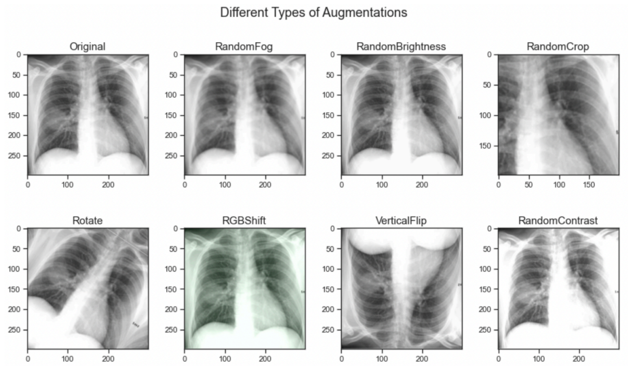

Fig. 3 The visualization of the augmented images from different augmented techniques.

three recent deep learning and transfer learning methods (modified VGG16, modified Restnet50, and modified CNN). Among deep convolutional-based models, we obtain the best performance of over 95% accuracy, 88% recall, 91% precision, and 90% F1-measure with the modified CNN model. In addition to quantitively measurements, we introduce a gradient class activation mapping (Grad-CAM) technique to visualize damaged lung regions for interpreted diagnosis. Our most significant contributions are highlighted below:

- We investigate the issue of covid-19 detection on chest X-ray films, which will aid the radiologist in promptly detecting the pandemic's early spread.
- We offer a unified framework that combines feature extraction with three modified deep learning models trained by transfer learning on the chest X-ray dataset to compare the performance in classifying chest X-ray pictures into positive and negative classes.
- The effectiveness of our proposed framework on actual datasets is demonstrated through a comparative analysis and evaluation process.

II. Related Work

Chen et al. [10] proposed the CNN model to identify lung nodules from chest radiographs. They employed the cunning edge detector to get rid of the rib crossing. Then, to classify the nodule with improved sensitivity and fewer false-positive classifications, the SVM classifier with a Gaussian kernel was applied. Karim et al. [6] proposed A Deep Neural

TABLE I

THE STATISTIC OF TRAIN, VALIDATION, TEST DATASETS USED IN OUR EXPERIMENTS

Class	Train	Validation	Test
Positive	2654	261	701
Negative	7287	844	2061
Total	9941	2762	1105

TABLE II

A REPRESENTATION OF CONFUSION MATRIX

Actual ↓Predict →	Negative (0)	Positive (1)
Negative (0)	True Negative (TN)	False Positive (FP)
Positive (1)	False Negative (FN)	True Positive (TP)

Network-based automated technique for COVID-19 detection in CXR pictures. They utilize layer-wise relevance propagation and gradient-guided class activation maps to highlight class-distinguishing regions. Ozturk et al. [8] presented DarkCovidNet, a CNN model for the automatic detection of COVID-19 utilizing chest radiographs. The authors employed the model design for DarkNet-19 [9] as an initial point rather than creating the model from scratch.

III. Methodology

This section presents an overview of our proposed framework for COVID-19 detection via a Deep Convolutional Neural Network from Chest X-ray Medical Images. As shown in Figure. 2, our framework consists of three main modules: (1) Data Management to pre-process image data from X-ray images as proper inputs for deep learning models. (2) Model learning to identify key attributes and develop a variety of models for fine-tuned models. (3) Model Classification compares various models for the COVID-19 classification task. Each component is elaborated in detail in the following sections.

A. Data Management

In this study, a public dataset from the COVID-19 RADIOGRAPHY DATABASE [1] (the COVID-19 Dataset Award by Kaggle Community) is used for experiments. We select two types of X-ray images consisting of those people with COVID (positive cases) and those people who are not (negative cases).

Data-Preprocess: Each image in the dataset initially has an image height and width equal to 299; the image dimension is set to 3; the maximum RGB value of the image is 255, and the minimum RGB value of the image is 5. We then resized

[1]https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database

the X-ray images to have image height and width equal to 70 while the image dimension is set to 3. Then, an image normalization stage has been used to consider the significant variation in image appearances, such as brightness and contrast). In this step, the pixel intensities are scaled and normalized to a range of [0, 1].

Large volumes of data are required to produce robust and generalized deep learning models. However, there is a shortage of medical imaging data, and labeling the information is expensive. As a result, we used the ImageDataGenerator TensorFlow library[2] to apply different types of augmentation techniques to the dataset. In this study, we apply image augmentation techniques such as random brightness, random crop, random rotate, RGBShift, random vertical flip, and random contrast operations. The augmentation outputs are shown in Figure 3.

At last, we divide the dataset into 70, 10, and 20 percent portions for training, validation, and testing, respectively, in our experiment. Table I displays the statistic of total 13808 samples in the datasets used in our investigations.

B. *Model Learning*

This section provides a details deep learning models used in our framework. we conduct experiments with deep learning methods by exploiting the Scikit-learn Python library[3]. For all models, we utilize Early-Stopping callback to stop training when there is no improvement on validation loss for 4 epochs. Each implementation detail is described in following section.

Modified VGG16 Model: The VGG16 deep model performed admirably on large and small datasets after being initially trained on the ImageNet dataset. This network contains a modest 3 × 3 receptive field and 16 convolutional layers. It has one Softmax classifier, three fully connected layers, five of kernel size 2× 2 max-pooling layers. All hidden layers use the ReLU activation function. In this study, we added AveragePooling2D layers, a New 128 Fully-connected layer, a dropout layer, and a new-Softmax layer to the original VGG16 model. The following hyperparameters are initialized after adding additional layers: learning rate is set to 0.001, batch size is set to 256, dropout factor is set to 0.2, and training optimizer is set to Adam. Finally, a new updated VGG16 model is obtained, from which detailed features are extracted.

Modified RestNet50 Model: Similar to VGG16 pre-trained model, RestNet50 model is trained on ImageNet dataset. However, this network contains receptive fields and 50 convolutional layers. All hidden layers use the ReLU activation function. In this study, we added AveragePooling2D layers, a New 128, and 64 Fully-connected layer, a dropout layer, and a new-Softmax layer to the original RestNet50 model. The following hyperparameters are initialized after adding additional layers: learning rate is set to 0.001, batch size is set to 256, dropout factor is set to 0.2, and training optimizer is set to Adam. Finally, a new updated RestNet50 model is obtained, from which detailed features are extracted.

Modified CNN Model: This model is intended to extract features from images, as well as to capture the local correlation of input data. There are two primary parts of the model: A group of trainable filters that may be applied to complete rows of the matrix are included in the convolutional layer (conv). Then, by lowering the complexity and parameters in the network, the max-pooling layer (max-pool) is employed to strengthen pattern identification and prevent overfitting data.

In this study, the convolutional layer and max-pooling parameters were adjusted to three couples with the ReLu activation function as follows:

- conv filters = 128, conv kernel_size = (3, 3), max-pool size= (2,2)
- conv filters = 64, conv kernel_size = (3, 3), max-pool size= (2,2)
- conv filters = 64, conv kernel_size = (3, 3), max-pool size= (2,2)

The output from the max-pooling layer is then reshaped into long vectors for further processing using the Flatten layer.

For all deep learning models in our framework, our objective is to reduce the cross-entropy loss function for each deep learning model in our system as follows:

$$\mathcal{L}(\theta) = -y log(\widehat{y_1})\text{-}(1\text{-}y)\log(1\text{-}\widehat{y_0}) \quad (1)$$

where θ is the network parameters and $\widehat{y_i}$ is the probability of belonging to class *i*.

C. *Model Classification*

After obtaining fine-tuned models, we use Accuracy, Precision, Recall, and F-Measure metrics to evaluate the classification performance of different models. These measurements are described in Table II and the following equations:

$$Accuracy = TP + TN/(TP + TN + FP + FN) \quad (2)$$

$$Precision = TP/(TP + FP) \quad (3)$$

$$Recall = TP/(TP + FN) \quad (4)$$

$$F - Measure = 2 * P * R/(P + R) \quad (5)$$

where TP denotes True Positive ,FP denotes False Positive, FN denotes False negative, P denotes Precision, and R denotes Recall. In addition to quantitively metrics, we display the areas of the input image that are crucial for predictions in order to provide visual explanations and interpretation. To do this, we included the Grad-CAM approach [7] into the pipeline to create a rough localization map of the highlighted locations.

IV. Experimental Result

From Table III, IV, and V, we can make the following observations. Firstly, our modified CNN model performs better than Modified VGG16 and Modified RestNet50 in accuracy metric by 2-3% differences. Secondly, the best recall and precision scores of 88.73% and 91.74% on test datasets are accomplished by modified CNN model. This indicates that it is vital to include convolutional and max pooling networks to distinguish the importance features to the prediction.

[2]https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/image/ImageDataGenerator

[3] https://scikit-learn.org/stable/

TABLE III
THE PERFORMANCE COMPARISON OF TRAIN DATA RESULTS FOR EACH MODEL.

Model Name	**Train Data**			
	Recall	***Accuracy***	***Precision***	***F1 Score***
Modified CNN	**0.9570**	**0.9782**	**0.9610**	**0.9590**
Modified VGG16	0.8647	0.9453	0.9254	0.8940
Modified Resnet 50	0.8892	0.9530	0.9317	0.9100

TABLE IV
THE PERFORMANCE COMPARISON OF VALIDATION DATA RESULTS FOR EACH MODEL.

Model Name	**Validation Data**			
	Recall	***Accuracy***	***Precision***	***F1 Score***
Modified CNN	**0.8659**	**0.9439**	**0.8933**	**0.8794**
Modified VGG16	0.8161	0.9258	0.8623	0.8386
Modified Resnet 50	0.8046	0.9285	0.8824	0.8417

TABLE V
THE PERFORMANCE COMPARISON OF TEST DATA RESULTS FOR EACH MODEL.

Model Name	**Validation Data**			
	Recall	***Accuracy***	***Precision***	***F1 Score***
Modified CNN	**0.8873**	**0.9511**	**0.9174**	**0.9021**
Modified VGG16	0.8402	0.9258	0.8636	0.8518
Modified Resnet 50	0.8459	0.9276	0.8657	0.8557

Additionally, we have comparable F1 score results, showing that modified CNN outperforms other deep learning models.

Deep learning is being improved in many ways to make it more logical and understandable. In numerous medical imaging applications, including deep-learning models, making the deep learning model easier to interpret is essential. Gradient Weighted Class Activation Mapping (Grad-CAM) is a technique that aids in learning more about the model when performing detection tasks and provides an explicable picture of deep learning models by creating a visual representation of any densely coupled neural network. As seen in Figure 4, we can see that the heat map is denser in one specific area, primarily in the center of the lung of the negative cases. The heat-map is more evenly distributed around the positive images, in contrast. In doing so, it could help the radiologist in making decisions besides the proposed framework's prediction.

V. CONCLUSION

This paper proposes an automated COVID - 19 screening framework via Deep Convolutional Neural Network with Chest X-ray Medical Images. We conduct extensive experiments on three transfer-learning deep learning models. The best experimental results are obtained from the modified CNN model with more than 88% recall, 91% precision, 95% accuracy, and 90% F-Measure, respectively. We firmly believe that our framework could help early screening of positive or negative cases on chest X-ray images. In the future, We plan to improve the performance by ensemble more state-of-the-art models. Also, we plan to design a proper attention network that could capture and interpret the critical representation contributing to the final prediction.

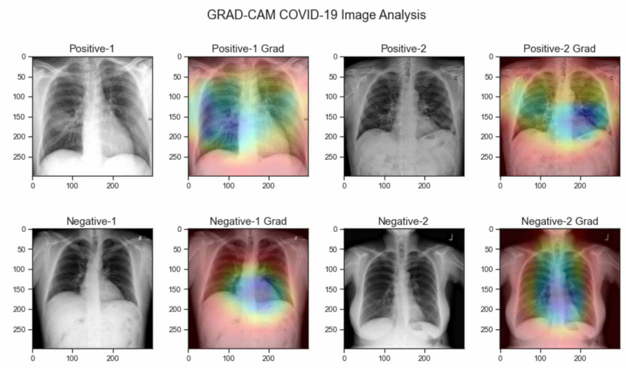

Fig. 4. The qualitative GRAD-CAM COVID-19 Image Analysis

REFERENCES

[1] S. Roy, A. Mitra, S. Roy, and S. K. Setua, "Blood vessel segmentation of retinal image using Clifford matched filter and Clifford Convolution," Multimedia Tools and Applications, vol. 78, no. 24, pp. 34839–34865, 2019.

[2] J. R. Zech, M. A. Badgeley, M. Liu, A. B. Costa, J. J. Titano, and E. K. Oermann, "Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study," PLOS Medicine, vol. 15, no. 11, 2018.

[3] A. Enkhbat, T. K. Shih, T. Thaipisutikul, N. L. Hakim, and W. Aditya, "HandKey: An efficient hand typing recognition using CNN for Virtual keyboard," *2020 - 5th International Conference on Information Technology (InCIT)*, 2020.

[4] E. Luz, P. Silva, R. Silva, L. Silva, J. Guimarães, G. Miozzo, G. Moreira, and D. Menotti, "Towards an effective and efficient deep learning model for COVID-19 patterns detection in X-ray images," Research on Biomedical Engineering, vol. 38, no. 1, pp. 149–162, 2021.

[5] P. Silva, E. Luz, G. Silva, G. Moreira, R. Silva, D. Lucio, and D. Menottu, "Efficient deep learning model for COVID-19 detection in large CT images datasets: A cross-dataset analysis," 2020.

[6] M. R. Karim, T. Dohmen, M. Cochez, O. Beyan, D. Rebholz-Schuhmann, and S. Decker, "Deepcovidexplainer: Explainable covid-19 diagnosis from chest X-ray images," 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2020.

[7] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," 2017 IEEE International Conference on Computer Vision (ICCV), 2017.

[8] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Computers in Biology and Medicine*, vol. 121, p. 103792, 2020.

[9] J. Redmon and A. Farhadi, "Yolo9000: Better, faster, stronger," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[10] S. Chen, K. Suzuki, and H. MacMahon, "Development and evaluation of a computer-aided diagnostic scheme for lung nodule detection in chest radiographs by means of two-stage nodule enhancement with support vector classification," *Medical Physics*, vol. 38, no. 4, pp. 1844–1858, 2011.