Diagnosing COVID-19 of Lung CT Scan by Using Convolution Neural Network

Mustafa Khalaf Hashem and Ahmed Slim Abbas

July 20, 2021

Diagnosing COVID-19 of lung CT scan by using convolution neural network

Mustafa Khalaf Hashem and Dr. Ahmed Slim Abbas

[1] Princeton University, Princeton NJ 08544, USA
[2] Springer Heidelberg, Tiergartenstr. 17, 69121 Heidelberg, Germany
lncs@springer.com

Abstract. The rapid spread and deaths caused by COVID-19 prompted us to find a way to diagnose it in a shorter period of time and with better accuracy. One of the ways to diagnose COVID-19 is through CT images because the disease has an effect on the patient's lungs. The proposed method is to classify the disease by parallel models after performing a preprocessing that involves resize and augmenting CT images using conventional data augmentation techniques and exposing the lungs as regions of interest. The three models (MuNet, NASNet, and MobileNetV2) were selected for the proposed model after comparing them with more than one model and obtaining the best accuracy. where the model got accuracy(0.97), Precision(0.94), Recall(0.99),(0.97) and F1 Score respectively using dataset 1 for covid. As for dataset 2 The proposed model obtained (0.9661) inaccuracy, (0.95) in Precision, (0.95) in Recall, and (0.95) in F1 Score for covid.

Keywords: Diagnosing COVID-19, CNN, CT scan.

1 INTRODUCTION

Coronaviruses consist of seven types, only four of which infect humans. The infection is a common cold in its most straightforward cases and reaches severe acute respiratory syndrome (SARS). Recently, a new type appeared, "COVID_19", which causes infectious diseases, according to what was stated by the World Health Organization[1].

The first appearance of the new type was in Wuhan city, China. From there, it spread all over the world[2].

Patients with this type suffer from dry cough, fever, fatigue, and even loss of smell and taste in its most severe cases. 1 in 5 patients suffers from complications such as pneumonia.

They appear in CT as ground-glass opacity (GGO), patchy multifocal consolidation, or crazy paving. 0.02 Of the infected die, especially the elderly and those with medical problems[1][3].

One of the methods of diagnosing this epidemic is RT-PCR. Recent reports have shown that CT is more sensitive to the diagnosis of the disease than PCR.

Since the cases of infection and suspicion are increasing, the mechanism for determining lung lesions manual is a daunting task. It requires intense effort, so we need a rapid automatic diagnosis to assess the disease[4][5][6].

Recently, Deep learning technology has become popular in medical image processing due to the rapid growth of artificial intelligence and its strong feature representation. [7]In particular, the convolutional neural network (CNN) has demonstrated high performance in image classification compared to image processing techniques in the ImageNet Visual Recognition Challenge at a Large Scale (ILSVRC) 2012. [7] The precision of a convolutional neural network increases the deeper it is, but at the same time, it causes over-fitting or vanishing gradient problems. To solve this problem, we resort to a parallel CNN. The proposed model consists of Transfer learning models. The results of the models are combined through ensemble Methods.

1.1 Related Works

1. Khumancha et al. (2019)[2] proposed two network models. The first is to detect nodules, and the other is to diagnose cancer using two datasets. Due to the first group's paucity, data augmentation of the data was used to train the CNN model for nodule diagnosis. After it obtained an accuracy of 89%, it was applied to the second database to detect nodules. The second model divides the tomography images of the second database and multiplies them with the original images, then counting the nodules in the form of cubes according to the first model's coordinates. The proposed model achieved 90% accuracy.
2. Sathish et al.(2020)[3] proposed a two-stage framework for lung segmentation and nodule detection. Lung segmentation training is on a modified version of SUMNet. Then, LeNet is used to treating the stains extracted from the cutting process after modifying the third convolutional layer of the kernel.
3. Li et al.(2020) [8] developed for COVID-19 diagnosis a tool transfer Learning 3D using a CT scan. The first part is pre-processed for 3D CT. Then the ROI (lungs) is extracted using U-Net. As for the second part depends on extracting the features maps by ResNet-50 model then be max-pooled. The third and final part is feeding a fully connected layer with the feature maps and used "soft-max" as an activation function for multi-class (COVID-19, CAP, and non-pneumonic). AUC recorded is 0.96.
4. Khalifa et al.(2020)[9] proposed a model consisting of three main blocks and binary classification. In the first block, GAN was used to increase the data. The second block's task is to train and validating using transfer learning models to minimize training time, mathematical calculations, and hardware resource usage. As for the third block task, it is the test where the Shufflenet achieved the highest accuracy.
5. Apostolopoulos et al.(2020)[10] evaluated the transfer learning models' performance for the detection of COVID-19 using X-ray images from two datasets. Group A consists of 224 photos confirmed with COVID_19, 700 pictures of bacterial infection, and 405 natural images. Group B: 224 with COVID_19, 714

pictures of pneumonia (bacterial with viral), and 504 pictures of normal. The best result obtained by binary classification MobileNet.

6. Loey et al.(2020)[11] proposed a model consisting of dataset preprocessing and then using transfer learning to train. This database consisted of 749 CT images. The test's accuracy was estimated using a confusion matrix, where transfer learning models were evaluated to diagnose COVID_ 19 using the original dataset and with the preprocessing procedure. Resnet50 achieved the highest accuracy of 82% with preprocessing.
7. Ko et al.(2018)[12] proposed a way to detect lung opacities using an ensemble of CNN on chest radiographs (CXR). the proposed model collects a results 2 Mask R-CNN and 3 RetinaNet models by majority weighted voting ensemble method where this method outperformed individual models.

1.2 Materials

In the next sections, we'll go over the dataset and methods we employed.

Dataset. Dataset 1 consists of (4153) CT images (2,167) of which contain COVID-19 (1986) CT images of a healthy person.

(1229) CT image of the intact image obtained from https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset?select=non-COVID. As (2924) obtained from https://www.kaggle.com/kerneler/starter-a-covid-multiclass-dataset-of-9486b3eb-d.

The dataset1 contains 2,924 CT images (2,167 with COVID) and 757 CT images of a healthy person. The images were taken from https://www.kaggle.com/kerneler/starter-a-covid-multiclass-dataset-of-9486b3eb-d**.**

1.3 Proposed method

The proposed method consists of three stages:

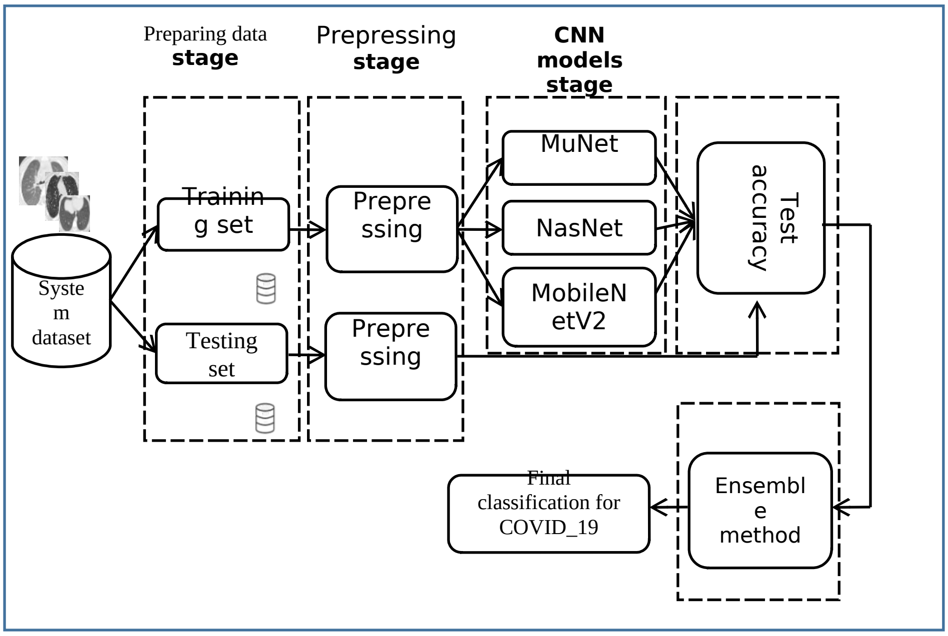

Fig 1: The overall proposed methodology structure for classification COVID_19

Pre-processing stage. The pre-processing stage includes all the preliminary processes that are performed straightforwardly on the dataset to get ready for the training and testing stage. The pre-processing stage consists of four steps include: Resize, segmentation, Data augmentation, and Normalization.

- **Resize.** Resize image into 224.
- **Segmentation lung.** Segmentation for Lung CT image is a fundamental step for lung image analysis, and it is a necessary step in providing an accurate lung. It aims to obtain the areas of interest, which are the lungs. To get a lung mask, CT images go through six: Binarization, Clear Border, Labeling, Erosion, Closure, and Filling Holes, as shown in fig (1). It starts with the binarization and ends with filling in the holes, and we finally get the binary lung mask[2].

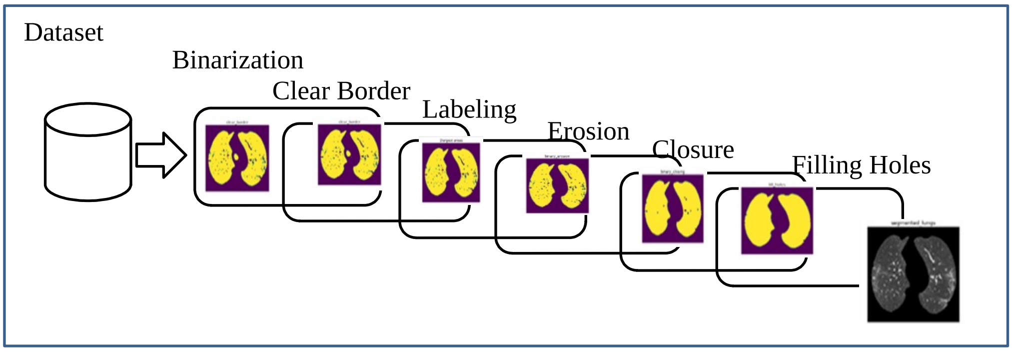

Fig 1: Segmentation operation

Binarization: The method of transforming a pixel image to a binary image is known as binarization. The picture is first changed to grayscale. Using the Otsu method.

Clear Border: The aim of Clear Border is to get rid of the blobs that are linked to the lungs' boundary. This is accomplished by using a straight border. Structures that are lighter than their surroundings and are bound to the picture boundary are suppressed by a clear border.

Labeling the Image: Labeling the Image at this point, the regions in the binary image that are made up of connected pixels are labeling. Different labels are given for different regions. After that, the two largest areas are preserved.

Erosion: This is one of the most widely used morphological image processing techniques. It's most commonly used on binary images. The operator's primary effect on a binary picture is to erode the boundaries of foreground pixel regions (i.e., white pixels, typically). As a result, foreground pixel areas shrink in size, while holes within those areas grow. This move is crucial because it removes the nodules from the blood vessels in the lungs.

Closure Operation: Closure Operation is a dilation accompanied by erosion that is used to close the wound. Dilation has the primary effect of gradually enlarging the boundaries of foreground pixels in a binary picture (i.e., white pixels, typically). Closing fills in gaps in regions while maintaining their original sizes. This move is necessary to keep the blobs attached to the lungs' walls.

- *Filling Holes:* Filling Holes is to fill the gaps in the lungs' binary mask.

- **Data augmentation**. The data augmentation process is one of the essential processes in deep learning, especially CNN, which requires a large number of data to be trained to obtain high accuracy. Therefore, the Data augmentation process is one of the processes used to solve the problem of data shortage that researchers face in this field. The data augmentation process is done by creating artificial differences in the image that create new images that are different from the existing ones. The differences are either by zooming or rotating the image to a certain degree or cropping the current images[11][13].

CNN parallel stage. This stage includes a model consisting of parallel CNN models. The main aim of it is to predict the presence or absence of Covid_19. CNN Parallel consists of a network built from scratch called MuNet and two pre-trained networks that are fed with CT images resulting from the pre-processing stage (NasNet, Mo-bileNetV2)[14][15]. As shown in the fig (1).

Table 1. The summary representation of MuNet

No of block	**Layers(type)**	**Shape of the output**	**Params_no**
1	conv2d (Conv2D)(3,3)	None, 224, 224, 32	896
	Activation('relu')	None, 224, 224, 32	0
	MaxPooling2D(2,2)	None, 112, 112, 32	0
	Dropout(0.2)	None, 112, 112, 32	0

	Batch_normalization	None, 112, 112, 32	128
2	conv2d_1 (Conv2D)(3,3)	None, 112, 112, 16	4624
	activation_1 (‘relu’)	None, 112, 112, 16	0
	MaxPooling2D_1 (2,2)	None, 56, 56, 16	0
	Dropout_1(0.2)	None, 56, 56, 16	0
	Batch_normalization_1	None, 56, 56, 16	0
3	Flatten	None, 50176	0
	Dropout_2(0.2)	None, 50176	0
4	Dense	None, 32	1605664
	activation_2 (‘relu’)	None, 32	0
	Dropout_3(0.2)	None, 32	0
	Batch_normalization_2	None, 32	128
5	Dense_1	None, 2	66
	activation_3('softmax')	None, 2	0
6	Total parameters: 1,611,570 Trainable parameters: 1,611,410 Non-trainable parameters: 160		

MuNET is used to categorize images into COVID and non-COVID binary classifications, as shown in Table 1.

NasNet. A model is used to design and optimize convolutional structures when using a significant dataset. The approach is inspired by Neural Architecture Search (NAS) framework, which optimizes architecture configurations using a reinforcement learning search method.

MobileNetV2. This model combines features from several filters to enhance the network's overall performance.

Convolution is typically performed by neural network architecture on a spatial and channel-by-channel basis. However, Inception architecture ignored spatial dimensions by performing a 1*1 convolution, followed through cross-channel and cross-spatial correlations using 3*3 and 5*5 filters[16].

The the inception architecture is 27 layers deep and uses numbers of features from numbers of filters to improve the network's overall performance. Inception architecture has produced excellent results when combined with other models.

Ensemble method. It is a method of combining predictions of multiple models trained on the same dataset and using the results of the grouped prediction.

The purpose of the ensemble methods is to improve the accuracy of results, as the accuracy of results, the higher they are, the better and safer, especially in the subject of research being a medical field and its direct impact on human life[17].

The individual models, when trained, do not give the same results, that is, each model has advantages and disadvantages that differ from the other model, so when they blood, they will complement one another[12][18][19].

Ensemble learning is named for the fact that it can reduce prediction variance and create good predictions than a single model. Ensemble learning collected the predictions of the three models used and by using equation (1).

$$\hat{y}_i = argmax \sum_{j=1}^{n} W_j \left(C_j (X) \right) \ldots\ldots (1)$$

1.3 Results and Evaluation

Dataset 1 contains 4,153 CT images of two poisonings, one Covid and the other non-Covid. During the pre-processing step, the data set was expanded to (4489) using data augmentation techniques.

Data set 1 was separated into a training set and a test set, with the training set accounting for 80% of the original set (3591) and 20% of the test set (898) CT scans.

The verification set accounts for 20% of the entire training set.

Classification accuracy, Precision, Recall, and F1 Score were used to find and evaluate the results. As shown in the following equations[18][20].

$$Accuracy = \frac{TruePositives + FalseNegatives}{Total\ Number\ of\ Samples} (2)$$

Or

$$Accuracy = \frac{TP + FN}{TP + TN + FP + FN}$$

$$Precision=\frac{TruePositives}{TruePositives+FalsePositives}(3)$$

$$Recall=\frac{TruePositives}{TruePositives+FalseNegatives}(4)$$

$$F_1=\frac{2*1}{\frac{1}{Precision}+\frac{1}{Recall}}(5)$$

According to the table (2) of Confusion Matrix.

Table 1: The Confusion Matrix CM

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">Predictive model</td></tr>
<tr><td>Yes</td><td>No</td></tr>
<tr><td rowspan="2">Actual class</td><td>Yes</td><td>True positive (TP)</td><td>True Negative (TN)</td></tr>
<tr><td>No</td><td>False positive (FP)</td><td>False Negative (FN)</td></tr>
</table>

After completing the preprocessing stage, the images will be entered into the three models in parallel. More than one model has been implemented and the most accurate model has been selected. The MUNet model was chosen for the first model with the proposed model after obtaining the highest accuracy compared with the two CNN models after obtaining an accuracy of (0.8444), which obtained (0.7435) and (0.7391), respectively, after three epochs. As for the second model, the NasNet model was chosen after comparing it with the VGG 16 and obtaining the highest accuracy (0.9415) compared to (0.7421), after ten epochs. As for the third model, the MobileNetV2 model was chosen. The optimization algorithm used is Adam.

After selecting the three models, we improved the accuracy by using a different learning rate over 20 epochs. As in the tables below

Table 2: illustrates the accuracy and loss of the three models with a learning rate of (0.01) within 20 epochs.

Model	Loss	Accuracy	val_loss	val_accura
MuNet	0.4818	0.7516	0.4629	0.7917
NASNet	0.6929	0.5210	0.6927	0.5212
MobileNet V2	0.6926	0.5210	0.6934	0.5215

Table 3:Results the Confusion Matrix metrics with a learning rate of (0.01).

Predictive	Precision	Recall	F1 Score	Support
COVID	0.00	0.00	0.00	1720
NonCOVID	0.52	1.00	0.62	1871

Table 4 are the confusion matrix metrics for the training set. The accuracy of the complete model is 0.52.

Table 4: illustrates the accuracy and loss of the three models with a learning rate of (0.001) within 20 epochs

Model	Loss	Accuracy	val_loss	val_accura
MuNet	0.2095	0.9117	0.4057	0.8263
NASNet	0.0690	0.9716	1.1799	0.7862
MobileNet V2	0.6922	0.5227	0.6927	0.5156

Table 5: Results the Confusion Matrix metrics with a learning rate of (0.001).

Predictive	Precision	Recall	F1 Score	Support
COVID	0.95	0.92	0.93	1714
NonCOVID	0.93	0.95	0.94	1877

Tables 5 and 6 did the adjustment of the learning rate from (0.01) to (0.001), where we noticed an increase in the scales and the accuracy of the model for the training group reached (0.95) and (0.78) for the test group.

Table 6: illustrates the accuracy and loss of the three models with a learning rate of (0.0001) within 20 epochs

Model	Loss	Accuracy	val_loss	val_accura
MuNet	0.1118	0.9638	0.8981	0.7138
NASNet	0.0721	0.9847	0.5877	0.7806
MobileNetV2	0.2729	0.9696	0.5814	0.8252

Table 7:Results the Confusion Matrix metrics with a learning rate of (0.0001).

Predictive	Precision	Recall	F1 Score	Support
COVID	0.94	0.99	0.97	1703
NonCOVID	0.99	0.95	0.97	1888

References

[1] “Q&A on coronaviruses (COVID-19),” [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/q-a-coronaviruses.

[2] M. B. Khumancha, A. Barai, and C. B. R. Rao, “Lung Cancer Detection from Computed Tomography (CT) Scans using Convolutional Neural Network,” *2019 10th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2019*. 2019, doi: 10.1109/ICCCNT45670.2019.8944824.

[3] R. Sathish, R. Sathish, R. Sethuraman, and D. Sheet, “Lung Segmentation and Nodule Detection in Computed Tomography Scan using a Convolutional Neural Network Trained Adversarially using Turing Test Loss,” *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2020-July. pp. 1331–1334, 2020, doi: 10.1109/EMBC44109.2020.9175649.

[4] C. Zheng *et al.*, “Deep Learning-based Detection for COVID-19 from Chest CT using Weak Label,” *medRxiv*. 2020, doi: 10.1101/2020.03.12.20027185.

[5] J. P. Kanne, B. P. Little, J. H. Chung, B. M. Elicker, and L. H. Ketai, “Essentials for radiologists on COVID-19: An update-radiology scientific expert panel,” *Radiology*, vol. 296, no. 2. Radiological Society of North America Inc., pp. E113–E114, Aug. 01, 2020, doi: 10.1148/radiol.2020200527.

[6] S. S. A. A. S. B. A. Gholamrezanezhad and Salehi, “Coronavirus Disease 2019 (COVID-19): A Systematic Review of Imaging Findings in 919 Patients.” American Journal of Roentgenology Coronavirus, 2020, doi: doi.org/10.2214/AJR.20.23034.

[7] J. Lee, S. K. Lee, and S. Il Yang, “An Ensemble Method of CNN Models for Object Detection,” *9th International Conference on Information and Communication Technology Convergence: ICT Convergence Powered by Smart Intelligence, ICTC 2018*. pp. 898–901, 2018, doi: 10.1109/ICTC.2018.8539396.

[8] L. Li *et al.*, “Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT.,” *Radiology*. p. 200905, 2020, doi: 10.1148/radiol.2020200905.

[9] N. E. M. Khalifa, M. H. N. Taha, A. E. Hassanien, and S. H. N. Taha, “The Detection of COVID-19 in CT Medical Images: A Deep Learning Approach,” 2020, pp. 117–130.

[10] I. D. Apostolopoulos and T. Bessiana, “Covid-19: Automatic detection from X-Ray images utilizing transfer learning with convolutional neural networks,” *Physical and Engineering Sciences in Medicine*. 2020, [Online]. Available: https://doi.org/10.1007/s13246-020-00865-4.

[11] M. Loey, G. Manogaran, and N. E. M. Khalifa, “A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images,” *Neural Computing and Applications*. 2020, doi: 10.1007/s00521-020-05437-x.

[12] H. Ko, H. Ha, H. Cho, K. Seo, and J. Lee, “Pneumonia Detection with Weighted Voting Ensemble of CNN Models,” *2019 2nd International Conference on Artificial Intelligence and Big Data, ICAIBD 2019*. pp. 306–310, 2019, doi: 10.1109/ICAIBD.2019.8837042.

[13] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *Journal of Big Data*, vol. 6, no. 1. 2019, doi: 10.1186/s40537-019-0197-0.

[14] B. Zoph, G. Brain, V. Vasudevan, J. Shlens, and Q. V Le Google Brain, “Learning Transferable Architectures for Scalable Image Recognition.”

[15] E. A. Mostafa Mehdipour Ghazia,∗, Berrin Yanikoglua, “Plant identification using deep neural networks via optimization of transfer learning parameters.” Neurocomputing.

[16] F. Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions.” 2017.

[17] A. Conneau, H. Schwenk, and Y. Le Cun, “Very Deep Convolutional Networks for Text Classification Alexis.” 2017.

[18] I. H. Witten, “Data Mining (Fourth Edition),” *Practical Machine Learning Tools and Techniques*. pp. 417–466, 2017.

[19] A. Yazdizadeh, Z. Patterson, and B. Farooq, “Ensemble convolutional neural networks for mode inference in smartphone travel survey,” *arXiv*. 2019.

[20] J. K. Kuhn, Max, “Applied Predictive Modeling.” Springer, 2013.