Multimodal covid network: Multimodal bespoke convolutional neural network architectures for COVID-19 detection from chest X-ray’s and computerized tomography scans

Thiyagarajan Padmapriya1 | Thiruvenkatam Kalaiselvi1 | Venugopal Priyadharshini2

1Department of Computer Science and Applications, The Gandhigram Rural Institute (Deemed to be University), Gandhigram, India
2Government Medical College, Omandurar Government Estate, Chennai, India

Correspondence
Thiruvenkatam Kalaiselvi, Department of Computer Science and Applications, The Gandhigram Rural Institute (Deemed to be University), Tamil Nadu, 624302, India.
Email: kalaiselvi.gri@gmail.com

Abstract
AI-based tools were developed in the existing works, which focused on one type of image data; either CXR’s or computerized tomography (CT) scans for COVID-19 prediction. There is a need for an AI-based tool that predicts COVID-19 detection from chest images such as Chest X-ray (CXR) and CT scans given as inputs. This research gap is considered the core objective of the proposed work. In the proposed work, multimodal CNN architectures were developed based on the parameters and hyperparameters of neural networks. Nine experiments evaluate optimizers, learning rates, and the number of epochs. Based on the experimental results, suitable parameters are fixed for multimodal architecture development for COVID-19 detection. We have constructed a bespoke convolutional neural network (CNN) architecture named multimodal covid network (MMCOVID-NET) by varying the number of layers from two to seven, which can predict covid or normal images from both CXR’s and CT scans. In the proposed work, we have experimented by constructing 24 models for COVID-19 prediction. Among them, four models named MMCOVID-NET-I, MMCOVID-NET-II, MMCOVID-NET-III, and MMCOVID-NET-IV performed well by producing an accuracy of 100%. We obtained these results from a small dataset. So we repeated these experiments in a larger dataset. We inferred that MMCOVID-NET-III outperformed all the state-of-the-art methods by producing an accuracy of 99.75%. The experiments carried out in this work conclude that the parameters and hyperparameters play a vital role in increasing or decreasing the model’s performance.

Keywords
artificial intelligence, chest X-rays, convolutional neural networks, coronavirus disease, COVID-19, CT scans, deep neural networks
INTRODUCTION

Coronavirus disease (COVID-19) first occurred in Wuhan city of China. It created a pandemic situation in the entire world. Due to the emergence of COVID-19, about 80,412,311 people were affected, and 1,758,886 people were found dead around the globe. COVID-19 spreads very fast worldwide, and it became a pandemic disease. No specific medicines have been discovered yet to cure COVID-19. It affects the day-to-day lives of humans and the economy of all the countries. It remains a challenge in the early diagnosis of COVID-19. Polymerase chain reaction (PCR) test is the most commonly used test for detecting COVID-19, but it produces significant false positives and false negatives, not reliable. Recently, X-ray and computerized tomography (CT) imaging of the chest have been used to detect COVID-19. Also, ultrasound thorax is a screening tool for patients in the Intensive Care Unit (ICU). One of the common symptoms of COVID-19 found in 85% of corona-affected patients is fever. Cough is found in 70% of corona-affected patients. About 43% of patients reported shortness of breath. There is also a possibility of abdominal problems for COVID-19 patients. Recently, this disease might be asymptomatic, but it still spreads to other people. The severity of COVID-19 may vary from mild to critical. Mild symptoms include no symptoms, mild cough, and fever. Severe symptoms may include dyspnea and anosmia. Acute symptoms of COVID-19 may consist of failure of the respiratory system or multiorgan loss.

Diagnosis of COVID-19 is made by real-time reverse transcription-polymerase chain reaction (RT-PCR). PCR tests produce a lower sensitivity of 65 to 95%, which shows that they may produce negative results even if they are infected. It takes more than 24 h to produce the test results. The role of medical imaging plays a vital role in the early diagnosis of COVID-19. Radiology plays a crucial role in diagnosing and identifying COVID-19 infection using X-ray and CT images. The radiological findings of COVID-19 show that the infections are ground glass patterned areas. These ground glass patterned areas affect the lower lobes, posterior segments in both the lungs. In China, CT is used to screen and diagnose COVID-19 due to an inadequate supply of test kits. The combination of laboratory results and clinical imaging features enables a better diagnosis. The accurate diagnosis of COVID-19 is essential for treatment planning. Handcrafted feature extraction techniques are required for conventional medical image classification approaches. Imaging studies of COVID-19 had reported some lung changes, such as opacities in the right infrahilar airspace and opacity in the left lower lung. Some researchers had found ground-glass opacities (GGO) or mixed ground-glass opacities along with vascular dilation in the lesions in most of the COVID-19 patients, interlobular septal thickening and signs of air bronchogram, peripheral focal which affects both the lungs in 75% of patients.

Artificial intelligence (AI) is a computer program or machine to think, learn, and act with human intelligence. It is also a field of study which tries to make computers “smart.” Artificial Intelligence (AI) plays a significant role in diagnosing COVID-19 in patients due to its outstanding throughput analysis and pattern recognition in CT and X-ray images. Artificial Intelligence-based tools help in the process of screening, predicting, and diagnosing Coronavirus disease (COVID-19). COVID19 is found using the imaging modalities such as chest X-rays (CXR) and chest computed tomography (CT) scans. To track the disease progression, frequent scanning of lung CTs is required. AI helps detect even minor changes in disease progression to provide necessary treatment to the patients.

Machine learning (ML) and deep learning (DL) are two broad areas of AI. ML needs human intervention where the features from images are extracted manually and fed into the system. In DL, feature extraction is performed automatically. Emerging deep learning-based approaches are considered effective and efficient for medical image analysis. Feature extraction is done automatically in DL-based techniques, and it is capable of learning more complicated patterns in image data. One of the significant disadvantages of DL models is the requirement of large amounts of labeled data for training the network. It also requires enormous computational resources. Transfer learning (TL) is a popular approach to DL in which a model that was developed and trained for a task is reused. This pre-trained model is a starting point for introducing a new model for a similar task. It gradually reduces training time, the requirement of substantial computational resources, and training data.

The recent works on COVID-19 chest X-ray images include developing DL models such as the COVIDXNet model with seven convolutional neural network (CNN) models. They have used VGG19, Google MobileNet, DenseNet models in their architectures. They have used 80 percent of X-ray images for training and 20 percent for testing. Their experiments conclude that VGG19 and DenseNet architectures produce better results in diagnosing COVID-19 disease. COVIDNet model was proposed for COVID-19 detection, which served as a multiclass classification model. This model belongs to the family of CNN models, and it is specially designed for COVID-19 detection in X-ray images. It provides good results compared to VGG19 and ResNet50 models in the case of sensitivity. Some researchers have developed artificial
intelligence-based methods for COVID-19 detection. A computer-aided CT diagnosis system for COVID-19 based on DL was developed by Song et al. Wang et al. developed a DL algorithm based on transfer learning for COVID-19 diagnosis. They have used a pre-trained Inception model for training the model. Zheng et al. developed software for COVID-19 detection using CT images of the lungs. They used U-Net architecture to segment the lungs, then fed the segmented results into a 3D deep neural network to diagnose COVID-19 disease. Xu et al. had segmented candidate infected regions from CT images using 3D deep architecture. Then they classified the type of infection using multiclass classification.

Barstugan et al. used ML methods to diagnose COVID-19 infections from chest CT images. They have used Gray Level Co-occurrence Matrix (GLCM), Local Directional Pattern (LDP), Gray Level Run Length Matrix (GLRLM), Gray-Level Size Zone Matrix (GLSZM), and Discrete Wavelet Transform (DWT) algorithms for feature extraction. They have used support vector machines (SVM) for classifying COVID-19 affected patients. The GLSZM method provided the best results among them. Chen et al. used residual attention U-Net architecture for segmenting COVID-19 chest CT images. They have done multiclass classification. In all the works on COVID-19 detection available in the literature, experiments are done for either X-ray images or CT images alone. Mukherjee et al. developed a tailored CNN architecture train using CT and X-ray images to predict COVID-19. They had trained and tested their model using a binary classifier and attained a prediction accuracy of 96.23%. The model architectures proposed for COVID-19 prediction by various authors are inferred in Table 1.

The proposed work has developed multimodal CNN architectures for COVID-19 detection from CXR's and CT scans. We have performed various experiments for fixing the parameters and hyperparameters for our bespoke CNN architectures. Experiments were carried out for setting suitable learning rates, optimizers, batch normalization, and the number of epochs. After fixing these hyperparameters, the models are trained and tested by varying the number of layers from two to seven. Some of the experimented proposed multimodal CNN architectures (MMCOVID-NET) outperformed the current works available in the literature.

In Section 2, the materials and metrics used for experiments are discussed. In Section 3, the architectures of the proposed models are detailed. In Section 4, the experimental results are analyzed and discussed. Section 5 concludes the work along with future enhancements.

### 2 | MATERIALS AND METRICS

In the proposed work, we have conducted experiments on an X-ray image dataset and CT image dataset collected from various hospitals. The X-ray images were collected from patients with COVID-19 and healthy individuals. The CT images were obtained from patients with COVID-19 and normal controls. The images were cropped to a size of 256x256 and normalized to have a mean of 0 and standard deviation of 1. The images were then divided into training and testing sets in a ratio of 80:20. The training set was used for training the models, and the testing set was used for evaluating the performance.

### Table 1: Existing architectures for COVID-19 detection

| S. no | Author and year | Methodology            | Type of network     | Imaging modality | Hyperparameters used in the network |
|-------|-----------------|------------------------|---------------------|------------------|-------------------------------------|
| 1     | Wang et al (2020) | Transfer learning      | Inception v3        | Chest CT         | Batch normalization, dropout (0.5)   |
| 2     | Chairmaine et al (2020) | Transfer learning  | ResNet23            | Chest CT         | Batch normalization                  |
| 3     | Linda et al (2020) | Deep learning          | CNN model named COVID-Net | Chest X-ray    | -                                   |
| 4     | Prabira and Santhi (2020) | Machine learning | SVM                 | Chest X-ray    | -                                   |
| 5     | Tulin et al (2020) | Deep learning          | CNN model named DarkCovidNet | Chest X-ray    | Batch normalization                  |
| 6     | Ali et al (2020) | Transfer learning      | ResNet50, ResNet101, ResNet152, InceptionV3 and Inception-ResNetV2 | Chest X-ray | -                                   |
| 7     | Arpan et al (2020) | Transfer learning      | DenseNet121         | Chest X-ray    | -                                   |
| 8     | Farooq et al (2020) | Transfer learning      | ResNet50, VGG16     | Chest X-ray    | -                                   |
| 9     | Lawrence et al (2020) | Transfer learning | ResNet50, VGG16     | Chest X-ray    | -                                   |
| 10    | Mukherjee et al (2020) | Deep learning         | Three Layered CNN Architecture | Chest CT + Chest X-ray | Image size, batch size, dropout, and number of epochs |
from internet repositories maintained by research and medical centers. We have employed two different datasets (Dataset 1 and Dataset 2). All the X-ray and CT images were in JPEG format. Several evaluation metrics were used to validate the performance of the proposed models. The detailed description of the datasets used for our experiments and evaluation metrics are explained below.

2.1 | Dataset 1

2.1.1 | X-ray image dataset

We have collected the COVID-19 X-ray image dataset from two different sources developed by Cohen\textsuperscript{29} and Wang et al.\textsuperscript{30} The dataset by Cohen consists of about 12 X-ray images for COVID-19 positive cases and 13 images for COVID-19 negative cases. Some of the COVID-19 positive and negative X-ray images are shown in Figure 1. The dataset by Wang et al. contains about 125 COVID-19 positive cases (43 female and 82 male) and 500 COVID-19 negative X-ray images. The average age of the patients is 55 years. We have used 137 COVID-19 positive cases to train and test the network and 513 COVID-19 negative cases. The number of images is inadequate to train the CNN model. So we have employed a data augmentation technique to increase the number of training images.

2.1.2 | CT Image Dataset

COVID-19 CT images were obtained from the database provided by Zhao et al.\textsuperscript{31} The COVID-19 CT dataset consists of 349 CT images that contain the clinical findings (ground glass patterned areas) of COVID-19 from 216 patients. This dataset also includes 397 CT images for COVID-19 negative cases. Some of the COVID-19 positive and negative CT images are shown in Figure 2.

2.2 | Dataset 2

This study makes use of open-source datasets that include 2500 chest CT-scans (1257 for covid and 1243 for non-covid) as well as 2500 chest X-ray scans (1257 for covid and 1243 for non-covid). The chest CT-scan dataset\textsuperscript{33} is derived from (kaggle.com/plameneduardo/sarscov2-ctscan-dataset/). The chest X-ray image dataset is available at (http://www.kaggle.com/praveengovi/coronahackchest-xraydataset), and we only used 2500 of the 5933 chest X-ray pictures for a balanced dataset. Figures 3 and 4 show the sample images from X-ray and CT scans for covid and non-covid along with their histograms to indicate the variations in their intensities. The variation in background and foreground intensities for positive and negative covid cases can be seen in Figures 3 and 4.

2.3 | Metric for model evaluation

We have used the following metrics for evaluating our model’s performance. The metrics can be defined as follows.

\[
\text{Accuracy} = \frac{\text{number of correct predictions}}{\text{total number of predictions}}
\]  
\[
\text{Sensitivity/Recall} = \frac{\text{TP}}{\text{TP + FN}}
\]
\[ \text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad \text{(3)} \]

\[ \text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \text{(4)} \]

\[
\text{Dice coefficient/F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad \text{(5)}
\]

TP is the number of covid images identified precisely. TN is the number of normal images detected precisely. FP is the number of covid images erroneously identified. FN is the number of normal images erroneously identified.

3 | METHODOLOGY

The framework of the proposed MMCovid-Net model is categorized into training and testing phases with the following steps.

Phase 1: Training the model

i. Organization of input images

ii. Data preprocessing

iii. MMCovid-Net model development

Phase 2: Testing the model
i. Loading the model MMCOVID-NET
ii. Model evaluation

### 3.1 Phase 1: Training the model

The flow diagram of the training phase is shown in Figure 5. The chest CT dataset is divided into COVID and non-COVID CT images. Likewise, the X-ray images are divided into COVID X-ray images and non-COVID X-ray images. We have used the Multiclassifier approach to train the CNN models. The datasets used for our experiments are small to train deep neural networks. Hence, we have employed data augmentation techniques such as rotation, flips (horizontal and vertical), shearing, and normalizing image data. The size of images in CT scans and CXR’s are varied. These images are resized to (200, 200, 1).

#### 4 MMCOVID-NET Model Development

MMCOVID-NET models are developed based on several experiments for parameters and hyperparameters selection. Initially, a two-layered CNN model is considered for parameters and hyperparameters selection. We have performed experiments choosing appropriate learning rates, optimizers, and the number of epochs. Learning rates have an impact on training accuracy in neural networks. We have selected learning rates such as 0.01, 0.001, and 0.0001 for our experiments. Among them, the learning rate with the value 0.0001 produced better results. So, we have fixed the learning rate to 0.0001 for all our experiments. The most commonly used optimizers in deep neural networks are stochastic gradient descent (SGD), Adam, and RMSprop. Among them, Adam and RMSprop optimizers produced notable results. So, we have chosen Adam and RMSprop for our experiments. We have performed some experiments for fixing the number of epochs for training our CNN models. We trained the model for 10, 30, 50, 75, 100 epochs, respectively. The model seemed under-trained in 10 epochs. The model produced better results while trained with 30 epochs. The training/testing accuracy decreases when the model is trained for more than 30 epochs, the training/testing accuracy decreases. So we fixed 30 epochs for introducing our models. The various experiments done with the above said combinations of learning rates and optimizers for 30 epochs as given in Table 2.

Note: The bolded rows indicates the combination of hyperparameters which produced good results.

From the above experiments, we finalized the learning rate to 0.0001. We have chosen Adam and RMSprop for
carrying out further investigations. The number of epochs is fixed to 30. After fixing these parameters, we experimented by varying the convolutional layers from two to seven. We have also used batch normalization (BN) for normalizing the input on a mini-batch basis. With a smaller batch size, the training accuracy increases. We have used a batch size of 8 for our experiments. The architecture of the five-layered CNN model is shown in Figure 6. The same architecture shown in Figure 6 is used for performing all the other experiments by varying the convolutional block from two to seven with and without batch normalization. The experiments are carried out by changing the number of layers, as shown in Figure 7.

4.1 | Phase 2: Testing the models

The prediction phase of the MMCovid-NET model is shown in Figure 8. The following are the steps that are employed in the testing phase.

The input can be chest CT or chest X-ray (CXR) images. During the prediction phase, the saved model is loaded for predicting COVID-19. The input image is evaluated to detect the presence of COVID-19 by the MMCovid-NET model. The output is covid positive if the input image contains covid patterns. Otherwise, the result is displayed as covid negative.
TABLE 2  Experiments for fixing hyperparameters

| S.no | Model hyperparameters | Accuracy | Sensitivity | Specificity | Precision | F1 score |
|------|-----------------------|----------|-------------|-------------|-----------|----------|
| 1    | Two layers + Adam + Learning Rate (0.01) | 70       | 43          | 78          | 49        | 45       |
| 2    | Two layers + Adam + Learning Rate (0.001) | 72       | 45          | 80          | 51        | 47       |
| 3    | Two layers + Adam + Learning Rate (0.0001) | 75       | 48          | 83          | 54        | 50       |
| 4    | Two layers + RMSprop + Learning Rate (0.01) | 75       | 48          | 83          | 54        | 50       |
| 5    | Two layers + RMSprop + Learning Rate (0.001) | 82       | 58          | 90          | 69        | 63       |
| 6    | Two layers + RMSprop + Learning Rate (0.0001) | 87.5     | 72          | 94          | 82        | 76       |
| 7    | Two layers + SGD + Learning Rate (0.01) | 68       | 40          | 75          | 47        | 42       |
| 8    | Two layers + SGD + Learning Rate (0.001) | 70       | 43          | 78          | 49        | 45       |
| 9    | Two layers + SGD + Learning Rate (0.0001) | 73       | 46          | 82          | 52        | 48       |

Note: The bolded rows indicates the combination of hyperparameters which produced good results.

FIGURE 6  Architecture of five layered MMCOVID-NET Model

Layer (type)      | Output Shape | Param # |
------------------|--------------|---------|
conv2d (Conv2D)   | (None, 200, 200, 64) | 1792    |
max_pooling2d (MaxPooling2D) | (None, 100, 100, 64) | 0       |
conv2d_1 (Conv2D)  | (None, 100, 100, 64) | 36928   |
max_pooling2d_1 (MaxPooling2) | (None, 50, 50, 64) | 0       |
conv2d_2 (Conv2D)  | (None, 50, 50, 64) | 36928   |
max_pooling2d_2 (MaxPooling2) | (None, 25, 25, 64) | 0       |
conv2d_3 (Conv2D)  | (None, 25, 25, 64) | 36928   |
max_pooling2d_3 (MaxPooling2) | (None, 12, 12, 64) | 0       |
conv2d_4 (Conv2D)  | (None, 12, 12, 64) | 36928   |
max_pooling2d_4 (MaxPooling2) | (None, 6, 6, 64) | 0       |
flatten (Flatten)  | (None, 2304) | 0       |
dense (Dense)      | (None, 128) | 295040   |
dense_1 (Dense)    | (None, 50)  | 6450    |
dense_2 (Dense)    | (None, 4)   | 204     |

Total params: 451,198
Trainable params: 451,198
Non-trainable params: 0

5 | EXPERIMENTAL RESULTS AND DISCUSSIONS

We have conducted various experiments based on the fixed parameters and hyperparameters that are mentioned in Section 3, such as learning rate (0.0001), optimizers (Adam & RMSprop), and the number of epochs (30). After fixing these parameters, we experimented by varying the number of layers from two to seven. The experiments were carried out in two different datasets (Dataset 1 and 2). In Dataset 1, the experiments were carried out by varying the number of layers with batch normalization. We carried out 24 experiments in total for COVID-19 detection. Among them, four models produced the best results. We have inferred some of the best results obtained from 24 experiments in Table 3. The
model with three convolution layers and Adam optimizer produced an accuracy of 100%, and it is named multimodal covid network I (MMCOVID-NET-I). The model with four convolutional layers, Adam optimizer, and batch normalization produced an accuracy of 100%, and it is named multimodal covid network II (MMCOVID-NET-II). With the combination of five convolution layers, RMSprop optimizer, and batch normalization, the model produced 100% accuracy, and it is named the multimodal covid network (MMCOVID-NET-III). The model with seven convolutional layers with RMSprop optimizer alone produced 100% accuracy. This model is named multimodal covid network IV (MMCOVID-NET-IV).

MMCOVID-NET models produced 100% accuracy in Dataset 1 since it is relatively small compared to Dataset 2. We repeated our experiments by varying the number of layers with batch normalization in Dataset 2, which contained 2500 X-ray and CT images. MMCOVID-NET-III model produced the highest accuracy of 99.75%. MMCOVID-NET-I, MMCOVID-NET-II, and MMCOVID-NET-IV models had the accuracies 92, 94, and 92, respectively. The best experimental results of the models obtained from Dataset 2 are inferred in Table 4.
Our experiments concluded that the model could learn effectively and efficiently when the learning rate is reduced from all the experiments. Adam optimizer produces the best testing accuracy in models with fewer layers. RMSprop produces the best testing accuracy in models with more layers. In our experiments, the MMCovid-NET-I model, designed with only three convolution blocks, gave 100% test accuracy when experimented in Dataset 1, and it produced about 92% when experimented using Dataset 2. Likewise, the models designed with four, five, and seven convolution blocks also produced an accuracy of 100% in Dataset

| S. no | Model description | Accuracy (In percentage) | Sensitivity | Specificity | Precision | F1 score |
|-------|-------------------|--------------------------|-------------|-------------|-----------|----------|
| 1     | 2 convolution layers + RMSprop | 87.5                    | 72          | 94          | 82        | 76       |
| 2     | 3 convolution layers + Adam (MMCOVID-NET-I) | 100 | 100 | 99 | 98 | 98 |
| 3     | 3 convolution layers + Adam + Batch Normalization | 87.5 | 72 | 94 | 82 | 76 |
| 4     | 4 convolution layers + Adam | 84 | 62 | 91 | 73 | 67 |
| 5     | 4 convolution layers + Adam + Batch Normalization (MMCOVID-NET-II) | 100 | 100 | 99 | 98 | 98 |
| 6     | 5 convolution layers + Adam | 87.5 | 72 | 94 | 82 | 76 |
| 7     | 5 convolution layers + RMSprop + Batch Normalization (MMCOVID-NET-III) | 100 | 100 | 99 | 98 | 98 |
| 8     | 5 convolution layers + RMSprop | 87.5 | 72 | 94 | 82 | 76 |
| 9     | 6 convolution layers + RMSprop | 87.5 | 72 | 94 | 82 | 76 |
| 10    | 7 convolution layers + RMSprop (MMCOVID-NET-IV) | 100 | 100 | 99 | 98 | 98 |

Note: The bolded rows indicates the combination of hyperparameters which produced good results.

| S. no | Model description | Accuracy (In percentage) | Sensitivity | Specificity | Precision | F1 score |
|-------|-------------------|--------------------------|-------------|-------------|-----------|----------|
| 1     | 2 convolution layers + RMSprop | 87.5 | 72 | 94 | 82 | 76 |
| 2     | 3 convolution layers + Adam (MMCOVID-NET-I) | 100 | 100 | 99 | 98 | 98 |
| 3     | 3 convolution layers + RMSprop + Batch Normalization | 87.5 | 58 | 96 | 53 | 46 |
| 4     | 4 convolution layers + Adam + Batch Normalization (MMCOVID-NET-II) | 94 | 93 | 99 | 96 | 96 |
| 5     | 4 convolution layers + RMSprop | 87.5 | 58 | 96 | 53 | 46 |
| 6     | 5 convolution layers + Adam | 87.5 | 58 | 96 | 53 | 46 |
| 7     | 5 convolution layers + RMSprop + Batch Normalization (MMCOVID-NET-III) | 99.75 | 100 | 99 | 98 | 98 |
| 8     | 7 convolution layers + Adam + Batch Normalization | 93.75 | 93 | 99 | 96 | 96 |
| 9     | 7 convolution layers + RMSprop + Batch Normalization | 94 | 93 | 99 | 96 | 96 |
| 10    | 7 convolution layers + RMSprop (MMCOVID-NET-IV) | 92 | 90 | 98 | 94 | 94 |

Note: The bolded rows indicates the combination of hyperparameters which produced good results.
1 whereas 92%, 94%, 92% respectively in Dataset 2. MMCOVID-NET-I model, trained using only three convolution blocks (hidden layers), produced the best accuracy of 100% in dataset 1. It is sufficient to vary the hyperparameters such as the number of epochs, optimizers, learning rates, etc., which decides the accuracy of neural networks with a small number of hidden layers. MMCOVID-NET-II model trained using four layers, Adam optimizer and Batch Normalization produced 100% accuracy in dataset 1 and 94% accuracy in dataset 2. We also found that the models when trained using batch normalization produced the best results when compared to other models. Among the four MMCOVID-NET models, the MMCOVID-NET-III model is a five-layered CNN, with RMSprop optimizer and Batch Normalization outperformed. The main advantage of using the RMSprop optimizer is that it uses adaptive learning rates, i.e., changes the learning rates over time. This enables the model to learn the complex patterns from the input. RMSprop optimizer helps in avoiding both vanishing and exploding gradients while training the neural networks. The AUC Curve for MMCOVID-NET-III is high when compared to other models. This model produced consistent and best results when experimented with datasets 1 and 2. RMSprop performs well in the case of multi classifier problems. In the literature, some of the existing deep neural networks such as ResNet, VGGNet, GoogleNet, and InceptionV3 with many hidden layers are used to train the model, which is computationally expensive in execution time and resources needed for performing the experiments. The Area under Curve is plotted for all the proposed multimodal covid network models to test the model's behavior. The AUC curves for the proposed models are shown in Figure 9. The AUC scores of the proposed models in dataset 1 are listed in Table 5.

The performances of proposed MMCOVID-NET architectures are compared with state-of-the-art methods.

**FIGURE 9** AUC-ROC curve for proposed multimodal CNN's (A) MMCOVID-NET-I (B) MMCOVID-NET-II (C) MMCOVID-NET-III (D) MMCOVID-NET-IV
for COVID-19 prediction, and their results are inferred in Table 6. There are two state-of-the-art methods,\textsuperscript{22} in the literature that trained their models using CT and X-ray images. The results are inferred in Table 6. We have used the same dataset (Dataset 2) used in\textsuperscript{34} for comparing our proposed MMCOVID-NET models. We have also implemented\textsuperscript{22} using this same dataset and compared our results. MMCOVID-NET-III model outperformed most of the state-of-the-art methods by producing an accuracy of 99.75%.

| S.NO | Model | AUC score |
|------|-------|-----------|
| 1    | Proposed MMCOVID-NET-I | 98.13     |
| 2    | Proposed MMCOVID-NET-II | 97.46     |
| 3    | Proposed MMCOVID-NET-III | 98.90     |
| 4    | Proposed MMCOVID-NET-IV | 97.19     |

### Table 6: Comparison of proposed MMCOVID-NET models with other AI-based models for COVID-19 detection

| S.no | Method/Model | Accuracy |
|------|--------------|----------|
| 1    | ResNet50 and VGG16\textsuperscript{34} | 99.87    |
| 2    | DenseNet121 and MobileNet\textsuperscript{34} | 99.87    |
| 3    | Xception and InceptionV3\textsuperscript{34} | 98.80    |
| 4    | ResNet50\textsuperscript{34} | 98.27    |
| 5    | VGG16\textsuperscript{34} | 98.93    |
| 6    | DenseNet121\textsuperscript{34} | 98.27    |
| 7    | MobileNet\textsuperscript{34} | 97.87    |
| 8    | Xception\textsuperscript{34} | 96       |
| 9    | InceptionV3\textsuperscript{34} | 98.27    |
| 10   | 3 Layered Tailored CNN network\textsuperscript{22} | 93       |
| 11   | Proposed MMCOVID-NET-III | 99.75    |

Note: The bolded rows indicates the combination of hyperparameters which produced good results.

### 6 | CONCLUSION

In the proposed work, four multimodal covid network models are developed for MMCOVID-NET-I, MMCOVID-NET-II, MMCOVID-NET-III, and MMCOVID-NET-IV COVID-19 detection after performing various experiments. These models were trained using chest X-ray (CXR) and chest CT scan images. Existing works were focused on predicting COVID-19 from either CT scans or CXR’s images, whereas our proposed models are trained using both CT and X-ray scans. These models can detect the presence of COVID-19 in both CXR and CT scan images. In dataset 1, our models MMCOVID-NET-I, MMCOVID-NET-II, MMCOVID-NET-III, and MMCOVID-NET-IV produced 100% accuracy in detecting COVID-19. In dataset 2, MMCOVID-NET-III produced an accuracy of 99.75% competitively with state-of-the-art methods. The experiments carried out in this work conclude that the parameters and hyperparameters play a vital role in increasing or decreasing the model’s performance.

### ACKNOWLEDGMENTS

The authors wish to thank Dr. R. Rajeswaran, Radiologist, Sri Ramachandra University Medical College, Chennai and C. Naresh Kumar, Consultant Radiologist, Dindigul Scans Private Limited, Dindigul, Tamil Nadu, India for their help in qualitative validation.

### CONFLICT OF INTEREST

The authors declare no conflicts of interest.

### DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in references [29], [30], [31], [33].

### ORCID

Thiyagarajan Padmapriya \(\text{https://orcid.org/0000-0003-4714-7870}\)

Thiruvenkatam Kalaiselvi \(\text{https://orcid.org/0000-0002-0197-2077}\)

### REFERENCES

1. Novel coronavirus—China. Online 2020, Last accessed on 25.12.2020. http://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/
2. Rizk AM, Zidan MA, Emara DM, Abd El-Hady MA, Wahbi MO. Chest ultrasound in assessing patients in ICU: how can it help? Egypt J Radiol Nucl Med. 2017;48(1):313-322.
3. Li K, Fang Y, Li W, et al. CT image visual, quantitative evaluation, and clinical classification of coronavirus disease (COVID-19). Eur Radiol. 2020;30:1-10.
4. Hosseiny M, Kooraki S, Gholamrezaeezad A, Reddy S, Myers L. Radiology perspective of coronavirus disease 2019 (COVID-19): lessons from severe acute respiratory syndrome and the Middle East respiratory syndrome. Am J Roentgenol. 2020;214(5):1078-1082.
5. Zu ZY, Jiang MD, Xu PP, et al. Coronavirus disease 2019 (COVID-19): a perspective from China. Radiology. 2020;296(2):E15–E25.
6. Dimitrovski I, Kocev D, Kitanovski I, Loskovska S, Džeroski S. Improved medical image modality classification using a combination of visual and textual features. Comput Med Imaging Graph. 2015;39:14-26.
7. Kong W, Agarwal PP. Chest imaging appearance of COVID-19 infection. Radiol Cardiothor Imag. 2020;2(1):e200028.
8. Yoon SH, Lee KH, Kim JY, et al. Chest radiographic and CT findings of the 2019 novel coronavirus disease (COVID-19): analysis of nine patients treated in Korea. Korean J Radiol. 2020;21(4):494-500.
9. Zhao W, Zhong Z, Xie X, Yu Q, Liu J. The relation between chest CT findings and clinical conditions of coronavirus disease (COVID-19) pneumonia: a multicenter study. Am J Roentgenol. 2020;214(5):1072-1077.
10. Li Y, Xia L. Coronavirus disease 2019 (COVID-19): role of chest CT in diagnosis and management. Am J Roentgenol. 2020;214:1-7.
11. Kanne JP, Little BP, Chung JH, Elicker BM, Ketai LH. (2020). Essentials for Radiologists on COVID-19: An Update—Radiology Scientific Expert Panel.
12. Manna S, Wruble J, Maron SZ, et al. COVID-19: a multimodality review of radiologic techniques, clinical utility, and imaging features. Radiol Cardiothor Imag. 2020;2(3):e200210.
13. Huynh BQ, Li H, Giger ML. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J Med Imag. 2016;3(3):034501.
14. Hemdan EED, Shouman MA, Karar ME. (2020). COVID-19x-net: a framework of deep learning classifiers to diagnose COVID-19 in x-ray images. arXiv preprint arXiv:2003.11055.
15. Wang L, Lin ZQ, Wong A. Covid-net: a tailored deep convolutional neural network design for detecting covid-19 cases from chest x-ray images. Sci Rep. 2020;10(1):1-12.
16. Song Y, Zheng S, Li L, et al. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. IEEE/ACM transactions on computational biology and bioinformatics. 2020;18(6):2775–2780.
17. Wang S, Kang B, Ma J, et al. A deep learning algorithm using CT images to screen for corona virus disease (COVID-19). European radiology. 2020;1-9.
18. Zheng C, Deng X, Fu Q, et al. Deep learning-based detection for COVID-19 from chest CT using weak label. medRxiv. 2020.
19. Xu X, Jiang X, Ma C, et al. Deep learning system to screen coronavirus disease 2019 pneumonia. Engineering. 2020;6(10):1122–1129. arXiv preprint arXiv:2002.09334.
20. Barstugan M, Ozkaya U, Ozturk S. (2020). Coronavirus (COVID-19) classification using ct images by machine learning methods. arXiv preprint arXiv:2003.09424.
21. Chen X, Yao L, Zhang Y. (2020). Residual attention U-net for automated multi-class segmentation of COVID-19 chest CT images. arXiv preprint arXiv:2004.05645.
22. Mukherjee H, Ghosh S, Dhar A, Obaidullah SM, Santosh KC, Roy K. Deep neural network to detect COVID-19: one architecture for both CT scans and chest X-rays. Appl Intell. 2020;1-13.
23. Sethy PK, Behera SK. Detection of coronavirus disease (COVID-19) based on deep features. Preprints. 2020;2020030300.
24. Ozturk T, Talo M, Yildirim EA, et al. Automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput Biol Med. 2020;103792.
25. Narin A, Kaya C, Pamuk Z. (2020). Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. arXiv preprint arXiv:2003.10849.
26. Mangal A, Kalia S, Rajgopal K, Rangarajan K, Namboodiri V, Banerjee S, Arora C. (2020). CovidAID: COVID-19 detection using chest X-ray. arXiv preprint arXiv:2004.09803.
27. Farooq M, Hafeez A. (2020). Covid-resnet: a deep learning framework for screening of covid19 from radiographs. arXiv preprint arXiv:2003.14395.
28. Hall LO, Paul R, Goldgof DB, Goldgof GM. (2020). Finding covid-19 from chest x-rays using deep learning on a small dataset. arXiv preprint arXiv:2004.02060.
29. Cohen JP, Morrison P, Dao L. (2020). COVID-19 image data collection. arXiv preprint arXiv:2003.11597.
30. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. Chest-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2017:2097-2106.
31. Zhao J, Zhang Y, He X, Xie P. (2020). COVID-19-CT-dataset: a CT scan dataset about COVID-19. arXiv preprint arXiv:2003.13865.
32. Apostolopoulos ID, Mesias an TA. COVID-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med. 2020;43:1-640.
33. Angelov P, Almeida Soares E. SARS-CoV-2 CT-scan dataset: a large dataset of real patients CT scans for SARS-CoV-2 identification. MedRxiv. 2020.
34. Hilmizen N, Bustamam A, Sarwinda D. The multimodal deep learning for diagnosing COVID-19 pneumonia from chest CT-scan and X-ray images. 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI). IEEE; 2020, December:26-31.

How to cite this article: Padmapriya T, Kalaiselvi T, Priyadharsini V. Multimodal covid network: Multimodal bespoke convolutional neural network architectures for COVID-19 detection from chest X-ray’s and computerized tomography scans. Int J Imaging Syst Technol. 2022;1-13. doi:10.1002/ima.22712