



CovidConvLSTM: A fuzzy ensemble model for COVID-19 detection from chest X-rays

Subhrajit Dey ^a, Rajdeep Bhattacharya ^b, Samir Malakar ^c, Friedhelm Schwenker ^{d,*}, Ram Sarkar ^b

^a Department of Electrical Engineering, Jadavpur University, Kolkata, India

^b Department of Computer Science and Engineering, Jadavpur University, Kolkata, India

^c Department of Computer Science, Asutosh College, Kolkata, India

^d Institute of Neural Information Processing, University of Ulm, Ulm, Germany



ARTICLE INFO

Keywords:
Sugeno integral
COVID-19
Deep learning
ConvLSTM
Pneumonia
Fuzzy ensemble

ABSTRACT

The rapid outbreak of COVID-19 has affected the lives and livelihoods of a large part of the society. Hence, to confine the rapid spread of this virus, early detection of COVID-19 is extremely important. One of the most common ways of detecting COVID-19 is by using chest X-ray images. In the literature, it is found that most of the research activities applied convolutional neural network (CNN) models where the features generated by the last convolutional layer were directly passed to the classification models. In this paper, convolutional long short-term memory (ConvLSTM) layer is used in order to encode the spatial dependency among the feature maps obtained from the last convolutional layer of the CNN and to improve the image representational capability of the model. Additionally, the squeeze-and-excitation (SE) block, a spatial attention mechanism, is used to allocate weights to important local features. These two mechanisms are employed on three popular CNN models – VGG19, InceptionV3, and MobileNet to improve their classification strength. Finally, the Sugeno fuzzy integral based ensemble method is used on these classifiers' outputs to enhance the detection accuracy further. For experiments, three chest X-ray datasets, which are very prevalent for COVID-19 detection, are considered. For all the three datasets, it is found that the results obtained by the proposed method are comparable to state-of-the-art methods. The code, along with the pre-trained models, can be found at <https://github.com/colabpro123/CovidConvLSTM>.

1. Introduction

COVID-19 or severe acute respiratory syndrome coronavirus 2 (or SARS-CoV-2) has led to a global pandemic scenario in recent years and affected normal life for all people. COVID-19 is a highly contagious virus and spreads predominantly from person to person according to the reports of the World Health Organization (WHO). Patients infected with COVID-19 can carry the virus for a period ranging from two days up to two weeks before showing any symptoms. This is because an affected subject can spread the virus to others without even being aware that it is a carrier of the disease. Some of the common symptoms of the affected subjects are cough and cold, shortness of breath and fatigue. In certain cases, these symptoms turn into severe health complications like trouble in breathing, blue lips or face, pain in the chest, and also Pneumonia. Hence, early detection of the virus is very important for caring for the infected ones. One of the common ways of detecting

COVID-19 cases is by using the reverse transcription polymerase chain reaction (RT-PCR) test. However, the problem with this method is that it is costly to be used in developing countries like India, Bangladesh, and Sri Lanka. Moreover, this test is not completely accurate and as a result, it needs to be performed multiple times for assurance which makes this testing process more expensive. Though vaccines are now readily available in most parts of the world, however, to have a world with fully vaccinated people, we will need more time. Moreover, there is no assurance that fully vaccinated people will not get infected by COVID-19. Hence, early detection in a timely and cost-efficient way is of utmost need to fight this deadly disease.

Alternatively, specialist physicians use chest X-ray or computed tomography (CT) scan images to predict the presence of COVID-19. However, manual inspection of these images might lead to inaccuracies in detection due to the inter-/intra- personal observation and hence,

The code (and data) in this article has been certified as Reproducible by Code Ocean: (<https://codeocean.com/>). More information on the Reproducibility Badge Initiative is available at <https://www.elsevier.com/physical-sciences-and-engineering/computer-science/journals>.

* Corresponding author.

E-mail addresses: subhrajitdey.agt@gmail.com (S. Dey), rajdeep.cse17@gmail.com (R. Bhattacharya), malakarsamir@gmail.com (S. Malakar), friedhelm.schwenker@uni-ulm.de (F. Schwenker), raamsarkar@gmail.com (R. Sarkar).

this method becomes less reliable. Machine learning based techniques may help medical doctors in detecting COVID-19 cases using chest X-ray images or CT scan images (Santosh & Ghosh, 2021) or both (Mukherjee, Ghosh, Dhar, Obaidullah, Santosh, & Roy, 2021a). Ramdani, Allali, Chat, and El Haddad (2021) provide a comprehensive review for CT imaging for COVID-19 while Ardakani, Kanafi, Acharya, Khadem, and Mohammadi (2020) illustrate the results of 10 well-known convolutional neural network (CNN) models on CT images. Santosh (2020) is one of the earlier researchers who discussed the importance of using artificial intelligence (AI) - driven tools for COVID-19 detection and suggested using active learning combined with multitudinal or multimodal data in a cross-population train/test setup. Owing to the said reasons, developing a faster and more reliable technique to detect COVID-19 cases from the subject's chest X-ray or CT scan image becomes an omnipotent research area. Researchers worldwide are now using classical machine learning and deep learning models to predict COVID-19 cases from chest X-ray or CT scan images, for instance, Yang, Martinez, Visuña, Khandhar, Bhatt, and Carretero (2021). A choice between the two types of images, though, is governed by their success in the detection of diseases like breast cancer and lung diseases from the pathological images (Lakshmanaprabu, Mohanty, Shankar, Arunkumar, & Ramirez, 2019; Vaka, Soni, & Reddy, 2020). For example, in (Das, Santosh, & Pal, 2020; Jain, Gupta, Taneja, & Hemanth, 2021; Turkoglu, 2020), machine learning or deep learning approaches have been successfully applied for detecting COVID-19 cases from X-ray images. In this context, it is worth mentioning that getting a chest X-ray image is far cheaper than getting a CT scan image. Further, the computational need for processing a chest X-ray image is much lesser than that of a CT scan image. Additionally, current health guidelines recommend the use of chest X-ray images over CT Scan images (Dey, Bhattacharya, Malakar, Mirjalili, & Sarkar, 2021). Hence, considering all these factors, chest X-ray images are considered here for the detection of COVID-19.

In many previous research attempts related to COVID-19 detection, it has been observed that CNN based models (Dey, Bhattacharya, et al., 2021; Nour, Cömert, & Polat, 2020; Sahlol, Yousri, Ewees, Al-Qaness, Damasevicius, & Abd Elaziz, 2020; Turkoglu, 2020) yield more accurate results than the methods that use classification through classical machine learning algorithms with the help of hand-crafted features. This can be attributed to the fact that the filters used in convolutional layers of a CNN model help in extracting more accurate and robust features from image data which leads to improved classification results. In (e.g., Abbaszadeh & Hüllermeier, 2020; Nour et al., 2020; Sahlol, et al., 2020; Turkoglu, 2020), the features were fed directly to the classification models from the last layer of CNN models without improving their image representation capabilities.

However, a convolutional long short-term memory (ConvLSTM) network can be added at the end of the convolutional layers of a CNN model to obtain improved features. In the past, ConvLSTM networks were employed in cases where the input data are of spatio-temporal type (Majd & Safabakhsh, 2020; Mukherjee, Mondal, Singh, Sarkar, & Bhattacharjee, 2020; Rahman & Adjero, 2019; Shi, Chen, Wang, Yeung, Wong, & Woo, 2015; Zhu, et al., 2019). In the present work, this network is used to improve the representation of a single image as well as to take care of the spatial dependencies among different feature maps produced from a single input image. Additionally, a squeeze-and-excitation (SE) (Hu, Shen, & Sun, 2018) block which is a spatial attention technique is employed to give attention to the features in order to prioritize those which are important.

Apart from these, we find several COVID-19 techniques (Chandra, Verma, Singh, Jain, & Netam, 2020; Das, Ghosh, Thunder, Dutta, Agarwal, & Chakrabarti, 2021; Zhou, Lu, Yang, Qiu, Huo, & Dong, 2020) that focused on combining multiple CNN models rather than relying on a single model to come up with a strong classification model. However, in the literature, mostly ensemble algorithms in the form of weighted average (Das, Ghosh, et al., 2021) or majority voting (Chandra et al., 2020; Zhou, et al., 2020) have been used. The main problem of such

ensemble methods is that these are static in nature as these methods are dependent on Boolean sets. In contrast, the use of fuzzy integral based ensemble methods can help to get rid of this static binding as they combine the model outcomes dynamically and hence, they perform well as compared to the traditional ones.

Keeping the above facts in mind, in the present work, one of the well-known fuzzy integrals, called the Sugeno integral, is used to design a classifier ensemble. Three pre-trained models: VGG19 (Simonyan & Zisserman, 2015), InceptionV3 (Szegedy, et al., 2015) and MobileNet (Howard, et al., 2017) improved by using ConvLSTM networks and SE blocks are used as base classifiers. The pre-trained models for these complex networks are trained on a large dataset, known as ImageNet (Krizhevsky, Sutskever, & Hinton, 2012). Finally, two fully connected layers along with a softmax layer are added on the top of extracted features to build a classifier. These base classifiers are used to classify chest X-ray images of normal subjects, COVID-19 and Pneumonia infected patients. The entire model is named as CovidConvLSTM. During testing, the base classifiers are used to generate prediction scores first, and then these scores are combined using the Sugeno fuzzy integral to generate the final class prediction.

In a nutshell, the main highlights of the current work are as follows:

- Proposed a method to combine the outcomes of three deep learning based classifiers each using fine-trained models of VGG19, InceptionV3, and MobileNet, by applying the Sugeno fuzzy integral based ensemble method.
- For applying the Sugeno integral method, the fuzzy measure values are set experimentally.
- Along with the use of pre-trained models as feature extractors, each classifier also has a ConvLSTM layer and an SE block.
- Grad-CAM analysis is performed on an individual classifier to visualize its feature learning procedure.

The rest of the paper is organized as follows: in Section 2, some important research works are discussed that have taken place so far for COVID-19 detection, while in Section 3 the motivation behind our work is mentioned. Then in Section 4, the overall architecture of the present CovidConvLSTM model and the key components used in it are described. The Section 5 demonstrates the experiments performed and the detailed results obtained by the proposed method. Finally, in Section 6, the paper is concluded.

2. Related work

Several machine learning and deep learning models have been applied for COVID-19 detection using chest X-ray images. Most of these models relied on some vanilla CNN models (majority of which utilized the transfer learning concept) to extract features. These methods were able to perform well with the use of the deep features. For example, Goyal and Singh (2021) designed fusion and normalization features based recurrent neural network (RNN)-long short-term memory (LSTM) (abbreviated as F-RNN-LSTM) that relied on only handcrafted features comprised of gray-level co-occurrence matrix (GLCM), histogram of oriented gradients (HOG), intensity and geometric features. However, this model was empowered with a deep learner known as RNN-LSTM model. Also, Panetta, Sanghavi, Agaian, and Madan (2021) proposed a Fibonacci-p patterns-based shape-dependant feature descriptor. In some cases, handcrafted features were coupled with deep features. For example, Senan, Alzahrani, Alzahrani, Alsharif, and Aldhyani (2021) stacked GLCM, local binary pattern (LBP), and deep features extracted using a self-designed and scratch-trained CNN model. Some of the methods that extracted deep features used shallow learners like support vector machine (SVM) (Ismael & Şengür, 2021; Kedia, Katarya, et al., 2021; Nour et al., 2020), k-nearest neighbors (Nour et al., 2020), and XGBoost (Das, Roy, Malakar, Velásquez, & Sarkar, 2021) in order to reduce the training cost or get rid of the overfitting problem of deep learners. These models used either some state-of-the-art pre-trained

CNN models like ResNet50, ResNet101, VGG16, VGG19, DenseNet121, and DenseNet169 as a feature extractor (Das, Roy, et al., 2021; Ismael & Şengür, 2021; Kedia et al., 2021) or trained the CNN model from scratch (Nour et al., 2020).

However, most of the researchers relied only on CNN based models. For example, Hemdan, Shouman, and Karar (2020) reported the results of using seven different pre-trained CNN models to differentiate between COVID-19 and normal cases. Besides proposing one of the first open-source network designs for detection of COVID-19 from chest X-ray images, the authors in Wang, Lin, and Wong (2020) also introduced COVIDx which is among the largest datasets in terms of the number of COVID-19 CXR images that are publicly available and it has also been used in the present work. In contrast to this, Mukherjee, Ghosh, Dhar, Obaidullah, Santosh, and Roy (2021b) used a shallow-CNN tailored architecture, with a relatively smaller number of CNN parameters, to efficiently detect COVID-19 images. The authors Khan, Shah, and Bhat (2020) proposed a method utilizing a pre-trained Xception model. In another work, Roy, Shai, Ghosh, Bej, and Pati (2021) proposed a novel network consisting of an U-Net based architecture along with hybrid residual and dense connections which, aided by alpha-trimmed average pooling, proved to be very efficient in the rapid detection of COVID-19 cases. A transfer learning based method was proposed by (Brunese, Mercaldo, Reginelli, & Santone, 2020) wherein the chest X-rays were used to detect pulmonary disease and COVID-19, and additionally point out the areas in the image which are especially symptomatic of COVID-19 if the image belongs to that class.

Most of the mentioned methods did not employ any notable pre-processing or data augmentation mechanism to boost their result. However, there are some researchers who used data augmentation methods explicitly to include variations in the training data, or concentrated on image enhancement to improve the quality of chest X-ray images, or both. For example, Bashar, Latif, Ben Brahim, Mohammad, and Alghazo (2021) proposed an optimized approach in three stages — initially, image enhancement was performed, followed by data augmentation, and finally, the images were fed to transfer learning algorithms using some standard CNN architectures through which the images were classified. Chowdhury, et al. (2020) used image augmentation but did not employ image enhancement. The authors in Horry, et al. (2020) introduced a semi-automated pre-processing technique for chest X-ray images to create a trustworthy dataset followed by a deep learning framework for the classification task. In this context, it is to be noted that Rahman, et al. (2021) studied, in detail, the effect of various image enhancement techniques on COVID-19 detection using some standard CNN models. Also, Goyal and Singh (2021) used histogram equalization to enhance the quality before applying an adaptive image segmentation method to extract the region of interest, which was then used for classification purposes.

Some of the deep learning assisted methods used feature selection techniques, classifier ensemble methods, and CNN-LSTM models. These methods were primarily introduced to improve the end performance of the methods. (Turkoglu, 2020) first extracted features using a pre-trained AlexNet model and then applied ReliefF to obtain an optimal feature subset. The classification of the chest X-ray images was done by using an SVM classifier. In the work Dey, Chattopadhyay, et al. (2021), the authors used two pre-trained CNNs combined with a manta ray foraging based golden ratio optimizer (MRFGRO) to effectively detect COVID-19 from chest CT images. Goel, Murugan, Mirjalili, and Chakrabarty (2021) proposed a CNN model for COVID-19 detection, however, they additionally used the grey wolf optimizer (GWO) algorithm to optimize the hyperparameters for the training of the CNN layers. In another work, Sahlol, et al. (2020) used the Inception-Net model to extract features, and then applied the fractional order marine predators algorithm to obtain a near optimal feature subset. Another CNN based model was proposed by Togaçar, Ergen, and Cömert (2020) where first, the features were extracted using MobileNet and SqueezeNet, and then these extracted features were optimized using

social mimic optimization. The authors of Canayaz (2020) used pre-trained CNN models like AlexNet, VGG19, GoogleNet, and ResNet as feature extractors, and later, they used metaheuristic algorithms like particle swarm optimization (PSO) and GWO for the feature selection task.

Classifier ensemble scheme was utilized by some researchers, mostly using either majority voting (Chandra et al., 2020; Zhou, et al., 2020) or weighted average (Das, Ghosh, et al., 2021; Paul, Basu, Mahmud, Kaiser, & Sarkar, 2022) scheme. Zhou, et al. (2020) used three pre-trained models of AlexNet, GoogleNet and ResNet as feature extractors. Outcomes of these three classifiers were then combined using an approach called relative majority voting. They used CT scan images for COVID-19 detection. Chandra et al. (2020) used a two-stage classification approach by performing a majority voting based classifier ensembling of five supervised classification algorithms. Three pre-trained models of DenseNet201, ReNet50V2, and Inception v3 were used by Das, Ghosh, et al. (2021), outputs of which were later combined using a new weighted average ensembling method. Apart from these, the use of LSTM layers can be seen in Naeem and Bin-Salem (2021) where multilevel features were extracted using scene context gist features, scale invariant feature transform, and CNNs. After that LSTM was used to detect the features. Similarly, Islam, Islam, and Asraf (2020) proposed another LSTM network for COVID-19 classification where a CNN was trained from scratch for feature extraction after which LSTM layers were applied for classification. The authors in Aslan, Unlersen, Sabancı, and Durdu (2021) proposed a hybrid architecture by developing a CNN-based BiLSTM network that also utilized transfer learning and was effective for early detection of COVID-19. It is noteworthy to mention here that all these LSTM based models used LSTM or biLSTM as classifier whereas in the proposed work, ConvLSTM is used to learn correlation among feature maps generated from a CNN model.

Fuzzy integrals and fuzzy measures, as described in Grabisch, Sugeno, and Murofushi (2010), take classifier uncertainties into account because of which using fuzzy integral based ensembling methods results in more dynamism and an increase in classification performance. In Du, Zare, Keller, and Anderson (2016), the authors proposed a method capable of combining multiple two-class classifiers by learning fuzzy measures using an evolutionary algorithm. The learned fuzzy measures were then combined by Choquet integral to fuse the classifiers' outcomes for achieving greater classification performance. In Dey, Bhattacharya, et al. (2021), the authors proposed an ensemble model composed of CNNs as base learners, fused with a Choquet integral to correctly detect COVID-19, Pneumonia, and Normal cases. Also, in Abbaszadeh and Hüllermeier (2020), a method for binary classification was proposed where, as the aggregation function, the Sugeno fuzzy integral method was used to combine the different local evaluations into a single global evaluation method.

In many of the recent research articles like in Hu et al. (2018), the authors have shown that representations that are produced by the CNNs can be improved by integrating learning methods into the network. Such use helps capture the spatial dependencies among the feature maps. One way of achieving this is by using ConvLSTM networks which were initially introduced for applications like precipitation nowcasting (Shi, et al., 2015), estimation of biological age from physical activity (Rahman & Adjeroh, 2019), human action recognition (Majd & Safabakhsh, 2020) etc. In the research work of Rahman and Adjeroh (2019), the authors used ConvLSTM for biological age prediction by using data from wearable devices. Here, convolutional layers were used for the extraction of features but as the input data was a temporal sequence of daily activity records, the application of LSTM was also useful. It helped in establishing an association in each of the daily records of the locomotive activity data which was collected over some time. In this way, by establishing a relation among each record, the researchers were able to find clues that helped them to predict the biological age. Another research work Majd and Safabakhsh (2020) used the correlation CNN model for human action recognition from

video data. Here, the authors fed each frame of the video to the convolutional layers which resulted in the extraction of features from each frame, and later concatenated all the features from the other frames of the video. After that the model used the ConvLSTM layers to understand the correlation among the frames of the video. The problem with the above research works and also other works which used ConvLSTM layers is that they did not use it to learn correlation among the feature maps extracted by the convolutional layers of a CNN model. Establishing correlation among features results in capturing the contextual data hidden in them. Hence, based on these facts, in the present work, ConvLSTM networks are used on the pre-trained models generated feature maps by treating them as spatiotemporal data and passing them through ConvLSTM to improve the interdependency among the feature maps.

This learning ability can further be enhanced by introducing an attention mechanism among the features so that the important features are given a higher priority and to that end, SE blocks are incorporated into the proposed model. They were introduced in [Hu et al. \(2018\)](#) and subsequently used in various types of applications like in [Chen, Zhang, Zhong, Chen, and Yu \(2019\)](#) where the researchers proposed a 3D Near-infrared facial recognition system along with an SE block. In another case ([Li, Liu, Cui, Guo, Huang, & Hu, 2020](#)), EEG seizure detection framework was introduced which used a novel channel-embedding spectral-temporal SE block. An application of the SE block can also be seen in remote sensing which was proposed in [Gu, Sun, Zhang, Fu, and Wang \(2019\)](#). These works bear testament to the effectiveness of the SE block and hence, the same has been used in the present work to design base classifiers.

3. Motivation

In the light of the current pandemic situation, an efficient means for the detection of COVID-19 is a raging need. While there are other methods like the RT-PCR test, as mentioned before, the method is both costly and not full-proof. Hence, the primary motivation of this work is to design an efficient as well as an inexpensive decision support system for COVID-19 detection which can be used as an alternative to the RT-PCR test. Several methods have used the power of different vanilla CNN models which have, in recent history, proven to be effective to solve the mentioned problem. To this end, here, an initial goal of this work is to improve the performance of such vanilla CNN models and thus, a unique combination of deep learning tools is utilized. These deep learning tools, through their strengths, can improve image representation and thus, the classification ability. Later, by utilizing the benefit of the fuzzy integral based aggregator over traditional aggregators, an improved COVID-19 detection model is designed. The use of a classifier ensemble method helps in alleviating the chance of erroneous classification that can happen by relying on a single classifier.

Initially, the concept of transfer learning is considered for training the vanilla CNN models used here. Subsequently, a ConvLSTM layer and an SE block are used on the top of three efficient CNN models, namely VGG16, Inception V3, and MobileNet to improve their image representation capability, and thus their classification power. The ConvLSTM layer is introduced primarily to learn spatial dependencies among the feature maps generated by a CNN model and thus, it helps to obtain a more correlated feature map. Further, to capture and increase the focus on the important features in the output feature map, whereas an SE block is used to increase attention in those areas. Also, Sugeno fuzzy integral method is used here to fuse the outputs from the three improved CNN models. All these mentioned steps cumulatively produce a robust and efficient model where each of the said components has their own role to play, supported by proper logic, and thus generate competitive results when compared to some state-of-the-art methods.

4. Proposed method

In this work, an ensemble method of three classifiers by using the Sugeno integral method is proposed for COVID-19 detection from chest X-ray images. These classifiers are based on the pre-trained models of VGG19, InceptionV3, and MobileNet, and in all three classifiers, ConvLSTM layers are used for establishing spatial dependency encoding on the extracted features. Next, SE blocks are used to suppress the less useful features obtained from the previous steps, and then the chest X-ray images are classified using two fully connected (FC) layers with the generated features. Lastly, the output predictions from the base classifiers (i.e., pre-trained models + ConvLSTM + SE Blocks + 2 FC layers) are combined using the Sugeno fuzzy integral method to generate the final prediction results. The architecture of the entire method is shown in [Fig. 1](#). In this section, the basic concepts that have been used in this work like CNN, ConvLSTM network, SE block, transfer learning, and Sugeno integral method are discussed. Also, the overall architecture of the proposed method is described.

4.1. Convolutional neural network

CNN ([LeCun, Haffner, Bottou, & Bengio, 1999](#)) is a deep learning model that is used in various applications like image classification, object detection etc. The reason for this is that they are able to extract image features in a hierarchical fashion. If, in a network, a number of convolutional layers are used, then with the increase in depth of layers, the feature maps generated by the layers get more complex. In applications like COVID-19 detection where the differences in chest X-ray images of different classes is not that obvious, this capability of convolutional layers to capture complex image features is very important. A convolutional layer is able to capture image features by using learnable filters. After passing an image through these learnable filters, feature maps are generated which contain spatial data of the image. The better the quality of spatial information present in the feature maps, the better it is in terms of the image representational capability of the model. Optionally, the feature maps are flattened and passed through an appropriate approximation function for classification. However, in the present work, the CNN models are used as a feature extractor. Thus, the main role of a CNN is to transform the image taken in as an input into a form that can easily be analyzed without losing any useful information required for prediction.

4.2. Convolutional LSTM layer

ConvLSTM layer was introduced to deal with spatio-temporal data as shown in [Shi, et al. \(2015\)](#) i.e., it was used to deal with data having space and time related features like video, precipitation nowcasting data, sensor based human activity data, etc. Based on their ability to handle spatio-temporal data or time series data, they were used in various types of applications like prediction of biological age from daily activity records ([Rahman & Adjero, 2019](#)), human action recognition from video input ([Majid & Safabakhsh, 2020](#)), and gesture recognition from video data ([Zhu, et al., 2019](#)). In all of the mentioned applications, the ConvLSTM layers help in capturing the temporal dependencies present in the input data. They do this by using LSTM networks which are known for passing and storing information from previous steps to the present step of the underlying time series data. Hence, they establish an encoding in the input data. The internal working structure of the ConvLSTM layer is shown in [Fig. 2](#).

In many recent research articles like ([Hu et al., 2018](#)), it has been shown that the image representational capability of the convolutional layer can be significantly improved by increasing the interdependency among the features extracted by the convolutional layer. Hence, in this work, ConvLSTM layers are employed on feature maps generated from a pre-trained CNN model so that it can establish a spatial dependency encoding on the extracted feature maps in the same way it does on

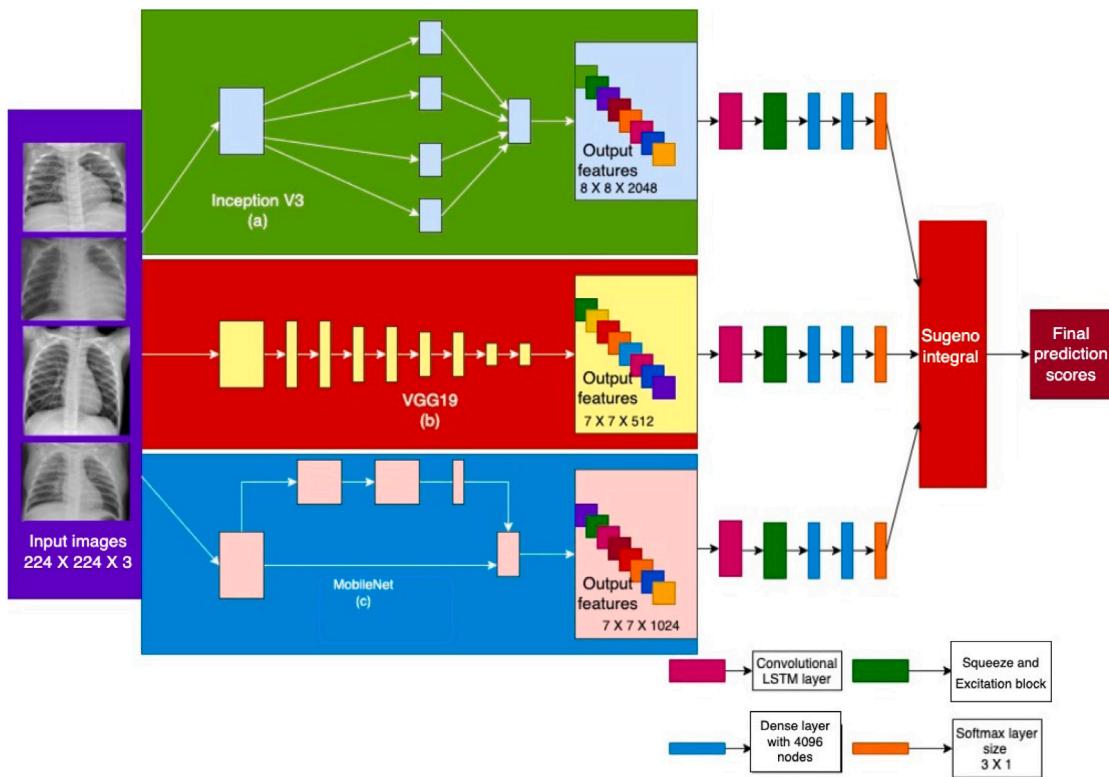


Fig. 1. The proposed Sugeno fuzzy integral based ensemble model with ConvLSTM networks and the SE block used for COVID-19 detection from chest X-ray images. Here, (a) denotes the CNN architecture which uses the pre-trained Inception V3 model, (b) denotes the CNN architecture with pre-trained model of VGG19 and (c) denotes the CNN architecture using pre-trained MobileNet model.

the temporal data. In a typical LSTM, the most important part is the cell state c_t , which is used to store information. If the input gate i_t gets activated, then the input value gets stored, and if the forget gate f_t gets activated, then the previous state c_{t-1} gets forgotten. Moreover, the output cell o_t controls whether the current cell state c_t will get converted to the final hidden state h_t . This is how a vanilla LSTM model works. However, in case of ConvLSTM layer, the inputs x_1, x_2, x_3, \dots , the cell states c_1, c_2, c_3, \dots , the hidden states h_1, h_2, h_3, \dots and the gates i.e., i_t, f_t, o_t all are 3D tensors. To understand the working principle of the ConvLSTM layer, first the inputs and gates are considered as vectors in a grid-like fashion in the spatial space. For a particular cell, the ConvLSTM layer predicts the next state of the cell by collecting the inputs and the last state of the local entities of the particular cell. One can understand this by looking into the following equations.

$$\begin{aligned} i_t &= \sigma(w_{x_i} * x_t + w_{h_i} * h_{t-1} + w_{c_i} \odot c_{t-1} + d_i) \\ f_t &= \sigma(w_{x_f} * x_t + w_{h_f} * h_{t-1} + w_{c_f} \odot c_{t-1} + d_f) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tanh(w_{x_c} * x_t + w_{h_c} * h_{t-1} + d_c) \\ o_t &= \sigma(w_{x_o} * x_t + w_{h_o} * h_{t-1} + w_{c_o} \odot c_{t-1} + d_o) \\ h_t &= o_t \odot \tanh(c_t) \end{aligned}$$

Here '*' represents the convolution operation, whereas the ' \odot ' represents the Hadamard product. To maintain the equal number of rows and columns, padding is employed throughout the operation.

4.3. Squeeze and excitation block

The concept of the SE block was introduced by Hu et al. (2018). Such blocks can be placed in large CNNs to establish attention among the features. By doing so, an SE block creates a system where the most important features are highlighted more, whereas the features which are not that important are suppressed. In this way, the performance of a deep learning model can be improved by suppressing redundant

or non-important features from being fed to the classification model as noisy data that may hamper the performance of the overall model. The working procedure of the SE blocks and also the layers used for creating an SE block are shown in Fig. 3.

As can be seen from Fig. 3, the term ratio, also known as reduction ratio, helps to control the computation cost and capacity of the SE blocks. SE block improves the CNN model's image representational quality by using a mechanism which can perform feature recalibration. Due to this, the model is able to use global information, and thus giving priority to those features which are important and suppressing the ones which are not. Also, placing the SE blocks at different layer depths can give different results.

4.4. Pre-trained CNN models

There are many large and complex CNN models that have been proposed in the past which have produced state-of-the-art results in various image classification problems. Some examples of these include AlexNet, VGG19, DenseNet and ResNet. These models were trained on the ImageNet dataset which houses approximately 20,000 categories of images. However, the problem is that not everyone has access to large computational power, and also for some applications, it is not possible to get access to large datasets. An example is the case of COVID-19 classification where images of chest X-rays of patients affected by COVID-19 are required. Hence, as an alternative, the concept of pre-trained CNN models is used. Here, layer weights of the large CNN models like AlexNet, DenseNet etc., are frozen after being trained on large datasets like ImageNet. Thus, researchers who use low computational devices and small datasets can remove the last classification layers of the pre-trained models, add their own custom layer, and use the above frozen layers as feature extractors. The CNN models used here are described hereafter.

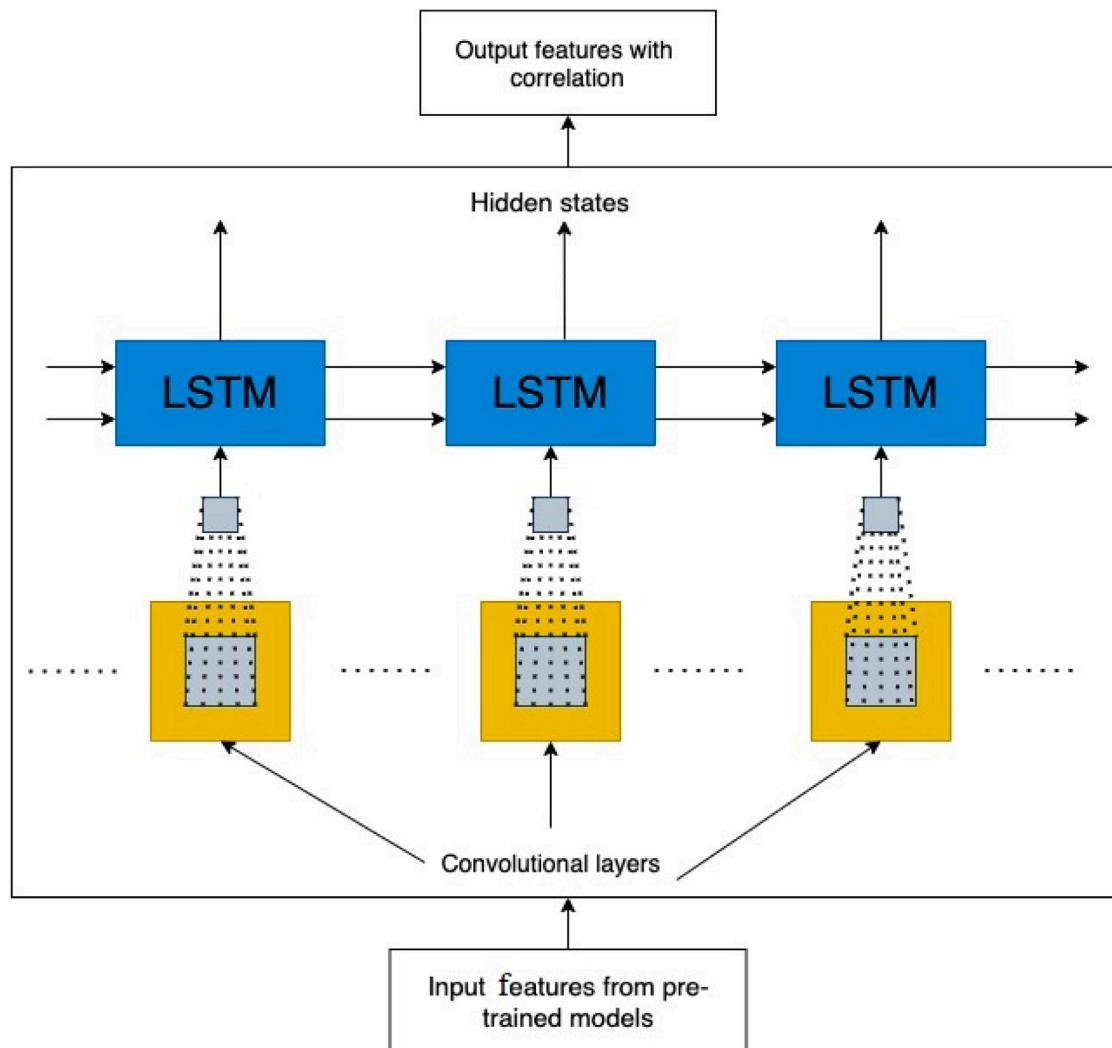


Fig. 2. The internal structure and the working procedure of the ConvLSTM layer. It shows that the convolutional layers being applied on the input feature maps which are then fed to LSTM units.

- **VGG19:** VGG19 is a deep and complex CNN model which is the successor to the AlexNet model. In total, there are 19 layers in VGG19 and all the convolutional layers present in it have a kernel size of 3 with strides of 1 which allows it to capture features which contain information about the whole image. Also, padding was included with convolutional layers to maintain the resolution. Maxpooling is used with a window of 2×2 along with strides of 2. After that, the rectified linear unit (ReLU) activation function is used to give non-linearity. At the end of the network, there are two dense layers and a softmax layer with 1000 classes.
- **Inception V3:** Inception architecture is different when compared to other CNN architectures in the sense that Inception Net architecture uses convolutional layers with multiple kernel sizes at the same level to extract different types of features. The filter sizes used are 1×1 , 3×3 and 5×5 . Subsequently, the outputs from the above layers are concatenated and sent to the next layers. Thus, InceptionNet focuses on making the overall network wider rather than deeper which helps in extracting improved features with less computational cost. There are various types of Inception-Net architecture and one of them is Inception V3 which has been used here as a feature extraction model. A smart factorization technique to convert 7×7 into three 3×3 convolutions is used in Inception V3 model that helps in reducing computational cost. Before the last layer in Inception V3 which has a convolution

of size 17×17 , it uses an auxiliary classifier which acts as a regularization layer. Inception V3 also uses label smoothing which acts as a regularization technique and prevents any logit values from becoming too large.

- **MobileNet:** MobileNet architecture was proposed to work efficiently on mobile devices and give results as accurate as larger and more complex architectures like VGG19. MobileNets use just one normal convolutional layer and for other layers, they use depthwise separable convolutions. Depthwise convolutions, unlike normal convolutions, perform convolutions on each image channel separately. Each has its own sets of weights. Depthwise separable convolutions may also have channel multipliers which means that for a multiplier value of m , each input channel will result in m number of output channels. After that, a pointwise convolution is performed which is nothing but a normal convolution with a kernel size of 1×1 which acts as a weighted average. Both the depthwise convolution and pointwise convolution together form the depthwise separable convolution. Though the end result of both the depth-wise separable convolutions and normal convolutions are similar, in the case of depthwise separable convolutions, the number of trainable parameters is less which results in less computational cost.

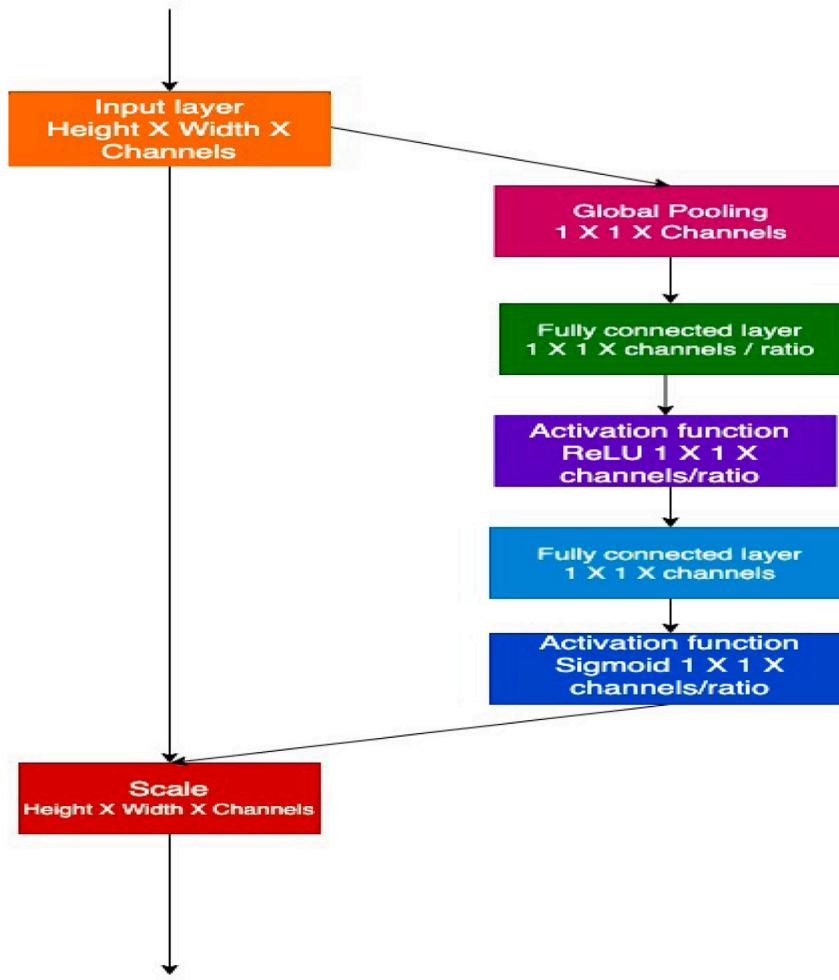


Fig. 3. The working principle of the SE block. Here the term ‘ratio’ is actually known as reduction ratio, which is set to 16 in the current work.

4.5. Sugeno fuzzy integral

Sugeno fuzzy integral is defined over a fuzzy set which defines inexact and subjective concepts for some situations where one cannot use probabilistic or deterministic measures to describe a situation. The fuzzy set concepts are widely used in many research problems that deal with pattern recognition (Mitra & Pal, 2005) and decision making (Abbaszadeh & Hüllermeier, 2020). The Sugeno fuzzy integral, which is defined over fuzzy sets, was introduced by Sugeno (1993). It combines data from various sources according to the value of their fuzzy measures. Before defining the Sugeno fuzzy integral, some related terminologies are first described here.

Fuzzy Set: In fuzzy systems, fuzzy sets are used to interpret and analyze imprecise information or data. A fuzzy set (say, A) is a pair (X, μ_A) , where X is a set, sometimes also called universal set, (most of the cases $X \neq \emptyset$, where \emptyset is empty set) whereas, $\mu_A : X \rightarrow [0, 1]$ is called membership function. Now, A can be defined as $A = \{(x, y) : x \in X \wedge y = \mu_A(x)\}$.

Fuzzy Subset: Let, $A = (X, \mu_A)$ and $B = (X, \mu_B)$ are two fuzzy sets. Now, A is called subset of B or A is said to be included in B (i.e., $A \subseteq B$) if and only if $\mu_A(x) \leq \mu_B(x), \forall x \in X$.

Monotonic Fuzzy Measure: Let, (X, A, μ) is a measurable space. The function $\mu : A \rightarrow [0, \infty]$ is a fuzzy measure if it satisfies the following properties:

- $\mu(\emptyset) = 0$ and $\mu(A) = 1$
- $\forall M, N \subseteq A$, if $M \subseteq N$, then $\mu(M) \leq \mu(N)$

- If $M_n \in A$, $\forall M_i \subseteq A$, $i = 1, 2, \dots, \infty$ and $M_1 \subseteq M_2 \subseteq M_3 \dots$ then

$$\lim_{i \rightarrow \infty} \mu(M_i) = \mu(\lim_{i \rightarrow \infty} M_i)$$

Here, $\lim_{i \rightarrow \infty} \mu(M_i)$ is the r-limit (Burgin, 2000) of the sequence $s = \{\mu(M_i) \in A : i \in \mathbb{N}\}$. A number a is called an r-limit of s (also represented by $a = r - \lim_{i \rightarrow \infty} \mu(M_i)$ or $a = r - \lim s$) if for any $\epsilon \in \mathbb{R}^+$, the inequality $d(a, \mu(M_i)) < r + \epsilon$ (where, $d(x, y) = |x - y| \forall x, y \in \mathbb{R}$) is valid.

The fuzzy measure actually represents the information about the sources that are being combined by the fuzzy integral. In other words, the fuzzy measures tell the integral about the importance or relevance of each source to the final result. So, for defining Sugeno integral, let us first assume that μ denotes a fuzzy measure on $A = \{a_1, a_2, \dots, a_n\}$ which denotes n entities and whose ranges are $\{g(a_1), g(a_2), \dots, g(a_n)\}$ where g denotes a function given by $g : A \rightarrow [0, 1]$ and also $g(a_1) \leq g(a_2) \leq \dots \leq g(a_n)$, then the Sugeno integral of the function $g : A \rightarrow [0, 1]$ with respect to the fuzzy measure μ is given by,

$$\int_{i=1}^n g(a_i) d\mu = \max_{1 \leq i \leq n} (\min(g(a_i), \mu(a_i))) \quad (1)$$

where $A_1 = \{a_1, a_2, \dots, a_n\}$, $A_2 = \{a_2, a_3, a_4, \dots, a_n\}$, $A_3 = \{a_3, a_4, \dots, a_n\}$ that means $A_x = \{a_x, a_{x+1}, \dots, a_n\}$, based on the condition $g(A_1) \leq g(A_2) \leq \dots \leq g(A_n)$. The membership values of $\{a_i\}$ (i.e., $\mu(\{a_i\})$) are the importance values provided by the experts of the problem domain. However, $\mu(\{A_x\})$'s are obtained using Sugeno λ -fuzzy measure whose characteristic equation is given by the equation

$$1 + \lambda = \prod_{i=1}^n (\mu(\{a_i\}) * \lambda + 1) \quad (2)$$

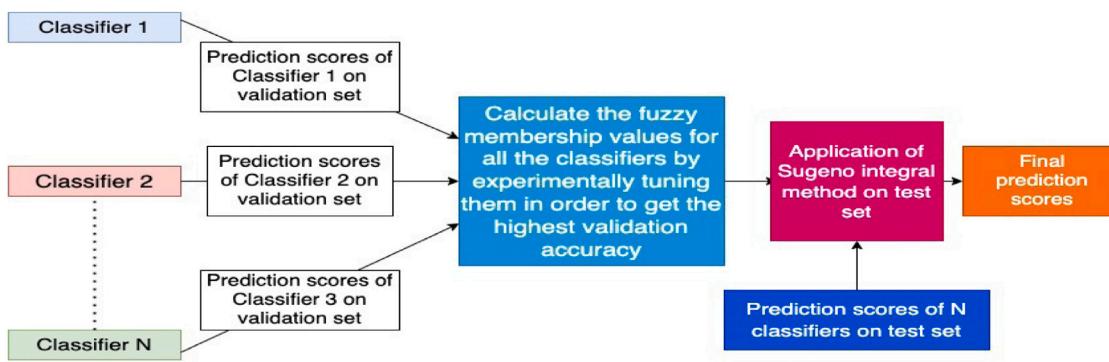


Fig. 4. Classifier ensembling using Sugeno fuzzy integral method. Here, the fuzzy membership values are calculated through experimental tuning of the validation accuracies of the classifiers.

Once the roots of the characteristic equation are obtained, one can decide the λ value. Now, using this λ value, any $\mu(A_x)$ value can be calculated by repetitive use of the following equation.

$$\mu(\{a_l, a_o\}) = \mu(\{a_l\}) + \mu(\{a_o\}) + \lambda * \mu(\{a_l\})\mu(\{a_o\}) \quad (3)$$

where $1 \leq l, o \leq n$.

In a classifier combination approach, p classifiers are used to classify q classes. Let, x_{ij} represent the confidence score of j th class of i th classifier. Now, for each j , Sugeno integral is used to obtain fuzzy confidence score. In this case for each $j = 1, 2, \dots, q$, $a_i = x_{ij}$ where $i = 1, 2, \dots, p$. The fuzzy membership values for each classifier (i.e., $\mu(\{a_i\})$) of Eq. (1) are set experimentally. To do this, a validation set is used to tune these values optimally. The rest of the fuzzy measure values for the classifier combination are calculated using Eq. (3) after finding out the value of λ from the Eq. (2). The ensemble of the classifiers using the Sugeno integral has been shown in Fig. 4.

4.6. Proposed system architecture

At first, 224×224 dimensional chest X-ray images are inputted to the base classifiers. Pre-trained models: VGG19, Inception V3, and MobileNet extract feature maps of dimension $7 \times 7 \times 512$, $8 \times 8 \times 2048$, and $7 \times 7 \times 1024$ respectively as can be seen from Fig. 1. Next, these feature maps are passed through ConvLSTM networks with multiple kernel sizes and the outputs so obtained are fused. The reason for using ConvLSTM blocks is that the convolutional layers, after extracting feature maps from an input image, send them to an LSTM layer which passes the most relevant information and discards the unnecessary information by encoding the spatial dependencies among the feature maps. Next, SE attention is employed on the feature maps from the ConvLSTM layer to give priority to those features which are important and suppress the ones which are not so important. For calculating the fuzzy measure values for the application of Sugeno fuzzy integral, a validation dataset is used. During testing, first a test sample is passed through the base classifiers to obtain prediction scores and then using these scores in Sugeno fuzzy integral we obtain final prediction label. The main steps of the proposed method are shown in Algorithm 1.

5. Results and discussion

The CovidConvLSTM is designed to perform classification of chest X-ray images of Normal subjects, COVID-19 and Pneumonia infected patients. In this section, the datasets and hyperparameters in use are first described, and then we illustrate the results and associated discussion. In the end, the results of the present method are compared to some state-of-the-art methods on the datasets considered here.

Algorithm 1: Key steps of the Proposed method

```

Input: Images for training  $I$   

        Images for testing  $S$   

Output: Normal or COVID-19 or Pneumonia cases.  

/* Model training phase and fuzzy measure value calculation */  

1 Resize training images  $I$  to  $224 \times 224 \times 3$ .  

2 Extract feature maps from  $I$  using pre-trained models of VGG19, InceptionV3 and MobileNet (as described in Section 3.4). Let the feature maps generated by VGG19, InceptionV3 and MobileNet be  $FS1$ ,  $FS2$  and  $FS3$  respectively.  

3 Apply the ConvLSTM (see Section 3.2) layer on the  $FS1$ ,  $FS2$  and  $FS3$ . Let the modified feature maps of  $FS1$ ,  $FS2$  and  $FS3$  be  $fs1$ ,  $fs2$  and  $fs3$  respectively.  

4 Utilize the SE block based attention mechanism (see Section 3.3) on the modified feature maps  $fs1$ ,  $fs2$  and  $fs3$ . Say, the feature maps  $fs1$ ,  $fs2$  and  $fs3$  be transformed to  $f1$ ,  $f2$  and  $f3$  respectively.  

5 Build classifiers  $C1$ ,  $C2$  and  $C3$  using two FC layers of 4096 nodes each on the feature maps and a softmax layer of  $3 \times 1$  on the top of  $f1$ ,  $f2$  and  $f3$  respectively.  

6 Split the images  $I$  into training and validation sets in the ratio of 9:1  

7 Train the classifiers  $C1$ ,  $C2$  and  $C3$  using the training set data.  

8 Use the validation set to generate prediction scores  $w_1$ ,  $w_2$  and  $w_3$  using  $C1$ ,  $C2$  and  $C3$  classifiers respectively.  

9 Determine the fuzzy measure values using experimental fine tuning on the prediction scores  $w_1$ ,  $w_2$  and  $w_3$ , such that the final validation accuracy is the highest.  

/* Model testing phase */  

10 Resize testing images  $S$  to  $224 \times 224 \times 3$ .  

11 Use the trained classifiers  $C1$ ,  $C2$  and  $C3$  to generate prediction scores  $P$  of size  $3 \times 1$  based on the testing samples.  

12 Apply Sugeno fuzzy integral to the prediction scores  $P$  using fuzzy measures  $F$  (see Eq. (1)) to calculate the final prediction scores  $y$ .  

13 Use prediction scores  $y$  to determine whether a given sample is a Normal or COVID-19 or a Pneumonia case.
  
```

5.1. Datasets description

The use of an appropriate dataset is very much necessary to prove the usefulness of any method. To this end, three publicly available datasets are used to evaluate performance of the proposed method. The

Table 1
Distribution of training and test sets for all three datasets used here.

Dataset	#Train samples			#Test samples		
	COVID-19	Normal	Pneumonia	COVID	Normal	Pneumonia
Kaggle pneumonia dataset+IEEE covid dataset	739	1072	3100	185	269	775
COVID-19 Radiography database	2531	7134	942	1085	3058	260
COVIDx CXR-3 database	16490	13992	–	200	200	–

first one is a combination of two publicly available datasets: chest-xray-pneumonia¹ and covid-chestxray-dataset². For this, in the case of the first dataset, the images were taken from the retrospective cohorts of pediatric patients who were one to five years old. They were from the Guangzhou Women and Children's Medical Center, Guangzhou, China. The second dataset is a repository of chest X-ray and CT scan images of patients suffering from COVID-19. For this research work, only chest X-ray images are considered. The second dataset is the COVID-19 Radiography database³ which was collated by a team of researchers from Qatar University in Doha, Qatar and University of Dhaka in Dhaka, Bangladesh along with their collaborators from Pakistan and Malaysia, and medical doctors. The third dataset, namely the COVIDx CXR-3 dataset⁴, is a two-class relatively large dataset consisting of over 30000 CXR images with a fairly even split between positive and negative samples. The dataset distributions for all the three datasets are provided in Table 1. During training of the individual base classifiers, 10% of the training samples are used as validation samples and the rest are used for classifier training.

5.2. Hyperparameters used

In the present set of the experiments, a similar set of hyperparameters is used to train all the three CNN based base classifiers used here. The input images are resized to 224×224 and during model training, the number of epochs is set to 50 such that the models do not overfit, and learning rate is set to 2e-4. Both of these parameters are kept small to avoid overfitting of the model. Root mean square propagation (RM-SProp) as optimizer and the rectified linear unit (ReLU) as activation function are used. Here, the CNN models are fine-tuned on the target set to generate feature maps which are then passed to the ConvLSTM layers with kernel size of 1×1 where the SE block has been used to provide attention to the feature maps. The value of ratio, or reduction ratio, is set to 16. After that there are two FC layers with 4096 nodes each and a softmax layer of size 3×1 is used to classify chest X-ray images with the help of the extracted features. Categorical-cross entropy as the loss function is used during classification.

5.3. Performance of different components of proposed model

In this subsection, the performance of different components of the proposed model is discussed. For this, a number of experiments on the said datasets are performed. However, for the sake of simplicity, the first dataset is exclusively used for demonstration purposes. In the first set of experiments, all the said pre-trained CNN models on the dataset are employed, and then ConvLSTM and the SE block are added. Thus, in total 9 classifiers are employed on the first dataset. The obtained test accuracies are shown in Fig. 5. From the results, it has been observed that the test accuracies of VGG19, Inception V3 and MobileNet are improved by 1.87%, 1.71% and 0.82% respectively. Thus, it can be safely said that the use of ConvLSTM layer and the SE block helps

in improving the base classifiers' performance. It is noteworthy to mention here that the final variants (i.e., after adding ConvLSTM and the SE block) are considered here as base classifiers for designing the Sugeno fuzzy integral based ensemble method. Now onward, VGG19 + ConvLSTM + SE block, Inception V3 + ConvLSTM + SE block, and MobileNet + ConvLSTM + SE block are known as base classifier 1, base classifier 2, and base classifier 3 respectively.

The performance of the base classifiers and the Sugeno fuzzy integral based ensemble method are also recorded in terms of precision, recall, F1-score, and test accuracy. Figs. 6(a), 6(b), and 6(c) show the class-wise as well as overall (average) precision, recall, and F1-score respectively for all three base classifiers and proposed ensemble method while Fig. 6(d) records the test accuracy for each of the said models. Among the base classifiers, it has been observed that the improved VGG19 model (i.e., base classifier 1) has the highest test accuracy, recall, and F1-score of 97.88%, 98.34% and 97.67% respectively, while a precision score of 98.00% is observed for improved MobileNet model (i.e., base classifier 3). However, Sugeno fuzzy integral aided ensemble method outperforms all the base classifiers (see Fig. 6(d)). Proposed Sugeno fuzzy integral based ensemble method outperforms the base classifier 1 in terms of accuracy by 0.74%, in terms of average recall by 0.34%, and in case of average F1-score by 1.00%. It also provides a precision score which is 0.34% more than that of base classifier 3. The precision-recall curve of all the three base classifiers are shown in Fig. 7, while their training nature in terms of training accuracy and train loss are shown in Fig. 8. From these curves, it has been seen that the area in all cases for all classes is very close to the ideal value of 1, thus it can be said that overall, the proposed model is well-trained to handle the present 3-class problem.

5.4. Grad-CAM visualization analysis

Gradient weighted class activation map (Grad-CAM) analysis helps us to understand the internal working of a CNN model. It does this by using a heat map that shows the important regions of the input image according to the model. Through this, one can verify that the convolutional model is highlighting those regions of interest which are required for the classification task. Fig. 9 depicts the original chest X-ray images along with their Grad-CAM analysis by our improved MobileNet model. It can be seen from the original X-ray images that they contain unwanted texts and other symbols that may affect the performance of the classifier but from the Grad-CAM visualization, it is to be noted that the model does not focus on the unwanted information available and only looks at those regions which actually helps in the classification task. Also, in Fig. 9, it can be seen that the COVID-19 chest X-ray image is not clear but still the classifier manages to classify it by focusing on the vital regions in lungs.

5.5. Comparison with other ensemble techniques

Here, performances of the Sugeno fuzzy integral based classifier ensemble technique are compared with three standard ensemble techniques which are average, weighted average, and majority voting. For the purpose of this comparison, like the previous case, the first dataset is considered. The experimental results are shown in Fig. 10. From this figure, it can be observed that the present classifier ensemble technique

¹ <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia/version/2>

² <https://github.com/ieee8023/covid-chestxray-dataset>

³ <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>

⁴ <https://www.kaggle.com/andyczao/covidx-cxr2>

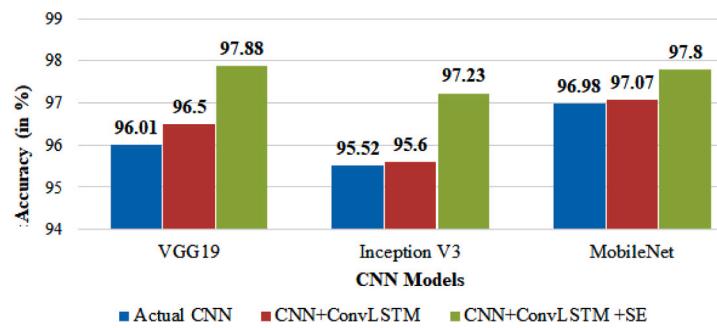


Fig. 5. Test accuracies of all three pre-trained CNN models (i.e., VGG19, Inception V3, and MobileNet and their modifications obtained by adding ConvLSTM and the SE block.

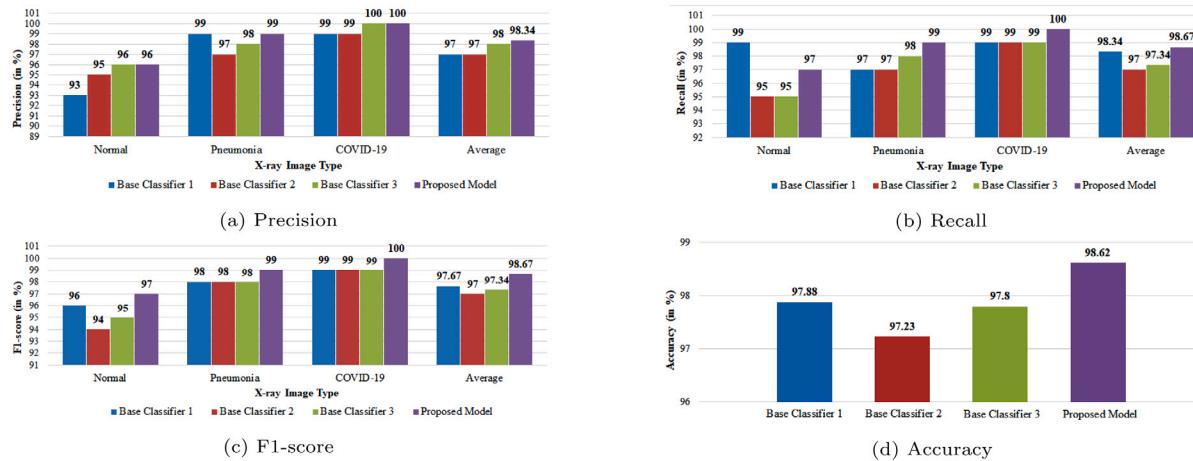


Fig. 6. Performances of three base classifiers (i.e., Base Classifier 1: VGG19 + ConvLSTM layer + SE block, Base Classifier 2: Inception V3 + ConvLSTM layer + SE block, Base Classifier 3: MobileNet + ConvLSTM layer + SE block in the charts) and Sugeno fuzzy integral based ensemble method (Proposed Model in the charts) on the test set of first dataset in Table 1. In (a), (b) and (c), precision, recall and F1-score values for each class (i.e., COVID-19, Pneumonia and Normal) and on average case respectively are shown.

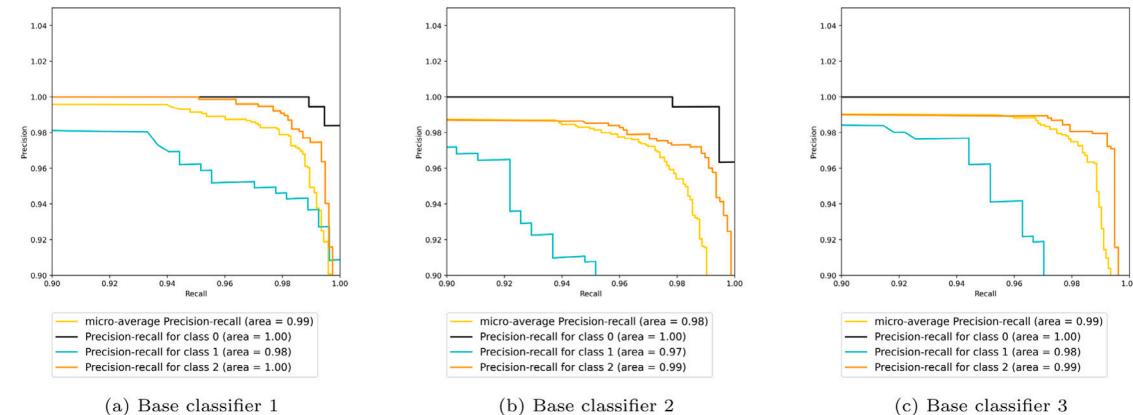


Fig. 7. Precision–Recall curve of the base classifier models: (a) VGG19 + ConvLSTM layer + SE block, (b) InceptionV3 + ConvLSTM layer + SE block model, and (c) MobileNet + ConvLSTM layer + SE block. Here, class 0 represents COVID-19, class 1 represents Normal, and class 2 represents Pneumonia.

outperforms others in terms of all the performance metrics. The Majority voting ensemble technique is the only one among other standard ensemble techniques that has performed close to the Sugeno fuzzy integral ensemble method. Sugeno integral method has outperformed majority voting ensemble technique by 0.25%, 0.33% and 0.33%, in terms of test accuracy, recall and F1-score respectively while providing the same precision score. The results indicate that Sugeno fuzzy integral based classifier ensemble method outperforms the other standard ensemble methods. The superior performance of current ensemble technique over the standard ensemble techniques is because Sugeno integral based method decides the final class level of a test sample using

different combinations of individual classifier's importance i.e., through fuzzy membership value μ of the classifiers whereas, standard ones decide the final class using individual classifier's prediction scores only.

Next, these two models (present and majority voting ensemble technique) are compared using their respective confusion matrices. Figs. 11(a) and 11(b) show confusion matrices of both the methods. In the case of prediction of true COVID-19 cases, the majority voting method predicts 185 cases correctly, which is the same as that of the Sugeno fuzzy integral based method. However, for Normal cases, the majority voting method correctly classifies 259 Normal cases, whereas the proposed ensemble method predicts 262 cases correctly. In the case

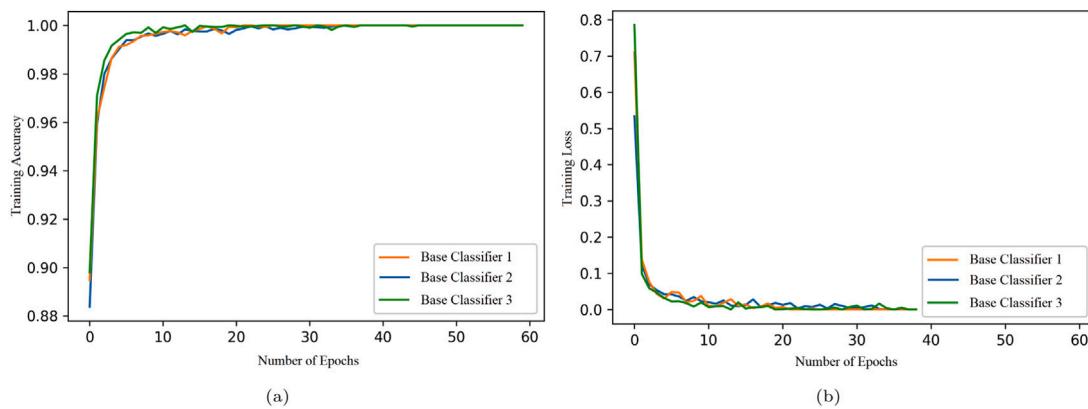


Fig. 8. Training nature of the base classifiers with respect to the number of epochs where, (a) represents the training accuracies and (b) represents the training losses.

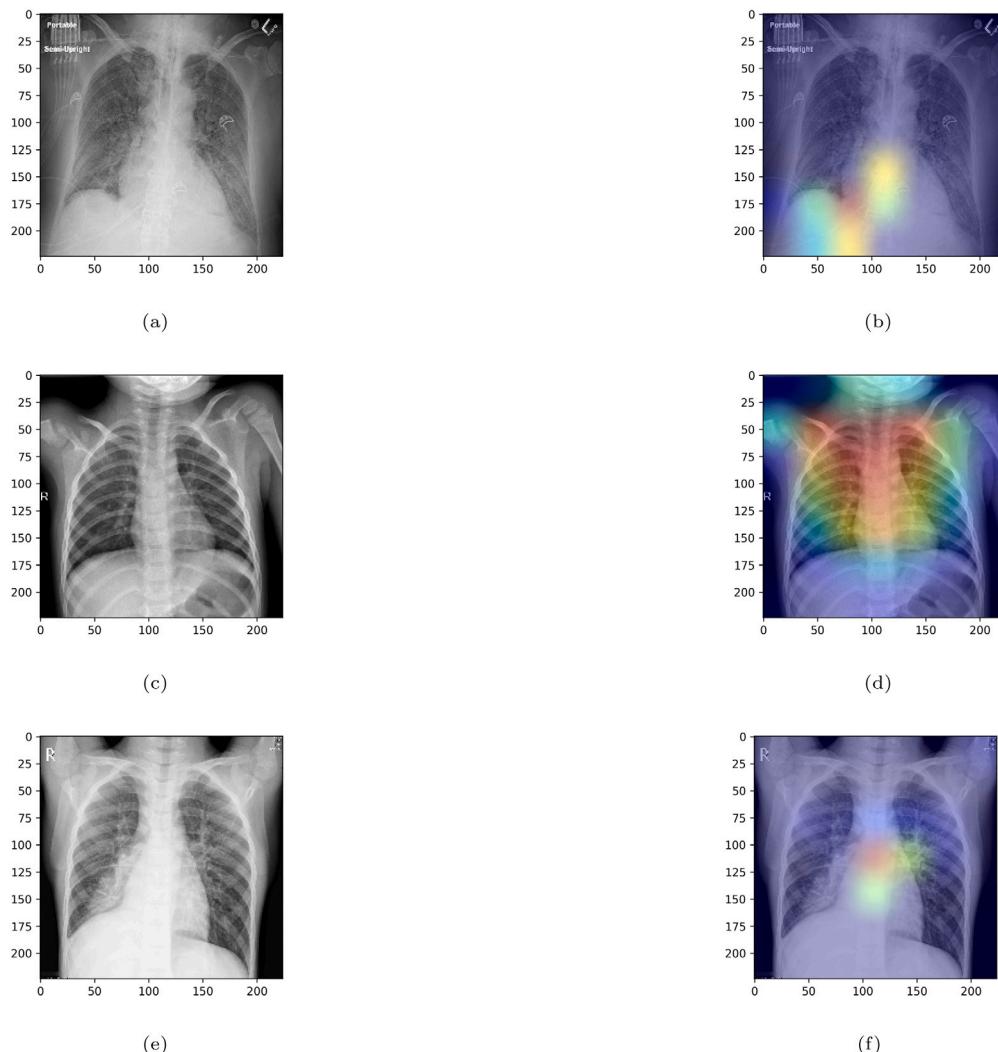


Fig. 9. Original chest X-ray images from the Kaggle pneumonia dataset + IEEE covid dataset (i.e., the first dataset in Table 1) with their Grad-CAM visualization generated using the model: MobileNet + ConvLSTM layer + SE block where, (a) original image of a COVID-19 infected patient, (b) Grad-CAM visualization image of the COVID-19 patient, (c) original image of a normal subject, (d) represents the Gram-CAM image of normal subject, (e) original image of a Pneumonia patient and (f) represents the Grad-CAM image of a Pneumonia patient.

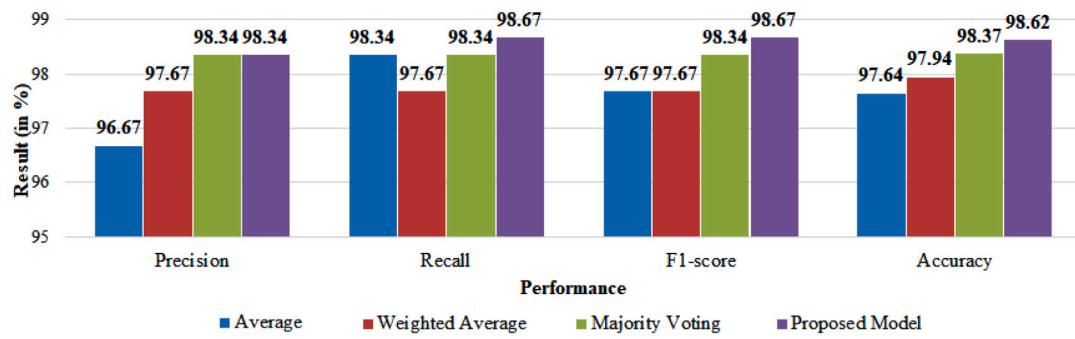


Fig. 10. Performance comparison of the present ensemble method with some standard ensemble techniques.

Two confusion matrices are shown side-by-side. Matrix (a) is for the majority voting ensemble and matrix (b) is for the proposed model. Both matrices have three columns: Predicted Covid Class, Predicted Normal Class, and Predicted Pneumonia Class. The rows represent True Covid Class, True Normal Class, and True Pneumonia Class. The numbers in the cells indicate the count of correctly predicted cases.

	Predicted Covid Class	Predicted Normal Class	Predicted Pneumonia Class
True Covid Class	185	0	0
True Normal Class	0	259	10
True Pneumonia Class	1	9	765

	Predicted Covid Class	Predicted Normal Class	Predicted Pneumonia Class
True Covid Class	185	0	0
True Normal Class	0	262	7
True Pneumonia Class	0	10	765

Fig. 11. Confusion matrices generated by (a) majority voting ensemble and (b) our proposed model with Sugeno fuzzy integral method.

of predicting true Pneumonia cases, both the methods correctly predict 765 cases.

5.6. Comparison with state-of-the-art methods

In this section, the performance of the present Sugeno fuzzy integral aided ensemble method is compared with several state-of-the-art methods proposed by Makris, Kontopoulos, and Tserpes (2020), Horry, et al. (2020), Hemdan et al. (2020), Ardakani et al. (2020), Aslan et al. (2021), Bashar et al. (2021), Chowdhury, et al. (2020), Das, Roy, et al. (2021), Goel et al. (2021), Islam et al. (2020), Ismael and Şengür (2021), Jain et al. (2021), Kedia et al. (2021), Khan et al. (2020), Mukherjee et al. (2021a), Naeem and Bin-Salem (2021), Panetta et al. (2021), Paul et al. (2022), Roy et al. (2021), Sedik, Hammad, Abd El-Samie, Gupta, and Abd El-Latif (2021), Senan et al. (2021), and Yang, et al. (2021) on all three datasets used here whenever applicable i.e., the results obtained by the methods on the respective datasets are cited or in few cases, their performances are evaluated using the proposed setup. The comparative results are recorded in Tables 2–4 for dataset 1, 2 and 3 (mentioned in Table 1) respectively. From Table 2, it is evident that the highest accuracy is achieved by the ResNet50 model (Senan et al., 2021), which is 0.08% more than the proposed method. However, in terms of precision, the proposed method outperforms ResNet50 (Senan et al., 2021) by 0.34%. The highest recall is achieved by the End-to-End CNN model (Mukherjee et al., 2021a) with 99.81% recall. The reason for the appreciable performance of the model in comparison to the state-of-the-art methods is the utilization of ConvLSTM layer and the SE block in each of the classifiers and also the application of Sugeno based ensemble method. Also, it can be noted that the efficiency of our Sugeno fuzzy integral based ensemble model from Fig. 12 which shows the confusion matrices of some of the state-of-the-art models. In case of the COVID-19 Radiography Dataset, performances of the various state-of-the-art models along with the

proposed method are depicted in Table 3. From the table, it can be observed the highest accuracy is achieved by the mAlexNet+ BiLSTM model Aslan et al. (2021), which is 0.06% more than the accuracy achieved by the proposed method. Also, in terms of average precision, recall, and F1-score, the mAlexNet+ BiLSTM model outperforms the proposed method by 0.25%, 0.63%, and 0.44% respectively. For the rest of the state-of-the-art models, the proposed CovidConvLSTM model can outperform all of them in terms of classification accuracy.

The performances of state-of-the-art models along with the present model on the COVIDx CXR-3 Dataset are shown in Table 4. From the table, it can be observed that the highest accuracy, average precision, recall, and F1-score is achieved by the proposed CovidConvLSTM model with values 86.75%, 88.71%, 86.75%, and 86.58% respectively. The second highest accuracy of 85.50% is achieved by the Optimized CNN model Bashar et al. (2021), which is 1.25% less than that of the proposed method. Also, in terms of accuracy, the CovidConvLSTM model outperforms the Optimized CNN model in terms of average precision, recall, and F1-score by 1.21%, 1.25%, and 1.58% respectively.

6. Conclusion

To fight the global outbreak of the COVID-19 disease, researchers all around the world are working to find ways to combat this pandemic. Hence, seeing the recent trend in the usage of deep learning models for early detection of COVID-19 cases, this paper comes up with a COVID-19 detection method which uses chest X-ray images as input data. The proposed method uses Sugeno fuzzy integral based classifier ensemble method to combine the outputs of three base classifiers – VGG19, InceptionV3, and MobileNet. Prior to that, the performance of base classifiers is enhanced by applying ConvLSTM layers and SE blocks. Due to the size of the available dataset being small and unavailability of the high-end computational resources, the power of transfer learning is leveraged. After performing various experiments, the results

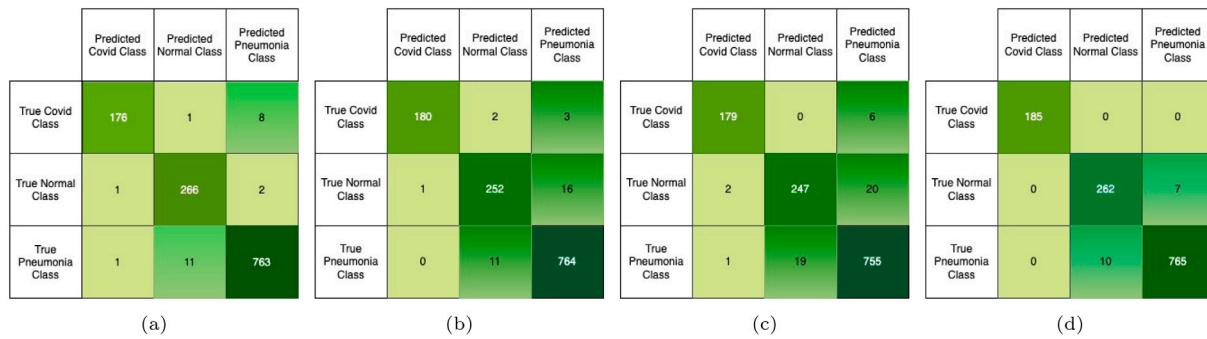


Fig. 12. Confusion matrices of some state-of-the-art CNN models: (a) VGG16 model, (b) DenseNet121 model, (c) CNN-LSTM and (d) Proposed method on dataset 1.

Table 2

Performance comparison of different state-of-the-art models along with ours on Kaggle pneumonia dataset + IEEE covid dataset (i.e., first dataset in Table 1).

Work ref.	Model used	Performance (in %) in terms of			
		Precision	Recall	F1-score	Accuracy
Hemdan et al. (2020) ^a	DenseNet121	97.37	96.52	96.94	97.31
Makris et al. (2020) ^a	VGG16	98.00	97.33	97.67	98.04
Horry, et al. (2020) ^a	ResNet50	85.67	90.67	86.67	88.77
Ardakani et al. (2020) ^a	ResNet101	94.33	91.00	92.67	94.38
Khan et al. (2020)	CoroNet	95.00	96.90	95.94	95.00
Jain et al. (2021) ^a	Xception	98.00	94.60	96.00	97.00
Ismael and Şengür (2021) ^a	End to End CNN	95.67	94.67	95.00	96.09
Das, Roy, et al. (2021) ^a	Bi Level Prediction	97.87	98.14	98.00	98.45
Goel et al. (2021)	OptCoNet	92.88	96.25	95.25	97.78
Paul et al. (2022) ^a	Inverted Bell Ensemble	97.21	97.81	97.50	97.97
Mukherjee et al. (2021a) ^a	End to End CNN	94.78	99.81	97.22	97.52
Bashar et al. (2021) ^a	Optimized CNN model	95.67	93.34	94.67	95.20
Senan et al. (2021) ^a	ResNet50 with GLCM and LBP features	98.00	98.67	98.67	98.70
Naeem and Bin-Salem (2021)	CNN-LSTM model	95.00	95.00	95.00	96.60
Islam et al. (2020) ^a	CNN-LSTM	95.96	95.33	95.64	96.09
Goyal and Singh (2021)	F-RNN-LSTM model	88.89	95.41	92.03	94.31
Proposed method	Base classifier 1	97.00	98.34	97.67	97.88
	Base classifier 2	97.00	97.00	97.00	97.23
	Base classifier 3	98.00	98.67	97.34	97.80
	CovidConvLSTM	98.34	98.67	98.67	98.62

^aIndicates that this method is evaluated on the present set-up.

Table 3

Performance comparison of the proposed method with some state-of-the-art models on the COVID-19 Radiography Dataset (i.e., second dataset in Table 1).

Work ref.	Model used	Performance (in %) in terms of			
		Precision	Recall	F1-score	Accuracy
Aslan et al. (2021)	mAlexNet+BiLSTM	98.77	98.76	98.76	98.70
Aslan et al. (2021)	mAlexNet	98.16	98.26	98.20	98.14
Kedia et al. (2021)	CoVNet-19	98.34	98.34	98.34	98.20
Sedik et al. (2021)	ConvLSTM	94.67	97.09	95.64	95.96
Panetta et al. (2021)	Classical Fibonacci p-pattern	97.78	96.90	97.32	97.79
Panetta et al. (2021)	Fibonacci p-pattern	97.20	96.76	96.69	98.03
Yang, et al. (2021)	Fast.AI ResNet	97.00	97.00	97.00	97.00
Paul et al. (2022) ^a	Inverted Bell Ensemble	97.24	97.25	97.24	97.64
Mukherjee et al. (2021a) ^a	End to End CNN	94.15	98.82	96.45	96.92
Roy et al. (2021)	CoWarriorNet	94.66	91.33	92.66	97.80
Bashar et al. (2021) ^a	Optimized CNN model	97.00	93.67	95.67	96.55
Islam et al. (2020) ^a	CNN-LSTM	94.25	90.89	92.40	94.35
Senan et al. (2021) ^a	ResNet50 with GLCM and LBP features	97.00	97.67	97.67	98.01
Goyal and Singh (2021)	F-RNN-LSTM model	93.65	96.78	95.19	95.04
Proposed method	Base classifier 1	96.82	96.76	96.79	97.62
	Base classifier 2	94.39	95.09	94.74	95.31
	Base classifier 3	97.58	97.15	97.35	97.95
	CovidConvLSTM	98.52	98.13	98.32	98.64

^aIndicates that this method is evaluated on the present set-up.

Table 4

Performances of CovidConvLSTM model and its sub-modules along with some state-of-the-art models on the COVIDx CXR-3 Dataset (i.e., third dataset in Table 1).

Work ref.	Technique	Performance (in %) in terms of			
		Precision	Recall	F1-score	Accuracy
Jain et al. (2021) ^a	Xception	79.50	74.00	73.00	74.25
Ardakani et al. (2020) ^a	ResNet101	80.50	76.50	76.00	76.75
Chowdhury, et al. (2020) ^a	CheXNet	82.50	75.00	73.50	74.75
Bashar et al. (2021) ^a	Optimized CNN	87.50	85.50	85.00	85.50
Senan et al. (2021) ^a	ResNet50	80.50	77.00	76.50	77.25
Islam et al. (2020) ^a	CNN-LSTM	70.47	70.25	70.16	70.25
Mukherjee et al. (2021a) ^a	End-to-End CNN	71.00	64.00	73.50	64.00
Proposed method	Base classifier 1	86.81	83.50	83.12	83.50
	Base classifier 2	85.64	80.50	79.76	80.50
	Base classifier 3	84.52	80.05	79.00	79.00
	CovidConvLSTM	88.71	86.75	86.58	86.75

^aIndicates that this method is evaluated on the present set-up.

demonstrate that the Sugeno fuzzy integral based classifier ensemble technique not only improves the performance of the individual classification models, but also outperforms standard classifier ensemble methods like majority voting and weighted average. Moreover, in most cases, the present technique outperforms state-of-the-art methods which are considered here for comparison while being evaluated on three datasets.

Although CovidConvLSTM performs well in comparison with other state-of-the-art COVID-19 detection methods, there is still room for improvement. As it can be seen, the proposed method has misclassified some of the test data samples which is alarming as the model deals with COVID-19 detection which can be a matter of life and death for the patient. The reason for this misclassification can be attributed to the individual classifiers' performance. This can be improved by using a set of stronger classifiers than the current ones and also by increasing the number of classifiers to be combined using the Sugeno fuzzy integral method. Apart from these, the use of some feature selection techniques on the extracted features can be another improvement to the present model. Further, in future, there is a plan to use this method for other image classification tasks as well.

CRediT authorship contribution statement

Subhrajit Dey: Methodology, Writing – original draft, Software, Conceptualization, Validation. **Rajdeep Bhattacharya:** Methodology, Writing – original draft, Conceptualization, Validation. **Samir Malakar:** Methodology, Writing – original draft, Software. **Friedhelm Schwenker:** Supervision, Formal analysis, Writing – review & editing. **Ram Sarkar:** Conceptualization, Writing – review & editing, Project administration, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank the Centre for Microprocessor Applications for Training, Education and Research (CMATER) research laboratory of the Computer Science and Engineering Department, Jadavpur University, Kolkata, India for providing us the infrastructural support.

References

- Abbaszadeh, S., & Hüllermeier, E. (2020). Machine learning with the sugeno integral: The case of binary classification. *IEEE Transactions on Fuzzy Systems*, 29(12), 3723–3733.
- Ardakani, A. A., Kanafi, A. R., Acharya, U. R., Khadem, N., & Mohammadi, A. (2020). Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. *Computers in Biology and Medicine*, 121, Article 103795.
- Aslan, M. F., Unlersen, M. F., Sabancı, K., & Durdu, A. (2021). CNN-Based transfer learning-BiLSTM network: A novel approach for COVID-19 infection detection. *Applied Soft Computing*, 98, Article 106912. <http://dx.doi.org/10.1016/j.asoc.2020.106912>.
- Bashar, A., Latif, G., Ben Brahim, G., Mohammad, N., & Alghazo, J. (2021). COVID-19 Pneumonia detection using optimized deep learning techniques. *Diagnostics*, 11(11), 1972.
- Brunese, L., Mercaldo, F., Reginelli, A., & Santone, A. (2020). Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. *Computer Methods and Programs in Biomedicine*, 196, Article 105608.
- Burgin, M. (2000). Theory of fuzzy limits. *Fuzzy Sets and Systems*, 115(3), 433–443.
- Canayaz, M. (2020). MH-COVIDNet: Diagnosis of COVID-19 using deep neural networks and meta-heuristic-based feature selection on X-ray images. *Biomedical Signal Processing and Control*, 64, Article 102257.
- Chandra, T. B., Verma, K., Singh, B. K., Jain, D., & Netam, S. S. (2020). Coronavirus disease (COVID-19) detection in chest X-Ray images using majority voting based classifier ensemble. *Expert Systems with Applications*, 165, Article 113909.
- Chen, Y., Zhang, Z., Zhong, L., Chen, T., Chen, J., & Yu, Y. (2019). Three-stream convolutional neural network with squeeze-and-excitation block for near-infrared facial expression recognition. *Electronics*, 8(4), 385.
- Chowdhury, M. E., Rahman, T., Khandaker, A., Mazhar, R., Kadir, M. A., Mahbub, Z. B., et al. (2020). Can AI help in screening viral and COVID-19 pneumonia? *IEEE Access*, 8, 132665–132676.
- Das, A. K., Ghosh, S., Thunder, S., Dutta, R., Agarwal, S., & Chakrabarti, A. (2021). Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network. *Pattern Analysis and Applications*, 1–14.
- Das, S., Roy, S. D., Malakar, S., Velásquez, J. D., & Sarkar, R. (2021). Bi-level prediction model for screening COVID-19 patients using chest X-Ray images. *Big Data Research*, 25, Article 100233. <http://dx.doi.org/10.1016/j.bdr.2021.100233>.
- Das, D., Santosh, K., & Pal, U. (2020). Truncated inception net: COVID-19 outbreak screening using chest X-rays. *Physical and Engineering Sciences in Medicine*, 43(3), 915–925.
- Dey, S., Bhattacharya, R., Malakar, S., Mirjalili, S., & Sarkar, R. (2021). Choquet fuzzy integral-based classifier ensemble technique for COVID-19 detection. *Computers in Biology and Medicine*, Article 104585.
- Dey, A., Chattopadhyay, S., Singh, P. K., Ahmadian, A., Ferrara, M., Senu, N., et al. (2021). MRFGRO: A hybrid meta-heuristic feature selection method for screening COVID-19 using deep features. *Scientific Reports*, 11(1), 1–15.
- Du, X., Zare, A., Keller, J. M., & Anderson, D. T. (2016). Multiple instance choquet integral for classifier fusion. In *2016 IEEE congress on evolutionary computation* (pp. 1054–1061). IEEE.
- Goel, T., Murugan, R., Mirjalili, S., & Chakrabarty, D. K. (2021). OptCoNet: An optimized convolutional neural network for an automatic diagnosis of COVID-19. *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies*, 51(3), 1351–1366.
- Goyal, S., & Singh, R. (2021). Detection and classification of lung diseases for pneumonia and Covid-19 using machine and deep learning techniques. *Journal of Ambient Intelligence and Humanized Computing*, 1–21.
- Grabisch, M., Sugeno, M., & Murofushi, T. (2010). *Fuzzy measures and integrals - theory and applications*. Heidelberg: Physica, 2000.

- Gu, J., Sun, X., Zhang, Y., Fu, K., & Wang, L. (2019). Deep residual squeeze and excitation network for remote sensing image super-resolution. *Remote Sensing*, 11(15), 1817.
- Hemdan, E. E.-D., Shouman, M. A., & Karar, M. E. (2020). COVIDX-Net: A framework of deep learning classifiers to diagnose COVID-19 in X-Ray images. arXiv preprint arXiv:2003.11055.
- Horry, M. J., Chakraborty, S., Paul, M., Ulhaq, A., Pradhan, B., Saha, M., et al. (2020). COVID-19 Detection through transfer learning using multimodal imaging data. *IEEE Access*, 8, 149808–149824.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision. arXiv preprint arXiv:1704.04861.
- Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7132–7141).
- Islam, M. Z., Islam, M. M., & Asraf, A. (2020). A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. *Informatics in Medicine Unlocked*, 20, Article 100412.
- Ismael, A. M., & Sengür, A. (2021). Deep learning approaches for COVID-19 detection based on chest X-ray images. *Expert Systems with Applications*, 164, Article 114054.
- Jain, R., Gupta, M., Taneja, S., & Hemanth, D. J. (2021). Deep learning based detection and analysis of COVID-19 on chest X-ray images. *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies*, 51(3), 1690–1700.
- Kedia, P., Katarya, R., et al. (2021). CoVNET-19: A deep learning model for the detection and analysis of COVID-19 patients. *Applied Soft Computing*, 104, Article 107184.
- Khan, A. I., Shah, J. L., & Bhat, M. M. (2020). CoroNet: A Deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. *Computer Methods and Programs in Biomedicine*, 196, Article 105581.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with deep convolutional neural networks. In *Advances in neural information processing systems*. Vol. 25 (pp. 1097–1105).
- Lakshmanaprabu, S., Mohanty, S. N., Shankar, K., Arunkumar, N., & Ramirez, G. (2019). Optimal deep learning model for classification of lung cancer on CT images. *Future Generation Computer Systems*, 92, 374–382.
- LeCun, Y., Haffner, P., Bottou, L., & Bengio, Y. (1999). Object recognition with gradient-based learning. In *Shape, contour and grouping in computer vision* (pp. 319–345). Springer.
- Li, Y., Liu, Y., Cui, W.-G., Guo, Y.-Z., Huang, H., & Hu, Z.-Y. (2020). Epileptic seizure detection in EEG signals using a unified temporal-spectral squeeze-and-excitation network. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(4), 782–794.
- Majid, M., & Safabakhsh, R. (2020). Correlational convolutional LSTM for human action recognition. *Neurocomputing*, 396, 224–229.
- Makris, A., Kontopoulos, I., & Tserpes, K. (2020). COVID-19 Detection from chest X-Ray images using deep learning and convolutional neural networks. In *11th Hellenic conference on artificial intelligence* (pp. 60–66).
- Mitra, S., & Pal, S. K. (2005). Fuzzy sets in pattern recognition and machine intelligence. *Fuzzy Sets and Systems*, 156(3), 381–386.
- Mukherjee, H., Ghosh, S., Dhar, A., Obaidullah, S. M., Santosh, K., & Roy, K. (2021a). Deep neural network to detect COVID-19: one architecture for both CT scans and chest X-rays. *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies*, 51(5), 2777–2789.
- Mukherjee, H., Ghosh, S., Dhar, A., Obaidullah, S. M., Santosh, K., & Roy, K. (2021b). Shallow convolutional neural network for COVID-19 outbreak screening using chest X-rays. *Cognitive Computation*, 1–14.
- Mukherjee, D., Mondal, R., Singh, P. K., Sarkar, R., & Bhattacharjee, D. (2020). EnsemConvNet: a deep learning approach for human activity recognition using smartphone sensors for healthcare applications. *Multimedia Tools and Applications*, 79(41), 31663–31690.
- Naeem, H., & Bin-Salem, A. A. (2021). A CNN-LSTM network with multi-level feature extraction-based approach for automated detection of coronavirus from CT scan and X-ray images. *Applied Soft Computing*, 113, Article 107918.
- Nour, M., Cömert, Z., & Polat, K. (2020). A novel medical diagnosis model for COVID-19 infection detection based on deep features and Bayesian optimization. *Applied Soft Computing*, 97, Article 106580.
- Panetta, K., Sanghavi, F., Agaian, S., & Madan, N. (2021). Automated detection of COVID-19 cases on radiographs using shape-dependent fibonacci-p patterns. *IEEE Journal of Biomedical and Health Informatics*, 25(6), 1852–1863. <http://dx.doi.org/10.1109/JBHI.2021.3069798>.
- Paul, A., Basu, A., Mahmud, M., Kaiser, M. S., & Sarkar, R. (2022). Inverted bell-curve-based ensemble of deep learning models for detection of COVID-19 from chest X-rays. *Neural Computing and Applications*, 1–15. <http://dx.doi.org/10.1007/s00521-021-06737-6>.
- Rahman, S. A., & Adjeroh, D. A. (2019). Deep learning using convolutional LSTM estimates biological age from physical activity. *Scientific Reports*, 9(1), 1–15.
- Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S. B. A., et al. (2021). Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. *Computers in Biology and Medicine*, 132, Article 104319.
- Ramdani, H., Allali, N., Chat, L., & El Haddad, S. (2021). Covid-19 Imaging: A narrative review. *Annals of Medicine and Surgery*, Article 102489.
- Roy, I., Shai, R., Ghosh, A., Bej, A., & Pati, S. K. (2021). CoWarriorNet: A Novel deep-learning framework for CoVid-19 detection from chest X-Ray images. *New Generation Computing*, 1–25.
- Sahlol, A. T., Yousri, D., Ewees, A. A., Al-Qaness, M. A., Damasevicius, R., & Abd Elaziz, M. (2020). COVID-19 Image classification using deep features and fractional-order marine predators algorithm. *Scientific Reports*, 10(1), 1–15.
- Santosh, K. (2020). AI-Driven tools for coronavirus outbreak: Need of active learning and cross-population train/test models on multitudinal/multimodal data. *Journal of Medical Systems*, 44(5), 1–5.
- Santosh, K., & Ghosh, S. (2021). Covid-19 Imaging tools: How big data is big? *Journal of Medical Systems*, 45(7), 1–8.
- Sedik, A., Hammad, M., Abd El-Samie, F. E., Gupta, B. B., & Abd El-Latif, A. A. (2021). Efficient deep learning approach for augmented detection of coronavirus disease. *Neural Computing and Applications*, 1–18.
- Senan, E. M., Alzahrani, A., Alzahrani, M. Y., Alsharif, N., & Aldhyani, T. H. (2021). Automated diagnosis of chest X-Ray for early detection of COVID-19 disease. *Computational and Mathematical Methods in Medicine*, 2021.
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., & Woo, W.-c. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in Neural Information Processing Systems*, 28.
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In *Proceedings of the 6th international conference on learning representations*.
- Sugeno, M. (1993). Fuzzy measures and fuzzy integrals—a survey. In *Readings in fuzzy sets for intelligent systems* (pp. 251–257). Elsevier.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1–9).
- Toğuçar, M., Ergen, B., & Cömert, Z. (2020). COVID-19 Detection using deep learning models to exploit social mimic optimization and structured chest X-ray images using fuzzy color and stacking approaches. *Computers in Biology and Medicine*, 121, Article 103805.
- Turkoglu, M. (2020). CoviDetectioNet: COVID-19 diagnosis system based on X-ray images using features selected from pre-learned deep features ensemble. *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies*, 1–14.
- Vaka, A. R., Soni, B., & Reddy, S. (2020). Breast cancer detection by leveraging machine learning. *ICT Express*, 6(4), 320–324.
- Wang, L., Lin, Z. Q., & Wong, A. (2020). COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Scientific Reports*, 10(1), 1–12.
- Yang, D., Martinez, C., Visuña, L., Khandhar, H., Bhatt, C., & Carretero, J. (2021). Detection and analysis of COVID-19 in medical images using deep learning techniques. *Scientific Reports*, 11(1), 1–13.
- Zhou, T., Lu, H., Yang, Z., Qiu, S., Huo, B., & Dong, Y. (2020). The ensemble deep learning model for novel COVID-19 on CT images. *Applied Soft Computing*, 98, Article 106885.
- Zhu, G., Zhang, L., Yang, L., Mei, L., Shah, S. A. A., Bennamoun, M., et al. (2019). Redundancy and attention in convolutional LSTM for gesture recognition. *IEEE Transactions on Neural Networks and Learning Systems*, 31(4), 1323–1335.