Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
A novel fusion based convolutional neural network approach for classification of COVID-19 from chest X-ray images

Anubhav Sharma a,*, Karamjeet Singh a,*, Deepika Koundal b

a Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, Punjab, India
b Department of Virtualization, School of Computer Science, University of Petroleum & Energy Studies, Dehradun, Uttrakhand, India

ARTICLE INFO

Keywords: COVID-19, Deep learning, Transfer learning, Confidence fusion, Chest X-ray

ABSTRACT

Coronavirus disease is a viral infection caused by a novel coronavirus (CoV) which was first identified in the city of Wuhan, China somewhere in the early December 2019. It affects the human respiratory system by causing respiratory infections with symptoms (mild to severe) like fever, cough, and weakness but can further lead to other serious diseases and has resulted in millions of deaths until now. Therefore, an accurate diagnosis for such types of diseases is highly needful for the current healthcare system. In this paper, a state of the art deep learning method is described. We propose COVDC-Net, a Deep Convolutional Network-based classification method which is capable of identifying SARS-CoV-2 infected amongst healthy and/or pneumonia patients from their chest X-ray images. The proposed method uses two modified pre-trained models (on ImageNet) namely MobileNetV2 and VGG16 without their classifier layers and fuses the two models using the Confidence fusion method to achieve better classification accuracy on the two currently publicly available datasets.

It is observed through exhaustive experiments that the proposed method achieved an overall classification accuracy of 96.48% for 3-class (COVID-19, Normal and Pneumonia) classification tasks. For 4-class classification (COVID-19, Normal, Pneumonia Viral, and Pneumonia Bacterial) COVDC-Net method delivered 90.22% accuracy. The experimental results demonstrate that the proposed COVDC-Net method has shown better overall classification accuracy as compared to the existing deep learning methods proposed for the same task in the current COVID-19 pandemic.

1. Introduction

In December 2019, a novel type of coronavirus (CoV) namely SARS-CoV-2 or what is commonly referred to as COVID-19, was identified amidst a widespread pneumonia outbreak in local hospitals of Wuhan, Hubei province in China. COVID-19 is considered to be of natural animal origin that started in the bat population and later broke the species barrier to infect humans, who became the carriers for the virus. The virus is responsible for respiratory infections leading to symptoms like fever (mild/severe), cough, dyspnea, bilateral infiltrates in chest imaging and other acute respiratory problems [1]. This virus has spread throughout the globe and has caused around 6 million deaths as of now since the start of the COVID-19 pandemic [2] in December 2019. This calls for new collaborations to accelerate the research and development of coronavirus vaccine and reliable, quick testing methods all across the world [3]. The World Health Organization (WHO) plays a critical role [4] in the on-going pandemic as many countries face shortages of testing kits, trained healthcare professionals, contact tracing protocols among other essentials and medical equipment. The new medical policies are designed for COVID-19 pandemic [5] which explore the medical facilities provided to CoV patients.

Many researchers have contributed to reduce the inefficiency in the current diagnosis of coronavirus. For further improving the diagnosis and better understanding of severe pediatric patients of COVID-19, authors [6,7] provided the epidemiological, laboratory findings, imaging data and clinical treatments to all healthcare professionals of Wuhan city, China. Jiang et al. [8] provided the necessary clinical study about the symptoms, complications and treatments required by the COVID infected patients. Later on, Tolksdorf et al. [9] introduced the concept of using syndromic surveillance data to assess COVID-19 patients analogous to influenza pneumonia patients for determining the condition of CoV affected patients. In [10], authors throw light towards the classification of pneumonia on the basis of chest radiology, rapid analysis of biological samples, severity scores and collection of other relevant information.

* Corresponding authors.
E-mail addresses: anubhav0399@gmail.com (A. Sharma), karamjeet.singh@thapar.edu (K. Singh), dkoundal@ddn.upes.ac.in (D. Koundal).

https://doi.org/10.1016/j.bspc.2022.103778
Received 20 January 2022; Received in revised form 9 April 2022; Accepted 27 April 2022
Available online 2 May 2022
1746-8094/© 2022 Published by Elsevier Ltd.
clinical data. Wang et al. [11] investigated the spread of CoV through the respiratory tract and later on, Bernheim et al. [12] concluded that the frequency of CT findings was related to the infection time course. From all these studies, it can be observed that the CT and X-ray imaging system are two effective tools for preliminary, quick identification and quantification of positive CoV cases. Therefore, a large number of researchers are working towards developing a quick, reliable and effective detection system for CoV disease from chest X-ray images by using modern computer science techniques like deep learning. This also has other advantages over traditional testing like easy scalability, safer sample collection, faster diagnosis.

Before COVID-19, in the recent times we have seen many applications of technology and computers in the medical image and signal processing [52-59] which have aided the professionals in better diagnosis certain ailments. Jaiswal et al. [13] proposed a deep learning-based approach to get a better clarity about the presence of pneumonia in chest X-ray images. Later, Chouhan et al. [14] suggested a transfer learning-based approach for the same. The recent progress of deep learning applications in medical image analysis has proven to be an effective tool for diagnosis of various diseases like breast cancer and many other pulmonary disorders [60-62]. Y-D. Zhang et al. [63] describe MIDCAN, A multiple input deep convolutional attention network that can handle CCT and CXR images simultaneously. They fused CCT and CXR images in order to improve their individual performance. S-H. Wang et al. [64] proposed a rank based average pooling module (NRAPM), which showed remarkable accuracy and precision values for COVID-19 diagnosis. COVID-Net [15] is a convolutional neural network (CNN)-based approach for detection of COVID-19 patients. The main goal is to provide an effective and reliable treatment to CoV patients which is achieved by patch-based convolutional neural network approach [16]. A similar approach [17] explores the different types of features and automatic detection of SARS-CoV-2 from chest X-ray images. Also, Ozturk et al. [18] have applied the existing DarkNet architecture for developing the DarkCovidNet model which achieves an accuracy of 98.08%, 87.02% for binary and multiclass classification tasks, respectively. Similarly, Ucar et al. [19] and Marques et al. [20] fine-tuned the Bayes SqueezeNet and EfficientNet models respectively, for providing automated diagnosis for SARS-CoV-2 patients. Rahimzadeh and Attar [21] developed a modified deep convolutional network by fusion Xception and ResNet50V2 to achieve multiclass accuracy of 91% on two open-source datasets. Hemdan et al. [22] introduced a deep learning technique to detect COVID-19 from X-ray images. The proposed framework included seven pre-trained models VGG19, MobileNetV2, InceptionV3, ResNet50, DenseNet201, Xception, and InceptionResNetV2. Among these seven tested classifiers, VGG19 and DenseNet201 achieved the highest accuracy of 90%. A plethora of works directed towards similar types of deep learning-based approaches are found in literature [23-28] and many other focus on transfer learning and combination of existing models [29-39] have shown promising results for binary and multiclass classification tasks. A direct diagnosis from the X-ray images by using the aforementioned techniques are helpful for various medical professionals to distinguish between flu-pneumonia and coronavirus infected patients. Alakus and Turkoğlu [40] extended deep learning applications in CoV detection by using laboratory findings rather than X-ray or CT-scan images of CoV patients. Later, Swapnarlekha et al. [41] provided a comparative performance analysis of various machine and deep learning-based models for predicting COVID-19 along with its application. In addition to aforesaid approaches, Karthik et al. [42] developed a channel-shuffled dual-branched convolutional neural network that uniquely learns X-ray patterns from the chest X-ray images. Authors in [50,51] employed U-net architecture and a meta-heuristic-based feature selection for the diagnosis of COVID-19 using X-ray images.

Our proposed methodology is a confidence fusion of VGG16 and MobileNetV2 which are pre-trained on ImageNet dataset in order to classify a sample chest X-ray image for 3-class and 4-class classification tasks. Various Models were tested, and the best performing models were selected for the process of fusion, hence creating a model that has the advantages from the unique feature selection process of both the architectures. This makes better predictions and achieves better performance than any single contributing model. Hybrid models also provide robustness and reduce the spread or dispersion of the predictions which might be an issue when using a single model.

The main contributions of the proposed method are three-fold and described briefly as follows.

- The proposed method tests various proven models like VGG16, VGG19, MobileNetV2, ResNet50, DenseNet to find the best performing fusion and provides a detailed analysis on the prepared dataset of X-ray images for a better classification of COVID-19 samples.
- The proposed novel hybrid CNN architecture integrates the two different types of features obtained from separate CNN architectures with help of confidence fusion method with the aim of providing better assistance and diagnosis to COVID-19 patients. As seen in Table 3, The proposed model can improve upon the performance of the parent models significantly for the evaluated performance metrics.
- The method was tested on multiple models and the best fusion model was selected for this study as described in Table 2.
- The proposed work would help in building the foundation for inventing the hybrid CNN architectures which helps the radiologists to correctly identify and distinguish between the different types of pneumonia and COVID infections from chest X-ray images.
- The proposed framework can be deployed as one of the useful techniques in the medical field for classification purposes. It also helps radiologists to diagnose and treat diseases at early stages.

The remainder of the paper is organized as follows: Section 2 describes concepts of Convolutional Neural Network, confidence fusion method, dataset description, details of building and training of the proposed methodology with parameters and necessary information about the developed deep learning application. Section 3 presents the experimental results and evaluation criteria including accuracy, precision, recall and F-score. Finally, the conclusion is presented in Section 4.

2. Materials and methodology

2.1. Convolutional neural networks - CNN

A Convolutional Neural Network is a deep learning algorithm that applies convolution, a mathematical operation as displayed in Fig. 1, between input data i.e., the image array and a set of learnable filters (kernels) to create feature maps that are able to extract local features like edges, sharp corners, shapes, gradients, etc. to solve various classification problems. As we move deeper into the network layers more complex feature maps are created.

Let the input image $I$ image matrix be of dimensions $(h \times w)$ and filters of size. Then the convolution operation on image $I$ and filter $F$ is given by Equation (1).

| Table 1 |
|---------|
| **Dataset analysis.** |
| **DATASET** | **Class** | **Number Of Samples** |
| D1 COVID-19 | 1784 |
| Healthy | 1755 |
| Pneumonia | 1345 |
| D2 COVID-19 | 305 |
| Healthy | 375 |
| Pneumonia Bacterial | 379 |
| Pneumonia Viral | 355 |
respectively, which can be inspected to better understand how and what feature/activation map. The result is fed to an activation function which produces the output for a given node. Fig. 2 .(a) and Fig. 2 .(b) visually depict various phases of feature extraction for MobileNetV2 and VGG16.

The dimensions of filters are generally taken as $3 \times 3$ and convolutional layers which enable the network to see more context information. The ‘16’ and ‘19’ refer to the count of weighted layers used in network configuration. In specific VGG16 consists of multiple kernels of size $3 \times 3$ and convolutional layers which enable the network to understand and recognize various complex features and patterns. The network also consists of average-pooling layers ($2 \times 2$) and 3 fully connected layers. All the hidden layers use ReLU as the activation function and Softmax activation is used to achieve the normalized classification vector. In recent times, the VGG architecture is being employed in various transfer learning approaches.

### Table 2

Comparative analysis of Models.

| METHOD                  | AVG ACCURACY | AVG PRECISION | AVG RECALL | AVG F-SCORE |
|-------------------------|--------------|---------------|------------|-------------|
| MobileNetV2 + VGG16     | 0.9648       | 0.9657        | 0.9659     | 0.9657      |
| MobileNetV2 + DenseNet  | 0.9608       | 0.9618        | 0.9620     | 0.9619      |
| MobileNetV2 + VGG19     | 0.9584       | 0.9598        | 0.9598     | 0.9597      |
| MobileNetV2 + Resnet50  | 0.9479       | 0.9498        | 0.9499     | 0.9499      |
| VGG16 + DenseNet        | 0.9627       | 0.9641        | 0.9642     | 0.9641      |
| VGG16 + VGG19           | 0.9574       | 0.9588        | 0.9597     | 0.9597      |
| VGG16 + Resnet50        | 0.9477       | 0.9505        | 0.9489     | 0.9495      |
| VGG19 + DenseNet        | 0.9580       | 0.9598        | 0.9597     | 0.9596      |
| VGG19 + Resnet50        | 0.9277       | 0.9332        | 0.9305     | 0.9306      |

**Table 3** represents the 3-Class Fold wise analysis.

| FOLD | METHOD                  | Precision | Recall | F-score | Accuracy |
|------|-------------------------|-----------|--------|---------|----------|
| Fold 1 MobileNetV2 | 0.9525 | 0.9502 | 0.9512 | 0.9484 |
| VGG16       | 0.9527 | 0.9522 | 0.9523 | 0.9509 |
| Fusion      | 0.9677 | 0.9663 | 0.9668 | 0.9656 |
| Fold 2 MobileNetV2 | 0.9476 | 0.9501 | 0.9487 | 0.9476 |
| VGG16       | 0.9579 | 0.957 | 0.9572 | 0.9558 |
| Fusion      | 0.9658 | 0.9676 | 0.9666 | 0.9656 |
| Fold 3 MobileNetV2 | 0.9461 | 0.9454 | 0.9458 | 0.9443 |
| VGG16       | 0.9445 | 0.9412 | 0.9427 | 0.9410 |
| Fusion      | 0.9611 | 0.959 | 0.96 | 0.959 |
| Fold 4 MobileNetV2 | 0.9379 | 0.9388 | 0.9383 | 0.9369 |
| VGG16       | 0.9537 | 0.9534 | 0.9535 | 0.9525 |
| Fusion      | 0.969 | 0.9699 | 0.9694 | 0.9689 |
| AVERAGE MobileNetV2 | 0.9460 | 0.9461 | 0.9460 | 0.9443 |
| VGG16       | 0.9522 | 0.95095 | 0.9514 | 0.9505 |
| Fusion      | 0.9659 | 0.9657 | 0.9657 | 0.9648 |

Conv = I * F  

(1)

The architecture of VGG [48] was proposed by Simonyan and Zisserman and achieved 92.7% accuracy which was among the top-5 in ImageNet. The two variants namely VGG16 and VGG19 have been widely used since. The ‘16’ and ‘19’ refer to the count of weighted layers used in network configuration. In specific VGG16 consists of multiple kernels of size $3 \times 3$ and convolutional layers which enable the network to understand and recognize various complex features and patterns. The network also consists of average-pooling layers ($2 \times 2$) and 3 fully connected layers. All the hidden layers use ReLU as the activation function and Softmax activation is used to achieve the normalized classification vector. In recent times, the VGG architecture is being employed in various transfer learning approaches.

### 2.1.2. MobileNetV2

MobileNet is a mobile architecture [43] which has considerable low complexity and size owing to the use of depth wise Separable Convolution, which makes it suitable to run on devices with low computational power. MobileNetV2 expanded the feature extraction and introduced an inverted residual structure. The model architecture consists of a convolutional layer followed by a series of residual bottleneck layers. Kernel size for all spatial convolution operations is taken as ReLU6 is used as the non-linearity along with batch normalization and dropout during the training phase. Each bottleneck block consists of 3 layers, starting with a $(1 \times 1)$ convolutional layer followed by the aforementioned $(3 \times 3)$ depthwise convolution layer and finally another $(1 \times 1)$ convolutional layer without ReLU6 activation. MobileNetV2 has wide applications in the current time due to its excellent feature extraction capabilities and small size.

### 2.2. Confidence fusion

Different layers of a convolutional neural network give different output. Outputs from multiple models at different stages can be used to produce a combined classification results which can often provide better results than the parent models. One such way is to fuse the output vectors [49] (categorical probability vectors) from the last Softmax classifier layer. Let $p_i$ be the output for a given CNN model and is represented as $[p_{i1}, p_{i2}, ..., p_{in}]$ where i represents the CNN under consideration and n represents the number of classification classes. The term $p_{ij}$ represents the probability that the given sample is classified as $j$th class. Average value for each class is calculated as in Equation (2).

\[ \bar{z}_j = \frac{1}{f} \sum_{i=1}^{f} p_{ij} \]

(2)

where $f$ is the total number of models being fused. Out of n classes the final predicted class $x$ is given by Equation (3).

\[ x = \arg \max_{i=1}^{n} p_{ij} \]

(3)

Fig. 1. Convolution operation with stride 1.
2.3. Dataset description

To test our proposed method, we collected a total of 4883 images for the 3 classes and is referred to as Dataset-1 (D1) [65] from Kaggle COVID-19 radiology database [66]. The database consists of COVID-19 X-ray images from padchest dataset [67], Italian Society of Medical and Interventional Radiology (SIRM) COVID-19 database [68], Germany medical school [69], and other publicly available databases available on github [70] well as images extracted from different publications. The Viral Pneumonia images are collected from Radiological Society of North America (RSNA) CXR dataset [71] and the Healthy images from Mendeley Data [72] and RSNA [71] datasets. For accurate evaluation of any deep learning model, an unbiased and balanced dataset is paramount. Hence, to counter the unbalanced dataset problem, we considered almost comparable number of samples for each class. Moreover, to make sure that the data did not contain similar samples/images from the same source, the images for each class were collected randomly from the sources. The second database referred to as Dataset-2 (D2) consists of 305 images for COVID-19 patients, 375 for healthy and 379, 355 for bacterial and viral pneumonia, respectively. This data has been directly taken from the dataset provided by Khan et al. [73].

The use of two separately prepared datasets provides a better clarity about the effects of size and count of images on the performance of the proposed method. The description of the number of image samples used for both the datasets are provided in Table 1. The visualization of X-ray images for each class as shown in Fig. 3.
2.4. Proposed methodology

2.4.1. Experimental setup and development

To evaluate the predictive performance, Precision, Recall and F-score of the proposed method, we have used Stratified K-Folds cross-validation ($k = 4$) technique. It also helps to avoid bias in the data selection. Hence, 75% of samples are used for training purposes and the remaining 25% of data is used for testing phase while preserving the percentage of samples for each class. This provides an efficient way of determining out of sample observations. For model building and evaluation, we have used Keras interface with a Tensorflow 2.0 backend with pre-trained models namely MobileNetV2 and VGG16. The models were executed and developed on Google Collaboratory Pro, using Google’s cloud GPU enabled notebook with 25 GB RAM and Tesla P100 GPU. The proposed framework was trained with Adam optimizer with loss function as the ‘Sparse_categorical_crossentropy’, learning rate of 0.001, batch size 32. Tests were performed for various parameter values and the above-mentioned set of values provided the best results. VGG16 was trained to 70 epochs and MobileNetV2 up to 25 epochs. Validation split value (Subset of training data taken for validation during training) was taken as 0.1. All the required session seed values were taken as 10.

2.4.2. Model building and architecture

In recent time’s deep learning algorithms, CNN models in particular have seen numerous breakthroughs. It is extremely expensive to train due to complex data models. Moreover, deep learning requires expensive GPUs. Since computational constraints have relaxed, and better hardware is more readily available these limitations are no longer a barrier which is a big step forward. High performance models which are publicly available can be effectively tweaked and used as a base for classification of COVID-19 cases. We expanded this approach by modifying the architecture of two models namely, MobileNetV2 and VGG16 which are then used for classification based on confidence fusion. These Exhaustive testing using various combination of models showed that the selected models had the best evaluation metrics. This shown in Table 2. The VGG16 and MobileNetV2 provided 2 very different architectures and unique feature extraction stages and show extremely promising results when compared with other models like VGG-19, ResNet. Upon fusion these shows quite significant improvement over the results of shown by either of its parent models for each of the evaluators.

The VGG architecture is a simple one with $3 \times 3$ convolutional layers stacked on top of each other with occasional pooling layers for reducing data dimensionality. The MobileNetV2 has an elaborate architecture with a fully convolution layer with 32 filters followed by 19 bottleneck layers [43]. The framework and architecture of our proposed work is presented in Fig. 4 and Fig. 5, respectively. Both the models are capable of extensive feature extraction and take fixed input of 3-channel (RGB) images of size $224 \times 224$. Hence, each image is resized to $224 \times 224$ pixels followed by reshaping of the image vector to $224 \times 224 \times 3$ and then, applying normalization operation. The fully connected output layers of the model used to make classifications are not loaded and the loaded layers are set as non-trainable. The extracted features vectors of size $7 \times 7 \times 1280$ and $7 \times 7 \times 1024$ from MobileNetV2 and VGG16 are passed to a separate GlobalAveragePooling2D which reduces the vector dimensionality to 1, followed by two fully connected layers.

The second fully connected layer supports Softmax activation function which outputs a vector of categorical probabilities for each class. The output vectors of both the models are used in the confidence fusion which is used to classify the image. The confidence vectors obtained for each passed sample by both the models are given by Equation (4) and Equation (5).

$$p_{\text{xy}} = [p_{\text{xy(COVID-19)}}, p_{\text{xy(Healthy)}}, p_{\text{xy(Pneumonia)}}]$$

$$p_{\text{mnv2}} = [p_{\text{mnv2(COVID-19)}}, p_{\text{mnv2(Healthy)}}, p_{\text{mnv2(Pneumonia)}}]$$

Equation (6) and finally Equation (7) are calculated for each given sample and $x$ represents the predicted class for the given sample. This is shown in Fig. 5.

$$f_x = \frac{(p_x + p_{x2})}{2}$$

$$x = \arg\max f_x, j \in \{ \text{COVID - 19, Healthy, Pneumonia} \}$$

2.5. Evaluation criteria

For multiclass classification, each classification class is representative of three types of data namely True Positive (TP)/ False Positive (FP) / False Negative (FN). True positive indicates the images that were classified correctly. False Positive contains samples that were falsely classified to the class under consideration whereas False Negative indicates the samples of the current class classified incorrectly to some other class. There are several performance metrics that provide different evaluations using the above mentioned characteristics. Following four metrics are considered to evaluate the performance of our proposed methods:

![Fig. 4. Framework for the proposed methodology.](image-url)
(a) **Precision:** It is the fraction of True classification among all samples predicted to belong to a particular class. For example, precision for covid refers to the proportion of people who test positive among all those who actually are predicted have the disease.

\[
\text{Precision} = \frac{TP}{TP + FP}
\]  

(b) **Recall:** It refers to the proportion of actual positives identified correctly. Recall for covid is the fraction of people who are predicted to have the disease to the actual number who had it.

\[
\text{Recall} = \frac{TP}{TP + FN}
\]

(c) **F-Score/F-Measure:** It is a metric used to balance both the concerns of precision and recall into one by providing a single number which is the weighted average of both.

\[
F - \text{score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
\]

(d) **Accuracy:** It is the fraction of samples classified correctly out of all the samples. For 3-class classification, accuracy is given as:

\[
\text{Accuracy} = \frac{T_{\text{COVID-19}} + T_{\text{Normal}} + T_{\text{Pneumonia}}}{\text{Totalsamples}}
\]

where \(T_{\text{COVID-19}}\) refers to true covid classifications and similarly for other classes.

3. **Comparative analysis of state-of-the-art deep learning methods**

The evaluation was performed using Stratified K-Folds cross-validation technique where 75% of the data was used for training purpose and the remaining 25% was used for testing purpose. Plots of accuracy and loss on the training and validation sets (10%) over training phase for Fold 2 are shown in Fig. 6. A good validation and testing accuracy is observed which help check for overfitting issues. The Classification performance for 3-classes: COVID-19, Healthy and Pneumonia for each fold are evaluated using various performance measures provided in Table 2. Fig. 7 shows the corresponding confusion matrices for both base models and the fused model which helps to obtain an idea about the general perforations of the proposed method. It is observed that, out of 154 COVID-19 images, only two images were misclassified among all 4-folds. The fused results also show better true positives and lesser false positives, false negatives that both the base models. Performance metrics for the proposed method, 4-fold Stratified K-Folds cross-validation technique and their average are tabulated in Table 3.

A significant improvement in the performance measures used for evaluation was observed in each fold when the base models and the fusion model were compared. This is also reflected in the corresponding
The proposed COVDC-Net method was tested on two different datasets and achieved an accuracy of 96.48% on 3-classes and 90.22% on 4-classes. This performance is superior in terms of precision, recall, F-score and accuracy as compared to existing deep learning-based approaches represented in Table 6. The ROC Curves for the used base models and the proposed fusion are shown in Fig. 8.

Shilby et al. [44] came up with R-CNN, a deep learning architecture with 97.36% accuracy on 2-classes having 183 COVID-19 and 13,619 (Normal + Pneumonia) images. Wang et al. [47] proposed a weakly supervised framework for COVID-19 classification and lesion localization task and reported a 90% classification accuracy on a dataset with 313 COVID-19 images and 229 Healthy/Normal chest X-rays. Later, Ozturk et al. [18] presented their model for binary and multiclass classification task and reported a 90% classification accuracy on a dataset with 125, 500, 500 samples for COVID-19, Pneumonia and Normal classes respectively. Transfer learning uses previously proven models and expands gained knowledge to other areas. Apostolopoulos et al. [17] used this idea and used various previous models and achieved 93.48% and 92.85% accuracy for VGG19 and MobileNetV2 respectively on a dataset of 224 COVID-19, 504 Healthy and 700 Pneumonia patients. Wang et al. [15] proposed a tailored convolutional network model which had an accuracy of 93.3% and a relatively small dataset of 53, 179, 179 images of COVID-19, Normal and Pneumonia, respectively. In addition, Law and Lin [45] implemented Transfer learning on a dataset with 1200 images of COVID-19 patients and found that VGG-16 gives superior performance metrics as compared to the other ResNet models. A Multi-model fusion study was performed by Cengil and Cinar [46], tested various models namely AlexNet, EfficientNet-b0, NASNet-Large and Xception to find the best performing amalgamations. Khan et al. [24] used pre-trained Xception architecture to get 95% 3-class and 89.6% 4-class accuracy on a dataset of 284 COVID-19, 310 healthy, 227 Viral Pneumonia and 330 Bacterial Pneumonia samples.

In this study, various X-ray samples collected from the aforementioned sources were analysed using the proposed COVDC-Net fusion method and investigated using 4 different metrics which are compared with other previous existing works. The important observations can be

Table 4
Class wise analysis.

| METHOD     | CLASS     | Precision | Recall   | F-score  |
|------------|-----------|-----------|----------|----------|
|MobileNetV2 | COVID-19  | 0.9474    | 0.9395   | 0.9434   |
|            | Normal    | 0.9258    | 0.9316   | 0.9286   |
|            | Pneumonia | 0.9648    | 0.9673   | 0.9659   |
|VGG16       | COVID-19  | 0.9584    | 0.94     | 0.9491   |
|            | Normal    | 0.9252    | 0.9515   | 0.9382   |
|            | Pneumonia | 0.9730    | 0.9614   | 0.967    |
|Fusion      | COVID-19  | 0.9711    | 0.9591   | 0.965    |
|            | Normal    | 0.9495    | 0.9618   | 0.9556   |
|            | Pneumonia | 0.9771    | 0.9762   | 0.9766   |
expressed as follows:

- The dataset imbalance as observed in many of the works render their findings inaccurate and not suitable for reference for practical implementations. This is observed in Shibly et al. [44,44], Wang et al. [15]. Our method uses a relatively balanced dataset which is randomly selected so as to ensure random nature of the sample and is evaluated using Stratified 4-fold Cross Validation method, hence providing consistent results.
- Considering various previous works, it is observed that many leave out certain metrics or present them vaguely. Hence these should not be used as a benchmark for practical purposes as it might have unanticipated repercussions. This can be seen in Apostolopoulos et al. [17] where performance metrics are not properly presented or missing for all the classes. We have presented the results for each class and for all the folds with meaningful results which are more suited to give a better understanding of how and where the improvements that are being observed. All the performance metrics can be seen in the Table 3 and show class wise results and improvement after the fusion of the parent models clearly.

- Most of the existing works focus on 3-class and binary classification tasks. Our study, extends the 3-class classification task to 4-class classification and continues to show significant improvements. It also performs better on comparison with, from where the dataset (D2) is borrowed which is used in the 4-class classification tasks.
- The model showed an improvement of 0.69% for 4-class classification task and 2.13% in overall accuracy for 3 class classification task on the dataset D2 [73] over the method proposed by Khan et al. [24].
- It was observed that the proposed method showed the quite significant improvements in accuracy, precision, recall for COVID-19 class classification of a COVID-19 case can have serious repercussions (as seen in Table 3 ), which is paramount for this study as missing for all the classes. We have presented the results for each class and for all the folds with meaningful results which are more suited to give a better understanding of how and where the improvements that are being observed. All the performance metrics can be seen in the Table 3 and show class wise results and improvement after the fusion of the parent models clearly.

### Table 6
Comparison of the proposed system with existing systems in terms of accuracy.

| AUTHOR         | CLASSES                  | TYPE        | MODEL                           | ACCURACY | PROS                                                       | CONS                                      |
|----------------|--------------------------|-------------|---------------------------------|----------|------------------------------------------------------------|-------------------------------------------|
| Shibly et al.  | 2 Class: (COVID-19: 183, Healthy: 13617) | Chest X-Ray | R-CNN                           | 97.36%   | faster R-CNN                                              | Limited data                             |
| Wang et al.    | 2 Class: (COVID-19: 313, Healthy: 229) | Chest CT    | DeCoVNet (UNet + 3D Deep Network) | 90.0%    | 10-folds cross-validation light-weight 3D CNN              | Limited data                             |
| Ozturk et al.  | 2 Class: (COVID-19: 125, Healthy: 500) | Chest X-Ray | DarkCovidNet                    | 98.08%   | The heatmaps produced by the model can be evaluated by an expert radiologist. High Binary Classification accuracy | Limited data Relatively low Multi-class accuracy |
| Apostolopoulos | 3 Class: (COVID-19: 125, Healthy: 500, Pneumonia: 500) | Chest X-Ray | VGG-19                          | 93.48%   | Multiple models used for testing                          | Limited no of evaluation metrics          |
| et al. [17]    | 3 Class: (COVID-19: 224, Healthy: 504, Pneumonia: 700) | MobileNet v2 | 92.85%                          |          |                                                           | Data-set imbalance                       |
| Law and Lin    | 3 Class: (COVID-19: 1200, Healthy: 1341, Pneumonia: 1345) | Chest X-Ray | COVID-Net                       | 93.3%    | Low architectural complexity                               |                                          |
| [45]           |                                             |             |                                 |          |                                                           |                                          |
| Cengil and Cinar | 3 Class: (COVID-19: 1525, Healthy: 1525, Pneumonia: 1525) | Chest X-Ray | AlexNet + EfficientNet- b0 + NASNetLarge + Exception | 95.9%    | 3 different datasets used i.e. robust Hybrid Model High Performance metrics | High model complexity                     |
| [46]           |                                             |             |                                 |          |                                                           |                                          |
| Khan et al.    | 3 Class: (COVID-19: 284, Healthy: 310, Pneumonia: 657) | Chest X-Ray | Cronet                          | 95%      | 4-Class Classification results High Accuracy for COVID-19 class | Limited data for COVID-19 Class           |
| [24]           |                                             |             | (Xception)                      |          |                                                           |                                          |
| Proposed       | 3 Class: (COVID-19: 1784, Healthy: 1755, Pneumonia: 1345) | Chest X-Ray | COVIDC-Net                      | 96.48%   | Balanced Dataset 4-Class, 3-Class Classification High Performance Metrics Achieved | Hybrid Methods are computationally expensive |
|                | 4 Class: (COVID-19: 305, Healthy: 375, Viral Pneumonia: 379, Bacterial Pneumonia: 355) |             |                                 | 90.22%   |                                                           |                                          |
4. Conclusion

In this study, a deep convolutional network-based method was proposed for efficient classification of the novel coronavirus from healthy and pneumonia infected patients. For which, we used pre-trained MobileNetV2 and VGG16 which have a very different architecture and feature selection process. On fusion of the models, after addition of some extra layers and considering several other crucial factors, the proposed method is able to benefit from various differentiating patterns and visual features for each pneumonia class. Therefore, the fusion performance observed was better than the individual performance of either of the models. These encouraging results show that such models can be used to diagnose other chest related diseases like tuberculosis and pneumonia. In real life scenarios, CNN is an effective method for detecting CoV-affected patients considering the sample collection is safer than a nasal swab and is much faster in terms of diagnosis. The proposed model provides a multiclass diagnosis with other pulmonary diseases which is a key factor as many of these might share the surprisingly similar symptoms and effects on the lungs. Other tests may get rendered useless in such cases and hence showing the supremacy of CNN models like COVDC-Net in practical applications in the coming time. Scaling these to existing hospitals and clinics with X-ray machines doesn’t require special infrastructure and testing equipment and is quick, which is another benefit of AI-Aided methods for diagnosis. In future, we plan to evaluate the proposed model on other wider sets of pulmonary diseases and make the model more robust. With an even more in depth understanding of the SARS-CoV-2 virus and better data, the models can be improved further.

CRediT authorship contribution statement

Anubhav Sharma: Conceptualization, Methodology, Software, Investigation. Karamjeet Singh: Data curation, Resources, Writing – original draft, Visualization. Deepika Koundal: Writing – review & editing.

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.

References

[1] K. McIntosh, Coronavirus disease 2019 (COVID-19), Epidemiology, virology, clinical features, diagnosis, and prevention, 2020.
[2] COVID-19 Weekly Epidemiological Update, Edition 84, published 22, WHO, 2022 https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19—22-march-2022 April 2022.
[3] Naming the coronavirus disease (covid-19) and the virus that causes it, WHO, 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019-technical-guidance/naming-the-coronavirus-disease-(covid-19)-and-the-virus-that-causes-it, (accessed March 25, 2022).
[4] E. Mahase, Coronavirus: Covid-19 Has Killed More People Than SARS and MERS Combined, Despite Lower Case Fatality Rate, BMJ 368 (2020), m641, https://doi.org/10.1136/bmj.m641.
[5] Statement on the fifth meeting of the International Health Regulations Emergency Committee regarding the coronavirus disease (COVID-19) pandemic, WHO, October 2020 https://www.who.int/news/item/30-10-2020-statement-on-the-fifth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic-2005 accessed March 25, 2022.
[6] D. Sun, H. Li, X-X. Lu, H. Xiao, J. Ren, F-R. Zhang, Z-S. Liu, Clinical features of severe pediatric patients with coronavirus disease 2019 in Wuhan: a single center’s observational study, World J. Pediatr. 16 (2020) 251–259. doi: 10.1007/s12634-020-00054-4.
[7] C. Huang, Y. Wang, X. Li, R. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q.-J. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, The Lancet 395 (10223) (2020) 497–506.
[8] F. Jiang, L. Deng, L. Zhang, Y. Cai, C.W. Cheung, Z. Xia, Review of the clinical characteristics of coronavirus disease 2019 (covid-19), J. Gen Intern Med. 35 (5) (2020) 1–5, https://doi.org/10.1007/s11606-020-05762-w.
[9] K. Tolkuddorf, S. Buda, E. Schuler, L.H. Wieder, W. Haas, Influenza associated pneumonia as reference to assess seriousness of coronavirus disease (COVID-19), Euro Surveill. 25 (11) (2020) 2000258, https://doi.org/10.2807/1560-7917.ES.2020.25.11.2000258.
[10] G. Mackenzie, The definition and classification of pneumonia, 2016, Pneumonia 8, 14, https://doi.org/10.1186/s41479-016-0012-z.
[11] W. Wang, Y. Xu, R. Gao, R. Lu, K. Han, G. Wu, W. Tan, Detection of SARS-CoV-2 in different types of clinical specimens, JAMA. 323 (18) (2020) 1843–1844, https://doi.org/10.1001/jama.2020.3786.
[43] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L-C Chen L, MobileNetV2: Inverted Residuals and Linear Bottlenecks, 2018, arXiv:1801.04381.

[44] K.H. Shibly, B.K. Law, L.P. Lin, Development Of A Deep Learning Model To Classify X-Ray Of COVID-19, Normal And Pneumonia-Affected Patients, in: IEEE International Conference on Image Processing (ICIP), 2021, pp. 1–6, https://doi.org/10.1109/ICIP48528.2021.9576049.

[45] B.K. Law, L.P. Lin, Development Of A Deep Learning Model To Classify X-Ray Of COVID-19, Normal And Pneumonia-Affected Patients, in: IEEE International Conference on Image Processing (ICIP), 2021, pp. 1–6, https://doi.org/10.1109/ICIP48528.2021.9576049.

[46] E. Cheng, A. Guan, The feature and content comprehension in the classification problem: An approach on COVID-19 disease detection, Int. J. Imaging Syst. Technol. 32 (2021) 26–40. Doi: 10.1111/jmii.12269.

[47] X. Wang, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, C. Zheng, A weakly-supervised framework for COVID-19 pneumonia classification using chest X-ray images, 2020, Expert Syst. Appl. 165, 113909. doi: 10.1016/j.eswa.2020.113909.

[48] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L-C Chen L, MobileNetV2: Inverted Residuals and Linear Bottlenecks, 2018, arXiv:1801.04381.

[49] S. Zheng, P. Qi, S. Chen, X. Yang X, Fusion methods for CNN-Based automatic modulation classification, IEEE Access Special Section on Artificial Intelligence for Physical-Layer Wireless Communications and Networks (2019) 66496 – 66504. doi: 10.1109/access.2019.2918136.

[50] P. Kalane, S. Patil, B. Patil, D.P. Sharma, Automatic detection of COVID-19 disease using U-Net architecture based fully convolutional network, 2021, Biomed. Signal Process. Control 64, 102257. doi: 10.1016/j.bspc.2021.102257.

[51] M. Canayaz, MH-COVIdNet: Diagnosis of COVID-19 using deep neural networks and meta-heuristic-based feature selection on X-ray images, 2021, Biomed. Signal Process. Control 64, 102257. doi: 10.1016/j.bspc.2021.102257.

[52] R. Dubey, M. Kumar, A. Upadhyay, R.B. Pachori, Automated diagnosis of muscle diseases from EMG signals using empirical mode decomposition based method, Biomed. Signal Process. Control 71 (2021). https://doi.org/10.1016/j.bspc.2020.101649.

[53] K. Simonian, A. Zisserman, Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations (ICLR), 2015.

[54] S. Zheng, P. Qi, S. Chen, X. Yang X, Fusion methods for CNN-Based automatic modulation classification, IEEE Access Special Section on Artificial Intelligence for Physical-Layer Wireless Communications and Networks (2019) 66496 – 66504. doi: 10.1109/access.2019.2918136.

[55] K. Simonian, A. Zisserman, Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations (ICLR), 2015.

[56] K. Simonian, A. Zisserman, Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations (ICLR), 2015.

[57] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[58] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[59] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[60] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[61] G. Altan, Deep Learning-based Mammogram Classification for Breast Cancer, Int. J. Comput. Biol. Med. 43 (2020) 635 – 649, https://doi.org/10.1016/j.ijcbm.2020.103792.

[62] G. Altan, Y. Kutlu, A. Gokcen, Chronic obstructive pulmonary disease severity analysis using deep learning on multi-channel lung sounds, Turk. J. Elec. Eng. Phys. Eng. Sci. Med. 43 (2020) 635 – 649, https://doi.org/10.1016/j.ijcbm.2020.103792.

[63] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[64] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[65] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[66] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[67] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[68] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[69] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[70] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[71] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[72] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[73] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[74] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[75] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.

[76] A. Upadhyay, M. Sharma, R.B. Pachori, R. Sharma, A non-parametric approach for multicomponent AM-FM signal analysis, Circuits, Syst. Process. 39 (2020) 6315–6355. doi: 10.1002/casp.2932.
chest X-ray, Pattern Recog. Lett. 150 (2021) 8–16, https://doi.org/10.1016/j.patcog.2020.06.021.

[64] S.-H. Wang, M.A. Khan, V. Govindaraj, S.L. Fernandez, Z. Zhu, Y.D. Zhang, Deep Rank-Based Average Pooling Network for Covid-19 Recognition, Comput., Mater. Continua 70 (2) (2022) 2797–2813, https://doi.org/10.32604/cmc.2022.020140.

[65] A. Sharma, K. Singh, D. Koundal, Dataset for COVDC-Net. https://github.com/sharma-anubhav/COVDC-Net. (Accessed November 2021).

[66] COVID-19 Radiography Database. https://www.kaggle.com/tawsifurrahman/covid19-radiography-database. (Accessed November 2021).

[67] BIMCV-COVID19, Datasets related to COVID19’s pathology course. https://bimcv.cipf.es/bimcv-projects/bimcv-covid19/#1590858128006-9e640421-6711. (Accessed November 2021).

[68] Italian Society of Medical and Interventional Radiology (SIRM) COVID-19 Chest X-Ray dataset. https://sirm.org/category/senza-categoria/covid-19/. (Accessed November 2021).

[69] H.B. Winther, H. Laser, S. Gerbel, S.K. Maschke, J.B. Hinrichs, J. Vogel-Clausen, F.K. Wacker, M.M. Höper, B.C. Meyer (2020). Dataset: COVID-19 Image Repository. Doi: 10.25835/0090041.

[70] J. P. Cohen, P. Morrison and L. Dao, COVID-19 image data collection. https://github.com/ieee8023/covid-chestxray-dataset. (Accessed November 2021).

[71] RSNA Pneumonia Detection Challenge, https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data. (Accessed November 2021).

[72] D. Kermany, K. Zhang, M. Goldbaum, Labeled Optical Coherence Tomography (OCT) and chest X-Ray images for classification, Mendeley data v2 (2018), https://doi.org/10.17632/ncsh9b7k9j.2.

[73] A.I. Khan, J.L. Shah, M.M. Bhat, Dataset for Covid-19 Classification 3, 4 class. https://github.com/drkhan107/CoroNet. (Accessed November 2021).