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.
Original Investigation

Stacking Ensemble and ECA-EfficientNetV2 Convolutional Neural Networks on Classification of Multiple Chest Diseases Including COVID-19

Mei-Ling Huang PhD. in Industrial Engineering & Management, Yu-Chieh Liao Master in Industrial Engineering & Management

Rationale and Objectives: Early detection and treatment of COVID-19 patients is crucial. Convolutional neural networks have been proven to accurately extract features in medical images, which accelerates time required for testing and increases the effectiveness of COVID-19 diagnosis. This study proposes two classification models for multiple chest diseases including COVID-19.

Materials and Methods: The first is Stacking-ensemble model, which stacks six pretrained models including EfficientNetV2-B0, EfficientNetV2-B1, EfficientNetV2-B2, EfficientNetV2-B3, EfficientNetV2-S and EfficientNetV2-M. The second model is self-designed model ECA-EfficientNetV2 based on ECA-Net and EfficientNetV2. Ten-fold cross validation was performed for each model on chest X-ray and CT images. One more dataset, COVID-CT dataset, was tested to verify the performance of the proposed Stacking-ensemble and ECA-EfficientNetV2 models.

Results: The best performance comes from the proposed ECA-EfficientNetV2 model with the highest Accuracy of 99.21%, Precision of 99.23%, Recall of 99.25%, F1-score of 99.20%, and (area under the curve) AUC of 99.51% on chest X-ray dataset; the best performance comes from the proposed ECA-EfficientNetV2 model with the highest Accuracy of 99.81%, Precision of 99.80%, Recall of 99.80%, F1-score of 99.81%, and AUC of 99.87% on chest CT dataset. The differences for five metrics between Stacking-ensemble and ECA-EfficientNetV2 models are not significant.

Conclusion: Ensemble model achieves better performance than single pretrained models. Compared to the SOTA, Stacking-ensemble and ECA-EfficientNetV2 models proposed in this study demonstrate promising performance on classification of multiple chest diseases including COVID-19.

KEY WORDS: COVID-19; Convolutional neural network; Ensemble learning; Stacking.

INTRODUCTION

COVID-19 causes disease in humans and vertebrates, and is a zoonotic infectious disease. Some confirmed patients may have severe pneumonia and respiratory failure (1). The most common tools for detecting lung infections are Chest X-ray and Computed tomography (CT). Chest X-ray is a very common noninvasive radiological test that has been widely used to screen for a variety of lung diseases such as COVID-19, pneumonia, pulmonary effusion, lung cancer, and emphysema (2). In clinical practice, chest X-ray images are often interpreted by radiologists, which is time-consuming and prone to errors in subjective assessments. A CT image is composed of a certain number of pixels with different grayscales arranged in a matrix. If the number of pixels is larger, the pixel value is smaller, and the image will be clearer. Although CT images provide very fine details, but it has more radiation than a chest X-ray image, and the equipment is relatively expensive (3).

In recent years, due to the rise of artificial intelligence, researchers have applied deep learning to detect COVID-19 by using chest X-ray images and CT images. Compared to traditional machine learning, deep learning can automatically extract features of images to reduce processing time (4). At present, there have been many researches applying convolutional neural network (CNN) on image recognition. This method has been proved to be a powerful image recognition...
technology and has been widely used for COVID-19 detection. Common CNNs include GoogLeNet, ResNet, Xception, DenseNet, MobileNet, and EfficientNet, etc. For example, Ieracitano et al. proposed a fuzzy logic-based convolutional neural network (CoNNet) and the classification accuracy is 81.00% in 2022 by using a total of 155 chest X-ray images (5). Khan et al. proposed two convolutional neural networks named DHL and DBHL in 2021, using a total of 6448 chest X-ray images. The results showed that the accuracy of the binary classification of COVID-19 and normal was 98.53% (6). Loey et al. proposed a Bayesian-optimized convolutional neural network in 2022 using a total of 10,848 chest X-ray images. The results showed that the three-category accuracy of COVID-19, general pneumonia and normal was 96.00% (7).

Hu et al. proposed a two-stage detection method in 2021 using a total of 11,092 chest X-ray images (8). The first stage is to train a CNN as a feature extractor, and the second stage uses extreme learning machines (ELMs) for real-time detection. In addition, the Chimp optimization algorithm is used to improve the results and increase the reliability of the network, and finally it is compared to the general CNN, Genetic algorithm optimized ELM, Cuckoo search optimized ELM, and Whale optimization algorithm optimized ELM. The results show that the proposed method has a binary classification accuracy of 99.11% for COVID-19 and non-COVID-19.

Thaseen et al. applied ensemble learning by combining ResNet, FitNet, IRCNN, MobileNet, and EfficientNet in 2022, using a total of 13,808 chest X-ray images (9). The results showed that the three-category accuracy of COVID-19, general pneumonia and normal was 99.00%. A hybrid convolutional neural network (HCNN) combining CNN and RNN on classification of three-category accuracy of COVID-19, general pneumonia and normal was proposed with the classification accuracy of 98.20% (10).

Musallam et al. proposed a convolutional neural network called DeepChest in 2022, using a total of 7512 chest X-ray images. The results showed that the three-category accuracy of COVID-19, general pneumonia and normal was 96.56% (11). A convolutional neural network named DenResCov-19 composed of DenseNet121 and ResNet50 networks was proposed, and DenResCov was combined with existing ResNet50, DenseNet121, VGG16, and InceptionV3 networks, using a total of 6469 chest X-ray images. The results showed that the four-category area under the curve (AUC-ROC) of COVID-19, general pneumonia, tuberculosis and normal was 99.60% (12).

The literature on COVID-19 detection using CT images such as Garg et al. used a series of models such as EfficientNet, DenseNet, VGG, and ResNet with a total of 20 convolutional neural networks in 2022, using a total of 4173 CT images. The results show that the binary classification accuracy of EfficientNet-B5 in detecting COVID-19 and non-COVID-19 is 98.45% (13). Lhasani et al. used transfer learning on VGG16, VGG19, Xception, InceptionV2, ResNet, DenseNet121, and DenseNet201 in 2021, combined with GradCam, using a total of 4986 CT images. The results show that in the binary classification of COVID-19 and non-COVID-19, DenseNet201+ GradCam achieves the best accuracy rate of 98.80% (14). Rahimzadeh et al. used ResNet50V2 as the backbone network and compared the model with ResNet50V2 and Xception after adding a feature pyramid network (FPN), using a total of 63,849 CT images. The results show that ResNet50V2+ FPN has an accuracy of 98.49% in COVID-19 and normal binary classification (15). Qi et al. first used five models of U-Net, LinkNet, R2UNet, Attention U-Net, and U-Net++ to segment CT images, and then used pretrained DenseNet121, InceptionV3 and ResNet50 for classification, using a total of Over 10,000 CT images. The results show that in the binary classification of COVID-19 and CAO, LinkNet performs best in lung segmentation with a Dice coefficient of 0.9830, while DenseNet121 with capsule network has a prediction accuracy of 97.10% (16).

Abdel-Baset et al. proposed a two-stage detection method in 2021. The first stage is to use the proposed GR-UNet to segment the area of lung infection, and then transfer learning is used as feature extraction; the second stage is to use the proposed GR-UNet to segment the lung infected area. The stage is to propose an infection prediction module that uses the infected location to make decisions about classification, using a total of 9,593 CT images. The results showed that the binary classification accuracy of COVID-19 and CAP was 96.80% (17). Ye et al. proposed a convolutional neural network named CIFD-Net in 2022, which can effectively handle the multi-region displacement problem through a new robust supervised learning, using a total of 45,167 CT images picture. The results showed that the binary classification accuracy of COVID-19 and non-COVID-19 was 91.19% (18).

Balaha et al. used transfer learning models including ResNet50, ResNet101, VGG16, VGG19, Xception, MobileNetV1, MobileNetV2, DenseNet121, and DenseNet169 in 2021, and added the Harris Hawks optimization to optimize hyper-parameters, and finally use fast classification stage and compact stacking stage to stack the best models into one, using a total of 15,535 CT images. The results show that in the binary classification of COVID-19 and non-COVID-19, the weighted sum method (WSM) is used to obtain an accuracy of 99.33% (19).

Qi et al. proposed a detection method named DR-MIL in 2021, which first treats a 3D CT image of a patient as a bag and selects ten CT slices as initial instances. The deep features were then extracted from the pretrained ResNet50 by fine-tuning and treated as a Deep Represented Instance Score. The bag with DRIS was input to K-Nearest Neighbor (KNN) to generate the final prediction, using a total of 241 patients CT images. The results showed that the binary classification accuracy of COVID-19 and CAP was 95.00% (20).

Some scholars used both chest X-ray images and CT images to detect COVID-19. Related studies such as Ahamed et al. fine-tune the pretrained ResNet50V2 in 2021, using a total of 4593 chest X-ray images and 3000 CT images. The results showed that the four-category accuracy for detecting COVID-19, viral pneumonia, bacterial pneumonia, and...
normal in chest X-ray images was 96.45%; the three-category accuracy for COVID-19, bacterial pneumonia and normal was 97.24%; COVID-19 and normal binary classification accuracy was 98.95%. The three-category accuracy for detecting COVID-19, CAP, and normal in CT images was 99.01%; the two-category accuracy for COVID-19 and normal in chest X-ray images was 96.45%; the three-category accuracy for COVID-19 and non-COVID-19 in CT images was 99.99% (21). Kumari et al. used four convolutional neural networks, InceptionV3, VGG16, Xception and ResNet50 in 2021, using a total of 2000 chest X-ray images and 2000 CT images. The results show that the binary classification accuracy of VGG16 in detecting COVID-19 and non-COVID-19 in chest X-ray images was 98.00%; the binary classification accuracy of Xception in detecting COVID-19 and non-COVID-19 in CT images was 83.00% (22). Ahsan et al. fine-tuned eight convolutional neural networks including VGG16, VGG19, ResNet15V2, Inception-ResNetV2, ResNet50, DenseNet201, MobilenetV2, and NasNetMobile in 2020, using a total of 400 chest X-ray images and 400 CT images. The results show that the binary classification accuracy of NasNetMobile in detecting COVID-19 and non-COVID-19 in chest X-ray images was 100.00%; the binary classification accuracy of NasNetMobile in detecting COVID-19 and non-COVID-19 in CT images was 95.20% (23). Jia et al. proposed a dynamic CNN modification method in 2021 and applied it to a fine-tuned ResNet, and finally compared the results with VGG16, InceptionV3, ResNet18, DenseNet121, MobileNetV3, and SqueezeNet, using a total of 7,592 chest X-ray images and 104,009 CT images. The results showed that the five-category accuracy of detecting COVID-19, viral pneumonia, bacterial pneumonia, tuberculosis and normal in chest X-ray images was 99.60%; in CT images, detecting COVID-19, non-COVID-19 and normal three-category accuracy was 99.30% (24).

Kassania et al. use eight models such as DenseNet, ResNet, MobileNet, InceptionV3, Xception, InceptionResNetV2, VGG and NASNet to extract features, and then extracted features were input to decision tree, random forest, XGBoost, AdaBoost, Bagging, and LightGBM. A total of 137 chest X-ray images and 137 CT images were used. The results show that DenseNet121+ Bagging combines chest X-ray images and CT images to detect COVID-19 and normal with a binary classification accuracy of 99.00% (25). Gour and Jain proposed an integrated stacked CNN in 2022. After fine-tuning VGG19 and Xception and generating five sub-models, all sub-models were stacked using a softmax classifier with a total of 3040 images Chest X-ray images and 4,645 CT images. The results showed that the three-category accuracy of detecting COVID-19, general pneumonia, and normal in chest X-ray images was 97.27%; and the two-category accuracy of detecting COVID-19 and normal in CT images was 98.30% (26). Kamil et al. used fine-tuned VGG19 in 2021, using a total of 977 chest X-ray images and 23 CT images. The results showed a 99.00% accuracy of the binary classification between COVID-19 and normal (27).

Saygılı applied Bag of Tree, K-ELM, KNN, and SVM with a total of 1125 chest X-ray images and 3228 CT images. The results showed that SVM has an accuracy of 85.96% in detecting COVID-19, general pneumonia and normal in chest X-ray images; K-ELM is accurate in detecting COVID-19 and non-COVID-19 in CT images. The accuracy was 98.88% (28).

Existing pretrained models are designed for general natural images and fine-tuned for classified image types, which is not specifically designed for COVID-19 detection. The general natural images are large and simple, while the images of COVID-19 have specific patterns and textures that differ significantly from natural images. Based on previous studies, we can find that the use of ensemble learning and the author’s self-proposed convolutional neural network has good performance in detecting COVID-19 in chest X-ray images and CT images. Ensemble learning mainly combines multiple existing convolutional neural networks, which can not only reduce the probability of misjudgment by a single model, but also improve the classification accuracy in less time. Ensemble learning solves the current need to detect COVID-19 without designing a new model to get good detection performance in the most time-saving way.

Challenging the possibility of improving accuracy on multiple chest disease diagnosis, this study proposes two classification models for multiple chest diseases including COVID-19. First, we obtained chest X-ray images and CT images from multiple public databases. Next, we select six pretraining models including EfficientNetV2-B0, EfficientNetV2-B1, EfficientNetV2-B2, EfficientNetV2-B3, EfficientNetV2-S and EfficientNetV2-M in the EfficientNetV2 series. The reason for choosing the previous models is the training speed and argument efficiency of EfficientNetV2 is better than some previous networks (29). To the best of our knowledge, this is the first study ensembles the series of EfficientNetV2 models for COVID-19 detection. In addition, this study proposes a self-designed model ECA-EfficientNetV2 based on ECA-Net and EfficientNetV2.

**MATERIALS AND METHODS**

Figure 1 represents the architecture of this study.

- **Step 1** Data extraction: The chest X-ray and CT images are collected.
- **Step 2** Image preprocessing: The size of all selected images is equalized and saved in PNG. The datasets are split into training, validation, and test subsets.
- **Step 3** Pretrained models: Six EfficientNetV2 models including EfficientNetV2-B0, EfficientNetV2-B1, EfficientNetV2-B2, EfficientNetV2-B3, EfficientNetV2-S, and EfficientNetV2-M are used.
- **Step 4** Proposed Ensemble-stacking model: The previous six EfficientNetV2 models are stacked.
- **Step 5** Proposed model: ECA-EfficientNetV2 model.
- **Step 6** Performance evaluation: Accuracy, precision, recall, F1-score, and AUC are recorded for each model.
Step 7 Comparison with SOTA: The results of models are compared to the related studies.

Data Extraction

Chest X-ray images and CT images are collected from five Kaggle datasets described as follows. The first Kaggle dataset is COVID-19 Radiography Dataset (30), which contains 3616 COVID-19, 1345 viral pneumonia, 6012 lung opacity and 10,192 normal for a total of 21,165 chest X-ray images (PNG). The second Kaggle dataset is Chest X-Ray Images (Pneumonia) (31), collected from pediatric patients aged 1 to 5 from Guangzhou Women and Children’s Medical Center with 4273 pneumonia and 1583 normal chest X-ray images (JPEG). The third Kaggle dataset is Tuberculosis Chest X-ray Dataset (32), which contains 3500 tuberculosis and 3500 normal chest X-ray images (PNG). Large COVID-19 CT scan slice dataset (33), containing 7593 COVID-19, 2618 chest abdominal pelvis and 6893 normal CT images for a total of 17,104 CT images (PNG), is the forth Kaggle dataset used in this study. The fifth Kaggle dataset is COVID-19 and Normal and PneumoniaCT_Images (34), containing 2035 COVID-19, 3309 pneumonia and 2119 normal CT images for a total of 7463 CT images (PNG).

The COVID-19 X-ray images was formed based on the previous datasets 1–3. The pneumonia images in the second dataset were split into groups of viral pneumonia and bacterial pneumonia. The total COVID-19 X-ray images contain five groups, which are 3616 COVID-19, 2780 viral pneumonia, 2838 bacterial pneumonia, 3500 Tuberculosis and 15,275 Normal for a total of 28,009 chest X-ray images (PNG).

The COVID-19 CT images was formed based on the previous datasets 4 and 5, which contains four groups including 9628 COVID-19, 2618 Chest Abdomen Pelvis (CAP), 3309 pneumonia, and 9012 Normal for a total number of 24567 chest CT images.

Image Preprocessing

To equalize the number of images for each group, 1200 images are randomly selected for each group for both X-ray and CT datasets in this study. The number of images in training, validation and test subsets are 900, 100, and 200, respectively for X-ray dataset; The number of images in training, validation and test subsets are 4500, 500, and 400, respectively for CT dataset. Ten-fold cross validation was performed in this study. Tables 1 and 2 display the number of images for COVID-19 X-ray Dataset (Dataset X-ray).

| Group                  | Training | Validation | Test |
|------------------------|----------|------------|------|
| COVID-19               | 900      | 100        | 200  |
| Bacterial pneumonia   | 900      | 100        | 200  |
| Tuberculosis           | 900      | 100        | 200  |
| Viral pneumonia        | 900      | 100        | 200  |
| Normal                 | 900      | 100        | 200  |
| Total                  | 4500     | 500        | 1000 |
each group for COVID-19 X-ray dataset (Dataset X-ray) and COVID-19 CT dataset (Dataset CT) used in the study.

The original images collected from the previous datasets differ in size and format. As the model requirement, all the images are preprocessed to in PNG with size of $224 \times 224$. Figures 2 and 3 display the examples of images after preprocessed for chest X-ray (Dataset X-ray) and CT (Dataset CT) images, respectively.

### Convolutional Neural Network

Convolutional Neural Network (CNN) is a feedforward neural network, mainly composed of multiple convolution layers, pooling layers and fully connected layers. Compared to other neural networks, convolutional neural networks have better performances on image or speech recognitions. The goal of training a convolutional neural network is to find the most appropriate weights in the process of multiple forward and reverse iterations. Transfer learning is having previously trained model on a larger database that we can directly apply the architecture and weights of the pretrained model to various studies to speed the efficiency of the training model. The Application module in Keras currently provides about 40 pretrained models, all of which are trained on the ImageNet dataset. We choose the current newer and efficient EfficientNetV2, which includes six series of models including EfficientNetV2-B0, EfficientNetV2-B1, EfficientNetV2-B2, EfficientNetV2-B3, EfficientNetV2-S and EfficientNetV2-M. The reason for not choosing EfficientNetV2-L is that the model has a large number of parameters and cannot be executed in our existing hardware resources.

### EfficientNetV2

There are some problems with the previous EfficientNet series models, such as (1) when the size of the image is large, the training speed for EfficientNet-B3 to EfficientNet-B7 is very slow. (2) Training is slow when Depthwise convolutions is used. (3) It is not the best choice to enlarge each layer with the same magnification. Constantly increasing image size, which also leads to large memory consumption, which in turn affects training speed. Therefore, Tan and Le (35) proposed a new convolutional neural network EfficientNetV2 in 2021, which uses a nonuniform scaling strategy and can increase the number of deeper layers and has faster training speed and better parameter efficiency as compared to the previous models.

In addition, Tan and Le (35) also used EfficientNet-B4 to do some experiments, and found that replacing the MBConv module with the Fused-MBConv module in the early stage can greatly improve the training speed. However, if each

![COVID-19](image1.png)  ![Viral Pneumonia](image2.png)  ![Bacterial Pneumonia](image3.png)

**Figure 2.** Chest X-ray images after preprocessing. (Color version of figure is available online.)

| Group        | Training | Validation | Test  |
|--------------|----------|------------|-------|
| COVID-19     | 900      | 100        | 200   |
| CAP          | 900      | 100        | 200   |
| Pneumonia    | 900      | 100        | 200   |
| Normal       | 900      | 100        | 200   |
| Total        | 3600     | 400        | 800   |
layer is replaced with the Fused-MBConv module, the number of parameters and FLOPs will be significantly increased, and the training speed will be relatively reduced. Therefore, Neural Architecture Search (NAS) was used to find the best combination of MBConv and Fused-MBConv. Figure 4 is the architecture of EfficientNetV2.

ECA-Net

Channel attention has greatly improved the performance of convolutional neural networks, and most scholars are currently working on developing more complex attention modules. The most representative method such as Hu et al. (36) used Squeeze-and-excitation (SE) module and proposed SE-Net in 2017. SE-Net first uses a global average pooling layer for each channel, and then uses two nonlinear fully connected layers and a sigma function to generate the weights for each channel. Although the SE module is widely used in some researches on the current channel attention module, it has been proved that dimensionality reduction will affect both the prediction performance for channel attention, and the efficiency of obtaining the weights between all channels.

Figure 3. CT images after preprocessing. (Color version of figure is available online.)

Figure 4. Architecture of EfficientNetV2. (Color version of figure is available online.)
Wang et al. (37) uses a lightweight and efficient channel attention (ECA) module, which only adds a small number of parameters, but can achieve significant performance gains. The ECA module does not use dimensionality reduction and operates through a one-dimensional convolution of size \( k \). According to the experimental results of ECA-Net on ImageNet-1K and MS COCO, ECA-Net has lower model complexity than the state-of-the-art methods. In addition, ECA-Net has better efficiency no matter in image classification, image segmentation or object detection.

The Proposed Stacking-Ensemble Model

Ensemble learning is a type of supervised learning that has been widely used in the fields of statistics, machine learning and deep learning. Compared to a single learning algorithm, the purpose of ensemble learning is to combine multiple algorithms or models to form a model with better predictive performance. Using an ensemble approach will yield better results when there are significant model-to-model differences (38,39) reviews on ensemble deep leaning. Details of ensemble classifiers with improved overfitting have been investigated by several studies, for instance (40,41). In recent years, due to the continuous improvement of computing power of computers, large-scale integrated models can be trained within a reasonable time, and have been applied on medical image recognition, face recognition, emotion recognition and financial decision-making, etc. Successful applications of ensemble classifiers could be seen in (42–45).

The main methods of ensemble learning can be divided into three categories: bagging, boosting and stacking. The main function of stacking is to combine multiple algorithms to make predictions, that is, the result integration of the voting method or the average method. Ensemble stacking is a powerful CNN ensemble method thanks to its unique meta-learning algorithm. Meta-learning algorithms work by taking the probabilities of the input from each sub-model and determining which model performs best in extracting features. The learning algorithms directly extend the learning of each sub-model and combine the best predictions from each model. If each model is unique, then each model learns differently. Stacking achieves better results than either trained model and is used to evaluate the error rate of bagging. This study chooses stacking as our approach for ensemble learning. The stacking-ensemble model developed in this study selects six pretrained models including EfficientNetV2-B0, EfficientNetV2-B1, EfficientNetV2-B2, EfficientNetV2-B3, EfficientNetV2-S and EfficientNetV2-M in the EfficientNetV2 series. Figure 5 presents the architecture of the proposed stacking-ensemble model in this study.

![Stacking-ensemble model](image_url)

Figure 5. Stacking-ensemble model. (Color version of figure is available online.)
The Proposed ECA-EfficientNetV2 Model

In addition to the proposed stacking-ensemble model, challenging the possibility of improving accuracy on multiple chest disease diagnosis, this study proposes a self-designed model ECA-EfficientNetV2 based on ECA-Net and EfficientNetV2. Motivation of choosing the previous two models as basis are described as follows:

1. The advantages of EfficientNetV2 includes: (1) The network is better than some previous networks in terms of training speed and number of parameters; (2) It proposes improved incremental learning methods that dynamically adjust regularization (e.g., dropout, data augmentation, and mix-up) according to the size of the training images; (3) Progressive learning is used to perform well on pre-trained ImageNet, CIFAR, cars, and flowers datasets.

2. The major advantages of ECA-Net are: (1) ECA is a lightweight attention module that only contains $k$ parameters ($k \leq 9$), which can be used to improve the performance of large convolutional neural networks; (2) The ECA module uses a dimensionality-free local cross-channel interaction method that adaptively selects suitable adjacent channels to compute attention; (3) The features extracted by the SE module between different channels are relatively similar, while the features extracted by the ECA module from different channels are different, which indicates that the features extracted by the ECA module are better for classification.

Based on the previous two points, the proposed ECA-EfficientNetV2 are designed and introduced as follows:

1. ECA-EfficientNetV2 uses the dilated convolution module as the first two layers. The advantage is that while the kernel size is increased, the parameters or calculation amount of the original model can be maintained. Each module contains two dilated convolutional layers and an activation function SELU, where the dilation rate of the first layer is set to two, and the dilation rate of the second layer is set to three.

2. To reduce both the number of parameters and the computation complexity, ECA-EfficientNetV2 replaces SE module with ECA module in MBEConv and Fused-MBEConv convolution modules, and renames as MBEConv and Fused-MBEConv.

3. Zero-padding is added to the convolutional layers in both the MBEConv and Fused-MBEConv modules to prevent the input image be affected by the kernel size. In addition, the original activation function SELU is changed to ReLU to overcome the problem of vanishing gradient.

4. After MBEConv and Fused-MBEConv, two general convolution modules are added. The internal parameters include Zero-padding, Stride, and SELU. In addition, we add a batch normalization after each convolutional layer, which makes training easier and more stable, and improves the performance of the neural network.

5. Use the global average pooling layer to improve the problem of a large number of parameters that occurs in the fully connected layer. After that, add a dropout layer before the classification layer to generate multiple results by continuously updating the weights, and finally remove the outliers to avoid the problem of overfitting.

The number of parameters for ECA-EfficientNetV2 is 5,706,965, which is much less than the number of parameters (117,753,253) used in EfficientNetV2-L. Figure 6 presents the architecture of the proposed Fused-MBEConv layer and MBEConv layers, and Figure 7 shows the architecture of the proposed ECA-EfficientNetV2.

Model Evaluation Measures

The performance indices include accuracy, precision, recall, F1-Score and the AUC. The mathematical formulas of these indicators are shown in Eqs. (1–5), respectively. Accuracy is the ratio of the number of correctly classified samples to the total number of samples; Precision is the ratio of the number of true positives to the total number of elements labelled to the positive class; Recall presents the number of true positives divided by the total number of true positives and false negatives; F1-Score is a measure of precision and recall. The higher the F1-score, the better the classification performance of the model; AUC denotes the measure at distinguishing between the positive and negative classes. Higher the AUC, better the model.
\[
\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}
\]
\[
\text{Precision} = \frac{TP}{TP + FP} \tag{2}
\]
\[
\text{Recall} = \frac{TP}{TP + FN} \tag{3}
\]
\[
F1 - score = \frac{2 \sum_{i=1}^{n} \text{Recall}_i \times \text{Precision}_i}{n} \tag{4}
\]
\[
AUC = \frac{\sum_{i=1}^{n} \text{rank}_i - \frac{TP \times (1 + TP)}{TP \times TN}}{TP \times TN} \tag{5}
\]

where \( c \) is the number of classes; TP (true positive) represents the number of positive categories that are correctly classified as positive, FP (false positive) represents the number of negative categories that are incorrectly classified as positive, TN (true negative) refers to the number of negative categories that are correctly classified as negative, and FN (false negative) refers to the number of positive categories that are incorrectly classified as negative.

**RESULTS**

The equipment used in the experiment is an Intel(R) Core (TM) i9-10900F 2.81 GHz CPU, NVIDIA GeForce RTX 3070 8G GPU. The whole experiment process is performed using Python 3.8 [Python Software Foundation, Fredericksburg, Virginia, USA], which contains Keras 2.6 and Tensorflow 2.6. Ten-fold cross validation is used to evaluate the performance of each model under batch size of 16; epochs count of 30; optimizer of Adam, learning rate of 1e-5 and dropout of 0.4.

**Performance Results on Dataset X-ray**

On dataset 1, Tables 3–6 report the results of Accuracies, Precisions, Recalls, F1-Scores, and AUCs in the test sets, and the average and standard deviation for EfficientNetV2-B0, EfficientNetV2-B1, EfficientNetV2-B2, EfficientNetV2-B3,
EfficientNetV2-S, EfficientNetV2-M, the proposed stacking-ensemble model, and the proposed ECA-EfficientNetV2 models. The best performance comes from the proposed ECA-EfficientNetV2 model with the highest Accuracy of 99.21%, Precision of 99.23%, Recall of 99.25%, F1-score of 99.20%, and AUC of 99.51%. Figures 8 and 9 show the examples of training accuracy and loss, confusion matrix and ROC curve for each model, respectively.
Performance Results on Dataset CT

On dataset 2, Tables 7–10 report the results of Accuracies, Precisions, Recalls, F1-Scores, and AUCs in the test sets, and the average and standard deviation for EfficientNetV2-B0, EfficientNetV2-B1, EfficientNetV2-B2, EfficientNetV2-B3, EfficientNetV2-S, EfficientNetV2-M, the proposed stacking-ensemble model, and the proposed ECA-EfficientNetV2 models. The best performance comes from the proposed ECA-EfficientNetV2 model with the highest Accuracy of 99.81%, Precision of 99.80%, Recall of 99.80%, F1-score of 99.81%, and AUC of 99.87%. Figures 10 and 11 show the examples of training accuracy and loss, confusion matrix and ROC curve for each model, respectively.

DISCUSSION

Dataset X-ray

The performances of 5 metrics among EfficientNetV2-B0, EfficientNetV2-B1, EfficientNetV2-B2, EfficientNetV2-B3, EfficientNetV2-S, EfficientNetV2-M, the proposed stacking-ensemble model, and the proposed ECA-EfficientNetV2 models for dataset X-ray were compared in Table 11. Two accuracies are above 98%, which come from Stacking-ensemble and ECA-EfficientNetV2 models, while the accuracies from six pretrained models are inferior. In fact, the highest two performance for five metrics from Stacking-ensemble and ECA-EfficientNetV2 models are significantly different from six pretrained models (p-value < 0.01). Although the best performance, the accuracy (99.21%), precision (99.23%), recall (99.25%), F1-score (99.20%), and AUC (99.51%), comes from the proposed ECA-EfficientNetV2 model, the differences for five metrics between Stacking-ensemble and ECA-EfficientNetV2 models are not significant. Comparing the standard deviations of five metrics from Stacking-ensemble and ECA-EfficientNetV2 models, those from the latter are relatively small, which stands the stability of the model.

Figure 12 shows examples of confusion matrix from EfficientNetV2-B0, EfficientNetV2-M, Stacking-ensemble, and ECA-EfficientNetV2 models. In EfficientNetV2-B0, 58 Bacterial Pneumonia are misclassified to Tuberculosis, and 42 Tuberculosis are misclassified to Bacterial Pneumonia. In EfficientNetV2-M, 40 Bacterial Pneumonia are misclassified to Tuberculosis, and 51 Tuberculosis are misclassified to Bacterial Pneumonia. Obviously, Bacterial Pneumonia and Tuberculosis are misclassified to each other from EfficientNetV2-B0 and EfficientNetV2-M models, while the wrong phenomenon is much improved in Stacking-ensemble and ECA-EfficientNetV2 models.

Dataset CT

The performances of 5 metrics among EfficientNetV2-B0, EfficientNetV2-B1, EfficientNetV2-B2, EfficientNetV2-B3, EfficientNetV2-S, EfficientNetV2-M, the proposed stacking-ensemble model, and the proposed ECA-EfficientNetV2 models for dataset CT were compared in Table 12. The accuracies from the previous eight models are greater than 93%; the highest two are from Stacking-ensemble and ECA-EfficientNetV2 models. Actually, the same results could be found in precision, recall, F1-score and AUC. The five metrics from Stacking-ensemble and ECA-EfficientNetV2 models are statistically significant from the six EfficientNetV2 models (p-value < 0.01). The best performance, the accuracy (99.81%), precision (99.80%), recall (99.80%), F1-score (99.81%), and AUC (99.87%), comes from the proposed ECA-EfficientNetV2, which are not significantly different from the Stacking-ensemble model. Same as we found in Dataset X-ray, the standard deviations of five metrics from ECA-EfficientNetV2 model are relatively small, which stands the stability of the model.

Figure 13 shows examples of confusion matrix from EfficientNetV2-B0, EfficientNetV2-M, Stacking-ensemble, and ECA-EfficientNetV2 models for Dataset CT. In EfficientNetV2-B0, 23 COVID-19 are misclassified to Normal, and 8 Normal are misclassified to COVID-19. In EfficientNetV2-M, 18 COVID-19 are misclassified to Normal, and 12 Normal are misclassified to COVID-19. In Stacking-ensemble, 7 COVID-19 are misclassified to Normal, and 1 Normal are
misclassified to COVID-19. In ECA-EfficientNetV2, 2 COVID-19 are misclassified to Normal, and 0 Normal are misclassified to COVID-19. Obviously, COVID-19 and Normal are misclassified to each other from EfficientNetV2-B0 and EfficientNetV2-M models, while Stacking-ensemble and ECA-EfficientNetV2 models greatly improved the wrong phenomenon.

Comparison With SOTA
The results of the proposed Stacking-ensemble and ECA-EfficientNetV2 models are compared to the related studies using stacking ensemble models for COVID-19 diagnosis in Table 13. Most of the related studies worked on classification of 2 to 3 groups, while this study focuses on classification of five groups in X-ray and classification of 4 groups in CT. Most of the accuracies are greater than 90% by using stacking models, which represents the favorable performance of stacking method.

Although the accuracy from our Stacking-ensemble model is slightly inferior than (8,9,12,19,21,23,24,25,27,28). As we have mentioned before, most of those related studies worked on classification of two groups, while this study is working on the classification of four groups in X-ray and five groups in CT. Even in this situation, the accuracies of our proposed Stacking-ensemble model are pretty close to the previous studies. In addition, the performance of the proposed ECA-EfficientNetV2 model dominates most of the related studies.

Figure 8. Examples of training accuracy and loss from each model. (Color version of figure is available online.)
The accuracies for X-ray and CT from ECA-EfficientNetV2 model are close to 100%, which illustrates the great classification capability for multiple groups on chest diseases.

To verify the performance of the two proposed Stacking-ensemble and ECA-EfficientNetV2 models, one more open public dataset COVID-CT (https://github.com/UCSD-AI4H/COVID-CT) was tested and the performance metrics were displayed in Table 14. There are 746 CT images including two groups of COVID-19 and Non-COVID-19 in dataset COVID-CT. Except the accuracy of 93.33% from Shaik and Cherukuri (53), the accuracies from other related studies are lower than 90%. The accuracies are 94.86% and 95.29% from our proposed Stacking-ensemble and ECA-EfficientNetV2 models, respectively, which are higher than all of the related studies compared in Table 14. Especially, the number of parameter in ECA-EfficientNetV2 model is 5,706,965, which is much less than those in (48,50,51).

Figure 9. Examples of confusion matrix and ROC curve from each model. (Color version of figure is available online.)
This study applies six pretrained MobileNetV2 models on COVID-19 diagnosis, and stacks ensemble the previous six models to achieve better classification results. The self-designed Stacking-ensemble and ECA-MobileNetV2 models were proposed to classify multiple chest diseases including COVID-19 in this study. Classification was executed on
| Fold | Stacking-ensemble model | ECA-EfficientNetV2 |          |          |          |          |          |          |          |          |
|------|-------------------------|---------------------|----------|----------|----------|----------|----------|----------|----------|----------|
|      | Accuracy | Precision | Recall | F1-Score | AUC      | Accuracy | Precision | Recall | F1-Score | AUC      |
| 1    | 99.00%   | 99.01%    | 99.24% | 99.89%   | 99.33%   | 99.38%   | 99.34%    | 99.35% | 99.37%   | 99.58%   |
| 2    | 98.62%   | 98.67%    | 98.54% | 98.98%   | 99.08%   | 99.88%   | 99.87%    | 99.85% | 99.89%   | 99.92%   |
| 3    | 99.38%   | 99.46%    | 99.74% | 99.25%   | 98.58%   | 98.75%   | 99.72%    | 99.71% | 99.77%   | 99.83%   |
| 4    | 98.75%   | 98.74%    | 98.93% | 98.17%   |          | 99.25%   | 99.27%    | 99.24% | 99.28%   | 99.50%   |
| 5    | 98.75%   | 98.78%    | 98.79% | 99.17%   |          | 99.88%   | 99.85%    | 99.87% | 99.84%   | 99.92%   |
| 6    | 99.88%   | 99.87%    | 99.89% | 99.92%   |          | 99.94%   | 99.92%    | 99.96% | 99.95%   | 99.98%   |
| 7    | 98.25%   | 98.33%    | 98.41% | 98.83%   |          | 100.00%  | 100.00%   | 100.00% | 100.00%  | 100.00%  |
| 8    | 98.75%   | 98.77%    | 98.75% | 98.17%   |          | 100.00%  | 100.00%   | 100.00% | 100.00%  | 100.00%  |
| 9    | 98.88%   | 98.87%    | 98.85% | 98.25%   |          | 100.00%  | 100.00%   | 100.00% | 100.00%  | 100.00%  |
| 10   | 98.13%   | 98.16%    | 98.15% | 98.75%   |          | 100.00%  | 100.00%   | 100.00% | 100.00%  | 100.00%  |
| Average | 98.84±  | 98.87±    | 98.93± | 98.92±   | 99.23±   | 99.81±   | 99.80±    | 99.80± | 99.81±   | 99.87±   |

Figure 10. Examples of training accuracy and loss from each model. (Color version of figure is available online.)
five groups using X-ray images and four groups using CT images. The experimental results of two proposed models were compared to six pretrained MobileNetV2 models. The proposed Stacking-ensemble achieves accuracies of 98.71% and 98.84% on X-ray and CT datasets, respectively. With stability, the proposed ECA-MobileNetV2 achieves the highest accuracies of 99.21% and 99.81% on X-ray and CT datasets, respectively.

The proposed Stacking-ensemble model is the first study ensembles the series of EfficientNetV2 models on multiple chest diseases including COVID-19 detection, which reduce prediction variance to the training data and improves overall
classification performance when compared to any individual EfficientNetV2 model. As compared to the proposed ECA-MobileNetV2 model, the stacking model is computationally expensive; it takes around 1.5 hours in this experiment while the execution time is only 17 minutes in ECA-MobileNetV2 model. The major contribution of the study is the proposed

| Models                     | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | AUC (%) |
|----------------------------|--------------|---------------|------------|--------------|---------|
| EfficientNetV2-B0          | 85.23±1.63   | 85.35±1.61    | 85.26±1.63 | 85.23±1.65   | 90.77±1.02 |
| EfficientNetV2-B1          | 86.63±2.00   | 85.68±2.02    | 85.66±2.00 | 85.61±2.01   | 91.02±1.25 |
| EfficientNetV2-B2          | 83.31±2.59   | 83.48±2.63    | 83.34±2.59 | 83.27±2.58   | 89.58±1.62 |
| EfficientNetV2-B3          | 85.78±0.96   | 85.97±0.99    | 85.84±1.02 | 85.77±0.96   | 91.11±0.60 |
| EfficientNetV2-S           | 87.41±0.47   | 87.50±0.51    | 87.54±0.49 | 87.42±0.51   | 92.15±0.31 |
| EfficientNetV2-M           | 88.65±1.05   | 88.68±1.01    | 88.73±1.07 | 88.64±1.05   | 92.88±0.66 |
| Stacking-ensemble          | 98.71±0.26   | 98.44±1.05    | 98.75±0.26 | 98.72±0.26   | 99.20±0.16 |
| ECA-EfficientNetV2         | 99.21±0.30   | 99.23±0.30    | 99.25±0.30 | 99.20±0.31   | 99.51±0.19 |

**p-value**                    | 0.00**     | 0.00**       | 0.00**      | 0.00**      |

**p < 0.01**

Figure 12. Confusion matrix of Dataset X-ray. (Color version of figure is available online.)
**TABLE 12. Performance Comparison on Dataset CT**

| Models              | Dataset CT |          |          |          |          |
|---------------------|------------|----------|----------|----------|----------|
|                     | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | AUC (%)  |
| EfficientNetV2-B0   | 94.54(-0.84) | 94.58(-0.83) | 94.54(-0.83) | 94.50(-0.85) | 96.36(-0.56) |
| EfficientNetV2-B1   | 93.60(1.67)  | 93.66(1.74)  | 93.67(1.61)  | 93.55(1.70)  | 95.73(1.11)  |
| EfficientNetV2-B2   | 94.41(1.48)  | 94.82(2.04)  | 94.47(1.57)  | 94.35(1.49)  | 96.28(0.99)  |
| EfficientNetV2-B3   | 95.10(0.73)  | 95.15(0.72)  | 95.15(0.86)  | 95.12(0.70)  | 96.73(0.49)  |
| EfficientNetV2-S    | 96.34(0.66)  | 96.35(0.67)  | 96.35(0.66)  | 96.34(0.67)  | 97.56(0.44)  |
| EfficientNetV2-M    | 95.66(0.65)  | 95.67(0.67)  | 95.58(0.57)  | 95.65(0.66)  | 97.11(0.43)  |
| Stacking-ensemble   | 98.84(0.51)  | 98.87(0.50)  | 98.93(0.54)  | 98.92(0.61)  | 99.23(0.34)  |
| ECA-EfficientNetV2  | 99.81(0.27)  | 99.80(0.27)  | 99.80(0.28)  | 99.81(0.27)  | 99.87(0.18)  |
| p-value             | 0.00**      | 0.00**      | 0.00**      | 0.00**      | 0.00**     |

**p < 0.01**

**Figure 13.** Confusion matrix of Dataset CT. (Color version of figure is available online.)
ECA-MobileNetV2 model combines the advantages of EfficientNetV2 and ECA-Net to demonstrate superior classification performance with least variance and training time. The results of this study imply that the architecture of the proposed ECA-MobileNetV2 model can be used to assist radiologists in the X-ray and CT scans on multiple chest diseases including COVID-19.

The 6,000 X-ray and 4,800 CT images used in this study were collected from five Kaggle datasets. More chest images are encouraged to achieve more robust classification.

| No. | Study(s) | Dataset                  | Architecture                  | Class | Accuracy  |
|-----|----------|--------------------------|-------------------------------|-------|-----------|
| 1   | Ieracitano et al. (2022) | 155 X-ray images. | CovNNNet                      | 2     | 81.00%    |
| 2   | Khan et al. (2021) | 6448 X-ray images. | DHL DBHL                       | 2     | 98.53%    |
| 3   | Loey et al. (2022) | 10,848 X-ray images. | CNN+ Bayesian                  | 3     | 96.00%    |
| 4   | Hu et al. (2021) | 10,848 X-ray images. | CNN+ ELMs+ ChOA               | 2     | 99.11%    |
| 5   | Thaseen et al. (2022) | 13,808 X-ray images. | Ensemble CNN model             | 3     | 99.00%    |
| 6   | Kumar et al. (2021) | 6000 X-ray images. | HCNN                          | 3     | 98.20%    |
| 7   | Musallam et al. (2022) | 7512 X-ray images. | DeepChest                      | 3     | 96.56%    |
| 8   | Mamalakis et al. (2021) | 6000 X-ray images. | DenResCov-19                   | 4     | 99.60%    |
| 9   | Garg et al. (2022) | 4173 CT images. | EfficientNet-B5                | 2     | 98.45%    |
| 10  | Laha Sini et al. (2021) | 4986 CT images. | DenseNet201+GradCam             | 2     | 98.80%    |
| 11  | Rahimzadeh et al. (2021) | 63,849 CT images. | ResNet50V2+ FPN                | 2     | 98.49%    |
| 12  | Qi et al. (2022) | 10,000 CT images. | U-Net+ DenseNet121             | 2     | 97.10%    |
| 13  | Abdel-Basset et al. (2021) | 9593 CT images. | GR-U-Net+ CNN                  | 2     | 96.80%    |
| 14  | Ye et al. (2022) | 45,167 CT images. | CIFD-Net                       | 2     | 91.19%    |
| 15  | Balaha et al. (2021) | 15,535 CT images. | CNN+HFO+CSS+CSS+WSM            | 2     | 99.33%    |
| 16  | Ahamed et al. (2021) | 4593 X-ray images. | ResNet50V2(fine-tuning)        | 4     | 96.45%    |
|     | Kumar et al. (2021) | 3000 CT images. | 3                              |      | 99.01%    |
| 17  | Kurniati et al. (2021) | 2000 X-ray images. | VGG16                          | 2     | 98.00%    |
| 18  | Ahsan et al. (2020) | 2000 CT images. | Xception                       | 2     | 83.00%    |
| 19  | Jia et al. (2021) | 400 X-ray images. | NasNetMobile                   | 2     | 100.00%   |
|     | Kumar et al. (2021) | 400 CT images. | 2                              |      | 95.20%    |
| 20  | Kassania et al. (2021) | 7592 X-ray images. | ResNet(fine-tuning)             | 5     | 99.60%    |
|     | 20       | 104,009 CT images. | 3                              |      | 99.30%    |
| 21  | Gour and Jain (2022) | 3040 X-ray images. | Ensemble CNN                   | 3     | 97.27%    |
| 22  | Kamil et al. (2021) | 4645 CT images. | 2                              |      | 98.30%    |
| 23  | Saygili (2021) | 977 X-ray images. | VGG19(fine-tuning)              | 2     | 99.90%    |
|     | 20       | 23 CT images. | 2                              |      | 98.30%    |
| 24  | Stacking-ensemble | 1125 X-ray images. | SVM                             | 3     | 85.96%    |
|     |          | 3228 CT images. | K-ELM                          | 2     | 98.88%    |
| 25  | ECA-EfficientNetV2 | 6000 X-ray images. | Stacking-ensemble              | 5     | 98.71%    |
|     |          | 4800 CT images. | 4                              |      | 98.84%    |
|      | ECA-EfficientNetV2 | 6000 X-ray images. | ECA-EfficientNetV2             | 5     | 99.21%    |
|      |          | 4800 CT images. | 4                              |      | 99.81%    |
performance. Since the related studies compared in the study were searched from Scopus and were limited, more comprehensive review is suggested in the future research.

REFERENCES

1. C. Pneumonia, “COVID-19,” vol. 2019, 2019.
2. Liu H, Wang L, Nan Y, et al. SDFN: Segmentation-based deep fusion network for thoracic disease classification in chest X-ray images. Computerized Med Imaging and Graphics 2019; 75:66–73. doi:10.1016/j.compmedimag.2019.05.005.
3. Lin L, Qin L, Xu Z, et al. Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on Chest CT. Appl Intel 2020; 2021:1–5. [Online]. Available: http://arxiv.org/abs/2003.13865 (accessed February 01, 2022).
4. Sharma N, Sharma R, Jindal N. Machine learning and deep learning applications—a vision. In: Global Transitions Proceedings; 2021. p. 0–8. doi:10.1016/j.gtp.2021.01.004.
5. Ieracitano C, Mammone N, Versaci M, et al. A fuzzy-enhanced deep learning approach for early detection of COVID-19 pneumonia from portable chest X-ray images. Neurocomputing 2022; 481:202–215. doi:10.1016/j.neucom.2022.01.055.
6. Khan SH, Sohail A, Khan A, et al. COVID-19 detection in chest X-ray images using deep boosted hybrid learning. Comp Biol Med 2021; 137:104816. doi:10.1016/j.compbiomed.2021.104816.
7. Loey M, El-Sappagh S, Mirjalili S. Bayesian-based optimized deep learning model to detect COVID-19 patients using chest X-ray image data. Comp Biol Med 2022; 142:105213. doi:10.1016/j.compbiomed.2021.105213.
8. Hu T, Khishe M, Mohammadi M, et al. Realtime COVID-19 diagnosis from X-Ray images using deep CNN and extreme learning machines stabilized by chimp optimization algorithm. Biomed Signal Proces Contri 2021; 68:102764. doi:10.1016/j.bspc.2021.102764.
9. Sakthivel R, Sumaiya Thaseen I, Vanithac M, et al. An efficient hardware architecture based on an ensemble of deep learning models for COVID-19 prediction. Sustainable Cities and Soc Feb. 2022:103713. doi:10.1016/j.scs.2022.103713.
10. Kumar M, Shaiya D, Kunup V, et al. COVID-19 prediction through X-ray images using transfer learning-based hybrid deep learning approach. In: Materials Today: Proceedings; Dec. 2021. doi:10.1016/j.matpr.2021.12.123.
11. Musallam AS, Sherif AS, Hussein MK. Efficient framework for detecting COVID-19 and pneumonia from chest X-Ray using deep convolutional network. Egyptian Info J 2021:2022. doi:10.1016/j.aj.2022.01.002.
12. Mamalakis M, Swift AJ, Vorseelaars B, et al. DenResCov-19: A deep transfer learning network for robust automatic classification of COVID-19, pneumonia, and tuberculosis from X-rays. Computerized Med Imaging and Graphics 2021; 94:102008. doi:10.1016/j.compmediimag.2021.102008.
13. Garg A, Salehi S, la Rocca M, et al. Efficient and visualizable convolutional neural networks for COVID-19 classification using Chest CT. Expert Systems with Appl 2022; 195:116540. doi:10.1016/j.eswa.2022.116540.
14. LAHSAINI I, EL HABIB DAHO M, CHIKH MA. Deep transfer learning based classification model for covid-19 using chest CT-scans. Pattern Recog Letters 2021; 152:122–128. doi:10.1016/j.patrec.2021.08.035.
15. Rahimzadeh M, Attar A, Sakhaei SM. A fully automated deep learning-based network for detecting COVID-19 from a new and large lung CT scan dataset. Biomed Signal Proces Contri 2021; 68. doi:10.1016/j.bspc.2021.102588.
16. Qi QQ, Qi SL, Wu YN, et al. Fully automatic pipeline of convolutional neural networks and capsule networks to distinguish COVID-19 from community-acquired pneumonia via CT images. Comp in Biol Med 2022; 141. doi:10.1016/j.compbiomed.2021.105182.
17. Abdel-Basset M, Hawash H, Moustafa N, et al. Two-stage deep learning framework for discrimination between COVID-19 and community-acquired pneumonia from chest CT scans. Pattern Recog Letters 2021; 152:311–319. doi:10.1016/j.patrec.2021.10.027.
18. Ye QH, Guq Y, Ding WP, et al. Robust weakly supervised learning for COVID-19 recognition using multi-center CT images. Appl Soft Comput 2022; 116:108291. doi:10.1016/j.asoc.2021.108291.
19. Balaha HM, El-Gendy EM, Saafan MM, CovH2SD: a COVID-19 detection approach based on Harris Hawks optimization and stacked deep learning. Expert Systems with Appl 2021; 186. doi:10.1016/j.eswa.2021.115805.
20. Qi SL, Xu CW, Li C, et al. DR-MIL: deep represented multiple instance learning distinguishes COVID-19 from community-acquired pneumonia in CT images. Comp Methods and Programs in Biomed 2021; 211. doi:10.1016/j.cmpb.2021.106406.
21. Ahamed KU, Islam M, Uddin A, et al. A deep learning approach using effective preprocessing techniques to detect COVID-19 from chest CT-scan and X-ray images. Comp Biol Med 2021;105014. doi:10.1016/j.compbiomed.2021.105014.
22. Kumari S, Ranjith E, Gujjar A, et al. Comparative analysis of deep learning models for COVID-19 detection. Global Trans Proc 2021; 2(2):559–565. doi:10.1016/j.gtp.2021.08.020.
23. M. M. Ahsan, M. M. Islam, et al., “Study of different deep learning approach with explainable AI for screening patients with COVID-19 symptoms: using CT scan and chest X-ray image dataset,” 2020. doi: https://doi.org/10.3390/make2040027.

TABLE 14. Comparisons on COVID-CT Dataset

| Dataset       | Study(s)                        | Accuracy | Precision | Recall | F1-Score | Parameter |
|---------------|---------------------------------|----------|-----------|--------|----------|-----------|
| COVID-CT      | Mishra et al. (2020)            | 88.30%   | —         | —      | 86.70%   | —         |
|               | Saqib et al. (2020)             | 80.30%   | 78.20%    | 85.70% | 81.80%   | —         |
|               | He et al. (2020)                | 86.00%   | —         | —      | 85.00%   | 14,149,480|
|               | Mobiny et al. (2020)            | —        | 84.00%    | —      | —        | —         |
|               | Polisinelli et al. (2020)       | 85.00%   | 85.00%    | 87.00% | 86.00%   | 12,600,000|
|               | Yang et al. (2020)              | —        | 89.10%    | —      | —        | 25,600,000|
|               | Cruz (2021)                     | 86.00%   | —         | 89.00% | 85.00%   | —         |
|               | Shaik and Cherukuri (2021)      | 93.33%   | 93.17%    | 93.54% | 93.29%   | —         |
|               | Stacking-ensemble               | 94.86%   | 94.79%    | 94.83% | 94.84%   | —         |
|               | ECA-EfficientNetV2              | 95.29%   | 95.15%    | 95.24% | 95.27%   | 5706,965  |
24. Jia G, Lam HK, Xu Y. Classification of COVID-19 chest X-Ray and CT images using a type of dynamic CNN modification method. Computers in Biol and Med 2021; 134(April):104425. doi:10.1016/j.compbiomed.2021.104425.

25. Kassania SH, Kassania PH, Wesolowskic MJ, et al. Automatic detection of coronavirus disease (COVID-19) in X-ray And CT images: a machine learning based approach. Biocybernetics and Biomed Eng 2021; 41(3):867-879. doi:10.1016/j.bbe.2021.05.013.

26. Goer M, Jain S. Automated COVID-19 detection from X-ray and CT images with stacked ensemble convolutional neural network. Biocybernetics and Biomed Eng 2022; 42(1):27-41. doi:10.1016/j.bbe.2021.12.001.

27. Kamil MY. A deep learning framework to detect COVID-19 disease via chest X-ray and CT scan images. Int J Electrical and Comp Eng 2021; 11(1):844-850. doi:10.11591/ijece.v11i1.pp844-850.

28. Saygil: A. A new approach for computer-aided detection of coronavirus (COVID-19) from CT and X-ray images using machine learning methods. Appl Soft Computing 2021; 105:107323doi:10.1016/j.asoc.2021.107323.

29. M. Tan and Q. v. Le, “EfficientNetV2: smaller models and faster training,” 2021, [Online]. Available: http://arxiv.org/abs/2104.00298. (accessed December 01, 2021)

30. “COVID-19 Radiography Database | Kaggle.” https://www.kaggle.com/tausifurrahman/covid19-radiography-database (accessed Nov. 04, 2021).

31. “Chest X-Ray Images (Pneumonia) | Kaggle.” https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia (accessed Nov. 04, 2021).

32. “Tuberculosis (TB) chest x-ray database | Kaggle.” https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset (accessed May 06, 2022).

33. “Large COVID-19 CT scan slice dataset | Kaggle.” https://www.kaggle.com/maedemaftouni/large-covid19-ct-slice-dataset (accessed Nov. 09, 2021).

34. “COVID-19&Normal&Pneumonia_CT_Images | Kaggle.” https://www.kaggle.com/datasets/anaselmay/covid19normalpneumonia-ct-images (accessed May 08, 2022).

35. M. Tan and Q. v. Le, “EfficientNetV2: smaller models and faster training,” 2021.

36. J. Hu, L. Shen, S. Albanie, et al., “Squeeze-and-excitation networks,” 2017, [Online]. Available: http://arxiv.org/abs/1709.01507

37. Q. Wang, B. Wu, P. Zhu, et al., “ECA-Net: efficient channel attention for deep convolutional neural networks,” 2019, [Online]. Available: http://arxiv.org/abs/1910.03151 (accessed February 01, 2022).

38. Opitz D, Maclin R. Popular ensemble methods: an empirical study. J Artificial Intelligence Res 1999; 11:169–198. doi:10.1613/jair.614.

39. M. A. Ganae and M. Hu. “Ensemble deep learning: A review,” arXiv preprint arXiv:2104.02395 (2021).

40. Sollich P, Krogh A. Learning with ensembles: how over-fitting can be useful. In: Proceedings of the 8th International Conference on Neural Information Processing Systems; 1995, p. 190–196.

41. Pourtaheri ZK, Zahir SH. Ensemble classifiers with improved overfitting. In: 1st Conference on Swarm Intelligence and Evolutionary Computation; 2016doi:10.1109/CSI.EC.2016.7482130.

42. Rajaraman S, Kim I, Antani SK. Detection and visualization of abnormality in chest radiographs using modality-specific convolutional neural network ensembles. PeerJ 2020; 8:e8993. doi:10.7717/peerj.8993.

43. Devnath L, Fan Z, Luo S, et al. Detection and visualisation of pneumoconiosis using an ensemble of multi-dimensional deep features learned from Chest X-rays. Int J Environ Res and Public Health 2022; 19(18):11193. doi:10.3390/ijerph19181193.

44. Rajaraman S, Cemir S, Xue Z, et al. A novel stacked model ensemble for improved TB detection in chest radiographs. Chapter 1, Medical Imaging, Boca Raton: CRC Press, 2019.

45. Devnath L, Luo S, Summons P, et al. Deep ensemble learning for the automatic detection of pneumoconiosis in coal worker’s Chest X-ray radiography. J Clin Med 2022; 11(18):5342. doi:10.3390/jcm11185342.

46. Mishra AK, Das SK, Roy P, et al. Identifying COVID19 from Chest CT images: a deep convolutional neural networks based approach. J Healthcare Eng 2020; 2020. doi:10.1155/2020/9843654.

47. M. Saqib, S. Anwar, A. Anwar, et al., “COVID19 detection from radiographs: is deep learning able to handle the crisis?,” no. June, pp. 1–14, 2020, [Online]. Available: www.preprints.org

48. He X, Yang Y, Zhang SH, et al. Sample-efficient deep learning for COVID-19 diagnosis based on CT scans. IEEE Transactions on Med Imaging 2020; 100Xx. doi:10.1101/2020.04.13.20063941.

49. A. Mobily et al., “Radiologist-level COVID-19 detection using CT Scans with detail-oriented capsule networks,” 2020, [Online]. Available: http://arxiv.org/abs/2004.07407

50. Poliselli M, Cirne L, Piacidi G. A light CNN for detecting COVID-19 from CT scans of the chest. Pattern Recog Letters 2020; 140:95–100. doi:10.1016/j.patrec.2020.10.001.

51. X. Yang, X. He, J. Zhao, et al., “COVID-CT-Dataset: A CT scan dataset about COVID-19,” 2020, [Online]. Available: http://arxiv.org/abs/2003.13865. (accessed December 05, 2021).

52. Hernández Santa Cruz JF. An ensemble approach for multi-stage transfer learning models for COVID-19 detection from chest CT scans. Intelligence-Based Med 2021; 5(February):100027doi:10.1016/j.ijerph.2021.100027.

53. Shaik NS, Cherukuri TK. Transfer learning based novel ensemble classifier for COVID-19 detection from chest CT-scans. Computers in Biol Med 2021; 141(June 2021):105127. doi:10.1016/j.compbiomed.2021.105127.