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 deep transfer learning-based convolution neural network model for COVID-19 detection using computed tomography scan images for medical applications

Nirmala Devi Kathamuthu, Shanthi Subramaniam, Quynh Hoang Le, Suresh Muthusamy, Hitesh Panchal, Suma Christla Mary Sundararajan, Ali Jawad Alrubaie, Musaddak Maher Abdul Zahra

A Department of Computer Science and Engineering, Kongu Engineering College (Autonomous), Perundurai, Erode, Tamil Nadu, India
B Institute of Research and Development, Duy Tan University, Da Nang, Vietnam
C School of Medicine and Pharmacy, Duy Tan University, Da Nang, Vietnam
D Department of Electronics and Communication Engineering, Kongu Engineering College (Autonomous), Perundurai, Erode, Tamil Nadu, India
E Department of Mechanical Engineering, Government Engineering College, Patan, Gujarat, India
F Department of Information Technology, Panimalar Engineering College (Autonomous), Poonamallee, Chennai, Tamil Nadu, India
G Department of Medical Instrumentation Techniques Engineering, Al- Mustaqbal University College, 51001, Hilla, Iraq
H Computer Techniques Engineering Department, Al-Mustaqbal University College, Hillah 51001, Iraq
I Electrical Engineering Department, College of Engineering, University of Babylon, Hilla, Babil, Iraq

ARTICLE INFO
Keywords:
COVID-19
Deep learning
CNN
InceptionV3
ResNet-50
DenseNet121
VGG16
VGG19
Transfer learning

ABSTRACT
The Coronavirus (COVID-19) has become a critical and extreme epidemic because of its international dissemination. COVID-19 is the world’s most serious health, economic, and survival danger. This disease affects not only a single country but the entire planet due to this infectious disease. Illnesses of Covid-19 spread at a much faster rate than usual influenza cases. Because of its high transmissibility and early diagnosis, it isn’t easy to manage COVID-19. The popularly used RT-PCR method for COVID-19 disease diagnosis may provide false negatives. COVID-19 can be detected non-invasively using medical imaging procedures such as chest CT and chest x-ray. Deep learning is the most effective machine learning approach for examining a considerable quantity of chest computed tomography (CT) pictures that can significantly affect Covid-19 screening. Convolutional neural network (CNN) is one of the most popular deep learning techniques right now, and its gaining traction due to its potential to transform several spheres of human life. This research aims to develop conceptual transfer learning enhanced CNN framework models for detecting COVID-19 with CT scan images. Though with minimal datasets, these techniques were demonstrated to be effective in detecting the presence of COVID-19. This proposed research looks into several deep transfer learning-based CNN approaches for detecting the presence of COVID-19 in chest CT images. VGG16, VGG19, DenseNet121, InceptionV3, Xception, and ResNet50 are the foundation models used in this work. Each model’s performance was evaluated using a confusion matrix and various performance measures such as accuracy, recall, precision, f1-score, loss, and ROC. The VGG16 model performed much better than the other models in this study (98.00 % accuracy). Promising outcomes from experiments have revealed the merits of the proposed model for detecting and monitoring COVID-19 patients. This could help practitioners and academics create a tool to help minimal health professionals decide on the best course of therapy.

1. Introduction
The World Health Organization (WHO) obtained the 2019 Coronavirus replacement on December 31, 2019, (COVID-19). WHO declares COVID-19 a global emergency on January 30, 2020. This virus is zoonotic, thus it began in animals after humans. After animal
interaction, the virus spreads to humans. When someone exhales, coughs, sneezes, or talks, the Coronavirus can spread by respiratory droplets [1]. It’s additionally feasible that the virus unfolds from bats to different animals like snakes and pangolins and finally to people.

A recently identified coronavirus causes Coronavirus Disease (COVID-19), a communicable one. Mostly the patients diagnosed with the COVID-19 virus will have mild to moderate respiratory signs and get well without further treatment. A situation called cytokine launch syndrome or a cytokine typhoon has precipitated a couple of COVID-19 results along with liver difficulties, pneumonia, breathing failure, cardiovascular disorders, septic shock, and more. When the immune machine is activated through an infection, inflammatory proteins called cytokines are launched into the circulation, which could damage human cells and organs. People above 65, those with underlying medical disorders such as cardiovascular disease, diabetes, cancer, etc., are more likely to become critically ill. While most COVID-19 patients recover and return to normal health, some individuals may experience symptoms for weeks or even months after their acute sickness has passed. Even those who aren’t in the hospital and have a minor sickness can have symptoms that last a long time.

Several methods are used to detect Coronavirus, such as scanning molecular, and serological testing. Molecular testing is used to seek evidence of an illness still present. A cotton swab is typically used for collecting a sample from the back of the throat for testing performed with polymerase chain reaction (PCR). This test reveals the genetic material of the virus. The molecular test can identify only existing COVID-19 instances and cannot detect whether people have recovered from COVID-19.

Serological testing can identify virus-fighting antibodies. Serological tests employ blood samples to diagnose infections with little to no symptoms [2]. COVID-19 survivors had these antibodies in their blood and tissues. X-rays, CT scans, swabs, and serological testing can visualise the chest’s interior structure [3]. A painless and rapid chest CT scan is used to diagnose pneumonia.

Early detection of COVID-19 may aid in developing a treatment strategy and disease containment decisions [4]. Detecting, isolating, and caring for individuals as soon as feasible is vital for effectively handling this pandemic. Using RT-PCR to identify COVID is dangerous since the swab tests reach the throat through the nose, causing coughing and spreading virus particles into the air, putting health professionals’ lives at risk. According to researchers, CT scans are far safer than traditional swab examinations. In addition, for a COVID-positive patient, a CT scan test should be performed after the RT-PCR test. Standard imaging for detecting COVID-19, a chest CT scan, is a rapid and painless procedure. According to current research, CT testing has a sensitivity of 98 percent for COVID-19 infection, more significant than RT-PCR testing’s 71 percent.

Current research intends to develop conceptual transfer learning enhanced CNN framework models for detecting COVID-19 with CT scans by utilizing image classification and deep learning models. For unique medical disorders like COVID-19, it isn’t easy to obtain sufficiently big data from publically available corpus to perform training of deep learning models. Hence, the proposed models are enhanced with Transfer learning. Transfer learning helps to train the pre-trained models, where these pre-trained models learned weights are utilized in the current model and save the training time. This method is an alternate way of training deep learning models. It is frequently used to build customized deep-learning models to support smaller datasets [5]. These models sometimes perform better than fully trained networks, where fine-tuning helps achieve better performance [6]. This paper will show how transfer learning detects COVID-19 in CT scan images. ResNet50, DenseNet121, VGG16, VGG19, Inception V3, and Xception were the basis models employed for detection and classification. Though with minimal datasets, these techniques were demonstrated to be effective in detecting the presence of COVID-19. This could help practitioners and academics create a tool to help minimal health professionals decide on the best course of therapy.

Using image classification and deep learning models, this research aims to improve conceptual transfer learning CNN framework models for diagnosing and detecting COVID-19 with CT scans. Deep learning models can increase COVID-19’s diagnosis speed and accuracy. The suggested system reduces over-fitting with batch normalisation and dropout. False negatives during the COVID-19 outbreak could propagate the virus.

The research provides a tailored pre-trained model for identifying COVID-19 using CT images, reduces false positive and false negative findings in the modelling process, and eliminates overfitting by applying batch normalisation and regularisation. Extensive tests are done to compare the proposed models’ performance with baseline and existing models.

2. Literature review

Every area of work entails the creation of objectives, which aid in resolving the organization’s issues and difficulties. Artificial intelligence is a domain that includes both machine learning and deep learning. It entails both studying and developing intelligent things [7]. Artificial neurons are built up through data processing layers to form deep neural networks (DNNs) [8]. These networks have a deep architecture known as a ‘deep neural network,’ which comprises numerous layers that interpret data into decisions [9]. CNN is a sub-type with well-known uses represented by images and videos. While information exchange might be challenging to handle at times, CNNs have prospered and gained popularity due to remarkable advances in image and video processing [10].

Deep learning has shown enormous potential in numerous real-world applications in various fields [11,12]. Object recognition is included in all of these potential applications. Deep learning algorithms are used in this innovative approach to object recognition. Deep learning of this type has produced excellent results in recognizing photo objects [13]. CT scans of the chest have also become used in clinical practice for diagnosing lung illness in COVID-19 patients [14].

People are prone to both pneumonia and COVID-19. Pneumonia kills about 800,000 children under five every year, with around 2,200 dying daily. Pneumonia affects almost 1,400 children per 100,000 children [15]. In 2013, the most common cause of death was lower respiratory tract infection, mainly pneumonia, according to a new study. India has the largest number of pneumonia deaths (0.297 million) in the world, according to the John Hopkins Bloomberg School of Public Health, and dysentery deaths in children under five in 2015. Pneumonia was also the leading cause of death in children under five in 2015 [16]. Aside from pneumonia, COVID-19 infection rates are quite high.

COVID-19 is a very transmissible disease caused by the SARS-CoV-2 virus. It is the world’s largest pandemic since 1918, affecting more than 2.9 million persons globally. Adults over 60, including those with health difficulties, should be aware of the increased risk of SARS-CoV-2 infection. Mostly pneumonia and Coronavirus have the lungs of the human body, and they are to be considered for further research with the support of medical experts. As a result, doctors encourage patients to keep track of their oxygen intake with an oxygen analyzer so that any irregularities can be discovered and addressed as soon as feasible. CNN’s perfect for such a situation.

Thoracic X-ray imaging, CT, and MRI are diagnostic radiology procedures for lung illness that are highly successful and cost-efficient chest X-ray imaging modalities. They are also accessible in hospitals and have lower dose exposures to individuals. Even for highly qualified and experienced doctors, detecting pneumonia and COVID-19 with X-ray images is difficult, as X-ray images exhibit comparable location features for various illnesses, including lung disease. Several studies have focused on COVID-19, pneumonia detection, and analysis with various algorithms. Most healthcare applications are gaining popularity with highlighted features supported by Artificial intelligence (AI), Machine
Advances in Engineering Software 175 (2023) 103317

Learning, and Deep Learning [17–19].

In 2016, Redmon et al. [20] introduced YOLO, which can support detecting images up to forty-five frames within a fraction of a second. In the same year, Liu et al. [21] published the SSD algorithm. All SSD and YOLO win when it comes to identifying speed, but SSD uses a multi-scale feature map to distinguish between them. Deep network images’ spatial resolution has been drastically reduced, making it more difficult to discover small targets challenging to detect at low resolution, affecting the detection accuracy. YOLO does not support multi-scale feature maps for solo diagnosis. It aids in smoothing extracted features and their clipping with a lower-resolution feature map. AI’s most appealing feature is that it may be used to classify unseen images in a training set.

As a result, a rapid and effective method of detection that can distinguish between the two types of pneumonia is required. Several papers have applied deep learning in recently described COVID-19 pneumonia detection techniques. Authors have employed a deep CNN network structure to detect COVID-19 with 93.96 percent accuracy, 99.13 percent precision, and 94 percent sensitivity, respectively. A tiny dataset with 339 instances are used for building the model based on deep transfer learning based ResNet50 and obtained a well-performed model with an accuracy of 96.2 percent. Wang [22] proposed five pre-trained models, and the eXception model performs the best effect with 96.75 percent accuracy. 1102 chest X-ray images contain COVID-19 and non-COVID-19 cases, where this dataset is appropriately divided as training and testing for building the model.

COVID-19 lung infection has been identified with chest CT images by applying segmentation. Based on a 3D U-Net architecture, the suggested method can easily detect irregular regions with poor contrast between lesions with an accuracy of 0.98 and a precision of 0.73. Segmentation is vital in identifying COVID-19 along with the 3D deep learning model. The authors of [23] created and tested the COVID-SegNet, segmentation method on CT images of the chest. The proposed network would comprise ASNs with characteristics that help identify COVID-19 limits and placements. The MultiResUNet model was introduced by [23] to slice COVID-19 from CT images.

COVID-19 was detected using CNN in a chest X-ray by the authors of [4]. COVID-19 detection has been done using a variety of approaches. Most investigations, including chest X-rays, have shown a 90–94 percent accuracy rate. The major goal of this study is to use the pre-trained models and their weight for proposed models, where the top layer is customized and learned details of pre-trained models are transformed into the proposed model. The customized new models are utilized to detect COVID-19 with CT images. Using residual networks and multiple instance learning, Xiao et al. [24] developed and validated a deep learning-based model (ResNet34). The binary classification model proposed by Panwar et al. [25] can detect COVID-19. The input pictures were classified using a fine-tuned VGG model. Grad-CAM technology is applied to represent color visualization, making the suggested deep-learning model more explainable. To identify patients with COVID-19, Jaiswal et al. [26] used a deep transfer learning (DTL) based model. The proposed model collected features from the ImageNet dataset using its training weights and a CNN. To detect the Coronavirus-infected person using chest CT radiograph digital images, Loey et al. [13] used five deep CNN such as AlexNet, VGGNet19, VGGNet16, ResNet50, and GoogleNet. To aid in identifying COVID-19, the authors employed CGAN and standard data augmentations to create extra pictures. Models robustness was tested with CT scans of community-acquired pneumonia (CAP), and other than pneumonia were used.

Convolutional networks often perform better on large-scale datasets; transfer learning is one technique to address this shortcoming in a small-scale dataset. The latter entails training a model with a large dataset for one job, then fine-tuning it with a small dataset for another [27].

Data-augmentation Accuracy of COVID-19 patient detection using chest X-rays is optimistic, leading to the development of deep learning methods [28].

In order to detect COVID-19 using CT images, the research work proposed by [29] has an effective image classification method utilizing EfficientNet with preprocessing to enhance the quality of the images. On Efficient Net, a number of transformation methods, including Wavelet, Laplace, Adaptive Gamma, and CLAHE, with outstanding performance were achieved.

Various deep transfer models and fusion classifiers, as well as deep feature concatenation techniques, were proposed by the authors [30–32] for COVID-19 detection. Moreover, the predicting power of an Artificial Neural Network (ANN) is utilized to model the thermal conductivity and viscosity of the nanofluid with experimental data [33]. Furthermore, the Group Method of Data Handling-type neural network (GMDH type NN) can support a stream of data in different layers to build optimized networks [34].

Table 1. The parameter used in the various references consists of small [35,36], medium [13,29,37–40] and large [27,29,41,42] data samples shown in Table 1. The parameter used in the various references consists of Accuracy, loss, precision, recall/sensitivity, f1-score, and AUC. Authors in [37] proposed that VGG11, ResNet18, ResNet50, and ResNet50 performed well, with the highest Accuracy of 95.98% and a lower loss of 0.57%. However, ResNet-50 has the lowest Accuracy of 60 in [38], and VGG-19 performed with 94.5% accuracy. DNN and BiLSTM and transfer learning models were proposed by [39], and BiLSTM outperformed them. Synthetic augmentation also plays an important role in diagnosing diseases [28]. Various flavors of DenseNet was proposed by many research works [27,29,35–39,43].

Learning in the field of medical image analysis has received a lot of attention. However, knowledge is still required, and the analysis incorporates a range of algorithms to improve, speed up, and produce an accurate diagnosis. When compared to prior approaches, deep learning algorithms have shown to be more efficient and precise in detecting pneumonia, COVID-19, and other lung diseases. Currently, doctors in hospitals may encounter patients who have pneumonia caused by the flu, others, and COVID-19 all at the same time. Segmentation is required to analyze the images and detect disease when working with lung images, COVID-19 identification is a challenging task that frequently necessitates looking at clinical pictures of patients. Machine learning [43]

The primary novel ideas presented in this paper include the following.

This study aims to minimize both the false-positive and false-negative rates. Reducing false positive and false negative results in the modeling process is critically important in medical research, especially for severe disorders like COVID-19. False negatives should be kept to a minimum for obvious reasons to prevent misclassifying any COVID-19 patients as non-COVID-19, which might seriously affect our society. Additionally, it’s crucial to reduce the number of false positives because misclassifying a non-COVID-19 as COVID-19 could result in unnecessary testing.

In order to address the domain shift problem, the transfer learning
approach is used in this proposed study. Using a pre-trained network’s expertise, the proposed models can reduce the number of training images and boost training speed. The excellent result of these pre-trained deep models can significantly increase COVID-19’s diagnosis speed and accuracy. Over-fitting is minimized by batch normalization and dropout in the proposed system.

Pre-trained models’ capacity for recognition has been demonstrated to influence the quantity of frozen layers. Using its trained parameters on the ImageNet dataset and a convolutional neural structure, the proposed framework is being used to extract features. Numerous tests are run to determine how well the proposed DTL method fits on COVID-19 CT scan images.

3. Proposed system

Deep learning is an emerging field with several successful applications in recent years. Researchers can use it in many domains. Artificial intelligence problems have become considerably easier to forecast and work on thanks to deep learning. Everyone nowadays favors CNN over other layers for image classification because of its high accuracy. Moreover, it has a strong potential for extracting features from images that aid in image classification. Most importantly, CNN has been widely utilized in image analysis since its inception. Recent advances in deep learning, particularly in the medical imaging domain, suggest that various Deep CNN architectures could be used. To begin with, such individual baseline models are thoroughly analyzed in this study. VGG16, VGG19, InceptionV3, ResNet50, DenseNet121, and Xception are some baseline models. The convolution components of all of these baseline models proposed in the ImageNet challenge are retained precisely the same as the standard models in this study, Chest CT images are given as input to the model. Images with different sizes are resized with uniform dimensions.

3.1. Dataset description

Images from CT scans of the lungs are included in the dataset. A CT scan is a type of X-ray used to diagnose sensitive inner organs precisely. The dataset consists of 2481 images obtained from Kaggle [44]. The data is separated into two categories: COVID and non-COVID. COVID patients’ CT images are included in the COVID category, while healthy people are included in the non-COVID category. The COVID class has 1,252 CT images, while the non-COVID class contains 1,229 CT images. 80 percent of lung CT scans are used to train the model, and 20 percent are used to test it. Table 2 presents these details. The proposed architecture is displayed in Fig. 1.

3.2. Preprocessing

Image pre-processing is the first and most crucial step when working with image data. Deep learning is used in this proposed work to create a binary classifier for Chest CT scans and predict the presence of COVID-19. Because the input images are of different sizes, they are all resized to $128 \times 128 \times 3$ to maintain uniformity. Artifacts in CT scans, such as beam hardening, noise, and scatter, diminish the model’s accuracy. The images were converted from BGR to RGB color formats, pre-processed, and scaled to 49152 dimensions ($128, 128, 3$). This is highly relevant when using the OpenCV library’s `imread()` method, because its color order is BGR, whereas Pillow thinks the colors are ordered as RGB.

Table 2

| Dataset | COVID  | non-COVID |
|---------|--------|-----------|
| Training | 1001   | 983       |
| Testing  | 251    | 246       |
| Total    | 1252   | 1229      |

Fig. 1. Proposed System Architecture.
photos were transformed into pixel matrices and normalized to a range of 0 to 1. After pre-processing, labels were assigned to each image based on the directory of origin.

Data augmentation helps the model to know a wider range of features while also increasing the dataset size, which prevents the model from overfitting. A randomly performed affine transformation; randomly performed flipping. The hue, brightness, and saturation values are randomly adjusted to each training image. Shearing and rotation make up the random affine transformation. Rotation is to be performed in the range of 0 to 30. The Shearing and zooming range can be 0.2, and the brightness level can be adjusted in the range of 0.75 to 1.25 are some image augmentation parameters.

### 3.3. Pre-trained models and transfer learning

More data is needed to train a neural network if we train the model from scratch. A tiny training dataset was applied with Transfer learning to generate an accurate and comprehensive feature set to manage. Since most of the COVID-19 dataset is considerably smaller and needs more time to build an efficient model. Therefore, the proposed models are created from the pre-trained models and customized to detect COVID-19 cases.

#### 3.3.1. Densenet121

Huang et al. [45] created the DenseNet design, which builds on ResNet by introducing dense connections, representing that each layer has a connection with all layers. Naturally, this highly linked architecture ensures that every layer receives network parameters from the layers above it and passes on its feature map to the layers below it. Another significant benefit of this structure is the possibility of reusing features while keeping the overall number of parameters minimal. Several DenseNet architectural versions that are frequently used, including the DenseNet121 design structure, are used in this study.

#### 3.3.2. Resnet50

The essential principle of ResNet topologies is described in He et al. work’s [46]. Convolutional and pooling layers were added in a stacking manner, with one layer above on top of another layer. This might cause network performance to decline due to the vanishing gradient problem. Therefore, shortcut networks can deal with this, and a residual block is used. The use of skip connections effectively eliminates the significant training error that is common in otherwise deep architectures. ResNet50 is a kind of ResNet architecture that consists of 50 layers.

#### 3.3.3. InceptionV3

The basic idea behind the InceptionV3 architectures is to address the issue of excessive changeability in the position of prominent sections in the images under examination by allowing the network to incorporate numerous kernel types on the same level, effectively ‘widening’ the network [47]. The Inception modules make it possible to have many kernels running simultaneously. The initial InceptionV1 was proposed with this essential concept. Inception V3 is an extended version of
Inception V2. This will address some critical concerns about the representational bottleneck [48]. It also deals with auxiliary classifiers with kernel factorization and batch normalization.

3.3.4. Xception

Depth-wise separable convolution layers are used in the Xception network. Xception network consists of mapping spatial and cross-channel correlations, and this can be completely decoupled in CNN feature maps. Inception’s fundamental design has survived Xception. The Xception model’s 36 convolution layers can be divided into 14 discrete modules. Each layer has a continuous residual link surrounding it after the first and last layers are eliminated. The input image is transformed into spatial correlations within every output channel to acquire cross-channel correlations.

3.3.5. VGG16 and VGG19

The VGG Net design, introduced by Simonyan and Zisserman [49] from the University of Oxford’s Visual Geometry Group, is one of the most familiar Deep CNN models, having won first and second place in the ILSVRC 2014 image classification tasks. To obtain Computer Vision tasks more accurately, the models are designed to include increased layers of CNN with few kernels. VGG architecture variations have been

![Fig. 3. Results of Proposed Models Accuracy Graph.](image-url)
widely applied for extracting deep image characteristics for further processing in many Computer Vision jobs, particularly in the medical domain. VGG 16 has sixteen layers, whereas VGG19 consists of nineteen layers.

3.4. Customized proposed model building

Following pre-processing, about 20% of the dataset was set aside for testing, leaving the remaining 80% for training. CNN is one of the deep neural network structures and is a big step forward in image identification. They’re most usually employed to examine visual imagery and are frequently involved in image classification behind the scenes. CNN was trained using Keras, a Python package with a TensorFlow backend, a deep learning framework. Because the Imagenet data set is large enough (1.2 million images) to generate a general model, it has been frequently used to build several architectures. To generalize outside of the ImageNet dataset, we apply transfer learning. This only happens when the model has been pre-trained. We also employ a fine-tuning framework to

Fig. 4. Results of Proposed Models Loss Graph.
capture changes to a pre-trained model.

Our model was built around the various pre-trained models, including Densenet121, InceptionV3, Xception Resnet50, VGG16, and VGG19. For instance, the VGG19 CNN with a 2D Max Pooling output layer of shape (7, 7, 512) and no parameters can be imported using the Keras applications class. Because we only have two classification classes, Covid19 and Normal (No-Covid19), transfer learning was utilized to modify the output layer to a binary classifier. The outputs from the previous layer were retrieved and passed to a Flatten layer as a parameter. Since this layer comes after a 2D Max Pooling output layer with the shape (7, 7, 512), our flattened array has the dimensions $7 \times 7 \times 512 = 25088$. As rectangular or cubic shapes cannot be used as direct inputs, this is usually done towards the end of the CNN. A Dropout layer is added as a regularization strategy to reduce overfitting and enhance

Fig. 5. Results of a Proposed Models Confusion matrix.
Fig. 6. Results of Proposed Models ROC - AUC.

Table 4
Comparison Results of COVID-19 class.

| Metrics    | DenseNet121 | InceptionV3 | ResNet50 | Xception | VGG16 | VGG19 |
|------------|-------------|-------------|----------|----------|-------|-------|
| Precision  | 98.37       | 94.38       | 80.62    | 96.64    | 98.39 | 96.41 |
| Recall     | 96.41       | 93.63       | 82.87    | 91.63    | 97.61 | 96.41 |
| F1-score   | 97.38       | 94.00       | 81.73    | 94.07    | 98.00 | 96.41 |
| Support    | 251         | 251         | 251      | 251      | 251   | 251   |
| ROC / AUC  | 97.48       | 92.11       | 89.71    | 94.39    | 97.86 | 97.16 |
n. d. Kathamuthu et al. Advances in Engineering Software 175 (2023) 103317

generalization error. So we want to make a binary classifier, and the likelihood of the layer’s outputs being dropped out was set to 0.5, which is a common value. This is based on the idea that extensive neural networks on short datasets tend to overfit the training data, decreasing validation accuracy. The customized proposed model structure is represented in Fig. 2.

Table 5
Comparison Results of non-COVID-19 class.

| Metrics      | Densenet121 | InceptionV3 | Resnet50 | Xception | VGG16 | VGG19 |
|--------------|-------------|-------------|----------|----------|-------|-------|
| Precision    | 96.41       | 93.55       | 82.01    | 91.89    | 97.58 | 96.34 |
| Recall       | 98.37       | 94.31       | 79.67    | 96.75    | 98.37 | 96.34 |
| F1-score     | 97.38       | 93.93       | 80.82    | 94.26    | 97.98 | 96.34 |
| Support      | 246         | 246         | 246      | 246      | 246   | 246   |
| ROC / AUC    | 97.48       | 92.11       | 89.11    | 94.51    | 97.93 | 97.30 |

Table 6
Comparison of Training and Validation Performance.

| Metrics      | Densenet121 | InceptionV3 | Resnet50 | Xception | VGG16 | VGG19 |
|--------------|-------------|-------------|----------|----------|-------|-------|
| Training Accuracy | 97.98       | 94.41       | 81.70    | 93.09    | 97.63 | 96.88 |
| Testing Accuracy  | 0.08        | 0.17        | 0.43     | 0.18     | 0.09  | 0.12  |
| Training Loss    | 97.38       | 93.96       | 81.29    | 94.16    | 98.00 | 96.38 |
| Testing Loss     | 0.10        | 0.17        | 0.17     | 0.17     | 0.10  | 0.14  |

4. Results and discussions

This proposed work evaluates the models under consideration using Precision, Recall, F1-score, and AUC / ROC curve. These markers are helpful when evaluating a medical screening system for COVID-19 detection.

4.1. Performance metrics

4.1.1. Confusion matrix

It is assumed that the confusion matrix will produce four values: true positive (TP), true negative (TN), false positive (FP), and false negative (FN) (FN). TP means “properly predicted.” COVID-19 cases, TN indicates accurately predicted normal cases, FN indicates COVID-19 cases wrongly categorised as normal cases, and FP indicates normal cases incorrectly classified as COVID-19 by the proposed model.

4.1.2. Precision

Precision is a metric that measures how robust enough the model recognizes the positive samples and is specific to the expected outcome. The better value of precision shows, the more precise the positive sample prediction. Eq. (1) represents the precision formula for calculation.

\[
\text{Precision} = \frac{TP}{TP + FP} \tag{1}
\]

4.1.3. Recall

Improved accuracy in predicting the target instance and decreased likelihood of missing a bad instance are both the result of a higher recall rate. The method to find recall is expressed in Eq. (2).

\[
\text{Recall} = \frac{TP}{TP + FN} \tag{2}
\]

4.1.4. Accuracy

Accuracy is defined as the proportion of correct predictions made relative to the total number of samples. Eq. (3) is used for calculating the Accuracy of the various models.

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}
\]

4.1.5. F1-Measure

The F1 score is offered as a composite metric to mitigate the negative effects of both precision and recall when evaluating classifiers. F1-score
Table 7
Comparison of proposed models with state-of-the-art models.

| Reference | Model A | Accuracy | Loss | Precision | Recall / Sensitivity | F1 Score | AUC |
|-----------|---------|----------|------|-----------|----------------------|----------|-----|
| [37]      | VGG11 - BN | 92.73 | 0.94 | - | - | - | - |
|           | Resnet18 | 93.39 | 0.86 | - | - | - | - |
|           | Resnet50 | 95.98 | 0.57 | - | - | - | - |
| [38]      | Cnet-10  | 82.1  | - | - | - | - | - |
|           | VGG19    | 94.5  | - | - | - | - | - |
|           | Densenet169 | 93.15 | - | - | - | - | - |
|           | Inception V3 | 53.4  | - | - | - | - | - |
|           | VGG16    | 89    | - | - | - | - | - |
|           | Resnet50 | 60    | - | - | - | - | - |
| [39]      | VGG16 – DNN | 91.84 | - | 89.21 | 69.6 | - | - |
|           | VGG16 – BiLSTM | 95.68 | - | 91.02 | 88 | - | - |
|           | Resnet-50 – DNN | 92.64 | - | 90.93 | 70.4 | - | - |
|           | Resnet-50 – BiLSTM | 93.28 | - | 89.58 | 76 | - | - |
|           | Densenet-121 – DNN | 88.32 | - | 88.89 | 50.4 | - | - |
|           | Densenet-121 – BiLSTM | 89.76 | - | 86.27 | 60 | - | - |
|           | Concatenated Deep -DNN | 95.84 | - | 98.59 | 82.4 | - | - |
|           | Concatenated Deep – BiLSTM | 97.6 | - | 96.86 | 91.2 | - | - |
|           | CNN-AD    | 85    | - | 87 | 82 | 84 | - |
|           | CNN-SA    | 95    | - | 95 | 94 | 94 | - |
| [27]      | InceptionV3 | 84.51 | - | - | - | - | - |
|           | Resnet     | 60    | - | - | - | - | - |
|           | New-Densnet | 95.98 | - | - | - | - | - |
|           | VGG16    | 94.37 | 91.49 | 96.41 | 93.89 | 94.96 | - |
|           | VGG19    | 96.37 | 96.96 | 95.96 | 95.96 | 96.33 | - |
|           | Xception | 91.95 | 91.4 | 90.58 | 91.00 | 91.82 | - |
|           | Inception, V2, Resnet | 94.57 | 93.36 | 94.62 | 93.98 | 91.82 | - |
|           | Densenet121 | 95.77 | 91.73 | 99.55 | 95.48 | 96.12 | - |
|           | Densenet201 | 97.38 | 95.26 | 99.1 | 97.14 | 97.54 | - |
| [41]      | VGG16    | 96.18 | 96.16 | 96.2 | 96.17 | 96.2 | - |
|           | VGG19    | 94.77 | 94.95 | 94.66 | 94.75 | 94.66 | - |
|           | Resnet50 | 92.15 | 92.13 | 92.12 | 92.14 | 92.18 | - |
|           | Resnet50 V2 | 97.79 | 97.77 | 97.84 | 97.78 | 97.84 | - |
|           | Resnet50 V3 | 91.55 | 91.62 | 91.67 | 91.55 | 91.67 | - |
|           | Xception | 94.77 | 94.75 | 94.83 | 94.77 | 94.83 | - |
|           | MobileNet | 97.38 | 97.41 | 97.35 | 97.38 | 97.35 | - |
|           | InceptionResnet V2 | 91.35 | 91.33 | 91.39 | 91.34 | 91.39 | - |
| [40]      | ANN      | 85.09 | 82.59 | 87.69 | 82.59 | 87.69 | - |
|           | ANFIS    | 88.11 | 88.14 | 88.47 | 88.14 | 88.47 | - |
|           | CNN      | 87.36 | 87.4 | 87.73 | 87.36 | 87.73 | - |
|           | DTL      | 90.75 | 92.59 | 89.60 | 90.75 | 92.59 | - |
|           | DTL-Proposed | 93.02 | 95.18 | 91.46 | 93.02 | 95.18 | - |
| [42]      | VGG16    | 89    | - | - | - | - | - |
|           | Densenet169 | 93.15 | - | - | - | - | - |
|           | Inception V3 | 53.4  | - | - | - | - | - |
|           | Inception Resnet | 90.90 | - | - | - | - | - |
|           | Resnet 50 | 60    | - | - | - | - | - |
|           | AlexNet  | 82    | - | - | - | - | - |
|           | Laplace transform | 75.9  | - | 68 | 72 | 64 | - |
|           | Adaptive gamma correction | 90.94 | - | 91 | 91 | 92 | - |
|           | Wavelet transform | 92.55 | - | 93 | 93 | 93 | - |
|           | CLAHE transform | 94.56 | - | 95 | 91 | 93 | - |
| [29]      | VGG16    | 86    | - | 86 | 86 | 86 | - |
|           | Densnet  | 88    | - | 88 | 88 | 88 | - |
|           | Resnet110 | 88    | - | 88 | 88 | 88 | - |
|           | Spike train-based | 76    | - | 76 | 76 | 76 | - |
| [41]      | AlexNet  | 76.38 | - | - | - | - | - |
|           | VGG16    | 78.89 | - | - | - | - | - |
|           | VGG19    | 73.87 | - | - | - | - | - |
|           | GoogleNet | 77.39 | - | - | - | - | - |
|           | Resnet50 | 81.41 | - | - | - | - | - |
| [42]      | VGG16    | 76    | - | - | - | - | - |
|           | Resnet18 | 74    | - | - | - | - | - |
|           | Resnet50 | 80    | - | - | - | - | - |
|           | Densenet121 | 79    | - | - | - | - | - |
|           | Densenet-169 | 83    | - | - | - | - | - |
|           | Efficientnet-b0 | 77    | - | - | - | - | - |
|           | Efficientnet-b1 | 79    | - | - | - | - | - |
|           | CRNet   | 73    | - | - | - | - | - |
| [29]      | VGG16    | 94.37 | 91.49 | 96.41 | 93.89 | 94.96 | - |
|           | VGG19    | 96.37 | 95.96 | 95.96 | 95.96 | 96.33 | - |
|           | Xception | 91.95 | 91.40 | 90.58 | 91.00 | 91.82 | - |
|           | Inception, V2, Resnet | 94.57 | 93.36 | 94.62 | 93.98 | 91.82 | - |
|           | DenseNet121 | 95.77 | 91.73 | 99.55 | 95.48 | 96.12 | - |
|           | DenseNet201 | 97.38 | 95.26 | 99.10 | 97.14 | 97.54 | - |

(continued on next page)
Table 8 (continued)

| Reference | Model | Accuracy | Loss | Precision | Recall / Sensitivity | F1 Score | AUC |
|-----------|-------|----------|------|-----------|-----------------------|----------|-----|
| (43)      | DenseNet201 + Grad-Cam | 98.18 | 97.76 | 98.20 | 97.98 | 98.82 |
| VGG19     | 85.75 | 83.89 | 88.5 | 86.13 |
| ResNet101 | 86.25 | 93.41 | 78 | 86.25 |
| DenseNet169 | 88 | 92.70 | 82.5 | 88 |
| WideResNet50 2 | 90.75 | 92.23 | 89 | 90.59 |
| Stacked model | 90.75 | 91.37 | 90 | 90.68 |
| Proposed Models | | | | | |
| DenseNet121 | 97.38 | 0.1 | 97.39 | 97.39 | 97.38 | 97.48 |
| Inception V3 | 93.96 | 0.17 | 93.97 | 93.97 | 93.97 | 92.11 |
| ResNet50 | 81.29 | 0.43 | 81.32 | 81.27 | 81.28 | 89.41 |
| Xception | 94.16 | 0.17 | 94.27 | 94.19 | 94.17 | 94.45 |
| VGG16 | 98 | 0.1 | 97.99 | 97.99 | 97.99 | 97.90 |
| VGG19 | 96.38 | 0.14 | 96.38 | 96.38 | 96.38 | 97.23 |

Table 8

Accuracy of Proposed models with state-of-the-art models.

| Models | Proposed Accuracy | State-of-the-art models - Accuracy |
|--------|-------------------|-----------------------------------|
| DenseNet121 | 97.38 | (39) – 89.76 (27) – 95.98 (35) – 88 (36) – 79 (29) – 95.77 |
| Inception V3 | 93.96 | (27) – 84.51 (42) – 53.4 (49) – 60 (36) – 80 |
| ResNet50 | 81.29 | (49) – 60 (36) – 80 |
| Xception | 94.16 | (29) – 91.5 (29) – 91.95 |
| VGG16 | 98 | (38) – 89 (39) – 95.68 (29) – 94.37 (41) – 96.18 (42) – 89 (35) – 86 (13) – 78.89 (36) – 76 (29) – 94.37 |
| VGG19 | 96.38 | (34) – 94.5 (38) – 96.37 (39) – 94.77 (41) – 73.87 (43) – 96.37 (44) – 85.75 |

is calculated using Eq. (4).

\[
F_1 \text{- score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
\]

4.1.6. ROC Curve

It is not possible to have an AUC (Area under Curve) greater than 1, which is the area under the ROC (Receiver Operating Characteristics) curve. It is common practice to evaluate binary classifiers using the ROC curve and AUC.

4.2. Experimental results and analysis

4.2.1. Accuracy and loss graphs

The loss and accuracy graphs are the greatest approach to visualize our model training, as seen from the figure’s model accuracy plot below, it was capable of reaching identical validation and training accuracies after a few epochs, based on the F-1 Score. Time series graphs exhibit analyses of the suggested model’s loss and accuracy as a function of epoch.

4.2.2. ROC curve

The ROC curve displays our binary classifier’s diagnostic performance as the threshold varies. True positive rate (TPR) is plotted against false positive rate (FPR) at various threshold values to create the ROC curve. Recall sensitivity is TPR. False alarm probability (FPR) is calculated as (1 - specificity). ROC plots the fall-out function with sensitivity or recall. A good classifier has high recall and low false-positive rate. Fig. 4 shows that our model’s true-positive rate is over 0.8 and false-positive rate is less than 0.1, indicating an acceptable match.

4.2.3. Confusion matrix

The model’s performance was evaluated using a confusion matrix. Most photos are categorized as true positives and true negatives. Additionally, COVID-19 has a lower false negative rate than false positive rate. False positives and patient quarantine are preferable to misidentifying a positive case as unfavorable and allowing disease transmission.

The obtained accuracy results of the proposed models presented in Fig. 3 show the highest performance of VGG16 among various models proposed and the lowest performance of Resnet50. Other models obtained moderate results.

The obtained loss performance results of the proposed models are represented in Fig. 4, which shows the best results for VGG16 and moderated results for others except Resnet50.

The epoch vs. loss graph displays each epoch’s loss. Fig. 3 shows how loss values decrease as epochs increase. The obtained loss performance result of Resnet50 has the highest value. The lowest value for VGG16 shows the best performance of VGG16 among other models. The evidence from the loss graph shows the moderate performance of the models such as DenseNet12, Xception, VGG19, and InceptionV3.

Fig. 5 depicts the performance outcomes of the suggested model using the confusion matrix. It shows the high-performance results for VGG16 with a low misclassification rate. The FP (1.12%) and FN (0.80%) values of VGG16 yield the best prediction result for that model. Among the proposed prediction models, the Resnet50 has FP (8.65%) and FN (10.06%) as the highest misclassification, lowering that model’s prediction results. Fig. 6 depicts the acquired ROC findings of the proposed models. The results specifically highlight the superior performance of the VGG16 model in comparison to other models.

Fig. 5 depicts the outcomes of the proposed models, demonstrating the superior performance of VGG16 in comparison to other models. The proposed models’ performance indicators are shown in Tables 4 and 5. The proposed models’ training and validation Accuracy and loss are shown in Table 6. VGG16 had the best results of all the presented models in terms of Precision, Recall, F1-score, Accuracy, and ROC/AUC.

The results of contrasting the proposed models with the state-of-the-art models are presented in Table 7. Many different models were utilised, and their efficacy was measured in different ways. Each of the aforementioned models has its own method for determining if a CT scan image is positive or negative for COVID-19. Table 8 compares the proposed models to the state-of-the-art models.

The development of a customized model using DenseNet201 and the GradCam algorithm addressed by the authors [29] obtained the highest accuracy at 98.18%. The dataset used in the above article was imbalanced, consisting of 4986 CT images, including 1868 images of
COVID-19 confirmed and 3118 images of other lung diseases. The proposed models of this research article used the balanced dataset of 2481, including 1252 COVID-19 images and 1229 non-COVID-19 images. VGG16 performed well on all metrics, including Precision, Recall, F1-score, Accuracy, and ROC/AUC. Comparing our models to those in [29], DenseNet 121, Xception, VGG16, and VGG19 perform better. DenseNet 121 in [29] had 95.77% accuracy, but the proposed model has 97.38%. The suggested article’s xception model had 94.16% accuracy, while [43]’s had 91.95%. VGG16 and VGG19 models also performed well with the highest accuracy, 98% and 96.38%, respectively, than the models in [43].

Overall, VGG16 performed better than the other models tested. The next highest findings are supported by DenseNet121 and Resnet50, with the lowest results. The aforementioned findings demonstrate the potential utility of Deep transfer learning augmented CNN-based prediction models in the detection of COVID-19 in CT scans. The performance of Deep CNN models can be improved by using transformation-based bespoke models to levels above 90% in all performance criteria. With such a small FPR, the proposed method is practical for use in real-world screening settings.

5. Conclusion

The proposed study compares different CNN-based image classification methods for detecting COVID-19 in chest CT scan images using deep transformation enhancement. Conclusion of the present work is as below:

- Extensive testing shows that the proposed method yields excellent results, with accuracy rates of 90%+ across the board and a small percentage of false positives.
- According to the experimental findings, deep CNN-based techniques can significantly influence COVID-19 spread control by offering quick screening.
- With DL-based approaches already being widely employed in other medical imaging applications, it is high time they be implemented in the screening process for the COVID-19.
- According to analysis, VGG16 produces superior results because it contains smaller parameters and requires minimum training time. As a result, it outperforms other CNN models.
- This paper applied a ubiquitous and effective deep learning-based transformation to detect COVID-19 in suspected patients using CT scan images.

Declaration of Competing Interest

Current paper has no conflict of interests.

Data availability

Data will be made available on request.

References

[1] Tomar A, Gupta N. Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. Sci Total Environ 2020;728:138762.
[2] Kushwaha S, Bahl S, Bagha AK, Parmar KS, Javaid M, Haleem A, Singh RP. Significant applications of machine learning for COVID-19 pandemic. J Integr Manage 2020;5(9):453-79.
[3] Jiang Y, Guo D, Li C, Chen T, Li R. High-resolution CT features of the COVID-19 infection in Nanchong City: Initial and follow-up changes among different clinical types. Radiol infect dis 2020;7(2):71–7.
[4] Chadda A, Hassan I, Desrosiers C. Deep CNN models for predicting COVID-19 in CT and x-ray images. J med imaging 2021;8(3):014502.
[5] Arora V, Ng EYK, Leekha RS, Darshan M, Singh A. Transfer learning-based approach for detecting COVID-19 ailment in lung CT scan. Comput Biol Med 2021;135:104575.
training algorithms supported by sensitivity analysis: measure thermal conductivity and viscosity for liquid paraffin based nanofluid containing Al2O3 nanoparticles. J. Molecular Liquids 2019;276:850–60.

[34] Shahsavari A, Khanmohammadi S, Karimpour A, Goodarzi M. A novel comprehensive experimental study concerned synthesizes and prepare liquid paraffin-Fe3O4 mixture to develop models for both thermal conductivity & viscosity: a new approach of GMDH type of neural network. Int. J. Heat and Mass Transfer 2019;131:432–41.

[35] Ahuja S, Panigrahi BK, Dey N, Rajinikanth V, Gandhi TK. Deep transfer learning-based automated detection of COVID-19 from lung CT scan slices. Appl. Intel 2021; 51(1):371–85.

[36] He X, Yang X, Zhang S, Zhao J, Zhang Y, Xing E, Xie P. Sample-efficient deep learning for COVID-19 diagnosis based on CT scans. medRxiv 2020.

[37] Javadi Moghaddam S, Gholamalinejad H. A novel deep learning based method for COVID-19 detection from CT image. Biomed Signal Processing and Control 2021; 70:102987.

[38] Shah V, Kehy A, Shriddharani A, Punjabi M, Shah J, Mehendale N. Diagnosis of COVID-19 with bidirectional lstm network using deep features extracted from chest X-ray images.

[39] Automatic detection of COVID-19 with bidirectional lstm network using deep features extracted from chest X-ray images.

[40] Pathak Y, Shukla PK, Tiwari A, Stakn S, Singh S. Deep transfer learning based classification model for COVID-19 disease. Irbm; 2025.

[41] Shaik NS, Cherukuri TK. Transfer learning based novel ensemble classifier for COVID-19 detection from chest CT-scans. Comput. Biol Med 2022;141:105127.

[42] Ebenreer AS, Kamnan SD, Sivakumar M, Priya SJ. Effect of image transformation on EfficientNet model for COVID-19 CT image classification. Mater Today: Proceedings 2022;51:2512–9.

[43] Jangam E, Barreto AAD, Anavarapu CSR. Automatic detection of COVID-19 from chest CT scan and chest X-Rays images using deep learning, transfer learning and stacking. Applied Intel 2022;52(2):2243–59.

[44] https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset.

[45] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017. p. 4700–8.

[46] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770–8.

[47] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2015. p. 1–9.

[48] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 2818–26.

[49] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 2014.