Türkiye Klinikleri Biyoistatistik Dergisi
Turkiye Klinikleri Journal of Biostatistics Turkiye Klinikleri J Biostat. 2025;17(1):

ORIGINAL RESEARCH ORİJİNAL ARAŞTIRMA

DOI: 10.5336/biostatic.2024-106714

Evaluation of Computed Tomography Images by Convolutional Neural Network Architectures for COVID-19 Diagnosis: A Methodological Study

COVID-19 Tanısında Bilgisayarlı Tomografi Görüntülerinin Konvolüsyonel Sinir Ağları Mimarileri ile Değerlendirilmesi: Metodolojik Çalışma

Işıl ÜNALDI[a], Leman TOMAK[a]

[a]Ondokuz Mayıs University Faculty of Medicine, Department of Biostatistics, Samsun, Türkiye

ABSTRACT Objective: For the diagnosis of coronavirus disease-2019 (COVID-19) using computed tomography (CT) images, disease classification from images must be performed with high accuracy. This study aims to evaluate the classification success on CT images using fine-tuned convolutional neural network (CNN) architectures applied to the new dataset with transfer learning method, which provides accurate and fast diagnosis of COVID-19, by comparing their classification success with various performance measures. **Material and Methods:** The classification of the disease was performed with CNN architectures Xception, InceptionResNetV2, InceptionV3, ResNet50, VGG16, VGG19, MobileNetV2, DenseNet121, DenseNet169 and DenseNet201. The performance of the architectures was utilized with accuracy, sensitivity, specificity, precision, F1 score and Area Under the Curve (AUC). In total, the data included 7,593 CT images of 466 COVID-19 patients and 6,893 CT images of 604 nonCOVID-19 patients. **Results:** The architecture that had the highest performance was InceptionResNetV2, with an accuracy of 98.00%, sensitivity of 98.53%, specificity of 97.53%, precision of 97.25%, F1 score of 97.88%, and AUC of 98.03%. ResNet50 showed the lowest performance with accuracy of 97.14%, sensitivity of 98.75%, specificity of 95.70%, precision of 95.32%, F1 score of 97.01% and AUC of 97.23%. **Conclusion:** All architectures examined were found to work with performance measures above 95% in COVID-19 diagnostics. As a result, it was found that the pre-processing, fine-tuning, and hyperparameter optimization of the architectures we used are generalizable for COVID-19 and chest CT images, and that each of the architectures works with good performance.

Keywords: Coronavirus disease-2019; computed tomography; image classification; deep learning; convolutional neural networks

ÖZET Amaç: Bilgisayarlı tomografi (BT) görüntüleri kullanılarak koronavirüs hastalığı [coronavirus disease-2019 (COVID-19)] tanısı için görüntülerden hastalık sınıflandırmasının yüksek doğrulukla yapılması gerekmektedir. Bu çalışma, COVID-19'un doğru ve hızlı tanısını sağlayan, transfer öğrenme yöntemi ile yeni veri setine uygulanan ince ayarlı konvolüsyonel sinir ağları (KSA) mimarilerini kullanarak BT görüntüleri üzerinde sınıflandırma başarılarını çeşitli performans ölçütleri ile karşılaştırılarak değerlendirmeyi amaçlamaktadır. **Gereç ve Yöntemler:** Hastalığın sınıflandırılması KSA mimarilerinden Xception, InceptionResNetV2, InceptionV3, ResNet50, VGG16, VGG19, MobileNetV2, DenseNet121, DenseNet169 ve DenseNet201 kullanılarak gerçekleştirildi. Mimarilerin performansları doğruluk, duyarlık, seçicilik, kesinlik, F1 skoru ve Eğrinin Altındaki Alan [Area Under the Curve (AUC)] ile değerlendirildi. Kullanılan veri, 466 COVID-19 hastasının 7.593 BT görüntüsünü ve 604 COVID-19 olmayan hastanın 6.893 BT görüntüsünü içeriyordu. **Bulgular:** En yüksek performansa sahip mimari, %98,00 doğruluk, %98,53 duyarlık, %97,53 seçicilik, 97,25 kesinlik, 97,88 F1 puanı ve 98,03 AUC ile InceptionResNetV2 oldu. En düşük performansı %97,14 doğruluk, %98,75 duyarlık, %95,70 seçicilik, %95,32 kesinlik, %97,01 F1 skoru ve %97,23 AUC ile ResNet50 gösterdi. **Sonuç:** İncelenen tüm mimarilerin %95'in üzerinde performans ölçüleriyle çalıştığı görüldü. Sonuç olarak, kullandığımız mimarilerin ön işleme, ince ayar ve hiperparametre optimizasyonunun COVID-19 ve göğüs BT görüntüleri için genelleştirilebilir olduğu ve mimarilerin her birinin iyi performansla çalıştığı görülmüştür.

Anahtar Kelimeler: Koronavirüs hastalığı-2019; bilgisayarlı tomografi; görüntü sınıflandırma; derin öğrenme; konvolüsyonel sinir ağları

Correspondence: Işıl ÜNALDI
Ondokuz Mayıs University Faculty of Medicine, Department of Biostatistics, Samsun, Türkiye
E-mail: isil.unaldi@omu.edu.tr

Peer review under responsibility of Turkiye Klinikleri Journal of Biostatistics.

Received: 11 Nov 2024 ***Received in revised form:*** 12 Mar 2025 ***Accepted:*** 27 Mar 2025 ***Available online:*** 03 Apr 2025

The coronavirus disease-2019 (COVID-19) pandemic continues, with several variants that cause significant damage to health and economies worldwide. The Reverse Transcription-Polymerase Chain Reaction (RT-PCR) test is the diagnostic system for COVID-19 with severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) ribonucleic acid detected in a swab or saliva samples from the nose or throat from the patient. However, the sensitivity of RT-PCR may be low and the result may vary depending on the elapsed time.[1] However, this is a costly procedure requiring manual sampling and a new test kit for each new patient. They also have problems with performance measures. In addition, it has been observed that the RT-PCR test can give a high rate of false negative results in the early stages of the disease. For this reason, computed tomography (CT) and X-ray methods have started to be used for the diagnosis of COVID-19 in the early stages of the disease. Since these methods require less contact with the infected patient, they carry less risk of transmission.[2] Studies have proven that CT is a COVID-19 screening and diagnostic tool that complements the RT-PCR test. Based on the use of CT imaging in clinics and hospitals, disease classification from images needs to be performed with high accuracy.[3]

CT and X-ray, which are routine imaging tools for the diagnosis of pneumonia, are easy to perform and can provide rapid diagnosis. The ground-glass appearance and tissue changes seen in the lung are seen in almost all COVID-19 patients and can be easily recognized by CT and X-ray images. These features are also observed in patients with negative RT-PCR results but clinical symptoms. Especially for the diagnosis of COVID-19, CT and X-ray imaging are considered among the most effective methods. CT scanning is preferred over X-ray because of its versatility and 3-dimensional pulmonary view.[4-6] In this context, the use of machine learning and deep learning based rapid diagnosis systems over CT images is very important and necessary. Deep learning is a specialized and advanced subtype of machine learning. Machine learning method requires hand-crafted extraction of necessary features. Therefore, with the use of deep learning based methods, feature extraction has become automatic thanks to algorithms that are more practical and faster. With the studies carried out, early diagnosis and rapid treatment solutions of COVID-19 disease have been further developed with deep learning-based methods, providing clinicians with a more objective interpretation.[7] CT scans provide detailed visualization of the lungs and help radiologists diagnose COVID-19 in hospitals. However, a CT scan of a person contains hundreds of slides and diagnosing COVID-19 using such images can lead to delays in hospitals. Deep learning methods can help radiologists to quickly and accurately diagnose COVID-19 infection from these scans.[8]

Examination of medical images is done by experts such as radiologists and physicians. As medical data vary considerably from patient to patient, depending on the disease, labor becomes crucial in the diagnostic task. Moreover, interpretation may vary depending on the fatigue, speed and experience of clinicians with limited and subjective capacity. With the current pandemic, the ability of the health system to cope with the disease has been significantly tested, making it imperative to diagnose the outbreak quickly and accurately. For this, a good solution is to analyse medical images using an automated computing system with neural networks that can reach the accuracy of the human brain or even more. Furthermore, given the time factor, the amount of medical image analysis that an effective neural network can perform is greater than that of the human brain.[9] In this context, fine-tuned convolutional neural network (CNN) architectures that perform well in medical image analyses, are deep learning based, pre-trained, and can be used for new datasets with transfer learning methods have become a preferred method.

CT scans require some corrections to improve image quality during deep learning-based image analysis. The use of CNNs is also very important at this point. In this process called pre-processing, methods such as bringing the image size to a single standard, normalization, converting the image into sequences, sharpening, horizontal and vertical rotation, zooming, adjusting brightness, etc. are applied to ensure the most efficient analysis of the images.[10] Images must be fitted to the input size of CNN architectures in order to be trained and predictions can be made on them. For this reason, the input size of the architecture is checked and the images are rescaled.

In this study, it is aimed to detect the disease by using fine-tuned CNN architectures that provide accurate and fast detection of COVID-19 on chest CT images, which can be used for new datasets with pre-trained and transfer learning techniques, and to evaluate the classification success of the architectures by comparing various performance measures.

MATERIAL AND METHODS

In this study, an open-source large COVID-19 CT scan slice dataset, accessed via Kaggle, was used. Disease classification was performed using CNN architectures Xception, InceptionResNetV2, InceptionV3, ResNet50, VGG16, VGG19, MobileNetV2, DenseNet121, DenseNet169, and DenseNet201. The architectures were run with graphics processing unit support. Different data pre-processing, augmentation methods and hyperparameter adjustments were applied to the dataset and set as follows: rescale=1/255.0, horizontal flip and vertical flip=true, zoom range, width shift range, and height shift range=0.2, rotation range=360, batch size=32, epoch=100, minimum learning rate=0.0001, dropout=(0.5), input size=224×224, class_mode=categorical, number of classes=2, type of loss function=categorical crossentropy, type of activation function=rectified linear unit (ReLU), optimization method=adam optimizer. Batch normalization added between layers. The ReduceLROnPlateau, ModelCheckpoint, and EarlyStopping functions were used as callbacks during the training process. The performance of the architectures were utilized in terms of accuracy, sensitivity, specificity, precision, F1 score, and Area Under the Curve (AUC). Learning curves, confusion matrices and Receiver Operating Characteristic (ROC) curves of the architectures were obtained. Google Colaboratory was used to perform data analyses using Python. Python programming language was used together with TensorFlow and Keras libraries.

Dataset

Seven public datasets have been merged to form a large lung CT scan dataset for COVID-19 and are available as open access via Kaggle. The metadata of COVID-19 and nonCOVID-19 cases for these 7 separate datasets are given in Table 1.

TABLE 1: Metadata of each separate dataset

Dataset	Country	COVID-19 slices	COVID-19 cases	NonCOVID-19 slices	NonCOVID-19 cases
11	Iran	666	68	1,053	274
12	Italy	100	43	NA*	NA
13	Multiple	1,844	20	NA	NA
14	Multiple	34	17	NA	NA
15	Russia	785	50	5,080	254
16	China, Japan	349	213	NA	NA
17	Iran	3,815	55	760	76

*NA: Not available

These datasets have been frequently used in the COVID-19 diagnostics literature and have proven their effectiveness in deep learning applications. In this respect, this combined dataset is expected to increase the generalization success of deep learning methods. In total, the combined open source data included 7,593 CT images of 466 COVID-19 patients and 6,893 CT images of 604 nonCOVID-19 patients.[18] Chest CT samples of COVID-19 and nonCOVID-19 patients in the dataset are shown in Figure 1.

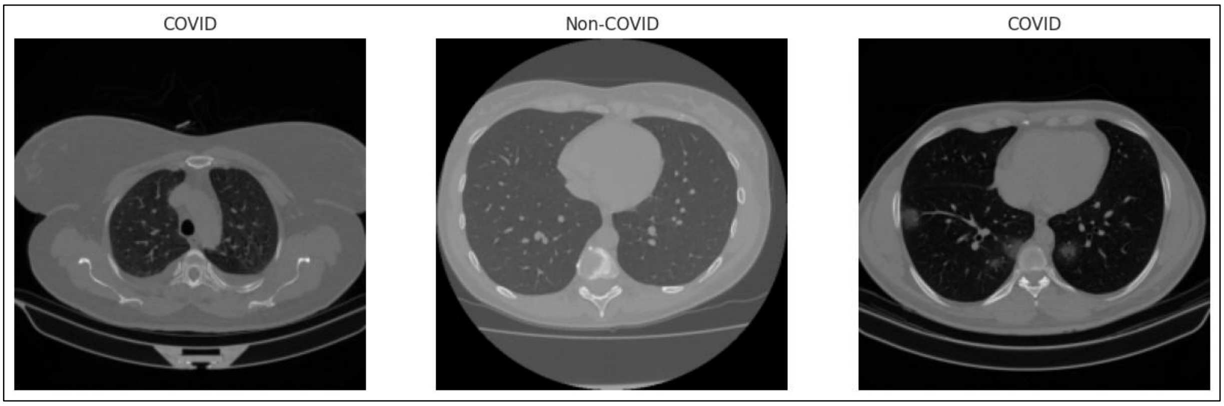


FIGURE 1: Sample images of 2 classes in the data set

Convolutional Neural Networks

Artificial intelligence systems based on deep neural networks that can detect, recognize, identify and classify objects in images are called CNN. CNN is a deep learning method that can learn directly from the input object without the need for a human for feature extraction by giving images as input to the network. Since this network is specifically designed to process a series of images, it is used in medical image analysis, image recognition, natural language processing and many other applications. More effectively than a normal neural network, it can automatically identify key features in the input without human intervention.[19]

CNN architecture is a multilayered structure that provides multilevel feature learning. In the first layers of the CNN, features such as texture and shape can be extracted, while in the deep layers, features can be extracted that are combined to accurately locate key elements. A CNN architecture consists of a succession of several basic building blocks: a convolution layer, pooling layer, fully connected layer, and an activation layer.

The convolution layers and pooling layers perform feature extraction, whereas the fully connected layer maps the extracted features to the final output for classification purposes. The convolution layer is very important in a CNN, which consists of a series of mathematical and special linear operations (convolution operation, etc.). In digital images, the kernel, a small set of parameters and an optimizable feature extractor are applied to each image location, while pixel values are stored in an array of numbers. This process, which makes CNNs efficient in image processing, can generate features from anywhere in the image. As the output of one layer feeds the next layer, the extracted features can become hierarchical and more complex.[20]

The transformation of input data into output through layers is called forward propagation and the optimization of parameters such as kernel is called training. By utilizing backpropagation and gradient descent optimization algorithms, the goal is to minimize the difference between the outputs and true labels. The performance of an architecture under specified kernels and weights is calculated by a loss function using the forward propagation algorithm on the training dataset, and the learnable parameters, that is, kernels and weights, are updated according to the loss value by backpropagation with the gradient descent algorithm.[21]

Training a deep learning model from scratch poses some challenges due to the scarcity of well-described medical images, especially those related to COVID-19. In this respect, pre-trained architectures are preferred for the classification task in this study. Fine-tuned Xception, fine-tuned DenseNet121, fine-tuned DenseNet201, fine-tuned InceptionResNetV2, fine-tuned InceptionV3, fine-tuned MobileNetV2, fine-tuned VGG16, fine-tuned VGG19, fine-tuned ResNet50, and fine-tuned DenseNet169 architectures were used for the classification of COVID-19.

Xception

Xception has an architecture similar to that of VGG. It is based on the inception modules of InceptionV3.[22] This is based on deeply separable convolution layers. This was inspired by the inception model. It had approximately 20 million learnable parameters. There were 36 convolution layers arranged in 14 blocks. All blocks except the first and last block have linear residual connections.[23,24]

DenseNet121

The DenseNet121 input size is a fixed 224×224. DenseNet121 consists of 121 layers containing over 8 million parameters. Batch normalization is applied for downsampling by dividing the same feature map dimensions with different filter numbers into Dense Blocks, and the layers between the blocks are called transition layers.[25]

DenseNet201

DenseNet201 utilizes the densified network, which provides easy-to-train and parametrically high-impact models due to the fact that features can be reused by different layers, increasing the diversity of the input of the next layer and improving performance. DenseNet201 has worked on various datasets such as CIFAR-100 and ImageNet with effective performances. DenseNet201 consists of 201 layers with 20 million parameters.[26]

InceptionResNetV2

InceptionResNetV2 is based on a common structure of the inception and the residual connection. In InceptionResNet, the residual links are combined with convolutional filters of different sizes. For InceptionResNet, residual connections are blended with convolution filters. These residual connections avoid distortion problems caused by deep structures and reduce the training time. InceptionResNetV2 has 164 layers and can therefore classify many images well.[27]

InceptionV3

InceptionV3 is an extended version of GoogLeNet that works with good classification performances in various medical applications. InceptionV3 reduces computational complexity by proposing an initialization model that assembles convolution filters of different sizes into a new filter by reducing the number of parameters.[28]

MobileNetV2

MobileNet provides a remarkable reduction in the computing power compared to other CNN models. With this feature, they have become suitable for use on mobile devices and computers with low computing power. MobileNet, similar to ReLU, contains nonlinear depth convolution layers, the fully connected layer and a softmax layer.[27]

VGG16

VGG16 is a model that shows the effect of network depth on performance. It has a fixed 224×224 image shape and 3×3 kernel size, this structure deepens the network, minimizing the number of parameters, so it can increase the learning rate.[29]

VGG19

VGG19 is a CNN architecture consisting of 19 layers, 16 convolution layers, and 3 fully connected layers. In this architecture, the number of parameters in the convolution layers is reduced with the help of filters. The VGG19 architecture contains approximately 138 million parameters.[30]

ResNet50

Having 50 layers, ResNet50 estimates the delta required to reach the final output from one layer to the next. ResNet reduces the vanishing gradient problem with its methods. ResNet helps prevent overfitting by allowing the weighting layer to be skipped.[31]

DenseNet169

DenseNet169 is a version of DenseNet. DenseNet merges feature maps from previous levels by transferring all feature maps to subsequent levels and combining them with newly created feature maps. DenseNet allows features to be reused, thereby alleviating the problem of vanishing gradients.[32]

RESULTS

The highest performing architecture was InceptionResNetV2, with accuracy of 98.00%, sensitivity of 98.53%, specificity of 97.53%, precision of 97.25%, F1 score of 97.88% and AUC of 98.03%. ResNet50 showed the lowest performance with accuracy of 97.14%, sensitivity of 98.75%, specificity of 95.70%, precision of 95.32%, F1 score of 97.01% and AUC of 97.23%. The performance results of the architectures are listed in Table 2.

TABLE 2: Performance measures obtained as a result of classification of CNN architectures

	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1 score (%)	AUC (%)
Xception	97.58	98.46	96.81	96.47	97.46	97.63
InceptionResNetV2	98.00	98.53	97.53	97.25	97.88	98.03
InceptionV3	97.86	97.94	97.79	97.51	97.73	97.87
ResNet50	97.14	98.75	95.70	95.32	97.01	97.23
VGG16	97.17	98.46	96.03	95.65	97.03	97.24
VGG19	97.69	98.31	97.14	96.82	97.56	97.72
MobileNetV2	97.17	96.55	97.72	97.41	96.98	97.14
DenseNet121	97.38	98.24	96.61	96.26	97.24	97.43
DenseNet169	97.90	96.77	98.89	98.73	97.74	97.83
DenseNet201	97.93	99.49	96.55	96.24	97.83	98.02

The learning curves of the CNN architectures are given in Figure 2.

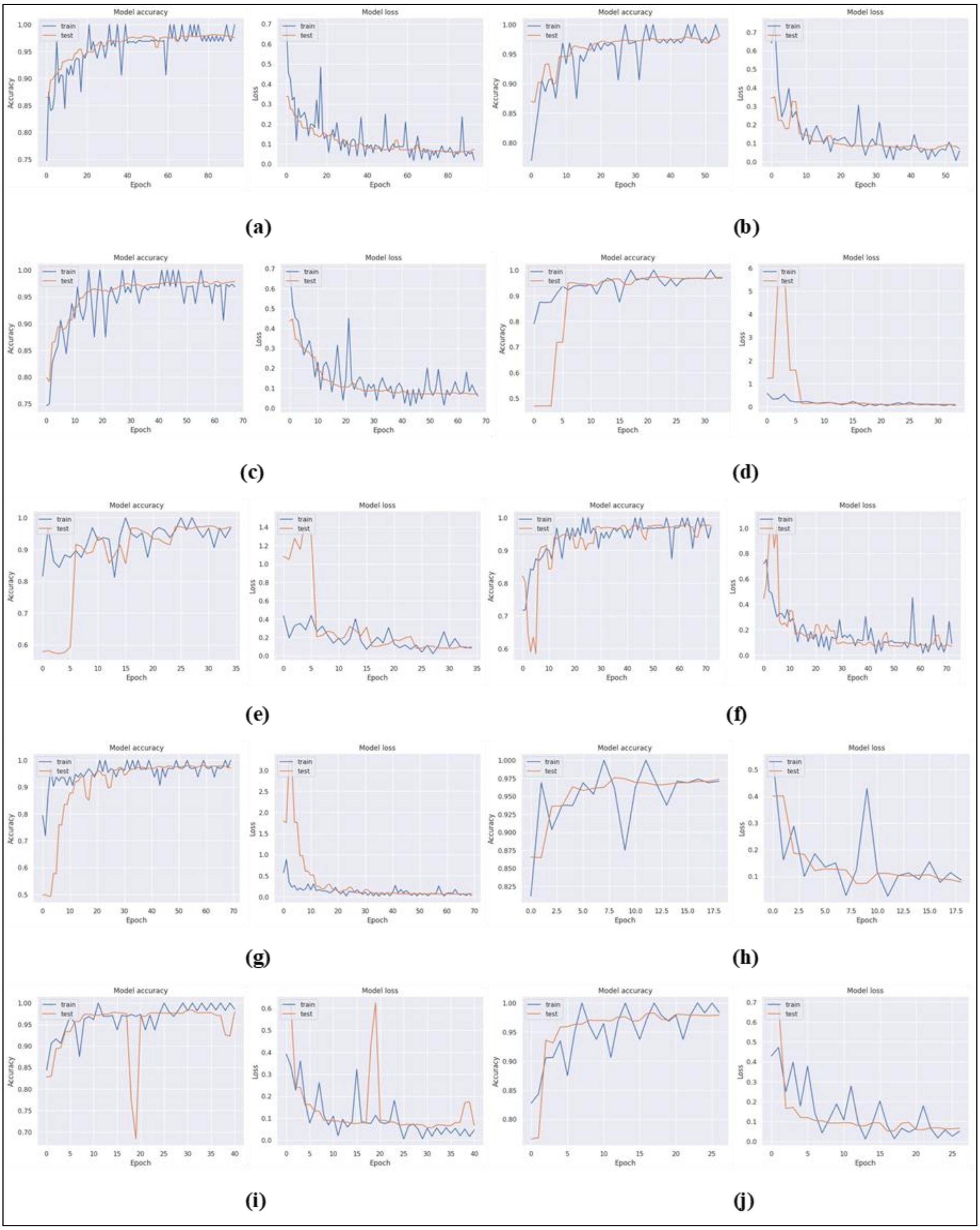


FIGURE 2: Learning curves of CNN architectures
(a) Xception; (b) InceptionResNetV2; (c) InceptionV3; (d) ResNet50; (e) VGG16; (f) VGG19; (g) MobileNetV2; (h)DenseNet121; (i) DenseNet169; (j) DenseNet201

The structure of the learning curves can be used to examine the behaviour of the architecture and to suggest adjustments that can be made to improve the training performance. When the learning curves were analysed, it was found that the training of each architecture was good, the accuracy curves increased together and the loss decreased together in the training and testing processes, and overfitting was prevented by the adjustments made.

The confusion matrices of the CNN architectures are given in Figure 3.

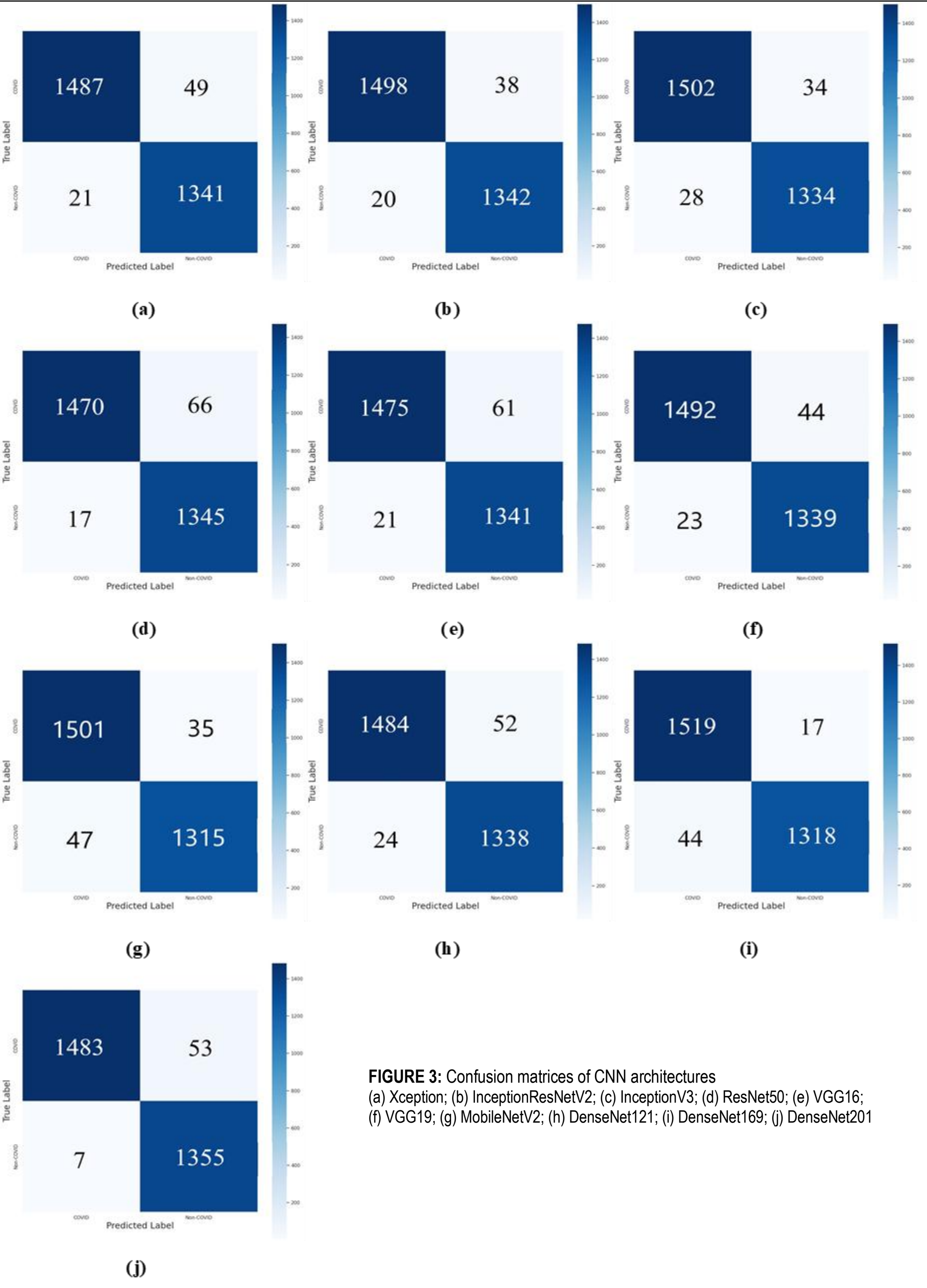
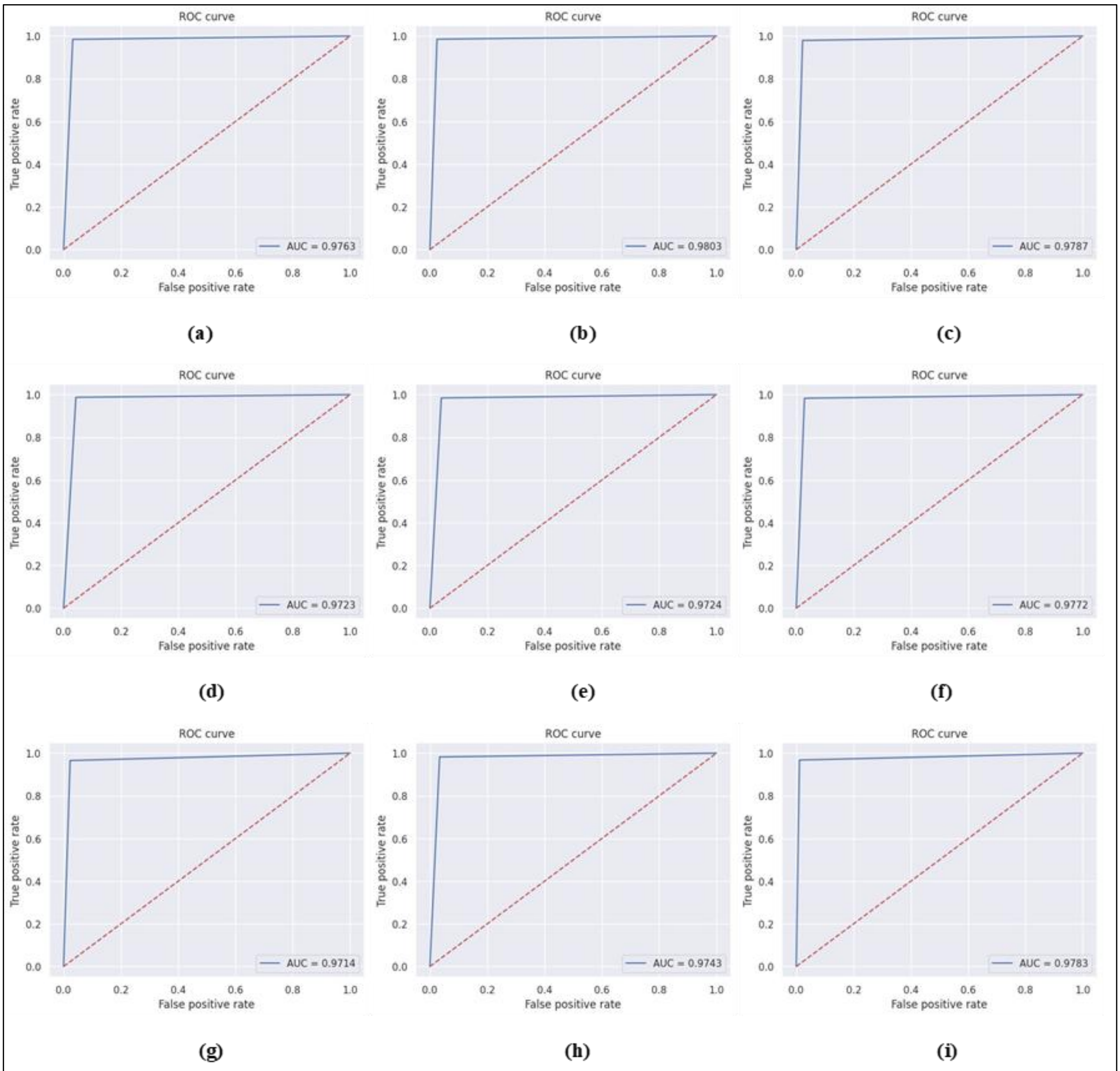


FIGURE 3: Confusion matrices of CNN architectures
(a) Xception; (b) InceptionResNetV2; (c) InceptionV3; (d) ResNet50; (e) VGG16; (f) VGG19; (g) MobileNetV2; (h) DenseNet121; (i) DenseNet169; (j) DenseNet201

The ROC curves of the CNN architectures are given in Figure 4.



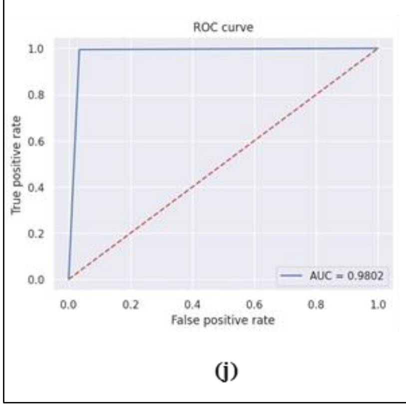


FIGURE 4: ROC curves of CNN architectures
(a) Xception; (b) InceptionResNetV2; (c) InceptionV3; (d) ResNet50; (e) VGG16; (f) VGG19; (g) MobileNetV2; (h) DenseNet121; (i) DenseNet169; (j) DenseNet201
ROC: Receiver operating characteristic; AUC: Area under the curve

DISCUSSION

All architectures examined in the study were found to work well in COVID-19 diagnostics, with performance measures above 95%. The performance results for each architecture were very similar. The structure of the learning curves can be used to examine the behavior of the architecture and to make suggestions for corrections to improve training and performance. When the learning curves obtained for each architecture are analyzed, it can be said that the dataset and the fine-tunings made are suitable and generalizable for this architecture because there is not much deviation in the curves of all architectures, which has achieved high performance measures. In addition, working with large datasets significantly improves the ability of architectures to learn patterns, particularly in the training phase in the detection of lung anomalies.

In the literature, there are various studies on the use of CNNs for diagnosing COVID-19 from medical images. Kaur and Kaur proposed a new fine-tuned pretrained MobileNetV2 architecture for the diagnosis of COVID-19 from CT images and used Xception, NasNetLarge, NasNetMobile, DenseNet201, DenseNet121, DenseNet169, InceptionV3, and InceptionResNetV2 architectures for comparison. Their analyses demonstrated accuracies of 89.65% for Xception, 89.16% for NasNetLarge, 85.71% for NasNetMobile, 82.75% for DenseNet201, and 79.80% for DenseNet121.[33] Malik et al. used CT, X-ray and cough sound images to diagnose nine types of lung diseases, including COVID-19, and proposed a new disease diagnose network called DCDD_Net, and used InceptionResNetV2, EfficientNet-B0, DenseNet201, and Xception architectures to compare the performance of this network. They achieved accuracy of 83.09%, precision of 85.78%, recall of 80.94%, F1 score of 82.88%, AUC of 97.90% with Xception and accuracy of 85.37%, precision of 87.85%, recall of 84.42%, F1 score of 84.88%, AUC of 98.22% with DenseNet201.[34] Keerthana et al. aimed to develop an inductive parameter transfer learning-based approach for COVID-19 and normal diagnoses from lung CT images. In their proposed approach, they performed feature extraction by taking DenseNet201, InceptionV3, Xception, VGG19, and ResNet50 as basic architectures, and then fine-tuned them by adding special layers. They used 2 different optimizers (Adam and RMSprop) for parameter tuning, and evaluated the results using 2 different datasets. They classified COVID-19 with a precision of 92%, recall of 86%, and F1 score of 89% for the Xception architecture and precision of 95%, recall of 94%, and F1 score of 94% for the DenseNet201 architecture on the Kaggle dataset using the Adam optimizer.[35] Huang and Liao used InceptionV3, ResNet50V2, Xception, DenseNet121, MobileNetV2, EfficientNet-B0 and EfficientNetV2 architectures with and without fine-tuning on CT and X-ray images for COVID-19 diagnosis, applied 5-fold cross validation and reported the results. They obtained the following results on CT images: accuracy of 93.46±0.80%, precision of 94.09±0.82%, recall of 94.06±0.90%, F1 score of 93.89±0.77% with Xception architecture, accuracy of 96.78±1.19%, precision of 96.43±1.15%, recall of 96.52±1.28%, F1 score of 96.48±1.35% with fine-tuned Xception architecture, accuracy of 93.12±2.21%, precision of 93.11±2.25%, recall of 93.07±2.23%, F1 score of 93.06±2.31% with DenseNet121 architecture, and accuracy of 96.58±1.07%, precision of 96.59±1.02%, recall of 96.53±1.09%, F1 score of 96.48±1.18% with fine-tuned DenseNet121 architecture.[36] Matsuyama performed classification on CT data consisting of 720 images with ResNet50 architecture. As a result of the analysis, accuracy of 92.2%, sensitivity of 90.4%, specificity of 93.3%, precision of 92.6%, F1 score of 91.5% was obtained with the method proposed in the study.[37] He et al. obtained accuracy of 86%, F1 score of 85%, AUC of 94% using DenseNet169 architecture with the Self-Trans method they proposed in the diagnosis of COVID-19 on CT images. With the transfer learning method, they achieved accuracy of 76%, F1 score of 76%, AUC of 82% using VGG16; accuracy of 80%, F1 score of 81%, AUC of 88% using ResNet50; accuracy of 79%, F1 score of 79%, AUC of 88% using DenseNet121.[38] Serte and Demirel ran their proposed deep learning model on CT images and achieved accuracy of 84%, sensitivity of 100%, specificity of 80%, AUC of 96% with their method based on ResNet50 architecture.[8] Yang et al. performed classification with DenseNet architecture on high-resolution CT images and obtained accuracy of 92%, sensitivity of 97%, specificity of 87%, F1 score of 93%, AUC of 98%.[39] Mishra et al. performed CT image classification with VGG16 and ResNet50 and achieved an average accuracy of 99.12%, sensitivity of 98.8%, specificity of 99.12%, precision of 99.12%, F1 score of 98.8% with VGG16 and accuracy of 99.62%, sensitivity of 99.6%, specificity of 99.6%, precision of 99.6%, F1 score of 99.6% with ResNet50.[40] Shaik and Cherukuri conducted their analyses on SARS-CoV-2 CT scan dataset and COVID-CT dataset using VGG16, VGG19, InceptionV3, ResNet50, ResNet50V2, InceptionResNetV2, Xception, and MobileNet architectures. ResNet50V2 achieved the highest accuracy with 97.79% on SARS-CoV-2 CT scan dataset and VGG16 achieved the highest accuracy with 92% on COVID-CT dataset.[41] Bouzid et al. used ResNet, DenseNet and EfficientNet architecture families for COVID-19 diagnosis using CT scans in their study. They created a new dataset by combining 7 different datasets for their analyses. They added additional cases from 2 different datasets. In their experiments, they obtained accuracy of 97.60%, recall of 96.39%, specificity of 98.20%, precision of 96.74% with ResNet50; accuracy of 98.01%, recall of 97.00%, specificity of 98.51%, precision of 97.11% with DenseNet121; accuracy of 98.48%, recall of 97.73%, specificity of 98.86%, precision of 97.75% with DenseNet201.[42] In their study, Appavu et al. focused on image classification with VGG19, ResNet50, VGG19, InceptionV3 architectures using a dataset containing a total of 2400 CT images, 1,200 slice COVID-19, 1,200 slice normal. They obtained their results as follows: accuracy of 94.19%, sensitivity of 99.22%, specificity of 88.39%, precision of 88.39% with VGG16; accuracy of 73.03%, sensitivity of 67.44%, specificity of 79.46%, precision of 79.46% with ResNet50; accuracy of 88.38%, sensitivity of 99.22%, specificity of 75.89%, precision of 75.89% with VGG19; accuracy of 87.97%, sensitivity of 86.15%, specificity of 90.09%, precision of 90.09% with InceptionV3.[43] Keshamoni et al. created a COVID-19 diagnostic system integrating X-ray, electrocardiography and CT images using CNN architectures and evaluated the performances with VGG19, ResNet50 and ImageNet architectures. The VGG19 architecture they proposed on CT data achieved accuracy of 99%, recall of 98.9%, precision of 99.23%, F1 score of 98.34%.[44] Mousavi and Hosseini used a dataset containing 10,239 CT images and evaluated the performances with EfficientNetB4, InceptionResNetV2, InceptionV3 architectures. They obtained accuracy of 99.073%, precision of 99%, recall of 99%, F1 score of 99% with EfficientNetB4; accuracy of 99.414%, precision of 99%, recall of 99%, F1 score of 99% with InceptionResNetV2; accuracy of 99.366%, precision of 99%, recall of 99%, F1 score of 99% with InceptionV3.[45] Rajpoot et al. used the SARS-CoV-2 CT scan dataset containing 2,482 images accessed from Kaggle in their study. On this dataset, they used VGG16 with accuracy of 92.77%, sensitivity of 92.77%, specificity of 92.77%, F1 score of 93.77%, AUC of 93.00%; ResNet50 with accuracy of 95.78%, sensitivity of 95.80%, specificity of 95.80%, F1 score of 95%, AUC of 96%; Xception with accuracy of 92.97%, sensitivity of 92.99%, specificity of 92.99%, F1 score of 92.97%, AUC of 93%; DenseNet169 with accuracy of 93.98%, sensitivity of 94.11%, specificity of 94.11%, F1 score of 93.97%, AUC of 94% were obtained.[46] Tello-Mijares et al. used AlexNet, GoogleNet, ResNet50, VGG16 and VGG19 architectures to evaluate lung damage in COVID-19 patients with CNN using CT images. As a result of their analysis, accuracy of 94.97%, sensitivity of 96.23%, specificity of 94%, precision of 94%, F1 score of 95.09%, AUC of 94.96% with AlexNet; accuracy of 93.59%, sensitivity of 92.98%, specificity of 94%, precision of 94%, F1 score of 93.62%, AUC of 93.59% with GoogleNet; accuracy of 94.28%, sensitivity of 90.58%, specificity of 98%, precision of 98%, F1 score of 94.13%, AUC of 94.33%; accuracy of 95.84%, sensitivity of 95.38%, specificity of 96%, precision of 96%, F1 score of 95.87%, AUC of 95.85% with VGG16; accuracy of 92.89%, sensitivity of 98.46%, specificity of 87%, precision of 89%, F1 score of 93.34%, AUC of 93% with VGG19.[47]

When this study is compared with the reviewed studies, it is found that InceptionResNetV2 achieves higher or similar performance measures than most architectures. The reasons for this difference were interpreted as the use of different datasets, different fine-tuning and hyperparameter adjustments, and the effects of dataset sizes on performance. As a result, it was found that the pre-processing, fine-tuning, and hyperparameter optimization of the architectures we used are generalizable for COVID-19 and chest CT images, and that each of the architectures works with good performance.

CONCLUSION

The results of the study showed that the CNN architectures used, the pre-processing methods applied to the image, the fine-tuning and the selected classifiers significantly affect the performance of the architectures. The most important feature that affects the success of deep CNN architectures is working with large datasets. Having more labelled training data and having a balanced dataset are the most important regulations affecting the success of deep learning methods. In this study, a balanced dataset was used. The data was of sufficient size. In addition, dropout layer addition, batch normalization, data augmentation, training with callbacks were preferred to control overfitting. As a result, learning curves were analysed and it was found that overfitting could be prevented.

In this study, it has been shown that the CNN method is an effective method that will help clinicians in the diagnosis of COVID-19 with CT images, and that the disease can be detected as soon as possible, the effectiveness of treatment can be increased and the negativities caused by the disease can be prevented if sufficient data and fine-tuned architectures suitable for the data are worked with.

Source of Finance

During this study, no financial or spiritual support was received neither from any pharmaceutical company that has a direct connection with the research subject, nor from a company that provides or produces medical instruments and materials which may negatively affect the evaluation process of this study.

Conflict of Interest

No conflicts of interest between the authors and/or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Authorship Contributions

Idea/Concept: *Işıl Ünaldı, Leman Tomak;* ***Design:*** *Işıl Ünaldı, Leman Tomak;* ***Control/Supervision:*** *Leman Tomak;* ***Data Collection and/or Processing:*** *Işıl Ünaldı;* ***Analysis and/or Interpretation:*** *Işıl Ünaldı, Leman Tomak;* ***Literature Review:*** *Işıl Ünaldı;* ***Writing the Article:*** *Işıl Ünaldı;* ***Critical Review:*** *Leman Tomak.*

REFERENCES

1. Gunraj H, Sabri A, Koff D, Wong A. COVID-Net CT-2: enhanced deep neural networks for detection of COVID-19 from chest CT images through bigger, more diverse learning. Front Med (Lausanne). 2022;8:729287. PMID: 35360446; PMCID: PMC8960961.
2. Demir FB, Yılmaz E. X-ray görüntülerinden COVID-19 tespiti için derin öğrenme temelli bir yaklaşım [A deep learning-based approach for detection of COVID-19 from X-ray images]. Avrupa Bilim ve Teknoloji Dergisi. 2021;(32):627-32. https://doi.org/10.31590/ejosat.1039522
3. Gunraj H, Tuinstra T, Wong A. COVIDx CT-3: a large-scale, multinational, open-source Benchmark dataset for computer-aided COVID-19 screening from chest CT images. ArXiv. 2022. https://doi.org/10.48550/arXiv.2206.03043.
4. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology. 2020;296(2):E32-E40. PMID: 32101510; PMCID: PMC7233399.
5. Dong D, Tang Z, Wang S, Hui H, Gong L, Lu Y, et al. The role of imaging in the detection and management of COVID-19: a review. IEEE Rev Biomed Eng. 2021;14:16-29. PMID: 32356760.
6. Islam MM, Karray F, Alhajj R, Zeng J. A review on deep learning techniques for the diagnosis of novel coronavirus (COVID-19). IEEE Access. 2021;9:30551-72. PMID: 34976571; PMCID: PMC8675557.
7. Clark S, Kamalinejad E, Magpantay C, Sahota S, Zhong J, Hu Y. A review of CNN on medical imaging to diagnose COVID-19 infections. 34th International Conference on Computer Applications in Industry and Engineering. CAINE. 2021;79:91-8. https://easychair.org/publications/paper/pWNx/open
8. Serte S, Demirel H. Deep learning for diagnosis of COVID-19 using 3D CT scans. Comput Biol Med. 2021;132:104306. PMID: 33780867; PMCID: PMC7943389.
9. Kugunavar S, Prabhakar CJ. Convolutional neural networks for the diagnosis and prognosis of the coronavirus disease pandemic. Vis Comput Ind Biomed Art. 2021;4(1):12. PMID: 33950399; PMCID: PMC8097673.
10. Riyadi S, Septiari TD, Damarjati C, Ramli S. Covid-19 detection based-on CT-scan images using inception deep learning. 2022 IEEE 12th International Conference on Control System, Computing and Engineering (ICCSCE) Penang, Malaysia. 2022. p.214-8. doi: 10.1109/ICCSCE54767.2022.9935633.
11. 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 Process Control. 2021;68:102588. PMID: 33821166; PMCID: PMC8011666.
12. Radiologists. [Internet]. COVID-19 CT segmentation dataset. © 2020 Artificial Intelligence AS [Cited: March 9, 2025]. Available from: [http://medicalsegmentation.com/covid19/].
13. Jun M, Cheng G, Yixin W, Xingle A, Jiantao G, Ziqi Y, et al. COVID-19 CT lung and infection segmentation dataset. Zenodo. 2020;20. https://doi.org/10.5281/zenodo.3757476.

14. Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M. COVID-19 image data collection: prospective predictions are the future. ArXiv preprint. 2020. arXiv:2006.11988.
15. Morozov SP, Andreychenko AE, Pavlov NA, Vladzymyrskyy AV, Ledikhova NV, Gombolevskiy VA, et al. Mosmeddata: cchest ct scans with COVID-19 related findings dataset. ArXiv preprint. 2020. arXiv:2005.06465.
16. Yang X, He X, Zhao J, Zhang Y, Zhang S, Xie P. COVID-CT-dataset: a CT scan dataset about COVID-19. ArXiv preprint. 2020. arXiv:2003.13865.
17. Afshar P, Heidarian S, Enshaei N, Naderkhani F, Rafiee MJ, Oikonomou A, et al. COVID-CT-MD, COVID-19 computed tomography scan dataset applicable in machine learning and deep learning. Sci Data. 2021;8(1):121. PMID: 33927208; PMCID: PMC8085195.
18. Maftouni M, Law ACC, Shen B, Kong Grado Z, Zhou Y, Yazdi NA. A robust ensemble-deep learning model for COVID-19 diagnosis based on an integrated CT scan images database. In Ghate A, Krishnaiyer K, Paynabar K, eds. IISE Annual Conference and Expo 2021. Institute of Industrial and Systems Engineers, IISE. 2021. p.632-7.
19. Taye MM. Theoretical understanding of convolutional neural network: concepts, architectures, applications, future directions. Computation. 2023;11(3):52. https://doi.org/10.3390/computation11030052.
20. Medium [İnternet]. Convolutional Neural Network (ConvNet yada CNN) nedir, nasıl çalışır? 2020. [Cited: September 26, 2024]. Available from: https://medium.com/@tuncerergin/convolutional-neural-network-convnet-yada-cnn-nedir-nasil-calisir-97a0f5d34cad
21. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights Imaging. 2018;9(4):611-29. PMID: 29934920; PMCID: PMC6108980.
22. Chollet F. Xception: deep learning with depthwise separable convolutions. CVPR. IEEE Computer Society. 2017. p.1800-7. https://ieeexplore.ieee.org/document/8099678
23. Varaich ZA, Khalid S. Recognizing actions of distracted drivers using inception v3 and Xception convolutional neural networks. IEEE. 2019. pp.1-8. doi: 10.23919/ICACS.2019.8689131.
24. Wani MA, Bhat FA, Afzal S, Khan AI. Basics of Supervised Deep Learning. Advances in Deep Learning. Berlin: Springer; 2020. p.13-29.
25. Shazia A, Xuan TZ, Chuah JH, Usman J, Qian P, Lai KW. A comparative study of multiple neural network for detection of COVID-19 on chest X-ray. EURASIP J Adv Signal Process. 2021;2021(1):50. PMID: 34335736; PMCID: PMC8314263.
26. Jaiswal A, Gianchandani N, Singh D, Kumar V, Kaur M. Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning. J Biomol Struct Dyn. 2021;39(15):5682-9. PMID: 32619398.
27. Khattab R, Abdelmaksoud IR, Abdelrazek S. Automated detection of COVID-19 and pneumonia diseases using data mining and transfer learning algorithms with focal loss from chest X-ray images. Applied Soft Computing. 2024;162:111806-6. doi: 10.1016/j.asoc.2024.111806.
28. Nguyen LD, Lin D, Lin Z, Cao J. Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation. 2018 IEEE International Symposium on Circuits and Systems (ISCAS). 2018. pp.1-5. doi: 10.1109/ISCAS.2018.8351550.
29. Issahaku FY, Liu X, Lu K, Fang X, Danwana SB, Asimeng E. Multimodal deep learning model for COVID-19 detection. Biomedical Signal Processing and Control. 2024;91:105906-6. doi:10.1016/j.bspc.2023.105906.
30. Toğaçar M, Ergen B, Cömert Z. Application of breast cancer diagnosis based on a combination of convolutional neural networks, ridge regression and linear discriminant analysis using invasive breast cancer images processed with autoencoders. Med Hypotheses. 2020;135:109503. PMID: 31760247.
31. Theckedath D, Sedamkar RR. Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks. SN Comput Sci. 2020;1(2). https://doi.org/10.1007/s42979-020-0114-9.
32. Asif S, Qurrat-ul-Ain, Awais M, Amjad K, Bilal O, Al-Sabri R, et al. A deep ensemble learning framework for COVID-19 detection in chest X-ray images. Netw Model Anal Health Inform Bioinform. 2024;13(1):30. https://doi.org/10.1007/s13721-024-00466-1.
33. Kaur J, Kaur P. A CNN transfer learning-based automated diagnosis of COVID-19 from lung computerized tomography scan slices. New Gener Comput. 2023;41(4):795-838. https://doi.org/10.1007/s00354-023-00232-3.
34. Malik H, Anees T, Al-Shamaylehs AS, Alharthi SZ, Khalil W, Akhunzada A. Deep learning-based classification of chest diseases using X-rays, CT scans, and cough sound images. Diagnostics (Basel). 2023;13(17):2772. PMID: 37685310; PMCID: PMC10486427.
35. Keerthana R, Gladston A, Nehemiah HK. Transfer learning-based CNN diagnostic framework for diagnosis of COVID-19 from lung CT images. The Imaging Science Journal. 2022;70(7):413-38. doi:10.1080/13682199.2023.2170768.
36. Huang ML, Liao YC. A lightweight CNN-based network on COVID-19 detection using X-ray and CT images. Comput Biol Med. 2022;146:105604. PMID: 35576824; PMCID: PMC9090861.
37. Matsuyama E. A deep learning interpretable model for novel coronavirus disease (COVID-19) screening with chest CT images. J. Biomed.Sci. Eng. 2020;13(7):140-52. doi: 10.4236/jbise.2020.137014.
38. He X, Yang X, Zhang S, Zhao J, Zhang Y, Xing E, et al. Sample-efficient deep learning for COVID-19 diagnosis based on CT scans. Medrxiv 2020:1-9. doi: 10.1101/2020.04.13.20063941.
39. Yang S, Jiang L, Cao Z, Wang L, Cao J, Feng R, et al. Deep learning for detecting corona virus disease 2019 (COVID-19) on high-resolution computed tomography: a pilot study. Ann Transl Med. 2020;8(7):450. PMID: 32395494; PMCID: PMC7210135.
40. Mishra NK, Singh P, Joshi SD. Automated detection of COVID-19 from CT scan using convolutional neural network. Biocybern Biomed Eng. 2021;41(2):572-88. PMID: 33967366; PMCID: PMC8084624.
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. PMID: 34915332; PMCID: PMC8665658.
42. Bouzid HAI, Berrani SA, Yahiaoui S, Belaid A, Belazzougui D, Djouad M, et al. Deep learning-based COVID-19 diagnosis: a thorough assessment with a focus on generalization capabilities. J Image Video Proc. 2024;40. https://doi.org/10.1186/s13640-024-00656-x
43. Appavu N, Babu CNK, Kadry S. COVID-19 classification in X-ray/CT images using pretrained deep learning schemes. Multimed Tools Appl. 2024;83:83157-77. https://doi.org/10.1007/s11042-024-18721-y
44. Keshamoni K, Rao LK, Rao DS. A multi-modal CNN-based approach for COVID-19 diagnosis using ECG, X-Ray, and CT. International Journal of Advanced Computer Science&Applications (IJACSA). 2024;15(6):1102-12. http://dx.doi.org/10.14569/IJACSA.2024.01506112
45. Mousavi M, Hosseini S. A deep convolutional neural network approach using medical image classification. BMC Med Inform Decis Mak. 2024;24:239. https://doi.org/10.1186/s12911-024-02646-5
46. Rajpoot R, Gour M, Jain S, Semwal VB. Integrated ensemble CNN and explainable AI for COVID-19 diagnosis from CT scan and X-ray images. Sci Rep. 2024;14(1):24985. PMID: 39443548; PMCID: PMC11499875.
47. Tello-Mijares S, Flores F, Woo F. Classification of severity of lung parenchyma using saliency and discrete cosine transform energy in computed tomography of patients with COVID-19. Int J Telemed Appl. 2025;2025:4420410. PMID: 39816708; PMCID: PMC11729514.