

A Deep Learning-Based System to Assist Radiologists in Detecting COVID-19 Disease from Chest Computed Tomography Images

Oğuzhan KATAR^{1*} , Erkan DUMAN² 

Abstract

The COVID-19 pandemic has had a significant negative impact on the world in various ways. In an effort to mitigate the negative effects of the pandemic, this study proposes a deep learning approach for the automatic detection of COVID-19 from chest computed tomography (CT) images. This would enable healthcare professionals to more efficiently identify the presence of the virus and provide appropriate care and support to infected individuals. The proposed deep learning approach is based on binary classification and utilizes members of the pre-trained EfficientNet model family. These models were trained on a dataset of real patient images, called the EFSCHE-19 dataset, to classify chest CT images as positive or negative for COVID-19. The results of the predictions made on the test images showed that all models achieved accuracy values of over 98%. Among these models, the EfficientNet-B2 model performed the best, with an accuracy of 99.75%, sensitivity of 99.50%, specificity of 100%, and an F1 score of 99.75%. In addition to the high accuracy achieved in the classification of chest CT images using the proposed pre-trained deep learning models, the gradient-weighted class activation mapping (Grad-CAM) method was also applied to further understand and interpret the model's predictions.

Keywords: COVID-19, Computed Tomography, EfficientNet, Classification, Grad-CAM.

Göğüs Bilgisayarlı Tomografi Görüntülerinden COVID-19 Hastalığını Tespit Etmede Radyologlara Yardımcı Derin Öğrenme Tabanlı Bir Sistem

Öz

COVID-19 pandemisinin dünya üzerinde çeşitli şekillerde önemli bir olumsuz etkisi oldu. Pandeminin olumsuz etkilerini azaltmak amacıyla bu çalışma, göğüs bilgisayarlı tomografi (BT) görüntülerinden COVID-19'un otomatik tespiti için bir derin öğrenme yaklaşımı önermektedir. Bu, sağlık uzmanlarının virüsün varlığını daha verimli bir şekilde tanımlamasını ve enfekte bireylere uygun bakım ve destek sağlamasını sağlayacaktır. Önerilen derin öğrenme yaklaşımı ikili sınıflandırmaya dayanmaktadır ve önceden eğitilmiş EfficientNet model ailesinin üyelerini kullanmaktadır. Bu modeller, göğüs BT görüntülerini COVID-19 için pozitif veya negatif olarak sınıflandırmak için EFSCHE-19 veri seti adı verilen gerçek hasta görüntülerinden oluşan bir veri seti üzerinde eğitildi. Test görüntülerinde yapılan tahminlerin sonuçları, tüm modellerin %98'in üzerinde doğruluk değerlerine ulaştığını gösterdi. Bu modeller arasında EfficientNet-B2 modeli %99,75 doğruluk, %99,50 duyarlılık, %100 özgüllük ve %99,75 F1 skoru ile en iyi performansı gösterdi. Önerilen önceden eğitilmiş derin öğrenme modelleri kullanılarak göğüs BT görüntülerinin sınıflandırılmasında elde edilen yüksek doğruluğa ek olarak, modelin tahminlerini daha iyi anlamak ve yorumlamak için gradyan ağırlıklı sınıf aktivasyon eşleşme (Grad-CAM) yöntemi de uygulandı.

Anahtar Kelimeler: COVID-19, Bilgisayarlı Tomografi, EfficientNet, Sınıflandırma, Grad-CAM.

¹Firat University, Department of Software Engineering, Elazığ, Turkey, okatar@firat.edu.tr

²Firat University, Department of Computer Engineering, Elazığ, Turkey, erkanduman@firat.edu.tr

¹<https://orcid.org/0000-0002-5628-3543> ²<https://orcid.org/0000-0003-2439-7244>

1. Introduction

Throughout history, our world has encountered various pandemics, including the black death, cholera, and Spanish flu. Currently, the global community is grappling with the COVID-19 pandemic, which originated in Wuhan, Hubei province of China (Tekin, 2021). On December 31, 2019, the World Health Organization (WHO) was informed of a cluster of pneumonia cases of unknown cause. Samples were collected from the affected patients and, upon analyzing the genome sequences of these samples, the WHO announced on January 12, 2020 that a novel coronavirus, subsequently named SARS-CoV-2, was responsible for the outbreak and the associated disease, COVID-19 (Issever et al., 2020). SARS-CoV-2 is believed to have initially transmitted from animals to humans due to its zoonotic characteristics (Ahmad et al., 2020). The virus rapidly spread globally following its emergence in Wuhan, leading to the declaration of a pandemic on March 11, 2020 (Hua and Shaw, 2020). As of July 2022, over 580 million people have been infected worldwide, with more than 6 million deaths attributed to COVID-19 (URL-1). The most prominent symptoms of COVID-19 include fever, dry cough, respiratory issues, severe sore throat, and diarrhea. Other commonly reported symptoms include loss of taste and smell, joint pain, and nasal congestion. In severe cases, which comprise 2%-3% of all COVID-19 cases, multiorgan dysfunction can lead to death (Madabhavi et al., 2020). Although vaccines have proven to be beneficial, they are not a definitive solution to preventing COVID-19. There have been instances of vaccinated individuals contracting the disease (Wadman, 2021). Therefore, measures such as creating isolation between healthy and infected individuals through raising awareness and promoting the use of masks are important in the efforts to curb virus transmission. Currently, COVID-19 detection and differentiation between healthy and infected individuals rely on molecular detection methods and radiological imaging techniques (Giri et al., 2021)

Molecular detection methods, such as real-time reverse transcription polymerase chain reaction (RT-PCR), are used to identify the virus. RT-PCR, which is also utilized in the detection of the Middle East Respiratory Syndrome Coronavirus (MERS-CoV), detects the presence of the virus and identifies it by analyzing the ribonucleic acid (RNA) component of the virus (Corman et al., 2020). To perform the test, a cotton swab is used to collect throat and nasal samples from the individual. The duration of time required to obtain results from RT-PCR tests can range from one hour to several days, depending on the specific PCR version. The reliability of the RT-PCR test result is dependent on various factors, including the sample collection method, the types of primers and probes used, the inclusion of appropriate controls, and the temperature value (Wu et al., 2020). However, the RT-PCR test may be less efficient and reliable due to limitations such as variations in viral load rates used by different kit manufacturers, shortages in kit supply, and the number of skilled health professionals

who will use the kit (Zali, 2021). Therefore, it may be useful to utilize alternative methods in the early stages of the disease in addition to or instead of the RT-PCR test.

To address the limitations of the RT-PCR test, it is advisable to examine radiological images prior to the test (Kommos et al., 2020). The lungs are often significantly affected in COVID-19 and SARS-CoV-2 infection can cause severe, potentially permanent damage to lung tissue (Ghaderzadeh et al., 2021). Therefore, lung tissue is often used to detect or monitor COVID-19 disease. Chest radiography or chest computed tomography (CT) scanning are commonly employed methods for radiologically diagnosing COVID-19 (Ng et al., 2020). Lung imaging is often the initial approach recommended by medical and healthcare protocols due to its expediency and simplicity (Fields et al., 2021). CT devices, which have detailed anatomical resolution, have been frequently used since the onset of the pandemic to monitor the effects of COVID-19 (Gundel et al., 2021). Radiologists examine chest CT scans obtained using these devices and create a report with recommendations for the patient's diagnosis. This report creation process is crucial and is therefore carried out with great care. During the current pandemic, radiologists are faced with the task of interpreting an excessive number of radiological images for COVID-19 diagnosis, which exceeds the standard limits (Chartrand et al., 2017). As a result, the time required to interpret the patient's report and initiate appropriate treatment increases significantly. In addition, the reporting process becomes prone to errors due to human factors such as fatigue from extended work hours. Computers with high processing capacity and the ability to be unaffected by human factors can assist radiologists in this process (Arsalan et al., 2020).

Artificial intelligence (AI)-based systems have been widely utilized in the detection of various diseases and for monitoring treatment progress (Owais et al., 2019). Currently, research involving the use and development of deep learning models for the classification of medical images based on disease is common (Mainak et al., 2019). AI-based systems assist physicians in tasks that can greatly impact human life, such as the detection of stroke, the classification of patients with kidney stone issues, and the identification of cancer cells (Chin et al., 2017; Yildirim et al., 2021; Khan et al., 2019). During the COVID-19 pandemic, numerous AI-based COVID-19 detection studies have been proposed in order to minimize social and economic damage and support physicians, similar to other studies. The main aim of classification studies is to determine, based on available data, whether there are positive cases of COVID-19 disease.

Bogu and Snyder (2021) collected heart rate data from real individuals using a smartwatch device, with durations ranging from 15 seconds to 1 minute. They divided this data into two classes based on COVID-19 disease status: infected and non-infected. Using deep learning methods, they classified individuals with COVID-19 disease based on their heart rate with 92% accuracy.

Pahar et al. (2021) classified cough recordings of individuals with positive and negative COVID-19 diagnoses, collected using smartphones. The dataset used in their study included both natural and forced cough samples. The researchers found that cough recordings with a positive diagnosis of COVID-19 were 15-20% shorter than other recordings. When input into seven different machine learning classifiers, the cough recordings achieved the highest success rate, with 95.33% accuracy, when classified using the ResNet-50 classifier. In addition to wearable sensors and smartphone technology, radiological images are frequently used in studies related to the detection of COVID-19 disease.

Wang et al. (2020) proposed a model called COVID-Net, a deep convolutional neural network (CNN) specifically designed for the detection of COVID-19 cases from chest radiographs. The model was trained using 13,975 samples collected from publicly available datasets, which were subjected to pre-processing steps such as cropping, rotating, and zooming. COVID-Net was trained for 22 epochs using ImageNet weights and Adam optimization with a batch size of 64. In the testing stages, COVID-Net was compared to the VGG-19 and ResNet-50 models, which were trained with the same dataset, and was found to be the best performing model with an accuracy rate of 93.3%.

Luz et al. (2022) classified X-ray images containing COVID-19 findings using models in the EfficientNet family. The models were trained using publicly available datasets collected from various sources and using cross-dataset evaluation methods. The dataset, which includes samples from three classes: COVID-19 positive, pneumonia, and normal, consists of samples with different resolution values. Among the models trained for 20 epochs, the EfficientNet-B3 model was the best performing model with an accuracy rate of 93.9%.

Song et al. (2021) developed a system that can assist clinicians in detecting COVID-19-infected patients from CT images. Their model, called DRE-Net, was trained on CT images from two hospitals in China. When the model's performance was evaluated using performance metrics after training, the researchers reported that DRE-Net achieved 86% accuracy. The gradient-weighted class activation mapping (Grad-CAM) method was used to show which areas of the image the model focuses on when classifying.

Bozkurt (2022), presents a framework called HANDEFU for the automatic early diagnosis of COVID-19 using chest x-ray images. HANDEFU is an interactive software that allows the user to select from a library of feature extraction techniques and classification methods to build their own model for diagnosis. The authors evaluated the performance of 27 different models using HANDEFU on an open-access dataset for COVID-19 diagnosis. The model with the highest performance, local binary pattern (LBP)+support vector machine (SVM), achieved an accuracy of 99.36%. This study is noteworthy for its development of a single framework that incorporates image preprocessing and diverse feature extraction and classification techniques.

Marques et al. (2020) developed a CNN using the EfficientNet architecture for the purpose of COVID-19 diagnosis. The CNN was evaluated using a stratified 10-fold cross-validation approach with images of COVID-19 patients, pneumonia patients, and healthy individuals. The results of binary classification (COVID-19 versus normal) showed an average accuracy of 99.62%, recall of 99.63%, precision of 99.64%, and F1 score of 99.62%. The results of multi-class classification (COVID-19, pneumonia, and normal) showed an average accuracy of 96.70%, recall of 96.69%, precision of 97.54%, and F1 score of 97.11%. The authors suggest that this CNN model could serve as a medical decision support system to assist healthcare professionals in COVID-19 diagnosis.

Wang et al. (2021) presented a method for detecting COVID-19 disease from chest CT images using a deep learning algorithm. Chest CT images from three different hospitals were labeled as 325 positives for COVID-19 and 740 negatives for COVID-19 by two radiologists, resulting in a dataset of 1065 total samples, which were sized at 299×299 px. After the learning rate was set at 0.01, the M-inception model was trained for 15,000 epochs. Upon completing performance evaluation, the researchers reported that the model achieved 89.5% accuracy.

These classification studies aim to interpret relevant radiological images quickly and reliably, regardless of human factors. In literature studies, deep learning models are typically trained using publicly available datasets. As a result, issues such as mislabeled images and unbalanced distributions in these datasets are often overlooked. To avoid the negative effects of public datasets, a special dataset was created that contains real patient scans. This dataset, called EFSCHE-19, includes chest CT samples with both positive and negative COVID-19 findings, and does not contain any personal information from any patient.

The primary objective of this study is to use a deep learning model to automatically classify COVID-19 and normal findings on chest CT images. The goal of this is to facilitate the early detection of the disease and to provide necessary information to physicians. The main contributions of this study are as follows:

- Evaluation of the performance of pre-trained EfficientNet models on a novel COVID-19 dataset.
- Demonstration of high classification performance of chest CT images on a uniformly distributed dataset without data augmentation.
- Use of the Grad-CAM algorithm to visualize predictions and increase confidence in the clinical use of deep learning models.
- Utilization of deep learning modules to perform tasks currently carried out manually by experts, resulting in reduced time and margin of error.

The remainder of this paper is organized as follows: Section 2 presents the proposed method for this study, including the dataset used, pre-trained deep learning models, training phases, pre-

processing, classification performance measures, and the Grad-CAM algorithm. Section 3 presents the parameters and environments used in the training phase, the numerical values of the model during the training phase, the test phase predictions, and performance values. The discussion and conclusion sections of the study are presented in Section 4 and Section 5, respectively.

2. Materials and Methods

A classifier system has been designed to use a deep learning model to predict whether chest CT images are positive or negative for COVID-19 using the binary classification method. Figure 1 illustrates the block representation of the proposed approach used in this study.

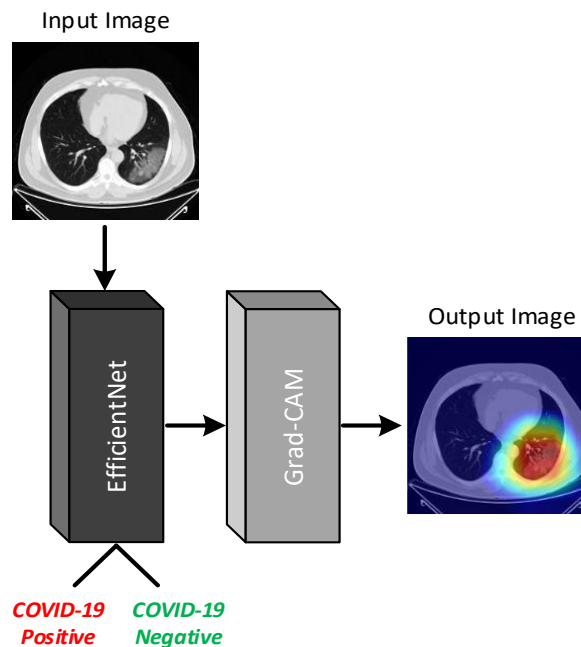


Figure 1. Block representation of the proposed approach.

The input layer of the deep learning model receives an image, which is then classified as positive or negative for COVID-19 disease in the output layer. A key factor in the success of this classification process is the availability of a well-trained deep learning model. To achieve a high success rate, it is essential to use an appropriate dataset for training the model. One option is to use publicly available datasets, which can be obtained at low cost and with minimal time investment. However, such datasets may suffer from issues such as uneven class distribution, mislabeled images, a narrow range of samples, and the use of different radiological scanning techniques, which can hinder the training of the model to the desired level of performance. To address these concerns, the present study used the EFSCCH-19 dataset for training, validation, and testing. This dataset is not included in the examples of publicly available datasets that are commonly used in the literature.

2.1. EFSCH-19 Dataset

This dataset was created with the approval of the ethics committee of Fırat University, Turkey. It is a binary classification dataset, comprising two classes labeled as COVID-19 positive and COVID-19 negative. The dataset includes a total of 4,000 chest CT images, with 2,000 samples in each class. All images have a resolution of 768×768px and are in the Digital Imaging and Communication in Medicine (DICOM) data format. To protect patient privacy, the layers containing personal information were removed from the images before they were provided to us. In radiological images, COVID-19 findings can be classified into two main categories: typical and atypical (Comert and Kiral, 2020). Typical findings include ground glass opacity, consolidation, and air bubbles. The EFSCH-19 dataset includes chest CT images of real patients with typical findings labeled as positive. These samples and the findings they contain have been verified by a radiologist. Figure 2 shows a random selection of samples from the EFSCH-19 dataset.

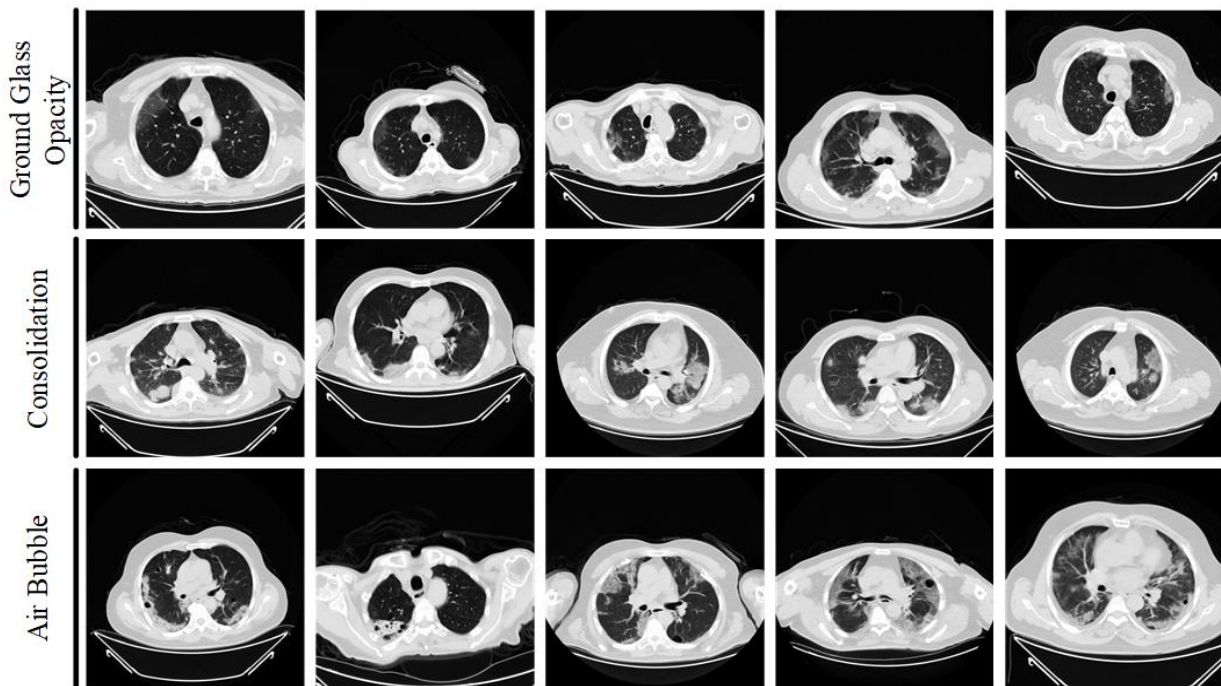


Figure 2. The EFSCH-19 dataset samples.

The radiologist's opinion was used to determine and label the classes to which the dataset samples belong. In the COVID-19 positive class, there are more images with the ground glass opacity finding than with other findings. However, care was taken to ensure that there is an equal number of samples between the two classes. The dataset includes chest CT images of healthy individuals labeled as COVID-19 negative. Figure 3 shows a selection of samples from the EFSCH-19 dataset with the COVID-19 negative label.

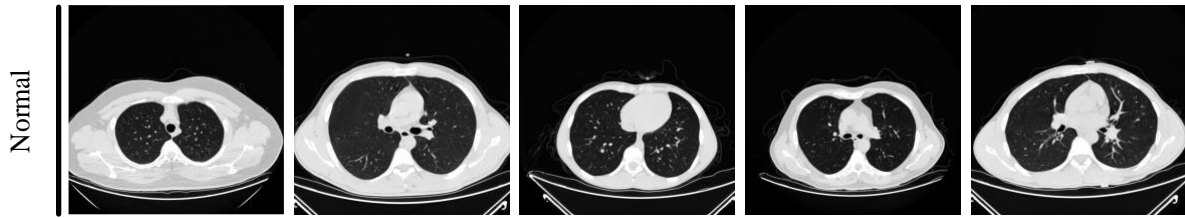


Figure 3. The normal chest CT images.

2.2. EfficientNet Models

CNNs are designed with a predetermined resource allocation and the number of layers is increased as additional resources become available to enhance accuracy. Increasing the number of layers and the depth or width of the network can also enhance accuracy by using a higher input resolution for training and evaluation. However, these techniques may not be adequate in terms of performance. Tan and Le (2019) proposed a new model scaling method that uses a combined coefficient to scale CNNs in a structured manner. This method employs fixed size scaling rather than traditional approaches such as increasing depth, width, or input resolution. As a result, the EfficientNet model family was developed. This family consists of eight models, named EfficientNet-B0 to EfficientNet-B7, with EfficientNet-B0 serving as the base model and the other models being scaled versions of it. Table 1 presents the architecture of the EfficientNet-B0 model. The basic network structure block diagram of the EfficientNet-B0 model is depicted in Figure 4.

Table 1. The architecture of EfficientNet-B0 (Tan and Le, 2019)

Stage	Resolution	Resolution	Channels	Layers
1	Conv3×3	224×224	32	1
2	MBCConv1,3×3	112×112	16	1
3	MBCConv6,3×3	112×112	24	2
4	MBCConv6,5×5	56×56	40	2
5	MBconv6,3×3	28×28	80	3
6	MBCConv6,5×5	14×14	112	3
7	MBCConv6,5×5	14×14	192	4
8	MBCConv6,3×3	7×7	320	1
9	Conv1×1 & Pooling & FC	7×7	1280	1

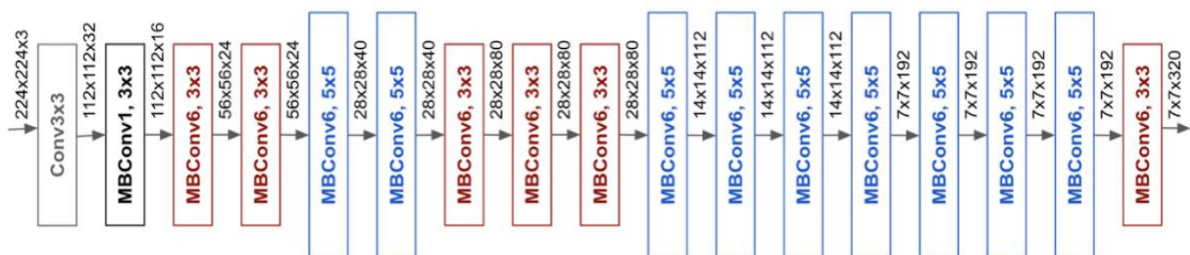


Figure 4. Block diagram of the EfficientNet-B0.

The EfficientNet model family consists of several models that have been derived using various scaling methods. In addition to the basic model, these models differ in the size of their input layers and the number of parameters they possess. The scaling methods responsible for these differences are depicted in Figure 5.

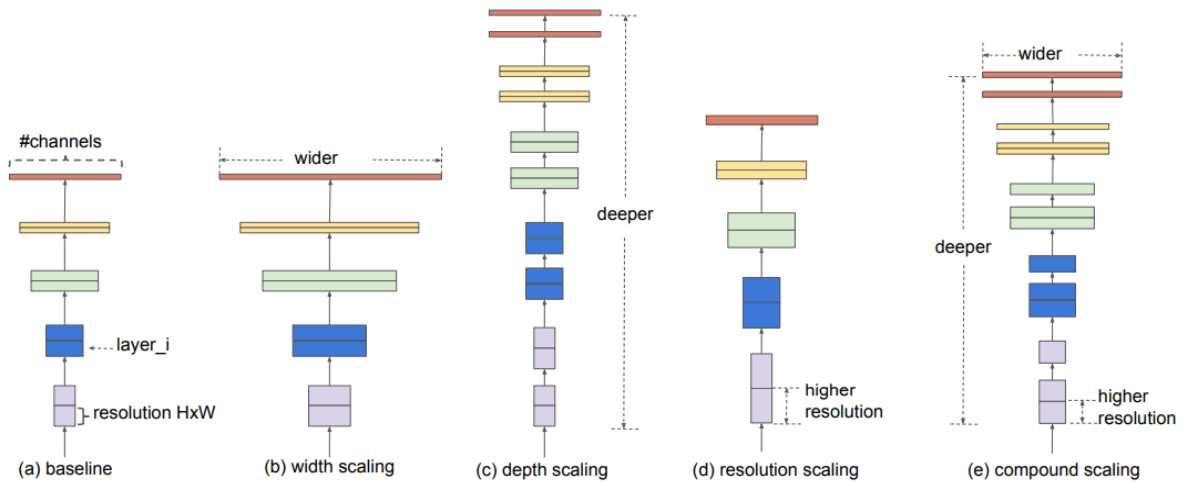


Figure 5. The Scaling methods for EfficientNet Models (Tan and Le, 2019).

The EfficientNet models are derived using scaling methods that involve increasing the depth, width, and resolution of the network, while simultaneously adjusting the input size and number of parameters to maintain constant efficiency. The basic model in the EfficientNet model family, along with the variations derived from it, are outlined in Table 2.

Table 2. The EfficientNet models' details (URL-2).

Model Name	Input Size	Dropout Rate	Depth Coefficient	Width Coefficient	Parameters
EfficientNet-B0	224×224	0.2	1.0	1.0	5.3 Million
EfficientNet-B1	240×240	0.2	1.1	1.0	7.9 Million
EfficientNet-B2	260×260	0.3	1.2	1.1	9.2 Million
EfficientNet-B3	300×300	0.3	1.4	1.2	12.3 Million
EfficientNet-B4	380×380	0.4	1.8	1.4	19.5 Million
EfficientNet-B5	456×456	0.4	2.2	1.6	30.6 Million
EfficientNet-B6	528×528	0.5	2.6	1.8	43.3 Million
EfficientNet-B7	600×600	0.5	3.0	2.0	66.7 Million

EfficientNet models have demonstrated state-of-the-art performance on various image classification benchmarks, including the ImageNet dataset (Deng et al., 2009). The ImageNet dataset, comprising 1.2 million training images and 50,000 validation images from 1000 classes, has consistently shown top results for EfficientNet models in the annual ImageNet Large Scale Visual Recognition Challenge (ILSVRC) (Russakovsky et al., 2015). Overall, EfficientNet models have exhibited high accuracy on the ImageNet dataset while also being comparatively efficient in terms of

the number of parameters and computational resources required. As such, EfficientNet models were employed as the classifier deep learning model in this study.

2.3. Training Phases

The chest CT images in the EFSCH-19 dataset cannot be utilized directly in the training phase due to their resolution and format. Therefore, it is necessary to apply pre-processing steps to the samples in the EFSCH-19 dataset. Following the pre-processing steps, all EfficientNet-B0 to EfficientNet-B7 models were trained while maintaining the parameters constant in order to determine the most suitable model for our study. Once the training phases have been completed, various performance metrics are examined and the most appropriate EfficientNet model to be used is determined. The block diagram of the method employed for training the models is provided in Figure 6.

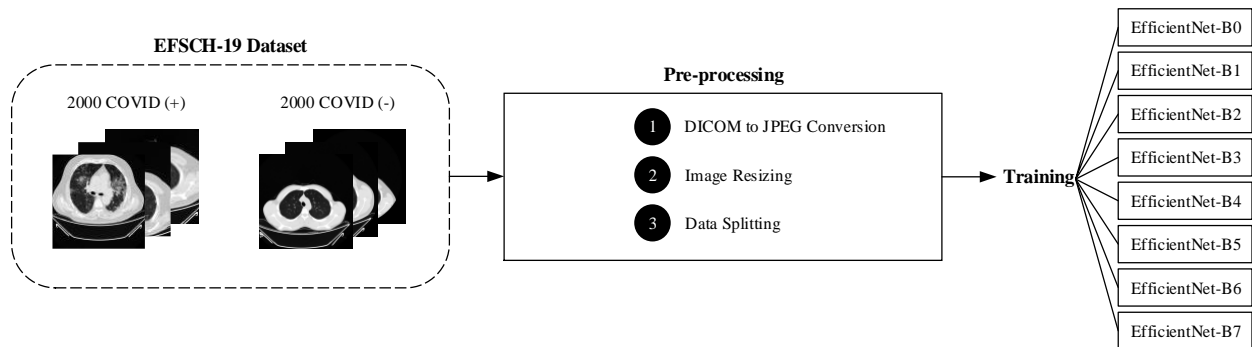


Figure 6. The proposed training method.

2.4. Pre-processing

All images in the EFSCH-19 dataset, which are in the DICOM data format, were converted to the JPEG format using the ImageJ application (Collins, 2007). This conversion process was carried out in a nearly lossless manner. Figure 7 illustrates the result obtained before and after the conversion process, with four times zoom.

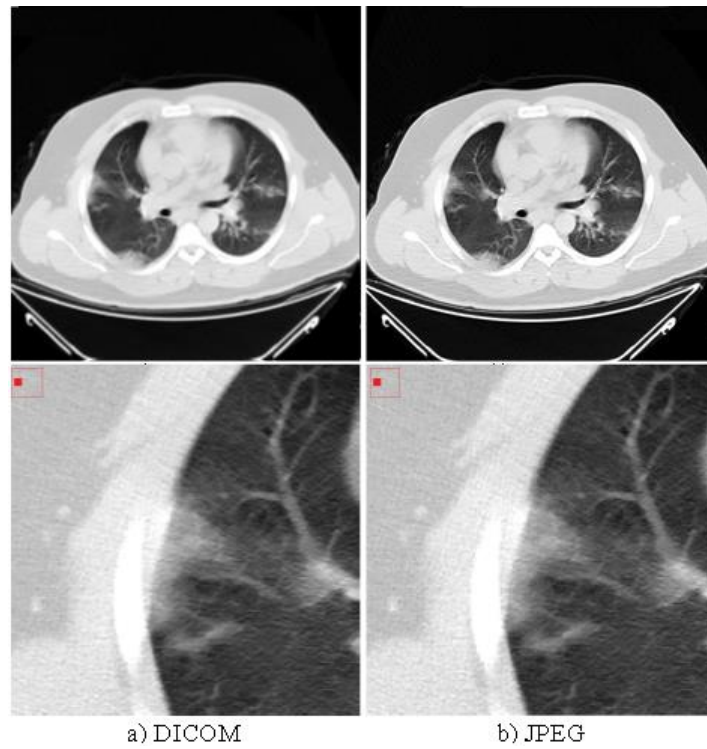


Figure 7. DICOM to JPEG conversion.

All samples were converted to the JPEG data format, and resizing was performed. Images with a resolution of 768×768 px were resized according to the input size of the model to be included in the training. The resizing process is illustrated in Figure 8.

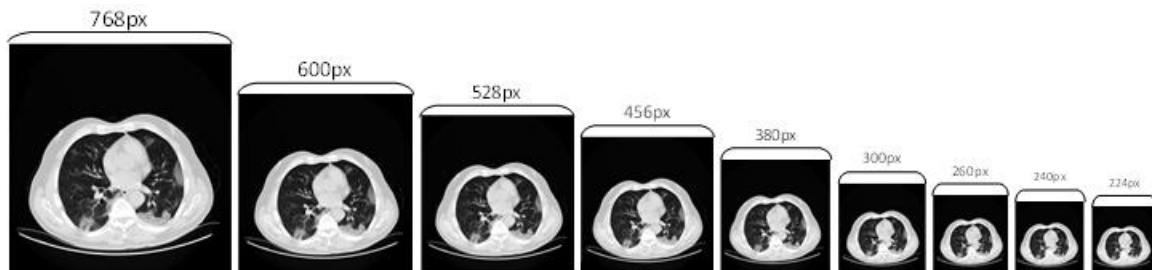


Figure 8. The resizing process.

Prior to initiating the training of the relevant models, 70% of the resized EFSCH-19 dataset samples were randomly divided for use in the training, validation, and testing phases, with 20% allocated for validation and 10% for testing. Detailed numerical information on the distribution of the dataset samples after data splitting is provided in Table 3. This numerical information is independent of the name of the trained model and applies to all eight models.

Table 3. Numerical distribution after the data splitting.

Phase	COVID-19 Positive Samples	COVID-19 Negative Samples
Train	1400	1400
Validation	400	400
Test	200	200

2.5. Evaluation Metrics

Confusion matrix-based performance measures are utilized to evaluate the performance of the deep learning model in classification studies. The confusion matrix provides information regarding the relationship between the class label predicted by the deep learning model for the input image and the known, actual class label of the input image. Deep learning models suitable for binary classification can only predict two different classes at their output, thus up to four different situations can arise in class prediction. These situations are as follows:

- True Positive (TP) is when an image with the COVID-19 positive label is predicted as COVID-19 positive.
- False Positive (FP) is when an image with the COVID-19 negative label is predicted as COVID-19 positive.
- False Negative (FN) is when an image with the COVID-19 positive label is predicted as COVID-19 negative.
- True Negative (TN) is when an image with the COVID-19 negative label is predicted as COVID-19 negative.

The confusion matrix can be visually expressed by placing the values of TP, FP, FN, and TN in a 2×2 matrix. Figure 9 illustrates the visual structure of the confusion matrix and the order in which the corresponding values are placed.

		Actual Class	
		Positive	Negative
Predicted Class	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

Figure 9. Structure of the confusion matrix.

If the number of examples of TP and TN cases is higher, it will indicate that the classification ability of the deep learning model is advanced. Several performance metrics have been standardized,

using the values of TP, FP, FN, and TN, to measure classification capability. These metrics and their corresponding mathematical equations are listed below.

$$\text{Accuracy (Acc)} = (TP + TN)/(TP + FP + FN + TN) \quad (1)$$

$$\text{Recall (Rec)} = TP/(TP + FN) \quad (2)$$

$$\text{Precision (Pre)} = TP/(TP + FP) \quad (3)$$

$$\text{Specificity (Spe)} = TN/(TN + FP) \quad (4)$$

$$\text{F1 Score (F1)} = (2 * (\text{Precision} * \text{Recall})) / ((\text{Precision} + \text{Recall})) \quad (5)$$

In addition to standardized statistical evaluation metrics, it is also important to compare the areas of focus for the deep learning model in its predictions.

2.6. Gradient-weighted Class Activation Mapping

The Grad-CAM algorithm visualizes, using the heat map technique (Selvaraju et al., 2017), the areas of the image that the classifier deep learning model considers when making a class prediction for the image provided as input. In this way, it can be determined whether the areas on which the classifier deep learning model focuses in correctly predicted images are, in fact, the areas that play an active role in determining the class. The proposed approach to the Grad-CAM algorithm is shown in Figure 10.

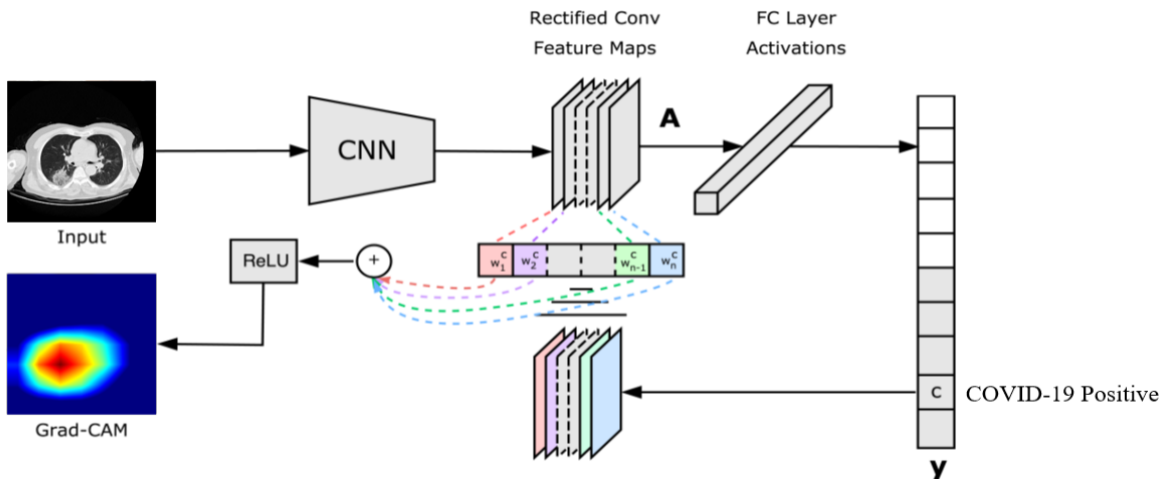


Figure 10. The proposed approach to the Grad-CAM algorithm.

The Grad-CAM algorithm operates by referencing the feature maps produced by the last convolution layer of a CNN. Therefore, the last convolution layer of the classifier deep learning model

should be referenced using the Grad-CAM algorithm. After referencing the last convolution layer, a heat map is created using gradients to highlight the important points of the class label in the image. The yellow and red areas in the generated heat maps indicate the pixel areas that the classifier deep learning model focuses on when making predictions.

3. Experimental Results

This section presents the results of EfficientNet models trained for the detection of COVID-19 disease from chest CT images. The following sections provide the evaluation metric values for the experimental findings.

3.1. Experimental Setups

In this study, members of the EfficientNet model family were used for training a deep learning model, utilizing the Keras library implemented in Python. The models included in the training utilized ImageNet weights rather than random starting weights, and the final layers of the models were modified to produce output in the range of 0 to 1, with the softmax activation function applied (The summary of the model is available at: <https://github.com/oguzhankatar/EfficientNet-B2.git>). The Adam optimization algorithm was utilized for training, with the early stopping function active for 50 epochs, a constant batch size of 16 and learning rate of 0.001, using the cross-entropy loss as the value to be minimized. The validation accuracy rate was tracked during training, and the early stop function was activated if the validation accuracy did not surpass the highest value for five consecutive epochs. These training processes were conducted in the Google Colab environment, with the aim of achieving the highest possible classification accuracy.

3.2. Results

The use of an early stopping function aims to prevent overfitting in machine learning models. Overfitting occurs when a model is too complex and performs exceptionally well on the training data, but fails to generalize to new, unseen data. The early stopping function monitors the performance of a model on a validation set during training and stops the training process if the model's performance on the validation set begins to decrease while the performance on the training set continues to improve. This indicates that the model is starting to overfit and the early stopping function ensures that the final model is the one that performed the best on the validation set.

In this study, the training of each model was set to a maximum of 50 epochs, but the early stopping function terminated the training at earlier epochs. The weights were saved following the training phases and the loss graphs for each model are provided in Figure 11. The analysis of loss graphs can facilitate the comprehension of the efficacy of a model's learning process and the identification of potential issues impacting the model's performance. As the model improves its accuracy in predicting outputs, a decrease in the loss value is anticipated during training. Conversely, an absence of decrease or an increase in the loss value may indicate ineffective learning and warrant further investigation.

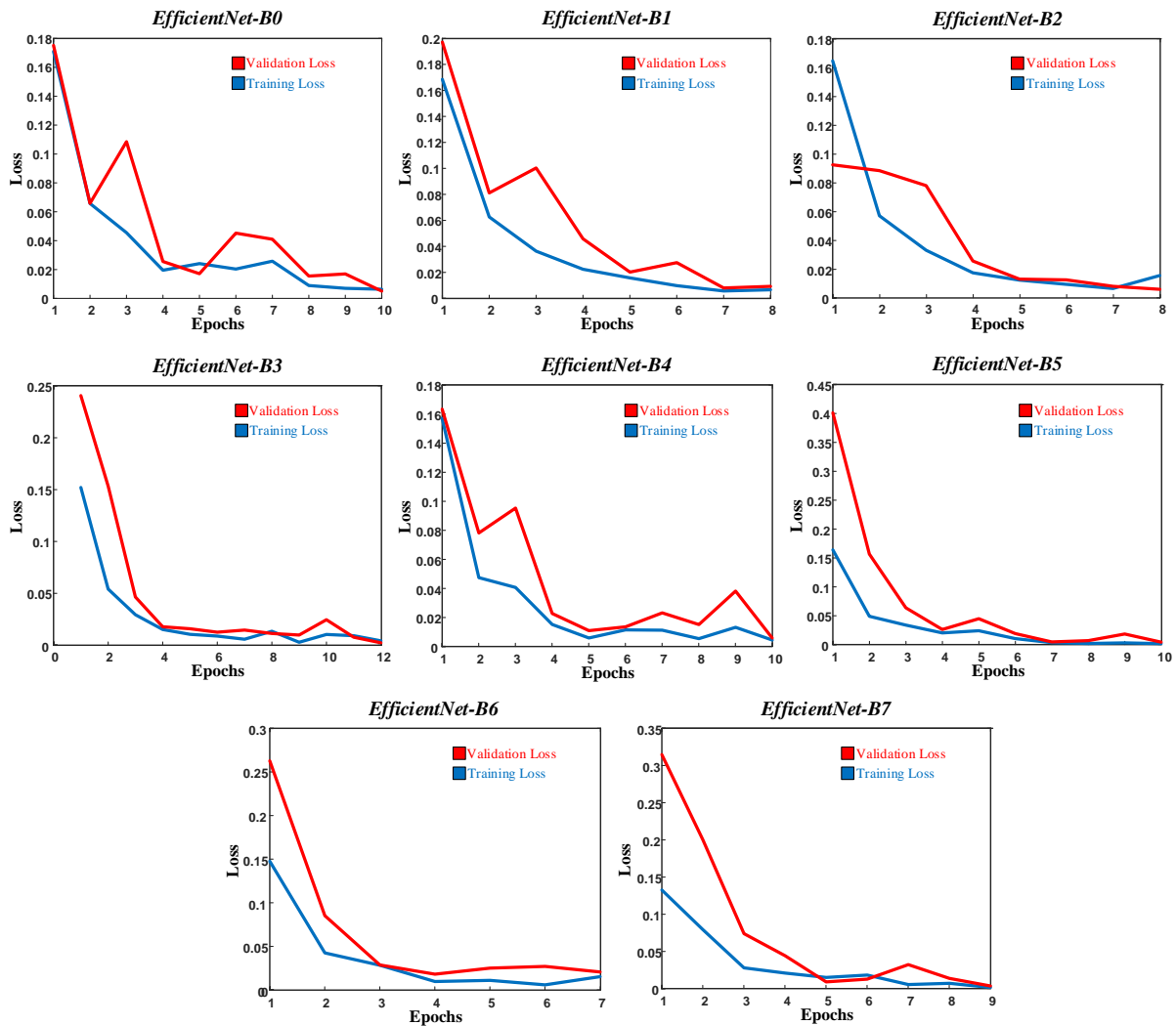


Figure 11. The loss graphs of each model.

The loss graphs of each model were examined and it was observed that the learning process occurred, as the loss values decreased. Therefore, no further research was necessary. The accuracy of the model on the training data is represented by the training accuracy, while the accuracy of the model on unseen data, also known as the validation data, is represented by the validation accuracy. Ideally, the training accuracy should be high and the validation accuracy should be close to the training

accuracy, indicating that the model is generalizing well to unseen data. If the training accuracy is significantly higher than the validation accuracy, this may indicate overfitting, where the model has learned patterns in the training data that do not generalize to unseen data. On the other hand, if the training accuracy is much lower than the validation accuracy, this may suggest underfitting, where the model is not learning enough from the training data. In Figure 12, the training accuracy and validation accuracy graphs obtained by each model during the training phase are shown.

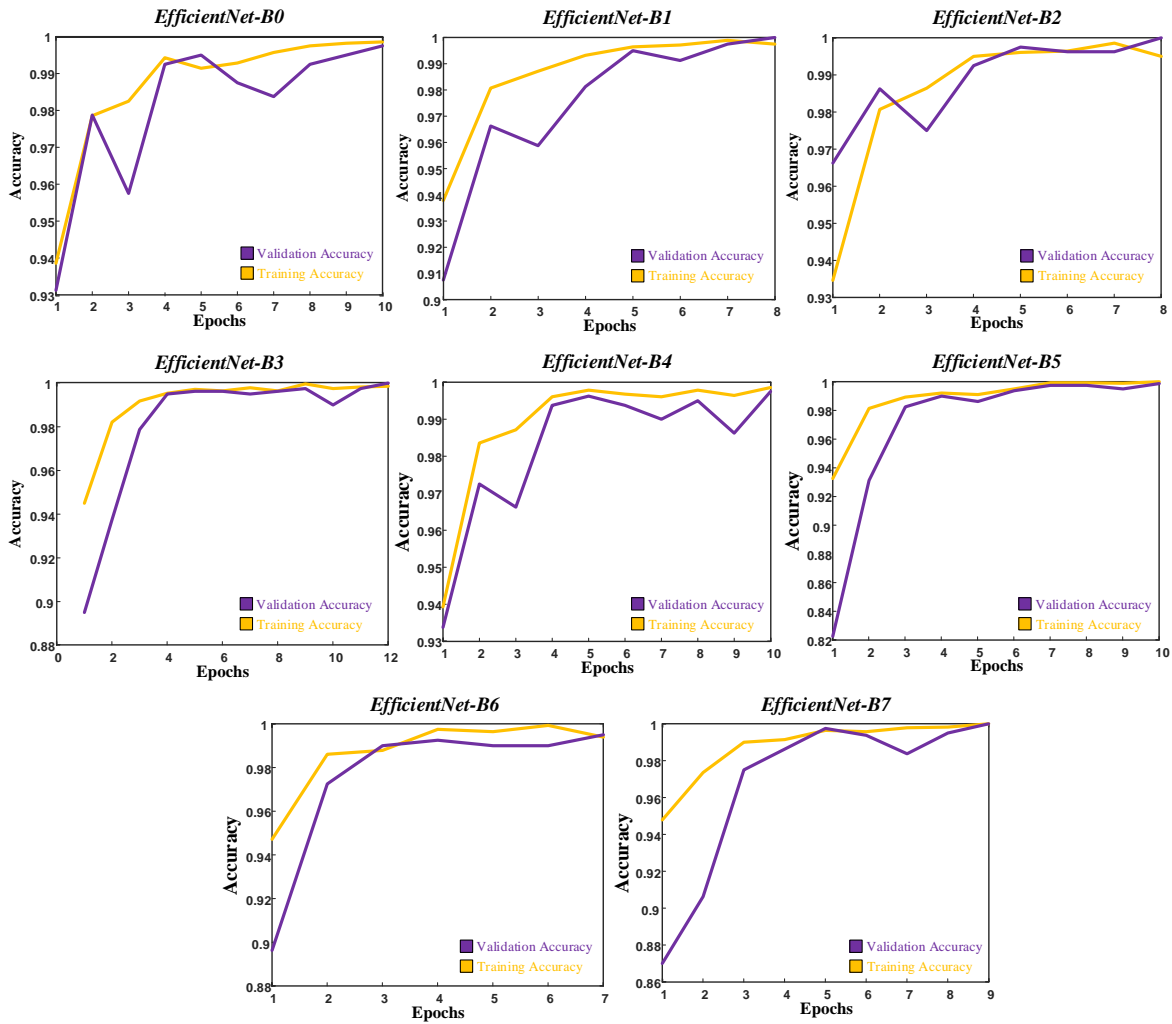


Figure 12. The accuracy graphs of each model.

Upon examining the accuracy graphs of the models, it was observed that they performed well on both the training data and unseen data. This indicates that the models have learned patterns in the training data that generalize well to unseen data, and are therefore able to accurately predict the output for new input data. This is generally a good indication of a well-performing model. However, in order to fully assess the performance of the models, input was only given to each model using images reserved for use in the testing phase. Each model made predictions for the corresponding test images. The confusion matrices resulting from these predictions are shown in Figure 13.

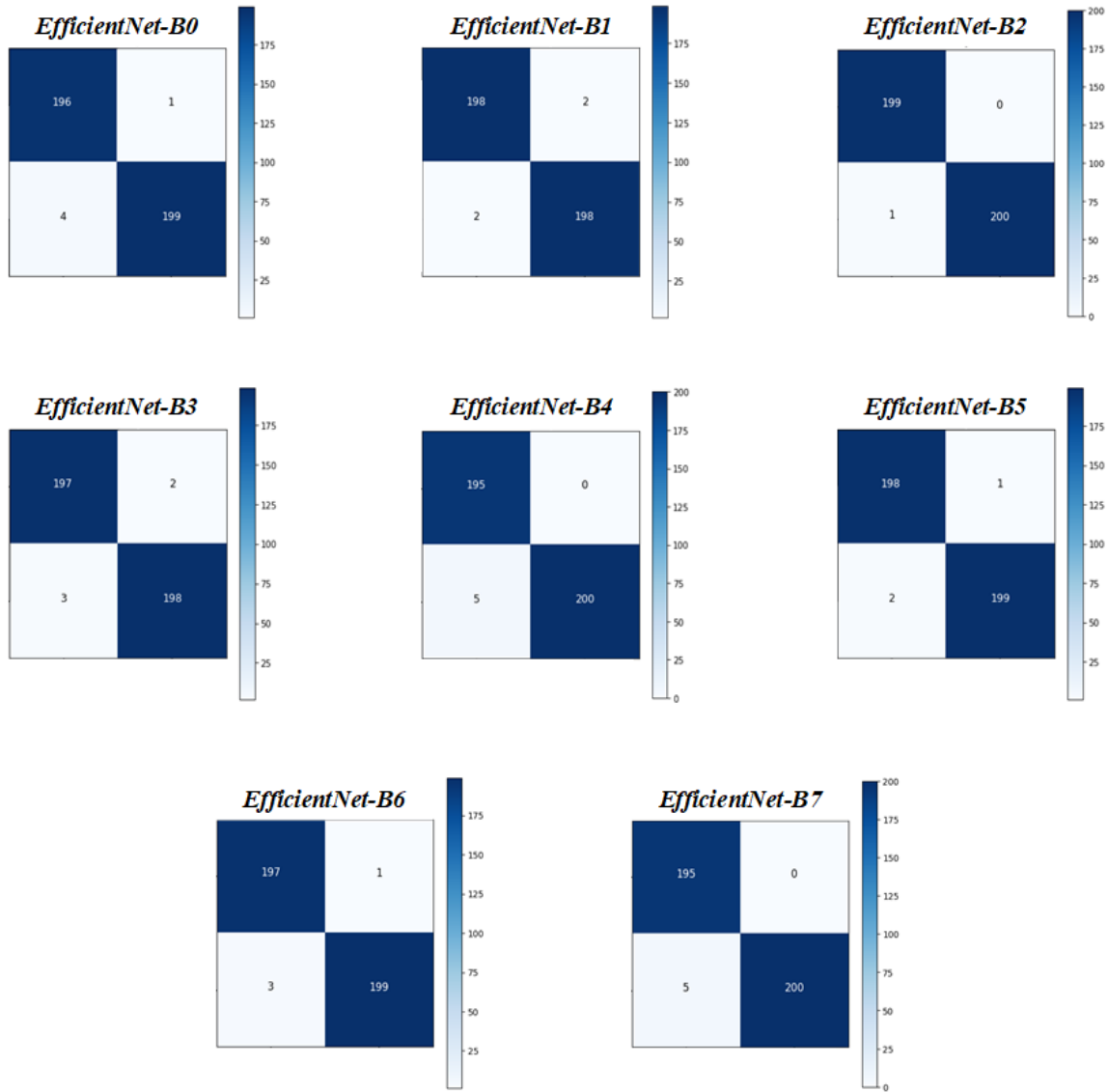


Figure 13. The confusion matrices of each model.

Upon examination of the predictions, it appears that the EfficientNet models are able to classify chest CT images as either COVID-19 positive or COVID-19 negative based on the findings they contain. The performance metric values achieved by the relevant models during the testing phase are presented in Table 4.

Table 4. The performance values of EfficientNet models.

Model Name	Accuracy	Recall	Precision	Specificity	F1 Score
EfficientNet-B0	0.9875	0.9800	0.9949	0.9950	0.9874
EfficientNet-B1	0.9900	0.9900	0.9900	0.9900	0.9900
EfficientNet-B2	0.9975	0.9950	1.00	1.00	0.9975
EfficientNet-B3	0.9875	0.9850	0.9899	0.9900	0.9875
EfficientNet-B4	0.9875	0.9750	1.00	1.00	0.9873
EfficientNet-B5	0.9925	0.9900	0.9950	0.9950	0.9925
EfficientNet-B6	0.9900	0.9850	0.9949	0.9950	0.9899
EfficientNet-B7	0.9875	0.9750	1.00	1.00	0.9873

Performance metrics, such as accuracy, precision, and recall, can be useful for evaluating the overall performance of a CNN, but they do not provide detailed information on how the CNN is arriving at its predictions. By visualizing the regions in the input that are most pertinent to a particular prediction, the Grad-CAM algorithm enables us to comprehend the specific features. This can be beneficial for debugging and improving the performance of the CNN, as well as for comprehending the internal functioning of the CNN. The Grad-CAM algorithm was applied to randomly selected samples from the test images during the prediction process, and the results are depicted in Figure 14.

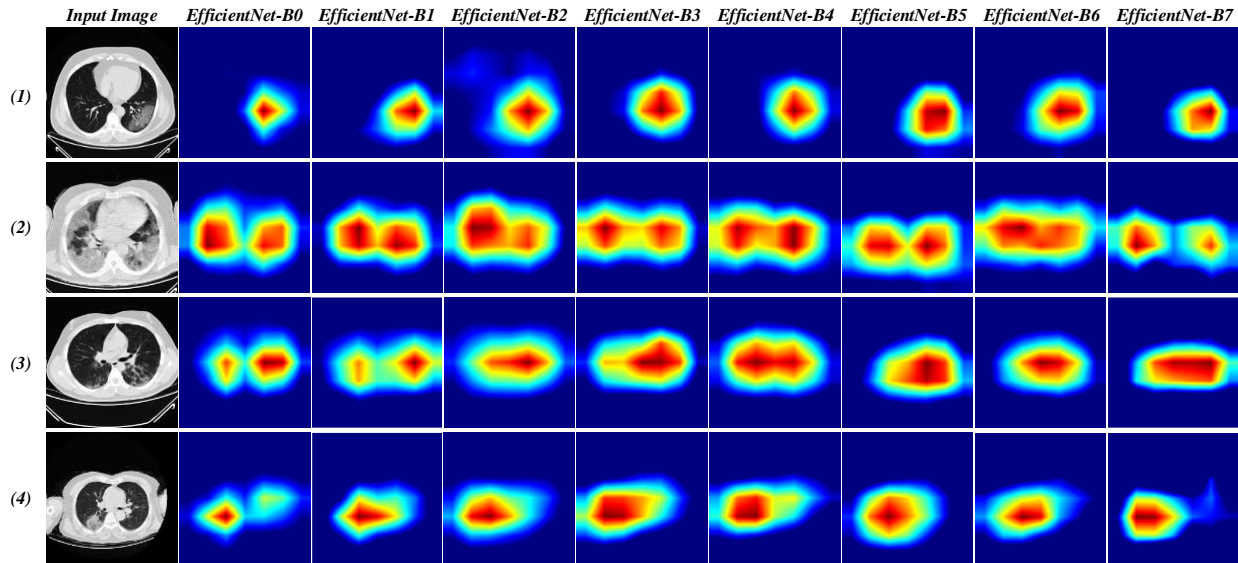


Figure 14. The Grad-CAM outputs of each model.

4. Discussion

Radiological images, such as CT scans and X-rays, can be used to detect COVID-19. However, detecting COVID-19 from radiological images can be challenging because the disease can present differently in different individuals and may not always produce visible abnormalities on images. Furthermore, the manual detection process is very time- and resource-intensive. Thus, it is important to be able to automatically detect COVID-19 without human intervention. Consequently, automatic detection of COVID-19 is a highly popular research topic, and there have been numerous deep learning-based studies on the subject using radiological images. Table 5 presents a curated selection of such studies. Ismael and Sengur (2021) employed deep learning approaches for the classification of COVID-19 and normal chest X-ray images. These approaches included deep feature extraction using pretrained CNNs such as ResNet18, ResNet50, ResNet101, VGG16, and VGG19, fine-tuning of these pretrained CNNs, and end-to-end training of a new CNN model. The deep features were classified using the SVM classifier with various kernel functions, including Linear, Quadratic, Cubic, and Gaussian. A dataset comprising 180 COVID-19 and 200 normal chest X-ray images was utilized

for experimentation. The results demonstrated the potential of deep learning in detecting COVID-19 based on chest X-ray images, with the highest accuracy score of 92.6% achieved by the fine-tuned ResNet50 model. Ozturk et al. (2020) proposed a new model for automatic COVID-19 detection using raw chest X-ray images for both binary (COVID vs. No-Findings) and multi-class (COVID vs. No-Findings vs. Pneumonia) classification. The model, which employed the DarkNet model as a classifier for the YOLO real-time object detection system and featured 17 convolutional layers with diverse filtering on each layer, achieved a classification accuracy of 98.08% for binary classification and 87.02% for multi-class classification. Shah et al. (2021) utilized a self-developed CNN model called CTnet-10 and several other CNN models, including DenseNet-169, VGG-16, ResNet-50, InceptionV3, and VGG-19, for the early diagnosis of COVID-19 using CT scan images. Among the models tested, VGG-19 demonstrated the highest accuracy of 94.52% in differentiating COVID-19 and non-COVID-19 CT scan images. Rahimzadeh et al. (2021) presented a fully-automated method for detecting COVID-19 from chest CT scan images. The method was evaluated using a new dataset consisting of 48,260 CT scan images from 282 normal individuals and 15,589 images from 95 COVID-19 patients. The proposed system was tested using Xception, ResNet50V2, and a novel model for single image classification and patient condition identification. The ResNet50V2 model achieved an accuracy of 98.49% on test images. Gupta et al. (2021) proposed an artificial intelligence model called InstaCovNet-19 for detecting COVID-19 and pneumonia in chest X-ray images. The model is an integrated stacked deep convolution network that leverages various pre-trained models, including ResNet101, Xception, InceptionV3, MobileNet, and NASNet, to compensate for a limited amount of training data. In evaluating the model on 3-class (COVID-19, pneumonia, normal) and 2-class (COVID, non-COVID) classification tasks, the authors achieved accuracies of 99.08% and 99.53%, respectively. Wang et al. (2021) developed an algorithm for detecting COVID-19 by collecting CT images of pathogen-confirmed COVID-19 cases and those diagnosed with typical viral pneumonia, and modifying the inception transfer-learning model. In an internal validation on the dataset, the algorithm demonstrated an accuracy of 89.5%. Katar and Duman (2021) developed a deep learning-based model for detecting COVID-19 from chest CT images using a dataset comprising 800 positive and 800 normal images. The model was evaluated on a test set of 400 randomly selected images and achieved an accuracy of 97.5%. Gupta and Bajaj (2023) proposed a framework for the automated screening of COVID-19 using chest CT-scan images and deep learning techniques. The framework leverages a publicly available CT-scan image dataset containing 1252 COVID-19 and 1230 non-COVID images, as well as two pre-trained deep learning models (MobileNetV2 and DarkNet19) and a newly designed lightweight deep learning model. The authors employed a repeated 10-fold holdout validation method for training, validation, and testing of the models and achieved the highest classification accuracy of 98.91% using transfer-learned DarkNet19. Zhou et al. (2021)

proposed an ensemble deep learning model (EDL_COVID) for rapid detection of COVID-19 from CT images. The model was trained on a dataset comprising 2500 CT images of COVID-19, 2500 images of lung tumors, and 2500 images of normal lungs, which were obtained from various sources. The model consisted of three deep CNN models (AlexNet, GoogleNet, and ResNet) initialized using transfer learning, which were used for feature extraction on all images, followed by a fully connected layer with a softmax classification algorithm. The authors constructed three component classifiers (AlexNet_Softmax, GoogleNet_Softmax, and ResNet_Softmax) and obtained the ensemble classifier EDL_COVID through the method of relative majority vote. The authors evaluated the performance of the ensemble model and found it had an accuracy of 99.05%.

Table 5. Comparison of Our Work With Some State-of-the-art Studies.

Study	Dataset Type	Number of Images	Model	Explainability	Performance
Ismael and Sengur (2021)	X-ray	1) <u>COVID-19</u> : 180 2) <u>Normal</u> : 200	ResNet-50	Black-box	Acc=92.63%
Ozturk et al. (2020)	X-ray	1) <u>COVID-19</u> : 125 2) <u>No-findings</u> : 500	DarkCovidNet	Grad-CAM	Acc=98.08%
Shah et al. (2021)	X-ray	1) <u>COVID-19</u> : 349 2) <u>Non-COVID-19</u> : 463	VGG-19	Black-box	Acc=94.52%
Rahimzadeh et al. (2021)	X-ray	1) <u>COVID-19</u> : 15,589 2) <u>Non-COVID-19</u> : 48,260	ResNet50V2	Black-box	Acc=98.49%
Gupta et al. (2021)	X-ray	1) <u>COVID-19</u> : 361 2) <u>Pneumonia</u> : 362 3) <u>Normal</u> : 365	InstaCovNet-19	Grad-CAM	Acc=99.08%
Wang et al. (2021)	CT	1) <u>COVID-19 Positive</u> : 325 2) <u>COVID-19 Negative</u> : 740	M-Inception-V3	Black-box	Acc=89.50%
Katar and Duman (2021)	CT	1) <u>COVID-19</u> : 1000 2) <u>Normal</u> : 1000	CNN	Black-box	Acc=97.50%
Gupta and Bajaj (2023)	CT	1) <u>COVID-19</u> : 1252 2) <u>Non-COVID-19</u> : 1230	DarkNet19	Black-box	Acc=98.91%
Zhou et al. (2021)	CT	1) <u>COVID-19</u> : 2500 2) <u>Normal</u> : 2500 3) <u>Tumor</u> : 2500	EDL_COVID	Black-box	Acc=99.05%
The proposed our study	CT	1) <u>COVID-19 Positive</u> : 2000 2) <u>COVID-19 Negative</u> : 2000	EfficientNet-B2	Grad-CAM	Acc=99.75%

In this study, we proposed the use of a pre-trained EfficientNet-B2 model for the automated detection of COVID-19 from chest CT images. The results of our study demonstrated a high accuracy of 99.75%, which exceeded the accuracy reported in most other studies. This demonstrates the ability of the model in our study to accurately classify the images into COVID-19 positive and COVID-19 negative categories with a high level of accuracy. The dataset used in this study was novel, consisting of 4,000 images, which may have contributed to the robustness and generalizability of the model. To provide insights into the model's decision-making process and identify any potential biases or weaknesses, we utilized the explainability method of Grad-CAM. In this study, the publicly shared dataset (COVID-CT-dataset) samples provided by Wang et al. (2020) were used to verify the validity

of the proposed method. All 349 COVID-19 and 397 Non-COVID CT images were used as input to the EfficientNet models trained with the EFSCH-19 dataset samples. The performance values resulting from the classification of 746 images by each model are presented in Table 6. The graphical representation of the values given in Table 6 is shown in Figure 15.

Table 6. Performance values of the EfficientNet models tested with the COVID-CT-Dataset.

Model Name	Accuracy	Recall	Precision	Specificity	F1 Score
EfficientNet-B0	0.9142	0.8589	0.9771	0.9771	0.9142
EfficientNet-B1	0.9517	0.9815	0.9140	0.9287	0.9466
EfficientNet-B2	0.9343	0.8989	0.9685	0.9703	0.9324
EfficientNet-B3	0.9651	0.9549	0.9713	0.9744	0.9631
EfficientNet-B4	0.9625	0.9496	0.9713	0.9743	0.9603
EfficientNet-B5	0.9477	0.9429	0.9456	0.9520	0.9442
EfficientNet-B6	0.9558	0.9540	0.9513	0.9573	0.9527
EfficientNet-B7	0.9437	0.9398	0.9398	0.9471	0.9398

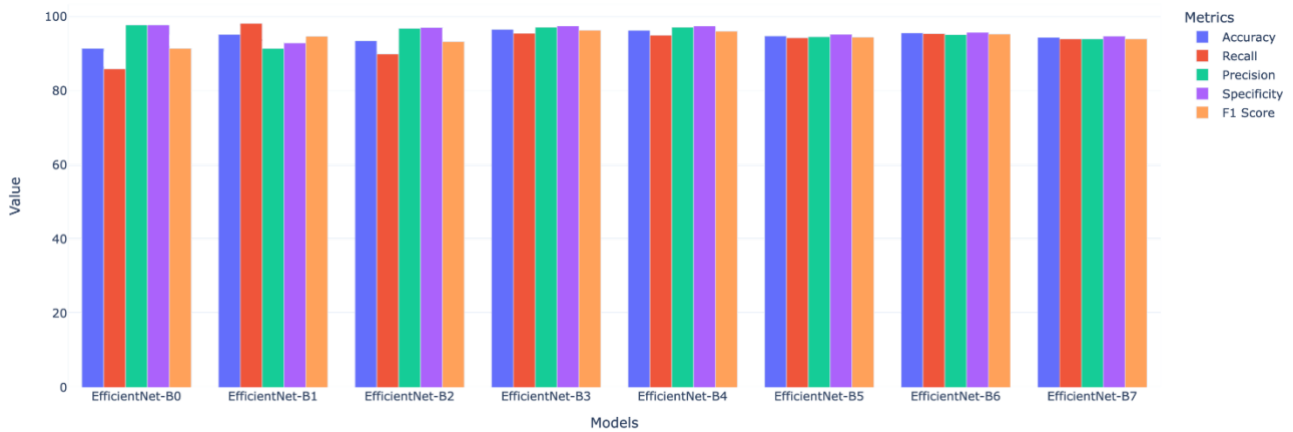


Figure 15. Bar chart of the performance of EfficientNet models.

Upon examination of the results in Table 6, it is observed that all models have a success rate of over 90%. However, with an accuracy rate of 96.51%, the EfficientNet-B3 model is more performant compared to the other models. While the EfficientNet-B2 model architecture was employed in our study, it is worth noting that it may not be optimal for all tasks and datasets. Alternative model architectures may potentially yield better performance in certain cases.

The limitations of this study are as follows:

- The dataset comprises 4,000 images that were collected from a single CT device, which may not be sufficient to adequately capture the variability and complexity of the findings analyzed.
- Identifying slices with signs of COVID-19 from chest CT scans is time-consuming.
- Our study only tested the performance of eight pre-trained models on the task of classifying CT images.

- The robustness of the models in the face of changes in lighting, background, or other factors that may affect the appearance of COVID-19 findings in images has not been evaluated.
- The model was only evaluated on a single split of the data, rather than using k-fold cross-validation to assess its performance more thoroughly.

The results of this study demonstrate the effectiveness of the EfficientNet-B2 model for automated detection of COVID-19 from chest CT images, but there are several areas of future work that could be pursued to further improve the performance and generalizability of the model. One potential direction for future work is to expand the dataset used for training and evaluation. The study used a novel dataset of 4,000 images, but it could be beneficial to further increase the size of the dataset to improve the robustness and generalizability of the model. This could involve collecting additional images, as well as increasing the diversity of the images in terms of factors such as patient demographics, clinical characteristics, and imaging modalities. Another avenue for future research is to evaluate the performance of alternative model architectures. While the EfficientNet-B2 model performed well in this study, it may not be optimal for all tasks and datasets. Future work could include testing the performance of other model architectures, such as those with a greater number of layers and parameters, to determine if they offer any benefits over the EfficientNet-B2 model. This could involve evaluating the models on the same dataset used in this study, as well as on other datasets to assess the generalizability of the results. Additionally, it would be beneficial to investigate methods for addressing any potential biases or weaknesses identified through explainability techniques. In this study, the explainability method of Grad-CAM was utilized to provide insights into the model's decision-making process, but further investigation may reveal other biases or weaknesses that could be mitigated. Finally, it may be interesting to explore the use of the model in other contexts or applications. While the present study demonstrated the effectiveness of the EfficientNet-B2 model for the automated detection of COVID-19 from chest CT images, it may be valuable to test the model on other imaging modalities, such as chest X-rays or MRIs, or to use it to detect other diseases or conditions.

5. Conclusion

The COVID-19 disease has negative impacts on our society in various ways. Classical laboratory methods used for the detection of the disease are insufficient due to their disadvantages. By using radiological imaging techniques instead of classical laboratory methods, the COVID-19 disease can be detected at an early stage. However, specialists are needed to detect the disease from radiological images. Due to the speed and ease of transmission, millions of people have been infected

with COVID-19. As a result, more than the normal number of patients are assigned to specialists who will examine the radiological images. Taking into account the maximum working time and other human characteristics, the time for diagnosis of the patient whose radiological image is taken increases significantly due to the intensity. There is a risk of transmission of the disease to other normal individuals in possible patients whose disease takes a long time to detect. Computers with the ability to process quickly and not be affected by human factors can assist experts in the examination of radiological images. In this study, a decision support system is proposed that can automatically detect COVID-19 disease from chest CT images. The Grad-CAM algorithm was used to explain the detection of COVID-19 disease by focusing on which areas in chest CT images. Eight different EfficientNet models trained with a normally distributed dataset called EFSCHE-19 can automatically classify chest CT images. Among these models, the EfficientNet-B2 model showed the highest performance for the test images. The margin of error is expected to be close to zero in this system, which was designed to support decisions that can directly affect human life. Therefore, improvements such as expanding the scope of the dataset used in the training of the model and increasing the number of samples it contains can be planned for future studies. With the use of such systems in the healthcare sector, early detection of diseases can be achieved and erroneous diagnoses caused by inexperienced personnel can be prevented.

Authors' Contributions

All authors contributed equally to the study.

Statement of Conflicts of Interest

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

The author declares that this study complies with Research and Publication Ethics.

References

- Tekin, A. (2021). Tarihten Günümüze Epidemiler, Pandemiler ve Ekonomik Sonuçları. *Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 40, 330-355.
- Issever, H., Issever, T., and Oztan, G. (2020). COVID-19 epidemiyolojisi. *Sağlık Bilimlerinde İleri Araştırmalar Dergisi*, 3(S1), 1-13.

- Ahmad, T., Khan, M., Musa, T. H., Nasir, S., Hui, J., Bonilla-Aldana, D. K., and Rodriguez-Morales, A. J. (2020). COVID-19: Zoonotic aspects. *Travel medicine and infectious disease*, 36, 101607.
- Hua, J., and Shaw, R. (2020). Corona virus (Covid-19)“infodemic” and emerging issues through a data lens: The case of china. *International journal of environmental research and public health*, 17(7), 2309.
- Madabhavi, I., Sarkar, M., and Kadakol, N. (2020). COVID-19: a review. *Monaldi Archives for Chest Disease*, 90(2).
- Wadman, M. (2021). SARS-CoV-2 infection confers greater immunity than shots. *Science*, 373(6559), 1067-1068.
- Giri, B., Pandey, S., Shrestha, R., Pokharel, K., Ligler, F. S., and Neupane, B. B. (2021). Review of analytical performance of COVID-19 detection methods. *Analytical and bioanalytical chemistry*, 413(1), 35-48.
- Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. K., and Drosten, C. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*, 25(3), 2000045.
- Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., and Zhang, Y. Z. (2020). A new coronavirus associated with human respiratory disease in China. *Nature*, 579(7798), 265-269.
- Zali, A., Sohrabi, M. R., Mahdavi, A., Khalili, N., Taheri, M. S., Maher, A., and Hanani, K. (2021). Correlation between low-dose chest computed tomography and RT-PCR results for the diagnosis of COVID-19: A report of 27,824 cases in tehran, Iran. *Academic Radiology*, 28(12), 1654-1661.
- Kommos, F. K., Schwab, C., Tavernar, L., Schreck, J., Wagner, W. L., Merle, U., and Longerich, T. (2020). The pathology of severe COVID-19-related lung damage: Mechanistic and therapeutic implications. *Deutsches Ärzteblatt International*, 117(29-30), 500.
- Ghaderzadeh, M., and Asadi, F. (2021). Deep learning in the detection and diagnosis of COVID-19 using radiology modalities: a systematic review. *Journal of healthcare engineering*, 2021.
- Ng, M. Y., Lee, E. Y., Yang, J., Yang, F., Li, X., Wang, H., and Kuo, M. D. (2020). Imaging profile of the COVID-19 infection: radiologic findings and literature review. *Radiology: Cardiothoracic Imaging*, 2(1).
- Fields, B. K., Demirjian, N. L., Dadgar, H., and Gholamrezanezhad, A. (2021, July). Imaging of COVID-19: CT, MRI, and PET. In *Seminars in Nuclear Medicine*, 51(4), 312-320.
- Gundel, S., Setio, A. A., Ghesu, F. C., Grbic, S., Georgescu, B., Maier, A., and Comaniciu, D. (2021). Robust classification from noisy labels: Integrating additional knowledge for chest radiography abnormality assessment. *Medical Image Analysis*, 72, 102087.
- Chartrand, G., Cheng, P. M., Vorontsov, E., Drozdal, M., Turcotte, S., Pal, C. J., and Tang, A. (2017). Deep learning: a primer for radiologists. *Radiographics*, 37(7), 2113-2131.
- Arsalan, M., Owais, M., Mahmood, T., Choi, J., and Park, K. R. (2020). Artificial intelligence-based diagnosis of cardiac and related diseases. *Journal of Clinical Medicine*, 9(3), 871.
- Owais, M., Arsalan, M., Choi, J., Mahmood, T., and Park, K. R. (2019). Artificial intelligence-based classification of multiple gastrointestinal diseases using endoscopy videos for clinical diagnosis. *Journal of clinical medicine*, 8(7), 986.
- Mainak, B., Venkatanareshbabu, K., Luca, S., Damodar, R. E., Elisa, C. G., Tato, M. R., and Andrew, N. (2019). State-of-the-art review on deep learning in medical imaging. *Frontiers in Bioscience-Landmark*, 24(3), 380-406.
- Chin, C. L., Lin, B. J., Wu, G. R., Weng, T. C., Yang, C. S., Su, R. C., and Pan, Y. J. (2017, November). An automated early ischemic stroke detection system using CNN deep learning algorithm. *8th International conference on awareness science and technology* (pp. 368-372).
- Yildirim, K., Bozdog, P. G., Talo, M., Yildirim, O., Karabatak, M., and Acharya, U. R. (2021). Deep learning model for automated kidney stone detection using coronal CT images. *Computers in biology and medicine*, 135, 104569.
- Khan, S., Islam, N., Jan, Z., Din, I. U., and Rodrigues, J. J. C. (2019). A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. *Pattern Recognition Letters*, 125, 1-6.
- Bogu, G. K., and Snyder, M. P. (2021). Deep learning-based detection of COVID-19 using wearables data. *MedRxiv*.
- Pahar, M., Klopffer, M., Warren, R., and Niesler, T. (2021). COVID-19 cough classification using machine learning and global smartphone recordings. *Computers in Biology and Medicine*, 135, 104572.
- Wang, L., Lin, Z. Q., and Wong, A. (2020). Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Scientific Reports*, 10(1), 1-12.

- Luz, E., Silva, P., Silva, R., Silva, L., Guimarães, J., Miozzo, G., and Menotti, D. (2022). Towards an effective and efficient deep learning model for COVID-19 patterns detection in X-ray images. *Research on Biomedical Engineering*, 38(1), 149-162.
- Song, Y., Zheng, S., Li, L., Zhang, X., Zhang, X., Huang, Z., and Yang, Y. (2021). Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. *IEEE/ACM Transactions on computational biology and bioinformatics*, 18(6), 2775-2780.
- Bozkurt, F. (2022). A deep and handcrafted features-based framework for diagnosis of COVID-19 from chest x-ray images. *Concurrency and Computation: Practice and Experience*, 34(5), e6725.
- Marques, G., Agarwal, D., and de la Torre Díez, I. (2020). Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network. *Applied soft computing*, 96, 106691.
- Wang, S., Kang, B., Ma, J., Zeng, X., Xiao, M., Guo, J., and Xu, B. (2021). A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). *European radiology*, 31(8), 6096-6104.
- Comert, S. S., and Kiral, N. (2020). Radiological Findings of COVID-19 Pneumonia/COVID-19 Pnomonisinin Radyolojik Bulgulari. *Southern Clinics of Istanbul Eurasia*, 31(SI), 16-23.
- Tan, M., and Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. *International conference on machine learning* (pp. 6105-6114).
- Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pp. 248-255
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A., and Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3), 211-252.
- Collins, T. J. (2007). ImageJ for microscopy. *Biotechniques*, 43(S1), S25-S30.
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE international conference on computer vision* (pp. 618-626).
- Ismael, A. M., and Şengür, A. (2021). Deep learning approaches for COVID-19 detection based on chest X-ray images. *Expert Systems with Applications*, 164, 114054.
- Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., and Acharya, U. R. (2020). Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Computers in biology and medicine*, 121, 103792.
- Shah, V., Keniya, R., Shridharani, A., Punjabi, M., Shah, J., and Mehendale, N. (2021). Diagnosis of COVID-19 using CT scan images and deep learning techniques. *Emergency radiology*, 28(3), 497-505.
- Rahimzadeh, M., Attar, A., and Sakhaei, S. M. (2021). A fully automated deep learning-based network for detecting covid-19 from a new and large lung ct scan dataset. *Biomedical Signal Processing and Control*, 68, 102588.
- Gupta, A., Gupta, S., and Katarya, R. (2021). InstaCovNet-19: A deep learning classification model for the detection of COVID-19 patients using Chest X-ray. *Applied Soft Computing*, 99, 106859.
- Katar, O., and Duman, E. (2021). Deep Learning Based Covid-19 Detection With A Novel CT Images Dataset: EFSCHE-19. *Avrupa Bilim ve Teknoloji Dergisi*, (29), 150-155.
- Gupta, K., and Bajaj, V. (2023). Deep learning models-based CT-scan image classification for automated screening of COVID-19. *Biomedical Signal Processing and Control*, 80, 104268.
- Zhou, T., Lu, H., Yang, Z., Qiu, S., Huo, B., and Dong, Y. (2021). The ensemble deep learning model for novel COVID-19 on CT images. *Applied soft computing*, 98, 106885.
- Yang, X., He, X., Zhao, J., Zhang, Y., Zhang, S., and Xie, P. (2020). COVID-CT-dataset: a CT scan dataset about COVID-19. *arXiv preprint arXiv:2003.13865*.
- URL-1: <https://www.arcgis.com/apps/dashboards/bda7594740fd40299423467b48e9ecf6>, (Date Accessed: 27 August 2022).
- URL-2: <https://keras.io/api/applications/>, (Date Accessed: 24 December 2022).