

Article

COVID-19 CXR Classification: Applying Domain Extension Transfer Learning and Deep Learning

KwangJin Park , YoungJin Choi  and HongChul Lee * 

School of Industrial Management Engineering, Korea University, Seoul 02841, Korea

* Correspondence: hclee@korea.ac.kr; Tel.: +82-2-3290-3389

Abstract: The infectious coronavirus disease-19 (COVID-19) is a viral disease that affects the lungs, which caused great havoc when the epidemic rapidly spread around the world. Polymerase chain reaction (PCR) tests are conducted to screen for COVID-19 and respond to quarantine measures. However, PCR tests take a considerable amount of time to confirm the test results. Therefore, to supplement the accuracy and quickness of a COVID-19 diagnosis, we proposed an effective deep learning methodology as a quarantine response through COVID-19 chest X-ray image classification based on domain extension transfer learning. As part of the data preprocessing, contrast limited adaptive histogram equalization was applied to chest X-ray images using Medical Information Mart for Intensive Care (MIMIC)-IV obtained from the Beth Israel Deaconess Medical Center. The classification of the COVID-19 X-ray images was conducted using a pretrained ResNet-50. We also visualized and interpreted the classification performance of the model through explainable artificial intelligence and performed statistical tests to validate the reliability of the model. The proposed method correctly classified images with 96.7% accuracy, an improvement of about 9.9% over the reference model. This study is expected to help medical staff make an integrated decision in selecting the first confirmed case and contribute to suppressing the spread of the virus in the community.

Keywords: COVID-19; chest X-ray; image classification; domain extension transfer learning; explainable artificial intelligence



Citation: Park, K.; Choi, Y.; Lee, H. COVID-19 CXR Classification: Applying Domain Extension Transfer Learning and Deep Learning. *Appl. Sci.* **2022**, *12*, 10715. <https://doi.org/10.3390/app122110715>

Academic Editors: Yujin Lim and Hideyuki Takahashi

Received: 26 September 2022

Accepted: 14 October 2022

Published: 22 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Background

The infectious coronavirus disease-2019 (COVID-19) or SARS-CoV-2 virus rapidly spread and was declared a pandemic by the World Health Organization (WHO) in March 2020 [1]. As of early October 2022, the recorded number of confirmed cases worldwide was 624,142,119, and the death toll was 6,552,954 [2]. COVID-19 is a viral disease that causes severe fever, a cough, shortness of breath, sore throats, headaches, diarrhea, a loss of taste, fatigue, and a loss of smell. Severe respiratory failure can lead to death [3].

Accordingly, screening for COVID-19 confirmed cases and responding to quarantine measures are essential. The early identification of confirmed cases through a polymerase chain reaction (PCR) test and X-ray is required for screening COVID-19 confirmed cases and implementing a suitable response (quarantine, treatment, and recovery). However, it has been pointed out that the PCR test is expensive, and it takes considerable time to confirm the test results [4]. The budgeting of the expenditure resulting from the high cost of testing and the input of a professional workforce for each testing location is critical. Therefore, the economic burden of government spending is high.

It was determined that COVID-19 is a respiratory lung disease that is also one of the mutations of pneumonia. Diagnosis using a chest X-ray (CXR) for the early detection of lung disease is effective because it requires less time and money [5]. Compared to CT, X-ray imaging is cheaper, faster, and more readily accessible, and it exposes the body to far less

hazardous radiation [6]. Therefore, X-ray images are an effective instrument for rapidly diagnosing pneumonia and COVID-19 in patients [7].

According to the WHO recommendation regarding chest imaging for COVID-19, chest X-ray images are recommended as complementary information for those who test negative in the initial PCR test but have suspicious COVID-19 symptoms [8]. As a result of an X-ray and CT of the patient with a false negative PCR test result, 70% of COVID-19 symptoms were confirmed by chest radiography or a chest CT [9,10]. Using X-ray images can help prevent and manage the prognosis of diseases by visualizing infected lungs along with a confirmation of lung infection [11,12].

In addition, this information can help professionals take timely decisions on admission and discharge criteria in terms of the infection stage, level of care, and patient care modalities and help to secure the need for ICU beds [3,9,13,14]. Accordingly, X-ray imaging, as an auxiliary tool, can help diagnose pneumonia and COVID-19 in patients [8,11,15]. Therefore, by using an AI model to accurately classify X-rays and enable the early screening of confirmed cases, it is possible not only to suppress the spread of disease in the community but also to reduce the cost of testing. It can also help to ensure a pre-emptive quarantine response during a pandemic that causes chaos [16,17].

Owing to the inherent qualities of machine vision, artificial intelligence (AI) and machine learning technology have achieved outstanding outcomes in the medical imaging area [18]. In an era in which COVID-19 is spreading across the world, machine learning researchers and computer scientists play a crucial role. Deep learning is one of the breakthroughs in AI. It collects the fine-grained features from the images [19]. It is a set of machine learning techniques primarily focused on automatically extracting and classifying features from images. Many challenges, including arrhythmia detection, skin cancer classification, breast cancer detection, brain disease classification, and pneumonia identification from chest X-ray images, have been successfully addressed with deep learning techniques [20]. Convolutional neural networks (CNNs) have shown promise in the analysis and classification of X-ray images using deep learning techniques [3].

Since the pandemic, research on COVID-19 disease detection and patient diagnosis by classifying images, such as X-ray, computed tomography (CT), and magnetic resonance imaging (MRI), using artificial intelligence (AI) has been actively conducted [21].

As COVID-19 spread worldwide, research was conducted to detect COVID-19 lesions on CXR images based on deep learning methodologies [16,22,23]. Owing to the severe lack of COVID-19 CXR images at the beginning of the pandemic, research predominantly focused on deep learning architectures and algorithms. To automatically identify COVID-19 lesions on CXRs, Ozturk et al. [20] proposed Dark-CovidNet to identify COVID-19 lesions in CXR. In addition, Mahmud et al. [24] proposed a CovXNet model to detect patterns of COVID-19 in the CXR images of patients. However, in these studies, the proposed models suffered from insufficient learning and testing owing to insufficient COVID-19 CXR data. These limitations significantly affected the generalization of the model performance.

In most studies using CXR images to detect and classify COVID-19 lesions, deep learning algorithms are proposed to overcome the limitations of the insufficient COVID-19 CXR image dataset, or transfer learning methodologies that are pretrained with natural images are used. However, these methodologies have to create a new algorithm when a new lung disease pandemic occurs, and in addition, a model pretrained with natural images rather than medical images has no choice but to show a low classification accuracy.

Therefore, we proposed an effective deep learning methodology for classifying COVID-19 CXR images. To effectively respond to a pandemic caused by a new lung disease such as COVID-19, it is necessary to train a model using a large dataset. However, in reality, we faced the problem of not being able to produce good results owing to the lack of COVID-19 data. To solve this, we used MIMIC-IV CXR, which contained information about 11 lung diseases, to match target domains with medical images rather than natural images to effectively classify images even without learning COVID-19 images between pretraining. We proposed a deep learning methodology based on domain extension transfer learning.

We improved the classification accuracy and generalization performance of the model using notarized MIMIC-IV data preprocessing and effective X-ray CLAHE. Based on domain extension transfer learning, the ResNet-50 model was used to effectively classify the COVID-19 CXR images through pretraining and fine-tuning. The numerical results of the classification performance of the model were visualized and interpreted using gradient-weighted class activation mapping (Grad-CAM) on the COVID-19 X-ray images containing the opinion of a radiologist. The reliability of the performance of the model was also evaluated through statistical tests.

The remainder of this paper is organized as follows. Section 2 describes the related works. Section 3 describes the dataset used in this study, the preprocessing methodology, and the deep learning methodology. Section 4 describes the hyperparameters applied to the proposed methodology, the performance evaluation indicators used in the results of the experiments, and the performance results of the final model. Section 5 presents the discussions, and Section 6 presents the conclusions and directions for future research.

2. Related Works

2.1. COVID-19 X-ray Image Classification Using Pretrained Models Based on Transfer Learning

To generalize the model performance, the classification of COVID-19 X-ray images using pretrained models based on transfer learning was performed. Apostolopoulos et al. [25] proposed a method for classifying COVID-19 CXR images using a model pretrained with natural images (ImageNet 1 k) based on the transfer learning methodology. Elpeltagy et al. [19] proposed a modified ResNet-50 pretrained model. It was trained and tested using the Mendeley dataset, which included two classes of COVID-19 and non-COVID-19 data. Owing to the five-fold cross-validation test, the average classification accuracy was 96.12%.

Sahinbas et al. [26] proposed a model using data that was directly collected from hospitals. Among the existing convolutional neural network (CNN) models, they conducted a study to detect COVID-19 lesions in X-ray images using VGG16, VGG19, ResNet, DenseNet, and Inception-V3 pretrained models. The VGG16 model showed a classification accuracy of 80%. Narin et al. [27] proposed a methodology for automatically screening COVID-19 lesions through a pretrained CNN model using 341 COVID-19 X-ray images collected by Cohen et al. as well as normal and pneumonia CXRs. The Inception-V3, ResNet-50, ResNet-101, ResNet-152, and Inception-ResNet-V2 models were used, with a five-fold cross-validation test showing that ResNet-50 achieved the highest classification performance of 99.7%.

2.2. Domain Extension Transfer Learning and Augmentations

In addition, studies have been conducted to diagnose COVID-19 using domain extension transfer learning strategies and contrast limited adaptive histogram equalization (CLAHE). Afshar et al. [28] showed that a pretrained model using the same kind of source domain (training data) as the target domain (test data) significantly improved the COVID-19 classification performance compared with applying transfer learning from ImageNet 1 k (pretrained with natural images). The experimental results demonstrated that the classification accuracy was improved by 2.5%, owing to the training and testing by matching domains with the proposed model. However, the sensitivity was 80%, showing a limitation of the study.

In addition, Basu et al. [29] matched the target and source domains with X-ray image data from scratch rather than using a pretrained model using natural images. Using AlexNet, VGGNet, and ResNet, image classification performances of approximately 82%, 90%, and 85%, respectively, were derived. In addition, Saiz et al. [30] verified that image conversion preprocessing using CLAHE has noticeable results when screening and diagnosing lung infections. The application of CLAHE resulted in a 1.6% and 9% improvement in the classification accuracy for normal and COVID-19 images in CXR image classification experiments, respectively.

As mentioned above, deep learning has been used to propose new algorithms and general transfer learning to overcome the limitations of insufficient data for COVID-19 image classification. However, effective deep learning methodologies are required to quickly and appropriately respond to COVID-19 and viruses that cause new lung diseases, such as COVID-19.

3. Materials and Methods

In this section, as shown in Figure 1, we proposed a framework that effectively classified COVID-19 CXR images through domain extension transfer learning based on ResNet-50 after applying the CLAHE technique that effectively detects lung disease lesions in MIMIC-IV CXR images.

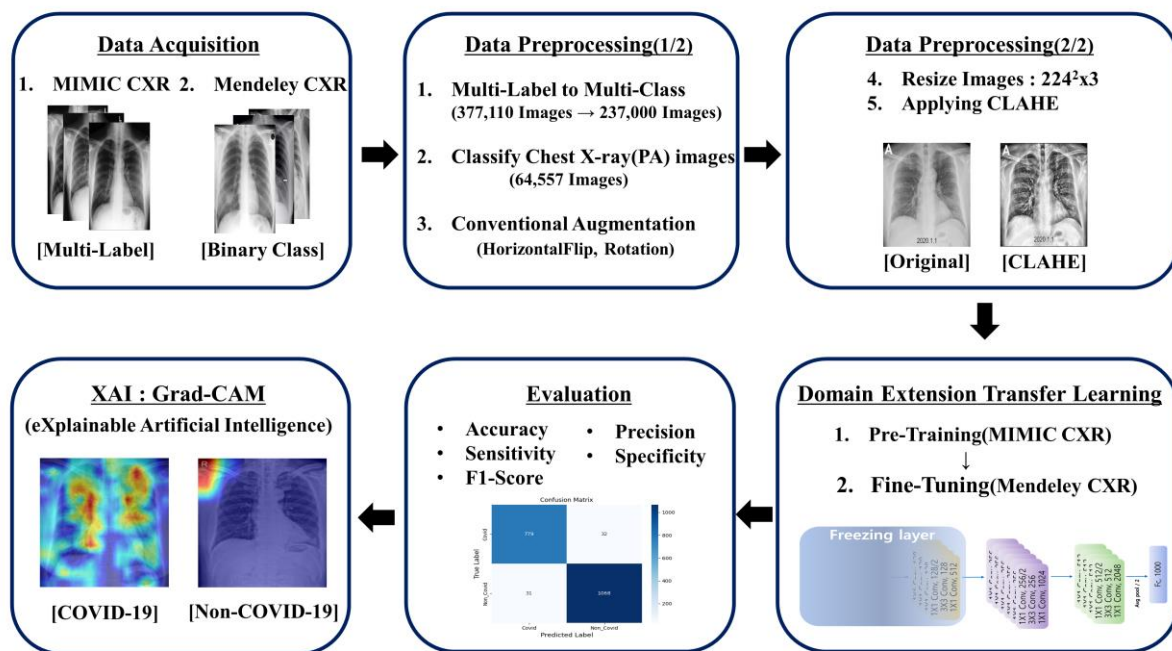


Figure 1. Architecture of the COVID-19 CXR classification framework.

3.1. Datasets

In this study, we used MIMIC-IV CXR data to apply domain extension transfer learning. The MIMIC-IV dataset CXR images were collected between 2011 and 2016 from intensive care patients treated at the Beth Israel Deaconess Medical Center in the United States [31]. A total of 377,110 X-ray images of healthy patients and those with various chest diseases in 12 classes were included in the dataset. The data used to classify the CXR images of COVID-19 confirmed patients were obtained from the Mendeley dataset, which comprises public data collected by El-Shafai et al. [32] in 2020. The Mendeley dataset has been cited more than 35 times in studies on COVID-19 detection published in overseas journals, with a total of 9544 X-ray images in the binary class. Each class comprises COVID-19 and non-COVID-19 images. In this study, we divided the training, validation, and testing of the Mendeley dataset into a ratio of 6:2:2 on the evaluation data of the model, as shown in Table 1.

There were many multilabel images owing to exploration data analysis (EDA) conducted on the images of the 377,110 MIMIC-IV dataset. For experimental purposes, the images were preprocessed as a dataset with only one class per image to remove inappropriate noise. Following the preprocessing procedure, we initially obtained 237,000 images. To conduct a more accurate classification of the COVID-19 X-ray images, we also classified 64,557 CXR images obtained from 237,000 images, known as posteroanterior (PA) images [33]. A PA X-ray image provides useful information for detecting lung diseases.

Therefore, in this study, 237,000 images were classified as MIMIC CXR, and 64,557 images were selected as PA images (MIMIC CXR(PA)). To pretrain the model, the MIMIC CXR and MIMIC CXR(PA) datasets were divided in the ratio of 8:1:1. The MIMIC CXR(PA) dataset was chosen for the final experiment because we determined that it was the only significant dataset. The MIMIC CXR(PA) dataset comprised 51,641 training images, 6458 validation images, and 6458 test images. 11 classes were used in the experiment because the pleural other class was absent from the CXR(PA) dataset. The detailed status of the MIMIC CXR(PA) dataset is described in Table 2.

Table 1. Summary of the Mendeley dataset.

Mendeley Dataset		Training	Validation	Testing
Class	Images			
COVID-19	4044	2426	809	809
Non-COVID-19	5500	3300	1100	1100
Total	9544	5726	1909	1909

Table 2. Summary of the MIMIC CXR(PA) ¹ dataset.

MIMIC CXR(PA) Dataset		Training	Validation	Testing
Class	Images			
Atelectasis	2105	1684	211	211
Cardiomegaly	3106	2484	312	312
Consolidation	285	228	28	28
Edema	587	469	58	58
Fracture	643	514	64	64
Lung Lesion	810	648	81	81
Lung Opacity	4113	3290	416	416
No Finding	48,577	38,861	4857	4857
Pleural Effusion	2177	1741	217	217
Pneumonia	1676	1340	167	167
Pneumothorax	478	382	47	47
Total	64,557	51,641	6458	6458

¹ MIMIC-IV: Medical Information Mart for Intensive Care-IV; PA: posteroanterior.

3.2. Preprocessing

Data Augmentation

The COVID-19 CXR image data were resized to the input size of the CNN model. Rotation and horizontal flip were suitable for data augmentation during classification of the COVID-19 CXR images. This approach was effective in improving model generalization performance when applied to the classification of COVID-19 CXR images.

According to research conducted by Reza et al. [34], the main idea of CLAHE involves mapping each pixel (Equations (1) and (2)), and it is based on the image distribution of grayscale. Research has demonstrated that a medical image is produced in a decent state by balancing the pixels of the image through image conversion. The study by Reza involved two equations on CLAHE. Equation (1) represents histogram equalization, and Equation (2) represents reducing bias by using contrast limit. Explanations of these equations are as follows. M and N are the number of pixels and grayscale in each region, respectively, and $f_{i,j}(n)$ is the histogram of the (i, j) region for $n = 0, 1, 2 \dots, N - 1$. An estimate of the

corresponding CDF, appropriately scaled by $(N - 1)$ for the grayscale mapping [34], is as follows:

$$f_{i,j}(n) = \frac{(N - 1)}{M} \times \sum_{k=0}^n h_{i,j}(k) \quad , \quad n = 1, 2, 3 \dots, N - 1 \quad (1)$$

To restrict the contrast to a specific level, clip limit β is used to cut the histogram to limit the slope of Equation (1). Clip limit is related to clip factor α [34] as follows:

$$\beta = \frac{M}{N} \left(1 + \frac{\alpha}{100} (s_{max} - 1) \right) \quad (2)$$

The collected X-ray images became cloudy over time, thereby making them unsuitable for lesion detection [35]. Therefore, CLAHE was applied in the pretreatment process. The improved quality of X-ray images enhanced the classification accuracy of COVID-19 CXR images by approximately 3.9%. Figure 2 shows the results of applying CLAHE to the Mendeley dataset. Unlike hazy X-rays, X-rays with CLAHE applied to the images show a ground-glass opacity called COVID-19 lesions. In both lungs, these lesions can be identified well with the naked eye.

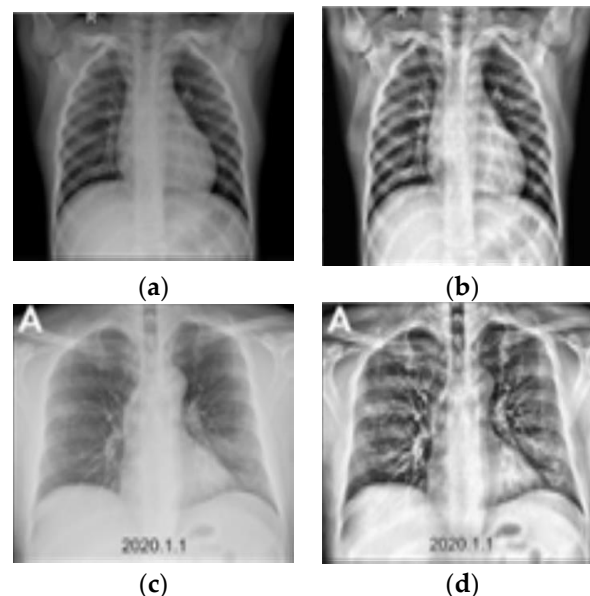


Figure 2. X-ray images: (a) non-COVID-19 (healthy) pre-CLAHE; (b) non-COVID-19 (healthy) post-CLAHE; (c) COVID-19 pre-CLAHE; and (d) COVID-19 post-CLAHE.

Table 3 shows the results obtained through signal-to-noise ratio (SNR) indicators to indicate whether image quality improves after CLAHE in this study. Comparing the existing image and the image after CLAHE revealed that the SNR value increased in both the original and resized images. Higher SNR value means that the image quality was relatively superior. Hence, CLAHE applied in this study affected model performance improvement.

Table 3. Image quality comparison after CLAHE with SNR.

		CLAHE		
COVID-19 X-ray		Before	After	Improvement
Resized	SNR	3.9 DN	4.6 DN	0.7 DN
	SNR (dB)	1.9 dB	2.2 dB	0.3 dB

3.3. Domain Extension Transfer Learning Pretraining and Fine-Tuning

We proposed an effective methodology for classifying COVID-19 X-ray images based on domain extension transfer learning which was conducted using a pretrained ResNet-50 model. Each pretraining process was carried out according to the target data using the MIMIC CXR and MIMIC CXR (PA) datasets.

To classify COVID-19 and non-COVID-19 CXR successfully, we matched the source domain and target domain in the medical area using a similar CXR dataset. Generally, data are limited, and it is expensive to get expert annotation in the medical area. Deep CNN training takes a long time as it needs a lot of computing and memory resources. In the absence of sufficient data, transfer learning (TL) offers a feasible option to fine-tune a CNN that has previously been pretrained on a huge collection of accessible labeled datasets from some other category. The collected COVID-19 dataset was deemed insufficient to train the CNN, which involves a sizable number of trainable parameters, from scratch to learn the complicated representative features for differentiating between COVID-19 and non-COVID-19 CXRs. Therefore, transfer learning was used to train final models for the target domain (D_T) and target task (T_T) with less annotated data by utilizing the information obtained from the source domain (D_S) and source task (T_S). The general idea of transfer learning is as follows.

A tuple of the form $\{\chi, P(X)\}$, where χ and $P(X)$ ($X = x_1, x_2, \dots, x_n \in \chi$) represent the feature space and associated marginal probability distribution can be used to describe a domain (D). A task T consists of a label space Y and a conditional probability distribution $P(Y|X)$ that is typically learned from the training data, given a domain $D = \{\chi, P(X)\}$. The idea of transfer learning is to learn the target conditional probability distribution $P(Y_T|X_T)$ in the target domains (D_T), using the knowledge gained from the source (D_S) by troubleshooting the tasks (T_S) [29]. Figure 3 shows the basic description of transfer learning.

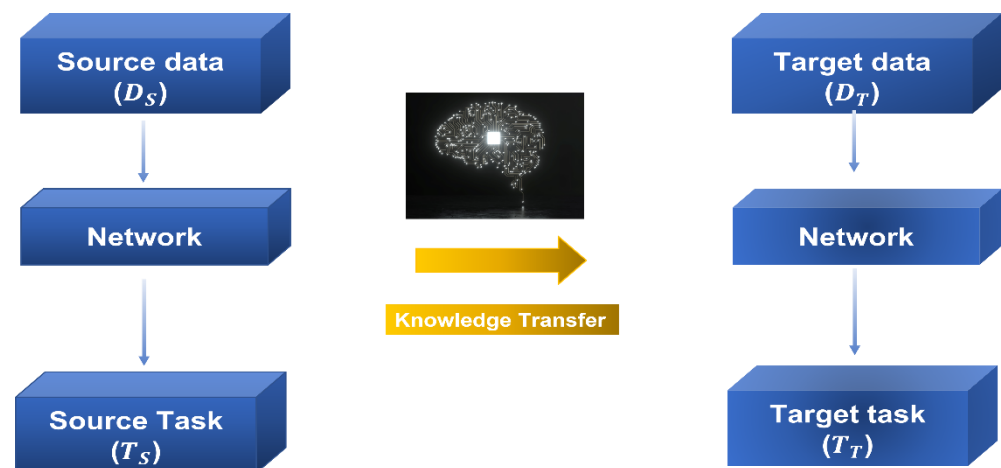


Figure 3. Transfer learning.

The recent research has shown that transfer learning has a very limited role to play when the source and target domains are quite different in nature, such as in natural and medical images, as the networks may obtain quite distinct high-level features in the two cases. Therefore, we used each MIMIC CXR and MIMIC CXR(PA) dataset as the source data (D_S) to pretrain the network first, and then by using pretrained model, we applied the knowledge obtained from solving source task (T_S) with source dataset to the target dataset (D_T), the Mendeley dataset. Therefore, we proposed a deep learning methodology that solves the target task (T_T) with domain extension transfer learning. Transfer learning was implemented to build up the final model to solve the target task including scarce target data (D_T) by utilizing the knowledge learned from the source data (D_S).

Fine-tuning is a deep learning technique to increase classification performance by fitting pretrained model weights to target data. Various fine-tuning strategies are also used in the field of imaging diagnostics [36]. In the biomedical field, there is a case where the classification performance was shown to improve by applying fine-tuning [37]. In this study, the ResNet-50 model was pretrained using MIMIC CXR and MIMIC CXR(PA) data based on domain extension transfer learning. Additionally, the performance of the models that were fitted to the target dataset was compared with that of ResNet-50, and the CNN models that had already been pretrained using natural images were pretrained using MIMIC CXR and MIMIC CXR(PA) data, respectively.

In this study, a fine-tuning strategy was used when the target domain dataset was insufficient and when the data in the source and target domains were similar. Here, only a few layers of the pretrained weights were fine-tuned, and the rest of the layers were frozen. As shown in Figure 4, the fine-tuning strategy was applied to the ResNet-50. The training strategy of the experiment involving fine-tuning from the 5th layer to the fully connected (FC) layer and freezing the remaining layers achieved the highest accuracy levels of COVID-19 X-ray image classification.

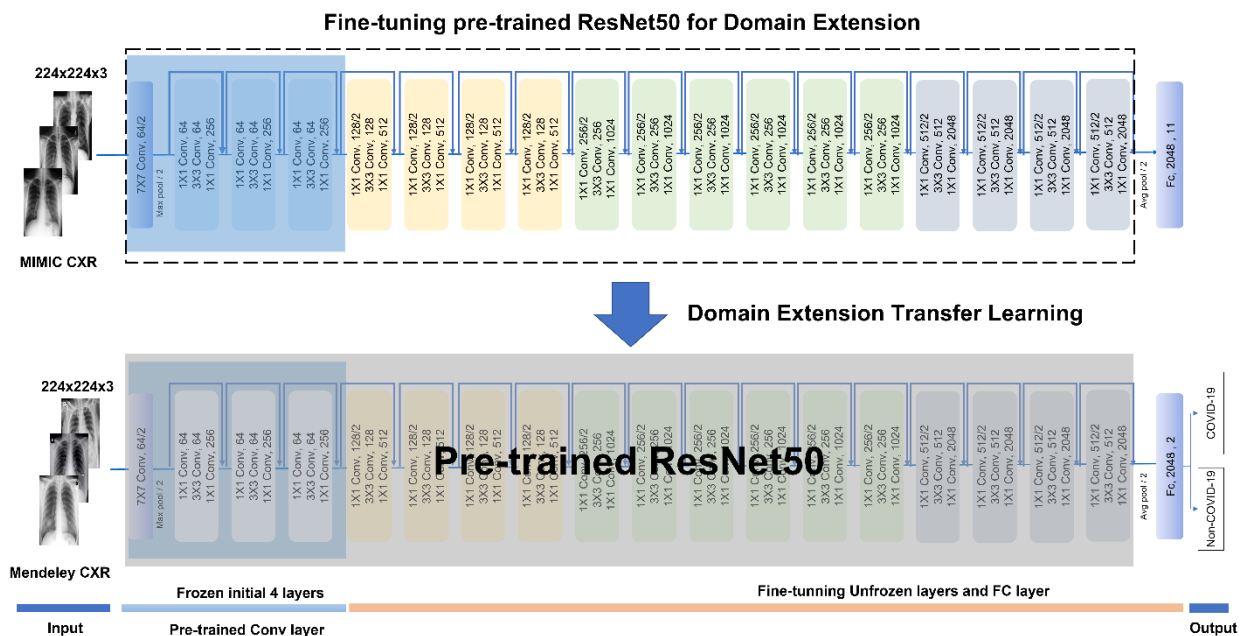


Figure 4. Fine-tuning strategy for domain extension transfer learning.

3.4. Classification Models

CNN Models for Image Classification

The CNN model used in this study was ResNet-50. In previous studies, ResNet-50 showed high performance in classifying COVID-19 CXR images.

ResNet was proposed by He et al. [38] in 2015. Contrary to the assumption that prediction performance improves as the neural networks of AlexNet and VGGNet deepen, performance degradation occurs owing to gradient vanishing and gradient exploding. To solve this problem, ResNet uses the skip/shortcut connection method. In the learning of the existing neural network, $H(X) = X$ was learned, but using residual functions, $H(X) = F(X) + X$ was learned, and residual learning was conducted so that $F(X) = 0$. Therefore, the problem of gradient loss was solved by suggesting a solution to approximate $H(X) = X$ (Equations (3)–(5), Figure 4). In the formula, X represents input, Y represents output, and $F(X)$ represents the residual function. ResNet is relatively more efficient than the existing CNN model because it uses an element-wise sum rather than the existing

element-wise product. Figure 5 describes the learning method of ResNet, which is called “Residual learning”.

$$F(X) := H(X) - X \quad (3)$$

$$H(X) = F(X) + X \quad (4)$$

$$Y = F(X, w_i) + X \quad (5)$$

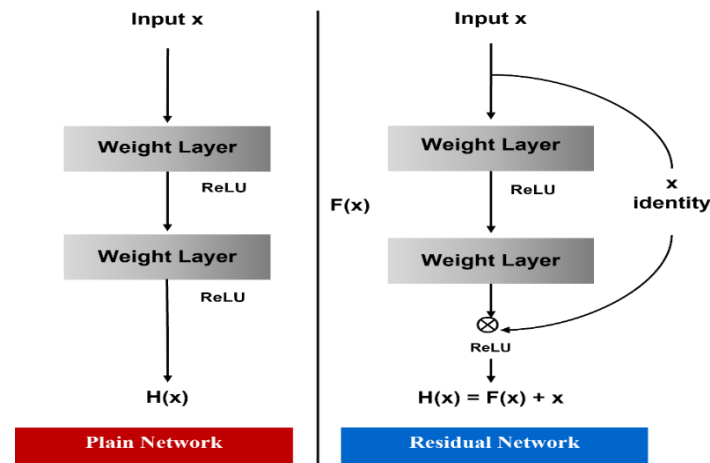


Figure 5. Residual learning.

To implement the proposed methodology, each image classification model was selected as a comparison group, and a comparative experiment was conducted to determine the effectiveness of ResNet-50. When the proposed methodology was applied to ResNet-50, the highest COVID-19 CXR classification accuracy was achieved. There was a high-performance difference in classification accuracy and precision, sensitivity, specificity, and the F1-score. Table 4 shows the results of applying the proposed methodology to each CNN model pretrained using natural images. The classification models used for the control group were VGG16 [39], DenseNet-121 [40], Inception-V3 [41], MobileNetV2 [42], and EfficientNet-B0 [43]. Figure 6 shows accuracy, inference time, MADDs, and parameters for each model.

Table 4. Performance of classification models after applying the proposed method on COVID-19 X-rays.

Models	Accuracy	Precision	Sensitivity	Specificity	F1-Score	Inference Time (MS)	MADDs(G)	Prams
ResNet-50	96.7%	96.2%	96.1%	97.2%	96.1%	45	4.08	23.5 M
VGG16	95.8%	95.2%	94.9%	96.5%	95.1%	225	15.46	134 M
DenseNet-121	95.6%	93.5%	96.2%	95.0%	94.8%	28	2.80	7.9 M
Inception-V3	95.9%	94.8%	95.6%	96.1%	95.2%	35	2.80	21 M
MobileNetV2	95.5%	94.3%	95.1%	95.8%	94.7%	82	0.3	3.5 M
EfficientNet-B0	96.3%	95.4%	95.9%	96.6%	95.6%	47	0.385	5.2 M

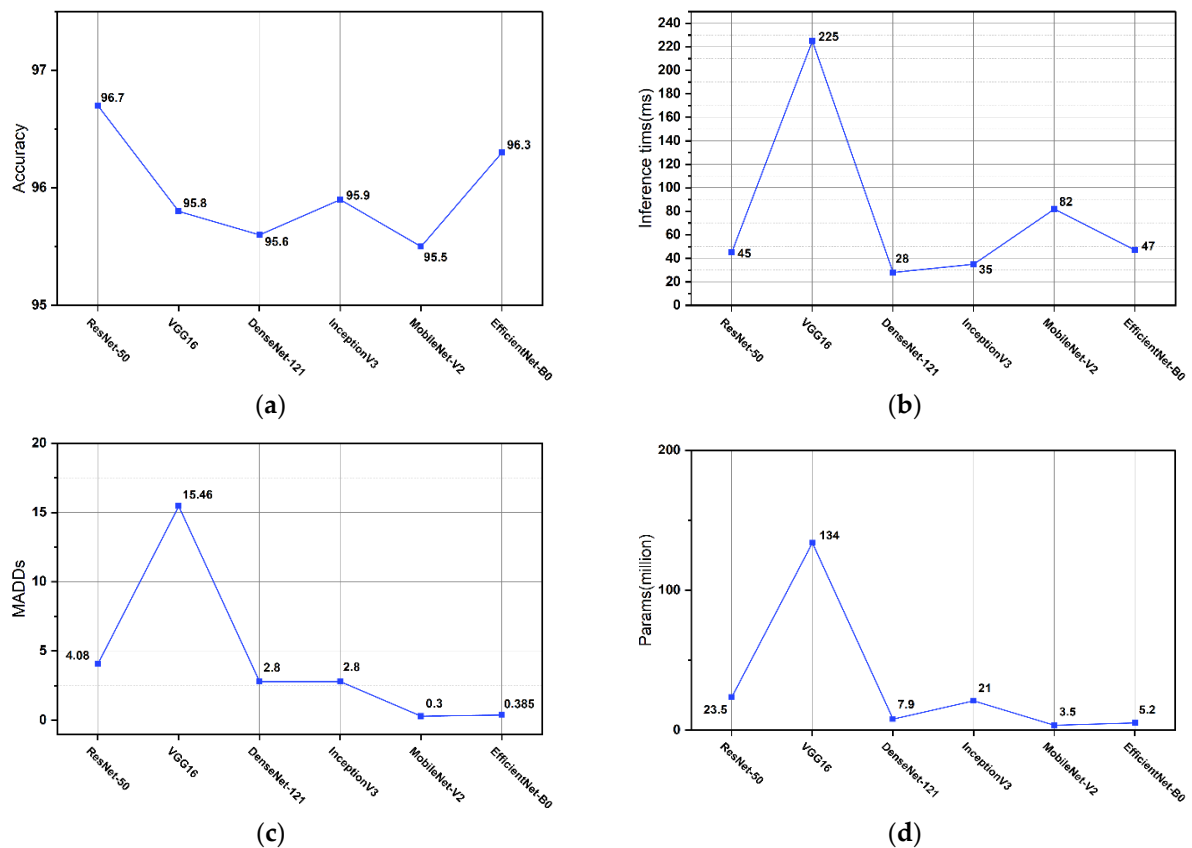


Figure 6. Accuracy, inference time, MADDs, and parameters for each CNN model after applying the proposed methodology. (a) Accuracy of each CNN model with the proposed methodology. (b) Inference time of each CNN model with the proposed methodology. (c) MADDs of each CNN model with the proposed methodology. (d) Parameters of each CNN model with the proposed methodology.

4. Results

4.1. Experimental Settings

The experimental setting comprised the AMD Ryzen 5 5600X 6-core processor (CPU), 32 GB RAM, GeForce 3060TI (NVIDIA), Python 3.7.11, and Pytorch 1.11.0. Cross-entropy was used as the loss function, and AdamW was used as the optimizer [44]. The optimal hyperparameters for the suggested methodology were determined by analyzing the training strategy for each stage using the pretraining and fine-tuning the settings of ConvNet, which shows a high performance in image classification [45]. To determine the optimal value, we used the cosine decay learning rate schedule. The hyperparameters used in the experiment were equally applied to all CNN model comparison groups, as listed in Table 5.

Table 5. Optimal hyperparameter values.

Configurations	Pretraining	Fine-Tuning
Input Size	224 ²	224 ²
Optimizer	AdamW	
Learning Rate	1×10^{-2}	1×10^{-5}
Weight Decay	0.05	1×10^{-8}
Optimizer Momentum	$\beta_1, \beta_2 = 0.9, 0.999$	
Batch Size	256	
Training Epochs	100	20
Learning Rate Schedule	Cosine Decay	
Warm-up Epochs	10	N/A
Warm-up Schedule	Linear	N/A
CLAHE	Clip Limit = 8, Grid Size = 15	

4.2. Evaluation

We used a confusion matrix to evaluate the image classification performance (Table 6). As evaluation indicators, accuracy, precision, sensitivity, specificity, F1-score (harmonic mean), Matthews correlation coefficient (MCC), and balanced accuracy were used, and each value was calculated as shown in Equations (6)–(12).

$$\text{Accuracy} = \frac{1}{N} \sum_{i=1}^2 N_{ii} / \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (6)$$

$$\text{Precision} = \frac{N_{jj}}{\sum_{i=1}^2 N_{ij}}, \text{ where } j = 1 / \frac{TP}{(TP + FP)} \quad (7)$$

$$\text{Sensitivity} = \frac{N_{jj}}{\sum_{i=1}^2 N_{ji}}, \text{ where } j = 1 / \frac{TP}{(TP + FN)} \quad (8)$$

$$\text{Specificity} = \frac{N_{jj}}{\sum_{i=1}^2 N_{ji}}, \text{ where } j = 2 / \frac{TN}{(FP + TN)} \quad (9)$$

$$\text{F1-score} = \frac{2 \times (\text{Precision}_i \times \text{Sensitivity}_i)}{\text{Precision}_i + \text{Sensitivity}_i} \quad (10)$$

$$\text{MCC} = \frac{(TP \times TN) - (FN \times FP)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad (11)$$

$$\text{Balanced accuracy} = \frac{\text{Sensitivity} + \text{Specificity}}{2} \quad (12)$$

Table 6. Confusion matrix of binary classification.

		Predicted Class		
		COVID-19	Non-COVID-19	Total
Actual Class	COVID-19	N_{11}	N_{12}	$\sum_{i=1}^2 N_{1i}$
	Non-COVID-19	N_{21}	N_{22}	$\sum_{i=1}^2 N_{2i}$
	Total	$\sum_{i=1}^2 N_{1i}$	$\sum_{i=1}^2 N_{2i}$	N

4.3. Evaluation Results

Table 7 shows the results of calculating the performance of the model to which the proposed methodology was applied using the performance evaluation indicators. Using the methodology proposed, a classification accuracy level of 96.7%, a precision level of 96.2%, a sensitivity level of 96.1%, a specificity level of 97.2%, an F1-score of 96.1%, an MCC level of 0.93, and a balanced accuracy level of 96.6% were obtained. The proposed model improved accuracy by 9.9%, precision by 13.1%, sensitivity by 9.6%, specificity by 10.2%, the F1-score by 11.9%, MCC by 0.2, and balance accuracy by 9.9% in each part. Compared with the baseline model, the classification performance of the model improved when the proposed methodology was applied. It was confirmed that the performance was significantly improved in sensitivity and specificity, which are important performance measures in the biomedical field.

Table 7. Performance of the baseline and proposed models.

Models	Accuracy ($\pm std$)	Precision ($\pm std$)	Sensitivity ($\pm std$)	Specificity ($\pm std$)	F1-Score ($\pm std$)	MCC ($\pm std$)	Balanced Accuracy ($\pm std$)
Baseline Model	86.8% (± 0.005)	83.1% (± 0.058)	86.5% (± 0.085)	87.0% (± 0.059)	84.2% (± 0.047)	73.4% (± 0.004)	86.7% (± 0.001)
Proposed Model	96.7% (± 0.025)	96.2% (± 0.016)	96.1% (± 0.018)	97.2% (± 0.092)	96.1% (± 0.097)	93.2% (± 0.003)	96.6% (± 0.008)
Improvement Ratio	9.9%	13.1%	9.6%	10.2%	11.9%	20%	9.9%

Figure 7 shows a confusion matrix for the actual and predicted values of the COVID-19 CXR binary classification. Figure 7a shows the classification result of the baseline model, and Figure 7b shows the classification result of the model to which the proposed methodology was applied. In the confusion matrix, the horizontal axis represents the predicted label value, and the vertical axis represents the actual label value. The values from the left-hand side to the right-hand side are from the good prediction of the actual label of the model. The baseline model showed a low accuracy for COVID-19 and non-COVID-19 images. However, the model to which the proposed methodology was applied showed a significant increase in the predicted label value of the actual label value. It can also be observed that the classification accuracy improved in all actual classes.

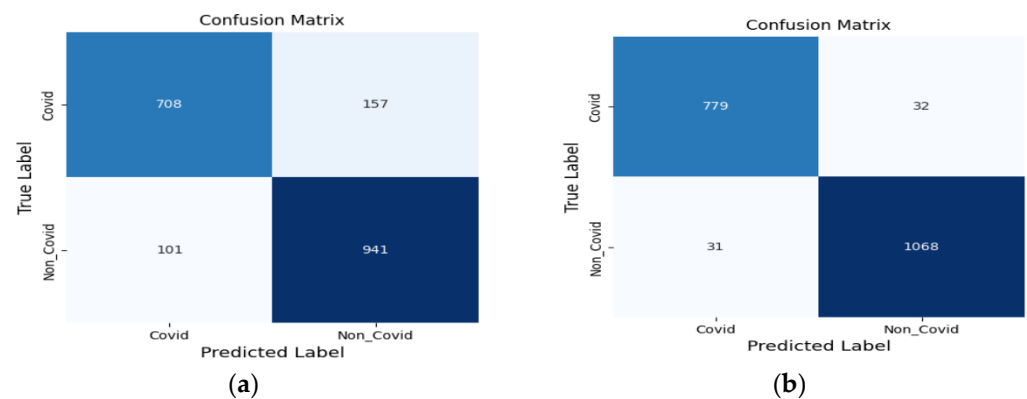
**Figure 7.** Confusion matrix for the COVID-19 CXR classification. (a) (Baseline) non-pretrained ResNet-50. (b) (Proposed) pretrained ResNet-50.

Table 8 shows the results of the comparative experiment conducted by applying the MIMIC CXR and MIMIC CXR (PA) data to ResNet-50. The baseline model results are from the image classification by training and testing the ResNet-50, using Mendeley datasets. The model to which the proposed methodology was applied used the ResNet-50, which was pretrained using natural images and pretrained using the MIMIC CXR (PA) data based on domain extension transfer learning, after which fine-tuning was performed using the Mendeley dataset. Model 1 is the result of fitting the ResNet-50 model pretrained using natural images to the Mendeley dataset. Models 2 and 3 are the results of applying domain extension transfer learning to the Mendeley datasets through ResNet-50 models pretrained using MIMIC CXR and MIMIC CXR (PA) data, respectively. Finally, Model 4 is the result of the proposed methodology using the MIMIC CXR dataset.

The experimental results demonstrate that the PA image datasets significantly improve the classification performance of CXR images and that the use of pretrained models is more effective in classifying images. Additionally, the proposed methodology for pretraining and fine-tuning the target domain rather than simply using a pretrained model is most effective in classifying COVID-19 CXR images.

Table 8. Results of comparative experiments.

Models	Accuracy ($\pm std$)	Precision ($\pm std$)	Sensitivity ($\pm std$)	Specificity ($\pm std$)	F1-Score ($\pm std$)	MCC ($\pm std$)	Balanced Accuracy ($\pm std$)
Baseline (Non-pretrained/Mendeley)	86.8% (± 0.005)	83.1% (± 0.058)	86.5% (± 0.085)	87.0% (± 0.059)	84.2% (± 0.047)	73.4% (± 0.004)	86.7% (± 0.001)
Proposed (Pretrained/MIMIC(PA)—Mendeley)	96.7% (± 0.025)	96.2% (± 0.016)	96.1% (± 0.018)	97.2% (± 0.092)	96.1% (± 0.097)	93.2% (± 0.003)	96.6% (± 0.008)
Model 1 (Pretrained/Mendeley)	95.8% (± 0.061)	94.8% (± 0.023)	95.4% (± 0.024)	96.1% (± 0.013)	95.1% (± 0.104)	91.3% (± 0.005)	95.8% (± 0.013)
Model 2 (Non-pretrained/MIMIC—Mendeley)	95.6% (± 0.027)	94.4% (± 0.014)	95.3% (± 0.023)	95.9% (± 0.037)	94.8% (± 0.013)	90.1% (± 0.002)	95.6% (± 0.018)
Model 3 (Non-pretrained/MIMIC(PA)—Mendeley)	95.6% (± 0.023)	94.5% (± 0.022)	95.0% (± 0.018)	96.0% (± 0.026)	94.8% (± 0.019)	90.1% (± 0.003)	95.5% (± 0.014)
Model 4 (Pretrained/MIMIC—Mendeley)	95.9% (± 0.032)	95.8% (± 0.023)	94.5% (± 0.005)	97.0% (± 0.005)	95.2% (± 0.005)	91.2% (± 0.002)	95.7% (± 0.012)

Figure 8 shows the training and validation accuracies and losses during the training. Table 9 shows the performance comparison of the proposed methodology with other studies. Compared to previous studies, our study showed significant performance improvements in sensitivity and specificity indicators when classifying COVID-19 and non-COVID-19 images. Compared to Elpeltagy et al. [19] who conducted experiments on the same dataset as ours, the model proposed in this study showed significant performance differences in precision and specificity before the five-fold test. After a five-fold test measuring the stability of the model within a small data set, the models showed significant performance differences in accuracy, precision, specificity, and the F1-score. The results of applying the proposed methodology to COVID-19 and non-COVID-19 X-ray imaging classification studies were effectively confirmed.

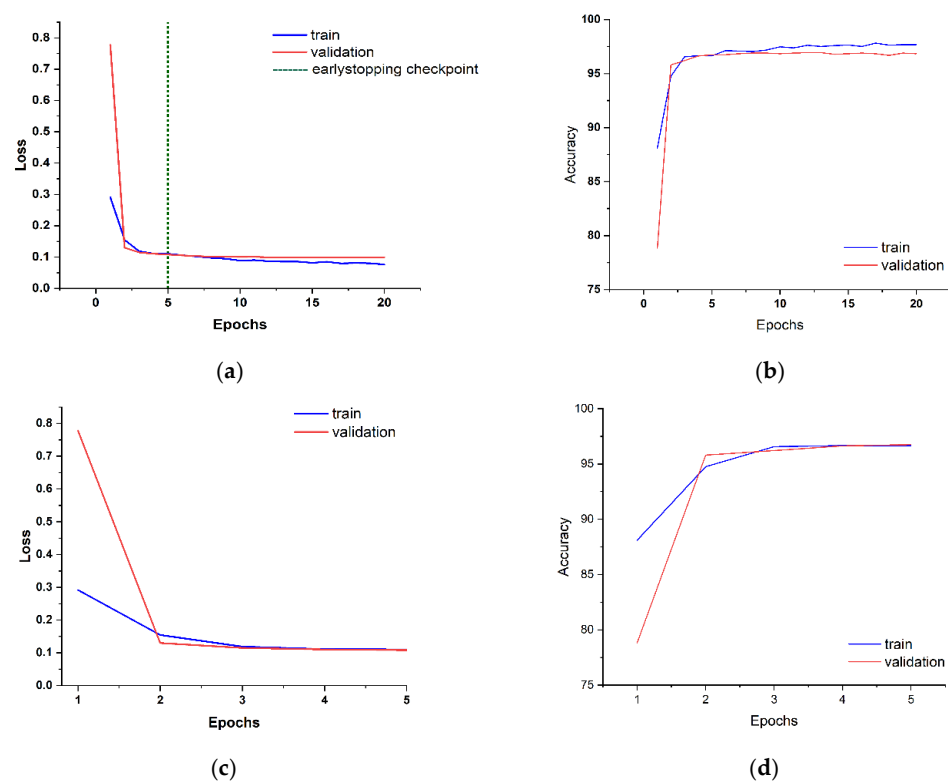


Figure 8. Accuracy and loss curves as a result of applying the proposed model on the Mendeley dataset. (a) Train and validation loss (20 epochs) (early stopping applied). (b) Train and validation accuracy (20 epochs). (c) Training and validation loss. (d) Training and validation accuracy.

Table 9. Performance comparison of the proposed methodology with other studies.

Author	Dataset	DL Model	All Data	COVID Data	Train All/COVID	Validation All/COVID	Test All/COVID	Sensitivity	Specificity
Proposed	COVID-19 and non-COVID-19	ResNet-50	74,101 (MIMIC and Mendeley)	5500	54,941/2426	7558/809	7558/809	96.1%	97.2%
[19]		ResNet-50	9545 (Mendeley)	5500	7636/4400	-/-	1909/1100	97.7%	94.9%
[20]		Dark-CovidNet	627	127	400/101	100/26	-/-	95.1%	95.3%
[24]		CovidXNet	610	305	244/244	61/61	-/-	97.8%	94.7%
[18]		RadiomiX	1381	181	1104/145	-/-	276/36	78.9%	91.0%
[46]		DenseNet-121	2724	1029	1059/526	328/177	1337/326	84.0%	93.0%
[47]		Inception-ResNet-v2	905	419	534/242	92/43	279/134	82.8%	84.3%
[48]		U-Net	5212	275	3285/657	597/120	1330/266	96.3%	93.6%

4.4. Statistical Test

In this study, the difference between the performance levels of the two models was verified through statistical tests for the model to which the baseline model and the proposed methodology were applied. The two-sample test was conducted, and the results are listed in Table 10. It was confirmed that the difference in the image classification performance between the baseline model and the model to which the proposed methodology was applied was statistically significant because the p -value was less than 0.05, and the 95% confidence interval did not include 0. We also confirmed that the difference in the classification performance between the baseline model and each model from model 1 to 4 was statistically significant because the p -value was less than 0.05, and the 95% confidence interval did not include 0.

Table 10. Two sample t-test on the study results.

<i>t</i> -Test		
Hypothesis	<i>p</i> -Value	95% c. i
$H_0 : \mu_{Base} = \mu_{Proposed, Model 1 to 4}$	$p < 0.05$	[9.879, 9.898]
$H_1 : \mu_{Base} \neq \mu_{Proposed, Model 1 to 4}$		[8.962, 8.981]
		[8.816, 8.835]
		[8.809, 8.828]
		[9.132, 9.151]

4.5. XAI

XAI addresses the problem of CNN models where they numerically show good classification performance but cannot intuitively interpret the results. We used Grad-CAM (a slope-based class activation map) among the XAI methodologies [49]. The Grad-CAM is an XAI methodology that follows CAM [50], and a generalized version of CAM utilizes information that influences the class prediction to demonstrate high performance in the artificial neural networks that handle diverse and complex tasks. In this study, the X-ray image diagnosed by the Korean Society of Radiology was used to visualize and interpret the image classification results of the Mendeley dataset for the model to which the proposed methodology was applied. The results of visualization and interpretation validated the classification performance of the model.

Figure 9 shows the results of the visualization through Grad-CAM in the COVID-19 and normal CXRs, which include expert opinions from the Korean Society of Thoracic Radiology (KSTR) [51] on the classification performance of the model to which the proposed methodology was applied and that of the baseline model. The non-COVID-19 CXR represents a CXR image of a normal lung of a male in his 50s, and the model to which the

proposed methodology was applied clearly shows that there is no ground-glass opacity related to COVID-19 around the lungs. The COVID-19 X-ray represents a CXR diagnosed by a specialist showing that ground-glass opacity is spread over the entire left and right lungs of a woman in her 50s. The results visualized by the model are the same as those diagnosed by domain experts. The baseline model provided an interpretation with a slightly larger error than the model to which the proposed methodology was applied.

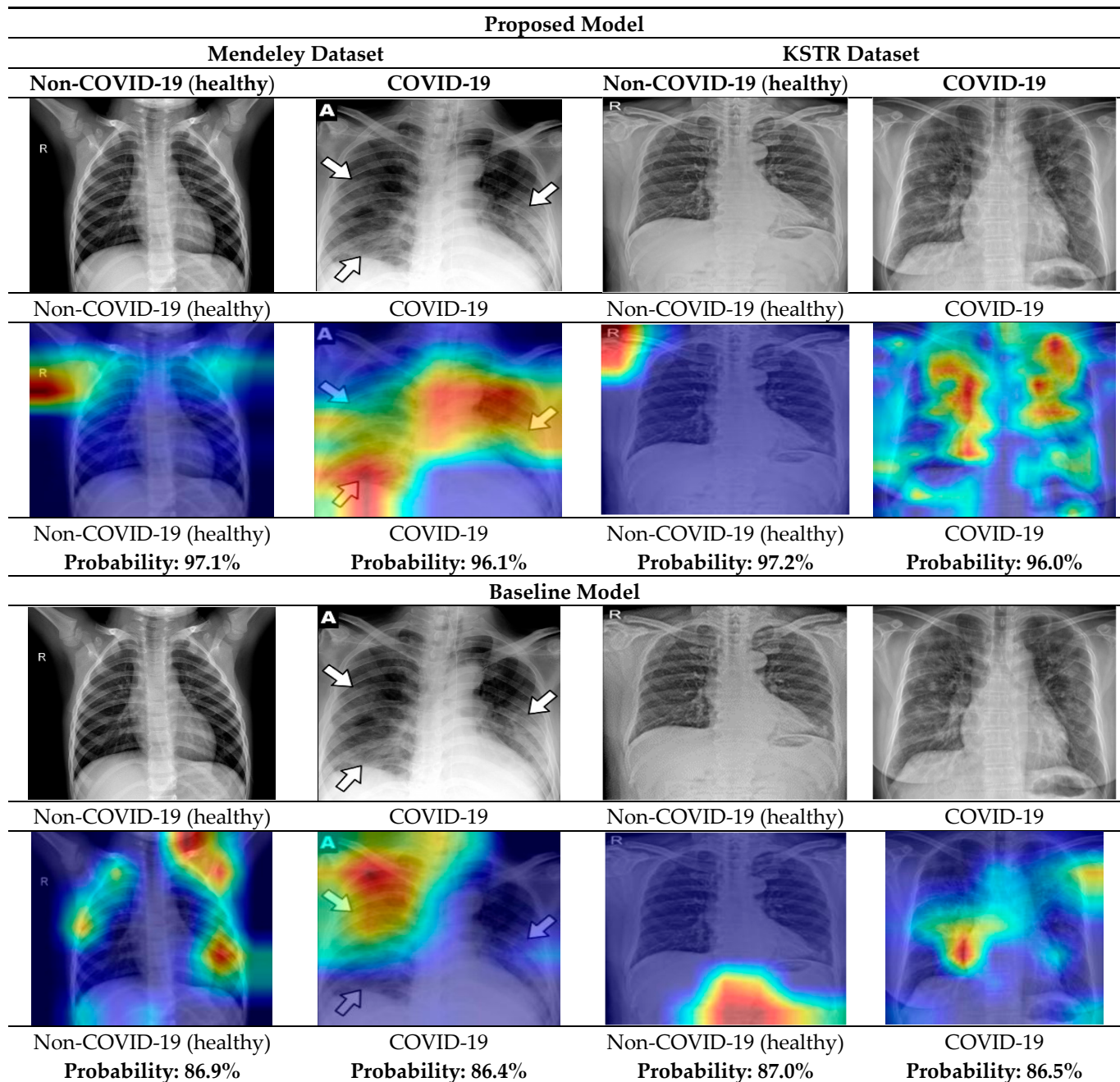


Figure 9. XAI on CXRs.

The Mendeley data provided the lesion locations within the image. As a result of applying Grad-CAM using the information, the location of the lesion was detected from the Mendeley data. The proposed model confirmed the superiority of the classification model by accurately visualizing lung lesions showing symptoms of COVID-19 with a high probability value of over 96.1% compared to the reference model.

5. Discussion

We proposed a deep learning methodology that can be used to effectively classify COVID-19 CXR images. The MIMIC-IV CXR and Mendeley datasets were used, and effective CLAHE was applied for detecting COVID-19 X-ray lesions. The classification model demonstrated improvements in the accuracy and generalization performance of the classification of COVID-19 CXR images through domain extension transfer learning based on ResNet-50. The classification model showed high-performance levels, with an accuracy of 96.7%, precision level of 96.2%, sensitivity level of 96.1%, specificity level of 97.2%, and an F1-score of 96.1%. As a result of applying domain extension transfer learning and the ImageNet pretrained ResNet-50 model, the performance of the classification model in terms of accuracy improved by 9.9%, precision improved by 13.1%, sensitivity by 9.6%, specificity by 10.2%, and the F1-score by 11.9%, when domain extension transfer learning was applied on the MIMIC CXR(PA) data.

To show that the proposed methodology is effective, we carried out comparisons of it with six state-of-the-art studies: Dark-CovidNet [20], CovidXNet [24], RadiomiX [18], DenseNet-121 [46], Inception-ResNet-v2 [47], and U-Net [48]. We also compared the proposed methodology with the study with the same dataset [19]. Among all the methods, the proposed methodology showed the highest performance of 97.2% in specificity, which is one of the important criteria in biomedical image classification.

Using Grad-CAM, one of the XAI methodologies, the numerical results of the proposed methodology were visualized to interpret the results of the performance of the model. A COVID-19 CXR image obtained from the Korean Society of Thoracic Radiology was cited and applied to Grad-CAM. Grad-CAM characterized the lung lesions, and the radiologist findings were used to confirm the superiority of the classification model. Furthermore, by verifying the reliability of the model performance through statistical tests, we showed that the proposed methodology can be used to effectively classify COVID-19 CXR images.

The results of applying the proposed methodology for matching the source domain and the target domain are as follows. The pretraining was conducted with the MIMIC-IV dataset without COVID-19 X-ray images. A fine-tuning strategy was performed to classify the COVID-19 images and non-COVID-19 images well. It was confirmed that the proposed methodology is effective through model performance figures and visual results.

We believe that if another new pandemic with similar lung diseases breaks out, we could apply the proposed methodology to a new scarce CXR image dataset. Therefore, we could quickly respond to screening for new viruses and prevent the spreading of viruses in the local society. In addition, by applying the proposed methodology not only would a quick quarantine response be possible but also a reduction in the burden of a large expenditure to support professionals, expensive inspection kits, and inspection sites. Moreover, the proposed methodology is able to provide visualized information to medical staff and professionals to help them to make good decisions regarding the level of treatment. Therefore, through the proposed methodology, we can build up a proper medical system for patients and the local society.

Recently, COVID-19 lesion detection and classification studies have been conducted using CT and MRI images as well as X-rays. The related studies have achieved more than 95% of the state-of-the-art. In future research, we will apply the domain extension transfer learning presented in this study to the areas of CT and MRI to confirm its applicability in the overall medical image classification field.

6. Conclusions

In this study, we proposed a model for classifying COVID-19 CXR images through domain extension transfer learning. It was confirmed that our proposed model was the most effective in classifying COVID-19 CXR images by performing pretraining and fine-tuning strategies on the target domain. However, there were the following limitations in carrying out the study:

1. The proposed methodology has limitations in that it is applied within the scope of CXR. By additionally collecting CT and MRI medical images and applying them to the proposed model, we are planning an integrated multimodal study by securing patient information related to COVID-19 symptoms and the diagnostic information of medical staff.
2. To prevent the increase in asymptomatic infections due to the spread of vaccines, we plan to expand the scope of the research by additionally collecting the biometric information of asymptomatic patients.
3. In addition, the proposed methodology has limitations in the scope of research applied only to the COVID-19 dataset. In the future, we plan to collect various new lung disease data and apply them to the proposed model for testing.

This study confirms that artificial intelligence significantly contributes to the achievement of pre-emptive and appropriate quarantine responses to COVID-19. It is expected to help reduce the social and economic burden caused by the pandemic and establish a medical system.

Author Contributions: Conceptualization, investigation, methodology, visualization, software, formal analysis, and writing—original draft preparation, K.P.; conceptualization, formal analysis, supervision, writing—review and editing, Y.C. and H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Brain Korea 21 Four (Grant No. 5199990914080).

Institutional Review Board Statement: In case of using MIMIC-CXR, “the collection of patient information and creation of the research resource was reviewed by the Institutional Review Board at the Beth Israel Deaconess Medical Center, who granted a waiver of informed consent and approved the data sharing initiative”.

Informed Consent Statement: Since the datasets were publicly available from PhysioNet (<https://physionet.org/content/mimiciv/2.0/> (accessed on 1 January 2022).) and Mendeley Data (<https://data.mendeley.com/datasets/8h65ywd2jr/3> (accessed on 1 January 2022)), informed consent was not applicable in our case.

Data Availability Statement: “MIT Laboratory-Medical Research Data, PhysioNet” at <https://www.physionet.org/content/mimic-cxr-jpg/> (accessed on 1 January 2022)). “Mendeley Data, Extensive COVID-19 X-ray and CT Chest Images Dataset” at <https://data.mendeley.com/datasets/8h65ywd2jr/3> (accessed on 1 January 2022). “Korean Society of Thoracic Radiology” at <https://kstr.radiology.or.kr/weekly/corona/> (accessed on 1 January 2022).

Acknowledgments: This work was supported by the Brain Korea 21 FOUR.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **2020**, *382*, 727–733. [CrossRef] [PubMed]
2. WorldOMeter. Coronavirus Live Statistics. Available online: <https://www.worldometers.info/coronavirus/> (accessed on 5 October 2022).
3. Misra, S.; Jeon, S.; Lee, S.; Managuli, R.; Jang, I.S.; Kim, C. Multi-Channel Transfer Learning of Chest X-ray Images for Screening of COVID-19. *Electronics* **2020**, *9*, 1388. [CrossRef]
4. Asrani, P.; Eapen, M.S.; Chia, C.; Haug, G.; Weber, H.C.; Hassan, M.I.; Sohal, S.S. Diagnostic approaches in COVID-19: Clinical 404 updates. *Expert Rev. Respir. Med.* **2021**, *15*, 197–212. [CrossRef]
5. Brenner, D.J.; Hall, E.J.; Phil, D. Computed Tomography—An Increasing Source of Radiation Exposure. *N. Engl. J. Med.* **2007**, *357*, 2277–2284. [CrossRef] [PubMed]
6. Wu, G.; Li, X. Mobile X-rays are highly valuable for critically ill COVID patients. *Eur. Radiol.* **2020**, *30*, 5217–5219. [CrossRef]
7. Akl, E.A.; Blazic, I.; Yaacoub, S.; Frija, G.; Chou, R.; Appiah, J.A.; Fatehi, M.; Flor, N.; Hitti, E.; Jafri, H.; et al. Use of Chest Imaging in the Diagnosis and Management of COVID-19: A WHO Rapid Advice Guide. *Radiology* **2021**, *298*, E63–E69. [CrossRef]
8. Akter, S.; Shamrat, F.M.J.M.; Chakraborty, S.; Karim, A.; Azam, S. COVID-19 Detection Using Deep Learning Algorithm on Chest X-ray Images. *Biology* **2021**, *10*, 1174. [CrossRef]

9. Gupta-Wright, A.; MacLeod, C.K.; Barrett, J.; Filson, S.A.; Corrah, T.; Parris, V.; Sandhu, G.; Harris, M.; Tennant, R.; Vaid, N.; et al. False-Negative RT-PCR for COVID-19 and a Diagnostic Risk Score: A Retrospective Cohort Study among Patients Admitted to Hospital. *BMJ Open* **2021**, *11*, e047110. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Okolo, G.I.; Katsigiannis, S.; Althobaiti, T.; Ramzan, N. On the Use of Deep Learning for Imaging-Based COVID-19 Detection Using Chest X-rays. *Sensors* **2021**, *21*, 5702. [\[CrossRef\]](#)
11. Furtado, A.; Andrade, L.; Frias, D.; Maia, T.; Badaró, R.; Sperandio Nascimento, E.G. Deep Learning Applied to Chest Radiograph Classification—A COVID-19 Pneumonia Experience. *Appl. Sci.* **2022**, *12*, 3712. [\[CrossRef\]](#)
12. Soda, P.; D'Amico, N.C.; Tessadori, J.; Valbusa, G.; Guarrasi, V.; Bortolotto, C.; Akbar, M.U.; Sicilia, R.; Cordelli, E.; Fazzini, D.; et al. AlforCOVID: Predicting the Clinical Outcomes in Patients with COVID-19 Applying AI to Chest-X-rays. An Italian Multicentre Study. *Med. Image Anal.* **2021**, *74*, 102216. [\[CrossRef\]](#)
13. Badawi, A.; Elgazzar, K. Detecting Coronavirus from Chest X-rays Using Transfer Learning. *COVID* **2021**, *1*, 403–415. [\[CrossRef\]](#)
14. Ramadhan, A.A.; Baykara, M. A Novel Approach to Detect COVID-19: Enhanced Deep Learning Models with Convolutional Neural Networks. *Appl. Sci.* **2022**, *12*, 9325. [\[CrossRef\]](#)
15. Teixeira, L.O.; Pereira, R.M.; Bertolini, D.; Oliveira, L.S.; Nanni, L.; Cavalcanti, G.D.C.; Costa, Y.M.G. Impact of Lung Segmentation on the Diagnosis and Explanation of COVID-19 in Chest X-ray Images. *Sensors* **2021**, *21*, 7116. [\[CrossRef\]](#)
16. Salvatore, C.; Interlenghi, M.; Monti, C.B.; Ippolito, D.; Capra, D.; Cozzi, A.; Schiaffino, S.; Polidori, A.; Gandola, D.; Ali, M.; et al. Artificial intelligence applied to chest X-ray for differential diagnosis of COVID-19 pneumonia. *Diagnostics* **2021**, *11*, 530. [\[CrossRef\]](#)
17. Le Dinh, T.; Lee, S.H.; Kwon, S.G.; Kwon, K.R. COVID-19 Chest X-ray Classification and Severity Assessment Using Convolutional and Transformer Neural Networks. *Appl. Sci.* **2022**, *12*, 4861. [\[CrossRef\]](#)
18. Guiot, J.; Vaidyanathan, A.; Deprez, L.; Zerka, F.; Danthine, D.; Frix, A.-N.; Thys, M.; Henket, M.; Canivet, G.; Mathieu, S.; et al. Development and validation of an automated radiomic CT signature for detecting COVID-19. *Diagnostics* **2020**, *11*, 41. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Elpeltagy, M.; Sallam, H. Automatic prediction of COVID-19 from chest images using modified ResNet50. *Multimed. Tools Appl.* **2021**, *80*, 26451–26463. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Ozturk, T.; Talo, M.; Yildirim, E.A.; Baloglu, U.B.; Yildirim, O.; Rajendra Acharya, U. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput. Biol. Med.* **2020**, *121*, 103792. [\[CrossRef\]](#)
21. Alghamdi, H.S.; Amoudi, G.; Elhag, S.; Saeedi, K.; Nasser, J. Deep learning approaches for detecting COVID-19 from chest X-ray images: A survey. *IEEE Access* **2021**, *9*, 20235–20254. [\[CrossRef\]](#)
22. El-Din Hemdan, E.; Shouman, M.A.; Karar, M.E. COVIDX-Net: A framework of deep learning classifiers to diagnose COVID-19 in X-ray images. *arXiv* **2003**, arXiv:2003.11055. [\[CrossRef\]](#)
23. Aggarwal, P.; Mishra, N.K.; Fatimah, B.; Singh, P.; Gupta, A.; Joshi, S.D. COVID-19 image classification using deep learning: Advances, challenges and opportunities. *Comput. Biol. Med.* **2022**, *144*, 105350. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Mahmud, T.; Rahman, M.A.; Fattah, S.A. CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. *Comput. Biol. Med.* **2020**, *122*, 103869. [\[CrossRef\]](#)
25. Apostolopoulos, I.D.; Mpesiana, T.A. COVID-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Phys. Eng. Sci. Med.* **2020**, *43*, 635–640. [\[CrossRef\]](#)
26. Sahinbas, K.; Catak, F.O. Transfer learning-based convolutional neural network for COVID-19 detection with X-ray images. In *Data Science for COVID-19*, 1st ed.; Kose, U., Gupta, D., de Albuquerque, V., Khanna, A., Eds.; Elsevier: London, UK, 2021; pp. 451–466. ISBN 9780128245361.
27. Narin, A.; Kaya, C.; Pamuk, Z. Automatic detection of Coronavirus Disease (COVID-19) using X-ray images and deep convolutional neural networks. *Pattern Anal. Appl.* **2021**, *24*, 1207–1220. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Afshar, P.; Heidarian, S.; Naderkhani, F.; Oikonomou, A.; Plataniotis, K.N.; Mohammadi, A. COVID-CAPS: A capsule network-based framework for identification of COVID-19 cases from X-ray images. *Pattern Recognit. Lett.* **2020**, *138*, 638–643. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Sanhita, B.; Mitra, S.; Saha, N. Deep learning for screening COVID-19 using chest x-ray images. In Proceedings of the 2020 IEEE Symposium Series on Computational Intelligence IEEE, Canberra, ACT, Australia, 1–4 December 2020; pp. 2521–2527.
30. Saiz, F.; Barandiaran, I. COVID-19 detection in chest X-ray images using a deep learning approach. *Int. J. Interact. Multimed. Artif. Intell.* **2020**, *6*, 4. [\[CrossRef\]](#)
31. Johnson, A.E.W.; Pollard, T.J.; Greenbaum, N.R.; Lungren, M.P.; Deng, C.; Peng, Y.; Lu, Z.; Mark, R.G.; Berkowitz, S.J.; Horng, S. MIMIC-CXR-JPG, a large publicly available database of labeled chest radiographs. *arXiv* **2019**, arXiv:1901.07042.
32. El-Shafai, W.; Abd El-Samie, F. Extensive COVID-19 X-ray and CT chest images dataset. *Mendeley Data* **2020**, *3*, 10. [\[CrossRef\]](#)
33. Martínez Chamorro, E.; Díez Tascón, A.; Ibáñez Sanz, L.; Ossaba Vélez, S.; Borrueal Nacenta, S. Radiologic diagnosis of patients with COVID-19. *Radiología* **2021**, *63*, 56–73. [\[CrossRef\]](#)
34. Reza, A.M. Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for real-time image enhancement. *J. VLSI Signal Process. Syst. Signal Image Video Technol.* **2004**, *38*, 35–44. [\[CrossRef\]](#)
35. Bashar, A.; Latif, G.; Ben Brahim, G.; Mohammad, N.; Alghazo, J. COVID-19 pneumonia detection using optimized deep learning techniques. *Diagnostics* **2021**, *11*, 1972. [\[CrossRef\]](#) [\[PubMed\]](#)

36. Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Imaging* **2018**, *9*, 611–629. [[CrossRef](#)] [[PubMed](#)]
37. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [[CrossRef](#)]
38. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; IEEE: Piscataway, NJ, USA, 2016. [[CrossRef](#)]
39. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2015**, arXiv:1409.1556.
40. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
41. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J. Rethinking the inception architecture for computer vision. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 2818–2826. [[CrossRef](#)]
42. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.
43. Tan, M.; Le, Q.V. EfficientNet: Rethinking model scaling for convolutional neural networks. *arXiv* **2019**, arXiv:1905.11946.
44. Loshchilov, I.; Hutter, F. Decoupled weight decay regularization. *arXiv* **2017**, arXiv:1711.05101.
45. Liu, Z.; Mao, H.; Wu, C.-Y.; Feichtenhofer, C.; Darrell, T.; Xie, S.A. A convnet for the 2020s. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–20 June 2022; pp. 11976–11986.
46. Harmon, S.A.; Sanford, T.H.; Xu, S.; Turkbey, E.B.; Roth, H.; Xu, Z.; Yang, D.; Myronenko, A.; Anderson, V.; Amalou, A.; et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nat. Commun.* **2020**, *11*, 4080. [[CrossRef](#)]
47. Mei, X.; Lee, H.-C.; Diao, K.-y.; Huang, M.; Lin, B.; Liu, C.; Xie, Z.; Ma, Y.; Robon, P.M.; Chung, M.; et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nat. Med.* **2020**, *26*, 1224–1228. [[CrossRef](#)]
48. Lokwani, R.; Gaikwad, A.; Kulkarni, V.; Pant, A.; Kharat, A. Automated detection of COVID-19 from CT scans using convolutional neural networks. *arXiv* **2020**, 13212. [[CrossRef](#)]
49. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2017; pp. 618–626.
50. Zhou, B.; Khosla, A.; Lapedriza, A.; Oliva, A.; Torralba, A. Learning deep features for discriminative localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2921–2929.
51. COVID-19 X-ray Case Study. Available online: <https://kstr.radiology.or.kr/weekly/corona/COVID-19Cases> (accessed on 2 May 2022).