Vol. 04, No. 1,(2020),021–029, **ISSN:2581-3242**

IJMLNCE JOURNAL

International Journal of Machine Learning and Networked Collaborative Engineering

Journal Homepage: http://www.mlnce.net/home/index.html

DOI : https://doi.org/10.30991/IJMLNCE.2020v04i01.003

Transfer Learning for Detecting Covid-19 Cases Using Chest X-Ray Images

Abdulmohsen N. Alotaibi*

Information Security Expert, Department of Information Protection, Saudi Aramco, Saudi Arabia

annalotaibi@hotmail.com, (https://orcid.org/0000-0002-4792-8207)

Abstract

The COVID-19 pandemic is a global health crisis that has already reached millions of people and led to the deaths of hundreds of thousands of people. COVID-19 has changed how we live, and it changed the way we interact and socialize with other humans, it caused us to alter the way we communicate to keep ourselves and others safe. Furthermore, COVID-19 has affected the economy, and lead a good number of companies to lay-off some of their employees to recover from the economic crisis caused by this virus. On top of that, the health systems are struggling with the high number of people that were infected or would like to be tested, That is why it is critical to developing a more efficient and fast way to detect and treat this illness. This paper utilizes a technique called transfer learning to help in detecting normal, COVID-19, and viral pneumonia cases from a given set of Chest X-Ray images. Four pre-trained models on ImageNet were chosen as the base model, which are ResNet50, VGG19, DenseNet121, and InceptionV3. The performance metrics of each fine-tuned model are overall similar. With an average recall, precision, f1-score, and accuracy of 97.42%, 97.42%, 97.23%, 98.3% respectively.

Keywords

Machine Learning

Deep Neural Network

Transfer Learning

COVID-19

Coronavirus

ImageNet

X-Ray

1. Introduction

Coronavirus disease 2019 (COVID-19) is a new type of coronavirus, that is very contagious and can spread quickly between humans[1]. There are multiple symptoms of the disease, which include fever, cough, sore throat, headaches, and can even cause a severe issue with breathing [2]. COVID-19 is very contagious, and truly changed the way we live, it made us more cautious of our surroundings, it also caused us to alter the way we usually socialize and communicate like humans. Furthermore, COVID-19 had caused catastrophic effects on the economy and forced many companies to lay-off some of their employees to cope with the economic collapse. According to scientific studies, COVID-19 has caused millions of people to be infected, and thousands of people have died because of it.

As of May 2020, around 3.5 million people were infected, and the virus has caused 250 thousand deaths [3]. Despite the low reported fatality rate, according to recent studies, COVID-19, has caused more deaths than ever reported for both SARS and MERS [7]. Figure 1 shows the geographic distribution of the number of COVID-19 cases, as of May 2020 [8].

* Corresponding author
Abdulmohsen N. Alotaibi
Email: annalotaibi@hotmail.com

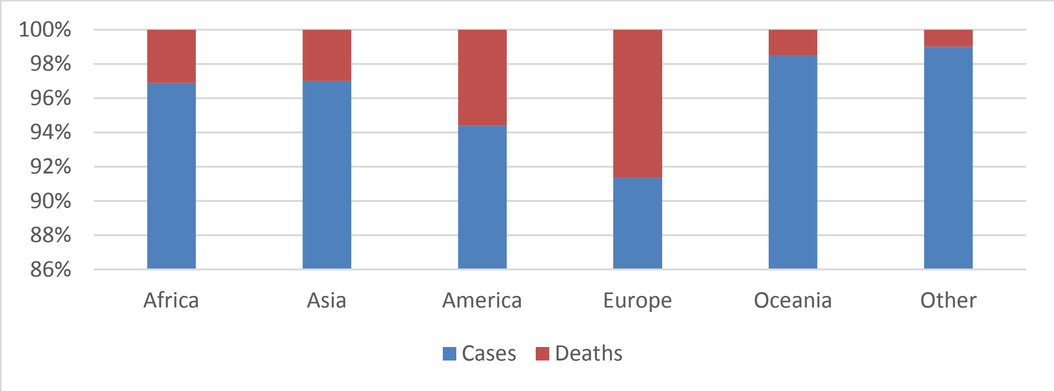

Figure 1:Geographic Distribution of the number of Reported COVID-19, as of May 2020 [8]

Due to the nature of COVID-19 being very contagious, and can quickly spread from one person to another, the health system, even in developed countries, is struggling to cope with the rapid increase in the number of COVID-19 cases. That is why it is essential to develop a fast and efficient way of detecting COVID-19 cases with a high accuracy rate.

Convolutional neural networks (CNNs) has been shown to detect features from images that can help distinguish the nature of the images [13]. They have been used in different computer vision problems, such as enhancing images with low lighting from endoscopy recordings [14], or even in segmentation approaches, such as identifying bone tissues from a given set of X-Ray images [15]. Transfer learning is an approach in deep learning, which utilizes existing pre-trained models on a task, and reuses them to solve another task. It is a well-known approach, mostly used in domains that have a limited supply of training data [16].

In this paper, we present a way of classifying COVID-19, standard, and viral pneumonia cases from X-Ray images by fine-tuning existing pre-trained models on the ImageNet. ImageNet is a project that aims to collect data for use in visual object recognition software. More than 14 million images were used in ImageNet and are organized according to the WordNet hierarchy [12]. Four pre-trained models were selected and fine-tuned: ResNet50, VGG19, DenseNet121, and InceptionV3. The overall results of the fine-tuned models are very promising, with an average recall, precision, f1-score, and accuracy of 97.42%, 97.42%, 97.23%, and 98.3%, respectively.

2. Literature Survey

A good number of studies were conducted, that utilized transfer learning techniques to overcome the issue of lack of data. They utilized existing pre-trained models that were trained on various image datasets, some of these studies focused on COVID-19, they have utilized transfer learning, and used different pre-trained models and fine-tuned it for this specific problem.

Ali et al. [9] developed a convolutional neural network (CNN) based on InceptionV3, ResNet50, and Inception-ResNetV2. The researchers utilized a dataset from an open-source GitHub repository, and they have achieved an accuracy of 98% for ResNet50 pre-trained model, this model had the highest accuracy between all utilized pre-trained models.

Muhammad et al. [4] collected the dataset used in this paper, and they utilized it to detect COVID-19. The researchers used AlexNet, ResNet18, DenseNet201, and SqueezeNet to classify chest X-Ray images into three classes, regular, COVID-19, and viral pneumonia. The classification problem was divided into two groups. The first group consisted of regular and COVID-19 classes, and the second group consisted of normal, viral pneumonia, and COVID-19 cases. The results showed that SqueezeNet produced the highest accuracy of 98.3% for the two groups.

Sarhang et al. [10] proposed a CNN model architecture and compared it with a fine-tuned AlexNet model. The researchers collected the dataset from five different sources, and their CNN showed promising results of around 94% accuracy for both X-ray images and C.T. scans. In contrast, AlexNet showed 98% accuracy of X-ray images and 82% accuracy for C.T. scans.

Ioannis et al. [11] collected datasets from different sources, and fine-tuned VGG19, Mobile Net, Inception, and Xception models to detect COVID-19 cases. Their results show that VGG19 and Mobile Net achieved the highest accuracy of 98.75% and 97.40%, respectively.

Shervin et al. [17] collected X-Ray images from two different datasets and merged them into one dataset, which contains 2,031 training data, and 3,040 testing data. They have used ResNet50, SqueezeNet, ResNet18, and DenseNet-121, as base models, and fine-tuned them to detect COVID-19 cases from given chest X-Ray images, they have achieved a sensitivity of 97.5%, and specificity of 90% on average.

Mohamed et al. [18] utilized a COVID-19 dataset created by a postdoctoral fellow from the University of Montreal; they organized the collected dataset into to sub-sets, one for training and one for testing, their collected dataset consists of four classes, COVID-19, normal, pneumonia bacterial, and pneumonia virus. They have introduced Generative Adversarial Network (GAN) to overcome the issue of lack of data, and to avoid overfitting. They utilized three deep transfer models, which are the AlexNet, GoogleNet, and ResNet18. They have divided the tests into three sections; the first section includes all four target classes. In contrast, in the second section, it contains only three selected target classes, and in the third section, it includes only two types. Their first test section, which consists of all four target classes, they have reported that GoogleNet achieved the highest results, with an 80.6% accuracy. In the second test section, with only three target classes, AlexNet reported the highest marks, with an accuracy of 85.2%. In the third and final section, which includes two target classes (COVID-19, and normal), GoogleNet achieved the highest results with an accuracy of 100%.

Lawrence et al. [19] collected their dataset from Radiopaedia, The Italian Society of Medical and Interventional Radiology (SIRM), and COVID-19 Chest X-Ray Dataset in GitHub. They used ResNet50 pre-trained model and fine-tuned it for COVID-19 cases, they achieved an accuracy of 89.2%, with a detection rate of 80.39% of the COVID-19 cases. They had a true negative rate (TNR) of 0.99 and an AUC of 0.95.

3. Proposed Methodology

In this paper, we propose a method of utilizing existing pre-trained models on ImageNet and constructing additional layers. In order to retrain them on a given COVID-19 dataset consisting of X-Ray images that belong to normal, viral pneumonia, and COVID-19 patients. In this paper, ResNet50, VGG-19, DenseNet121, and InceptionV3 were chosen as ImageNet pre-trained base models, then multiple layers were appended at the end of the model, along with an output layer.Figure2 shows a summary of the designs of the models utilized in this paper.

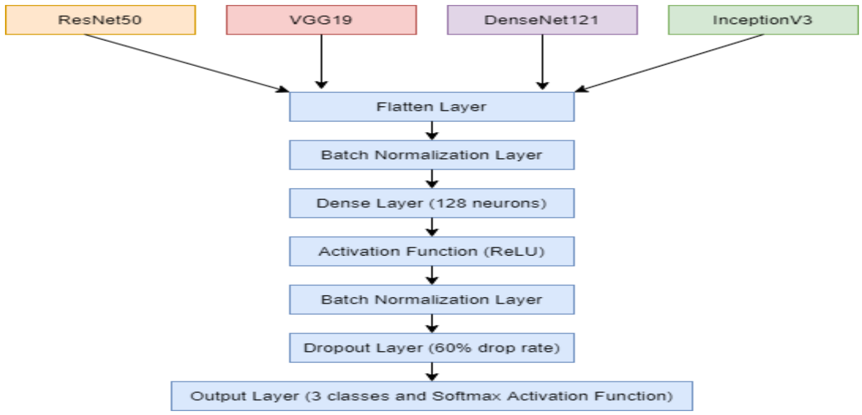

Figure 2:Summary of Models Design

Tensorflow was chosen to implement the approach. At first, the training, testing, and validation sets were defined, then data augmentation was applied to the training set. Finally, the pre-processing function is called on all images before feeding them to the model.

3.1. Experimental Data

The paper utilizes a COVID-19 Radiography Database [4]. This dataset consists of 219 COVID-19 cases, 1,341 Normal cases, and 1,345 Viral Pneumonia cases, as shown in Figure 3.

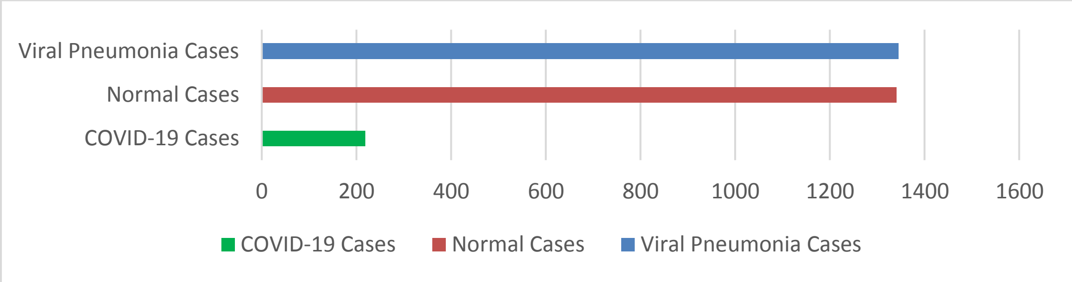

Figure 3:COVID-19 Radiography Database, Cases Distribution [4]

The authors of the dataset have collected six different datasets and merged them into one dataset. The dataset is created from various sources, first, is The Italian Society of Medical and Interventional Radiology (SIRM) which contains multiple COVID-19 cases, another dataset that the authors considered is The Novel Corona Virus 2019 Dataset, which has a good number of COVID-19 cases. Finally, the authors have looked into and considered different articles and sources, which include sources available on chest imaging at thread reader, Radiology Society of North America (RSNA) dataset from 2018, and Kaggle chest X-Ray dataset.

3.2. Data Augmentation

To avoid overfitting, the data augmentation process will be applied to the training set; the process will ensure that we have an equal number of samples for each class. The data augmentation process first calculates the number of samples of the majority class and multiplies it by 2 to get the number of new samples to generate (N). The function will loop over each image from each class, and applies data augmentation to it. The function will curve cyclically over the training set until N reaches zero for each class. For each image, one or more of the following functions is applied:

1-***Brightness increase:*** A value between [1.05, 1.1] is selected and is multiplied by each pixel value.

2-***Adjusting the image contrast:*** A gamma value between [1.05, 1.11] is selected.

3-***Elastic transform:*** The alpha value will be twice that of the image height, and the sigma value is 10% of the image height [5]. The image height for InceptionV3 is 299 pixels, and for the other models, it is 224 pixels.

4-***Zooming:*** The image is scaled by 105% to 110% of the original size.

5-***Translation:*** Left/Top or right/Bottom translation is randomly chosen to sample the pixels by 90% to 97% of the original image size.

6-***Rotation:*** A random rotation of [-5, 5] degrees is applied to the image.

To demonstrate the effect of the augmentation process on each original input of an X-Ray image, Figure 4 shows a sample of each class, as well as the results after applying the data augmentation process.

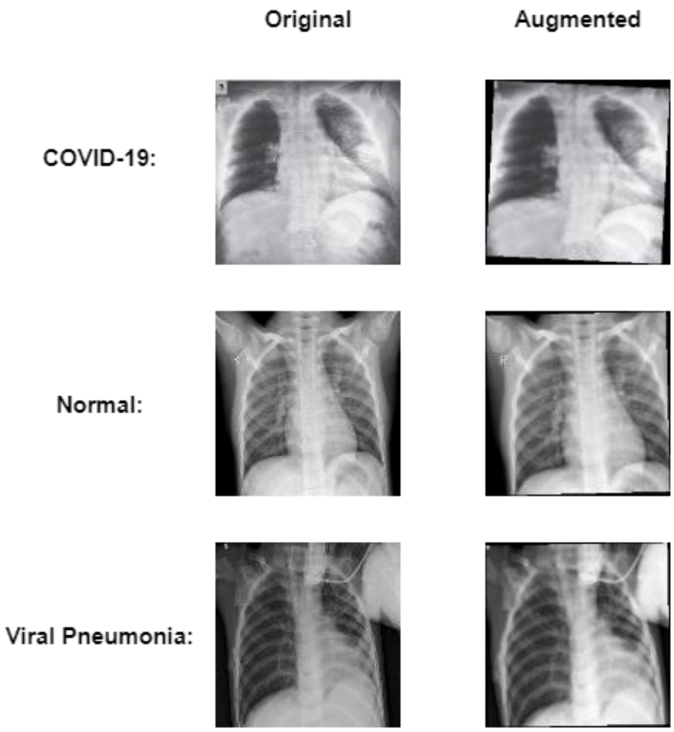

Figure 4:Dataset Original Image Example In Comparison to The Augmentation Process Results

3.3. Applying Pre-processing Steps

Before feeding the images into the model, each image is resized by the appropriate image size that the pre-trained model used during training. For ResNet50, VGG19, and DenseNet121, the picture was resized to 224x224. For InceptionV3, the image is resized to 299x299. Then each image is pre-processed based on the same pre-processing mode used when training the base models [6]. The "Caffe" mode converts the images from RGB to BGR, then zeros-center each color channel with respect to the ImageNet dataset, without scaling. The "T.F." scales the pixels between the values -1, and 1. The "Torch" mode will scale the pixels between 0 and 1, then normalize each channel concerning the ImageNet dataset. Table 1 shows a summary of the pre-processing function applied based on the base model.

Table 1:Pre-processing Steps Summary

	Resize	*Pre-processing mode*
ResNet50	224x224	Caffe
VGG19	224x224	Caffe
DenseNet121	224x224	Torch
InceptionV3	299x299	TF

3.4. Fine-tuning Pre-trained Models

The top layer was removed from all base models, and then the base model is fine-tuned by adding a flatten layer, followed by a batch normalization layer. Then a Dense layer is added with 128 neurons and Rectified Linear Unit (ReLU) activation function, this layer is followed by a batch normalization layer and a dropout layer with a dropout rate of 60%. Finally, a layer is added with three neurons, which is equal to the number of target classes, and Softmax was used as the activation function for the output layer. The training is done twice, once with all base model layers frozen, and the second run is executed after unfreezing the layers, starting from the selected convolutional layer. Table 2 summarizes the starting position in which the layers were made trainable for each base model used.

Table 2: Selected Position in Which the Layers Were Trainable

	Layer Position
ResNet50	165 and up
VGG19	17 and up
DenseNet121	423 and up
InceptionV3	299 and up

To ensure consistent results on each run, all batch normalization layers are by default unfrozen, and initially their moving average and moving variance are initialized to zeros. The model is compiled with an Adam optimizer with a starting learning rate of 0.001. Also, the ReduceLROnPlateau was utilized to allow for dynamic learning rate reduction; and will reduce the learning rate by a factor of 0.05, if five epochs have passed without improvements. An early stop callback was utilized to stop the training if the model did not improve after 10 epochs.

4. Result and Discussion

To evaluate the models, Stratified K-fold Cross Validation is utilized, with the number of folds (K)equals to 5. The evaluation metrics that were used where accuracy, precision, recall, and f1-score:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1 - score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (4)$$

Where T.P. is the true positive cases, F.N. is the false-negative cases, T.N. is the true negative cases, and finally, F.P. is the false-positive cases. The split training size in each fold is 2324, and the testing split size is 581. The test split is further divided into 50% as a testing set, and 50% as a validation set for the model. The testing set contained in each fold consists of 19 COVID-19 cases, 126 normal cases, and 145 viral pneumonia. And since we are using 5-folds cross-validation, then all confusion matrices produced for each fold are added together to get one summarized confusion matrix per model. Table 3 summarizes the confusion matrix output of each model.

Table 3: Summarized Confusion Matrix Per Model, where "C" is COVID-19, "N" is Normal, and "V" is Viral Pneumonia

		C	*N*	*V*
RESNET50	***C***	91	1	3
	N	0	617	13
	V	0	23	702

		C	*N*	*V*
VGG19	***C***	91	3	1
	N	0	620	10
	V	2	14	709

		C	N	V
DenseNet121	***C***	92	2	1
	N	0	628	2
	V	3	20	702

		C	N	V
InceptionV3	***C***	94	0	1
	N	2	608	20
	V	2	29	694

The recall, precision, f1-score, and accuracy are calculated from the summarized confusion matrix of each model shown in table 3. On average, the recall, precision, f1-score, and accuracy are 97.42%, 97.42%, 97.23%, and 98.3% respectively. The average performance metrics of each model are similar overall. Table 4 shows the recall, precision, f1-score, and accuracy of each model.

Table 4: Overall Performance Metrics for Each Model, where "R.C." is Recall, "P.R." is Precision, "F1" is F1-score, and "ACC" is accuracy. Note That "C" is COVID-19, "N" is Normal, and "V" is Viral Pneumonia

		R.C.	PR	F1	ACC
ResNet50	***C***	100%	96%	98%	99.72%
	N	96%	98%	97%	97.45%
	V	98%	97%	97%	97.31%
	Mean	**98%**	**97%**	**97.33%**	**98.16%**

		RC	PR	F1	ACC
VGG19	***C***	98%	96%	97%	99.59%
	N	97%	98%	98%	98.14%
	V	98%	98%	98%	98.14%
	Mean	**97.66%**	**97.33%**	**97.66%**	**98.62%**

		RC	PR	F1	ACC
DenseNet121	***C***	97%	97%	97%	99.59%
	N	97%	100%	98%	98.34%
	V	100%	97%	98%	98.21%
	Mean	**98%**	**98%**	**97.66%**	**98.71%**

		RC	PR	F1	ACC
InceptionV3	***C***	96%	99%	97%	99.66%
	N	95%	97%	96%	96.48%
	V	97%	96%	96%	96.41%
	Mean	**96%**	**97.33%**	**96.33%**	**97.52%**

The results show that the average recall, precision, f1-score, and accuracy are 97.42%, 97.42%, 97.23%, and 98.3%, respectively. It is clear from the results that there is great potential in using transfer learning techniques, to overcome the issue of lack of data. It also clarifies that utilizing machine learning and incorporating it in medicine for the detection and treating of patients with various diseases such as COVID-19 shows a great promise. The use of the transfer learning technique in this paper was due to the lack of datasets for COVID-19 because it is a relatively new virus, in theory, if we increased the volume of the real data. The models can be trained from scratch, and we should be able to achieve higher results.

The models discussed in this paper can be used in the medical field. They can be incorporated and used after receiving the X-Ray results directly for each patient. Further it can be used as an automated diagnosis tool, or even be used as a priority adjuster, where the patients that the model predicted to have COVID-19 with a high probability are treated first, depending on their condition.

5. Conclusion

COVID-19, has spread to over 3.5 million people, and even caused more than 250 thousand confirmed deaths, as reported on May 2020, it is very contagious, and it has changed the way we live, communicate, and socialize with others. Machine learning has shown a great promise in the field of medicine in recent years; it has been used to build diagnosis models, as well as models that can find a remedy for specific diseases. COVID-19 virus led us to explore and try to find a comfortable and yet efficient way of detecting if a person was infected with this virus. Since COVID-19 is relatively new, and we still lack publicly available data for further analysis and modeling. In this paper, we have utilized transfer learning, and fine-tuned existing models trained on ImageNet to help in detecting COVID-19, regular, and viral pneumonia cases from a given set of X-Ray images. The performance metrics of each fine-tuned model are similar overall with an average recall, precision, f1-score, and accuracy of 97.42%, 97.42%, 97.23%, and 98.3%, respectively.

References

[1]. World Health Organization, COVID-19 Health Topic - https://www.who.int/health-topics/coronavirus

[2]. WebMD, Coronavirus and COVID-19: What You Should Know - https://www.webmd.com/lung/coronavirus

[3]. World Health Organization, Coronavirus disease (COVID-19) Situation Report - 108 - https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200507covid-19-sitrep-108.pdf?sfvrsn=44cc8ed8_2

[4]. M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, M.B.I. Reaz, "Can A.I. help in screening Viral and COVID-19 pneumonia?" arXiv preprint, 29 March 2020 - https://arxiv.org/abs/2003.13145 - https://www.kaggle.com/tawsifurrahman/covid19-radiography-database (Version 1)

[5]. Elastic Transform for Data Augmentation - https://www.kaggle.com/bguberfain/elastic-transform-for-data-augmentation (Version 3)

[6]. ImageNet Input Pre-processing - https://www.tensorflow.org/api_docs/python/tf/keras/applications/imagenet_utils/preprocess_input

[7]. BMJ 2020;368:m641, Coronavirus: covid-19 has killed more people than SARS and MERS combined, despite lower case fatality rate - https://www.bmj.com/content/368/bmj.m641

[8]. COVID-19 situation update worldwide - https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases

[9]. Ali Narin , Ceren Kaya , ZiynetPamuk, "Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks" arXiv preprint, 24 March 2020 - https://arxiv.org/abs/2003.10849

[10]. HalgurdMaghdid, Aras Asaad, KayhanGhafoor, Ali Sadiq, "Diagnosing COVID-19 Pneumonia from X-Ray and CT Images using Deep Learning and Transfer Learning Algorithms" March 2020 - https://www.researchgate.net/publication/340374481_Diagnosing_COVID-19_Pneumonia_from_X-Ray_and_CT_Images_using_Deep_Learning_and_Transfer_Learning_Algorithms

[11]. Ioannis D. Apostolopoulos, TzaniBessiana, "Covid-19: Automatic detection from X-Ray images utilizing Transfer Learning with Convolutional Neural Networks" arXiv preprint, 25 March 2020 - https://arxiv.org/abs/2003.11617

[12]. The ImageNet Project Overview - http://image-net.org/about-overview

[13]. Alex Krizhevsky, IlyaSutskever, Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks" - https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[14]. Pablo Gómez, Marion Semmler, Anne Schützenberger, Christopher Bohr, Michael Döllinger, "Low-light image enhancement of high-speed endoscopic videos using a convolutional neural network," 21 March 2019 - https://link.springer.com/article/10.1007/s11517-019-01965-4

[15]. Cosmin CERNAZANU-GLAVAN, Stefan HOLBAN, "Segmentation of Bone Structure in X-ray Images using Convolutional Neural Network," February 2013 - https://www.researchgate.net/publication/269978717_Segmentation_of_Bone_Structure_in_X-ray_Images_using_Convolutional_Neural_Network

[16]. Karl Weiss, TaghiKhoshgoftaar, DingDing Wang, "A survey of transfer learning," December 2016 - https://www.researchgate.net/publication/303600638_A_survey_of_transfer_learning

[17]. ShervinMinaee, RaheleKafieh, Milan Sonka, ShakibYazdani, GhazalehJamalipour, Soufi "Deep-COVID: Predicting COVID-19 From ChestX-Ray Images Using Deep Transfer Learning" April 2020 - https://www.researchgate.net/publication/340806168_Deep-COVID_Predicting_COVID-19_From_Chest_X-Ray_Images_Using_Deep_Transfer_Learning

[18]. Mohamed Loey, FlorentinSmarandache وNourEldeen Mahmoud Khalifa "Within the Lack of Chest COVID-19 X-ray Dataset: A Novel Detection Model Based on GAN and Deep Transfer Learning" April 2020 - https://www.researchgate.net/publication/340840794_Within_the_Lack_of_Chest_COVID-19_X-ray_Dataset_A_Novel_Detection_Model_Based_on_GAN_and_Deep_Transfer_Learning

[19]. Lawrence O. Hall, Rahul Paul, Dmitry B. Goldgof, and Gregory M. Goldgof "Finding COVID-19 from Chest X-rays using Deep Learning on a Small Dataset" May 2020 - https://www.techrxiv.org/articles/preprint/Finding_COVID-19_from_Chest_X-rays_using_Deep_Learning_on_a_Small_Dataset/12083964

Author's Biography

Abdulmohsen N. Alotaibihas, a bachelor's degree in Computer Science, from Al-Imam Mohammad Ibn Saud Islamic University, Computer Science Department, Saudi Arabia.Abdulmohsen N. Alotaibispecializes in Information Security, and has many professional certifications which are CEH, GISF, GWEB, GCIH, ITIL v4 Foundation, and COBIT 5 Foundation. Abdulmohsen N. Alotaibi has been applying and learning Machine Learning for over three years now and has worked on multiple projects in various areas of Machine Learning.