An Investigation on Transfer Learning for Classification of COVID-19 Chest X-Ray Images with Pre-trained Convolutional-based Architectures

Mobina Abdoli Nemati
Electrical Engineering Department,
Iran University of Science and Technology
Tehran, Iran
Mobina_abdolinemati@elec.iust.ac.ir

Amirreza BabaAhmadi
School of Mechanical Engineering,
College of Engineering,
University of Tehran
Tehran, Iran
Babaahmadi.amir@ut.ac.ir

Abstract**—Medical image analysis techniques have been helpful for rapidly identifying COVID-19. Due to the difficulty of identifying critical visual features for many patients, machine learning techniques are promising for diagnosing infection. Given the variety of approaches to using convolutional-based architectures for medical image analysis tasks, we used the most well-known and powerful deep neural networks to provide a concise and reliable source of information on transfer learning-based methods, paving the way for future research. The article describes an investigation into the detection of COVID-19 in medical chest X-ray images using 27 pre-trained convolutional neural networks. This study utilized a binarized version of a few publicly available datasets from COVID-19 patients and healthy controls. The results of the 5-fold cross-validation demonstrate that architectures such as EfficientNet networks, MobileNet, Inception-ResNetV2, and NasNet-Large perform significantly better on the aforementioned binary pattern recognition task.**

Keywords—Deep Learning, Transfer Learning, COVID-19 Detection, Medical Computer Vision, Medical Image Analysis

I. Introduction

In medical image analysis, computer vision and deep learning-based methods have been chosen as reliable and robust methods in recent years. Manual image reading is time-consuming and prone to human error. As a result, the need for automated medical image analysis tools to analyze chest X-ray images is critical.

Numerous researchers have examined the application of computer vision and deep learning models to detect COVID-19 in medical images. The following are some of the most significant works:

The performance of integrated stacked deep convolutional neural networks on a multi-class problem was investigated in [1]. They concluded that ResNet101, Xception, Inception V3, MobileNet, and NasNet outperform other networks in detecting COVID-19, pneumonia, and normal cases. Reference [2] introduces a new architecture based on ResNet. They used a variety of kernel sizes to extract additional semantic features from the input images. They used a convolutional neural network to extract features from the images in another article [3] and then ensemble learning-based classifiers to classify their multi-resolution dataset. Reference [4] describes using a hybrid architecture based on CNN and LSTM to simultaneously detect the presence of COVID-19 and pneumonia cases in a particular dataset.

In [5], they developed Generative Adversarial Networks and Long Short-Term Memory networks in their novel architecture without utilizing feature extraction/selection layers. CNN and SVM were used in conjunction in [6] to extract features and classify COVID-19 from pneumonia and normal cases. DarkNet-53 was used as the backbone network in [7] to extract features via its skip connections. They examined 15 different pre-trained CNN architectures in [8]. The authors conclude that VGG19 outperforms all other examined networks. Multiple GAN architectures and a transfer learning approach were used in [9] to classify normal, pneumonia bacterial, and pneumonia virus. Multiple custom CNNs were designed in [10] to achieve high performance on the classification task for a specific dataset containing chest x-ray images.

In [11], three pre-trained convolutional architectures were evaluated: Inception V3, Xception, and ResNet. They reported that when compared to other networks, the Inception model provides the highest accuracy. They fine-tuned the final layer of ResNet18, ResNet50, SqueezNet, and DenseNet-121 in [12] and compared the results of those algorithms to four binary classifications of COVID-19 in normal cases. In [13], a hybrid model based on multiple architectures such as VGG-16, InceptionV3, and ResNet50 was used in conjunction with a voting layer to predict the final output. They examined the effectiveness of preprocessing data and its effect on the classification results of COVID-19 medical images in [14]. In [15], a tailored deep neural network based on VGG-19 and ResNet-50 was developed to classify COVID, pneumonia, and normal cases. Reference [16] introduced a new model (CoroNet). This model was created using the Xception architecture. They evaluated their newly introduced architecture's ability to classify pneumonia-bacterial, pneumonia-viral, COVID, and normal cases.

Numerous studies have been conducted to study the identification of COVID-19 images such as CT-Scan and X-ray medical images. However, a concise and comprehensive study examining the power of various deep neural networks in the domain of transfer learning appears to be uncommon among published articles. This paper is structured as follows: Section II discusses various deep learning architectures, delves into setting hyperparameters, and explores the training phase. Section III discusses the binarized dataset that we used to train various algorithms. Section IV describes the evaluation metrics used to assess the performance of the algorithms mentioned above. Section V discusses the results of the training and testing procedures and introduces suitable algorithms for detecting COVID-19 in chest x-ray images. Finally, Section VI concludes the paper and makes recommendations for future work and research.

II. DEEP NEURAL NETWORKS FOR TRANSFER LEARNING

Deep neural networks can extract semantic features from the data they process. The objective of this study was to compare the performance of 27 deep learning neural networks on a binary pattern recognition task. Numerous researchers have used well-known architectures in their work, including ResNet[17], MobileNet[18], VGG[19], Xception[20], EfficientNet[21], and DenseNet[22]. The networks discussed previously can be extremely useful for transfer learning tasks in medical computer vision and deep learning. TABLE I summarizes the key characteristics of the algorithms examined. Transfer learning is a technique that uses neural networks that have been trained on large datasets such as ImageNet.

This paper's transfer learning task froze all convolutional layers except the final one in our custom-designed dense layers to classify COVID-19 and normal cases. Dense layers were designed in the following manner: we used two fully connected layers for the binarized classification problem. The first layer contains 64 neurons, the second layer contains two neurons, the dropout is 0.5, the activation function is the ReLU function, and the final layer contains two neurons. The learning rate hyperparameter is 0.002, and the batch size is 64. For training and evaluation purposes, the number of epochs has been set to 120. This task incorporated the Adam optimization method.

TABLE I. PROPERTIES OF DEEP LEARNING ALGORITHMS FOR TRANSFER LEARNING-BASED CLASSIFICATION (IN THIS PAPER)

#	Algorithm	Depth	Size (Mb)	Parameters (millions)	Input Image Size
1	Inception V3	159	92	23.9	299x299
2	ResNet152V2	152	232	60.3	299x299
3	DenseNet121	121	33	8.06	224x224
4	DenseNet169	169	57	14.3	224x224
5	DesneNet201	201	80	20.2	224x224
6	EfficientNetb0	-	29	5.3	224x224
7	EfficientNetb1	-	31	7.8	224x224
8	EfficientNetb2	-	36	9.1	224x224
9	EfficientNetb3	-	48	12.3	224x224
10	EfficientNetb4	-	75	19.4	224x224
11	EfficientNetb5	-	118	30.5	224x224
12	EfficientNetb6	-	166	43.2	224x224
13	EfficientNetb7	-	256	66.6	224x224
14	MobileNet	88	16	4.2	224x224
15	ResNet50	50	98	25.6	224x224
16	ResNet101V2	101	171	44.6	299x299
17	ResNet152V2	152	232	60.3	299x299
18	VGG19	19	549	143.6	224x224
19	Xception	126	88	22.9	299x299
20	MobileNetV2	88	14	3.5	224x224
21	ResNet101	101	171	44.6	224x224
22	ResNet152	152	232	6.4	224x224
23	Inception-ResNetV2	572	215	55.9	299x299
24	NasNet-Large	-	343	88.9	331x331
25	NasNetMobile	-	23	5.3	224x224
26	ReNet50V2	50	98	25.6	224x224
27	VGG16	16	528	22.9	224x224

III. DATASET

We conducted our research using two publicly available datasets. Initially, we extracted 180 images from 118 COVID-19 cases via GitHub (https://github.com/ieee8023/covid-chestxray-dataset). We used another publicly available dataset to provide normal cases (https://www.kaggle.com/c/rsna-pneumonia-detection-challenge). This dataset contains 8851 non-pneumonia cases and 6012 pneumonia cases. Because this study aims to classify COVID-19 cases from the normal category, we considered only normal cases and excluded pneumonia samples. Since the number of COVID-19 samples is significantly less than the number of normal samples, we are left with an unbalanced dataset. In this case, the classifier cannot proceed with the classification correctly. The reality is that the classifier becomes biased in favor of the class with the most samples. We used the random undersampling technique to ensure that two classes had the exact sample count to address this issue. Additionally, we normalized the images in our dataset to eliminate the possibility of encountering both vanishing gradient and exploding gradient phenomena.

We used data-augmentation techniques to increase our dataset's robustness against possible image changes. We used a random rotation method for 0-45 degrees. This data augmentation improves training efficiency and prevents the model from overfitting. COVID-19 and normal samples from our dataset are depicted in Fig. 1. The dataset was divided into 80% for training and 20% for evaluation and testing.

Fig. 1. (a) Normal (Negative) case

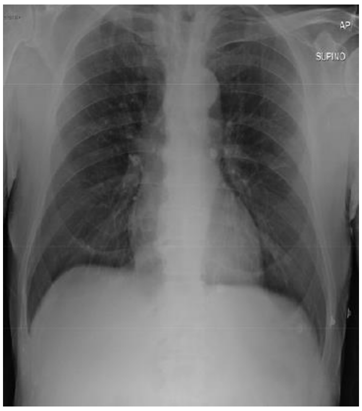

Fig 1. (b) COVID-19 (Positive) case

IV. Evaluation

This section defines specific criteria for determining the most efficient classifier for the given problem. True Positive indicates that the classifier can accurately predict COVID-19 cases, whereas True Negative indicates that the classifier can detect normal cases. False Positive indicates that the classifier is incorrectly predicting COVID-19 cases. False Negative indicates that COVID-19 was not detected in medical images. The following definitions apply to precision, sensitivity (recall), specificity, F-1 Score, and balanced accuracy.

$$precision = \frac{TP}{TP + FP} \tag{1}$$

$$sensitivity(recall) = \frac{TP}{TP + FN} \tag{2}$$

$$specificity = \frac{TN}{TN + FP} \tag{3}$$

$$F1 - Score = \frac{2 * precision * recall}{precision + recall} \tag{4}$$

$$Balanced - Accuracy = \frac{specificity + sensitivity}{2} \tag{5}$$

V. RESULTS AND DISCUSSION

TABLE II presents the results of 27 different deep neural network architectures. We used 5-fold cross-validation to evaluate and train transfer learning-based methods and reported the average training results on our medical image dataset.

TABLE II. RESULTS OF TRANSFER LEARNING ALGORITHMS

#	Algorithms	Accuracy	Precision	Sensitivity	Specificity	F1-Score	Balanced-Accuracy
1	InceptionV3	0.97826	0.97826	0.97826	0.97826	0.97826	0.97826
2	ResNet152 V2	0.98913	0.97826	1	0.97870	0.9890	0.98940
3	DenseNet121	0.98913	0.97826	1	0.97870	0.9890	0.98940
4	DenseNet169	0.97826	0.97826	1	0.97826	0.97826	0.97826
5	DenseNet201	0.97826	0.97826	1	0.97826	0.97826	0.97826
6	EfficeintNetb0	1	1	1	1	1	1
7	EfficientNetb1	1	1	1	1	1	1
8	EfficeintNetb2	1	1	1	1	1	1
9	EfficeintNetb3	0.98913	1	0.97870	1	0.98920	0.98940
10	EfficeintNetb4	0.98913	1	0.97870	1	0.98920	0.98940
11	EfficeintNetb5	1	1	1	1	1	1
12	EfficeintNetb6	1	1	1	1	1	1
13	EfficeintNetb7	0.98913	0.98826	1	0.97870	0.98900	0.98940
14	MobileNet	1	1	1	1	1	1
15	ResNet50	0.96739	1	1	1	1	1
16	ResNet101 V2	0.97826	0.97826	0.97826	0.97826	0.97826	0.97826
17	ResNet152 V2	0.98913	0.97826	1	0.97870	0.98900	0.98940
18	VGG 19	0.98913	0.97826	1	0.97870	0.98900	0.98940
19	Xception	0.97826	0.97826	0.97826	0.97826	0.97826	0.978260
20	MobileNet V2	0.98913	1	0.97870	1	0.98920	0.98940
21	ResNet101	0.95652	0.91300	1	0.92000	0.95450	0.96000
22	ResNet152	0.98913	0.97826	1	0.97870	0.98900	0.98940
23	Inception-ResNet V2	1	1	1	1	1	1
24	NasNet-Large	1	1	1	1	1	1

25	NasNetMobile	0.97826	0.97826	0.97826	0.97826	0.97826	0.97826
26	ResNet50V2	0.97826	0.97826	0.97826	0.97826	0.97826	0.97826
27	VGG16	0.98913	0.97826	1	0.97870	0.98900	0.98940

According to the results in TABLE II, the best models for classifying COVID-19 from standard samples are EfficientNetb0, EfficientNetb1, and EfficientNetb2. Additionally, these models can be used in mobile devices and low-computing-power applications. Intriguingly, the MobileNet architecture outperformed the next-generation MobileNetV2. MobileNet is another viable candidate to overcome the problem at hand. Two cumbersome architectures, Inception-ResNetV2 and NasNet-Large, provide exceptional performance. Both of these architectures are dependable and robust. The disadvantages of both Inception-ResNetV2 and NasNet-Large are that they incur high computational costs and consume significant memory space on devices. Due to their small memory utilization, low computational requirements, and acceptable performance, the EfficientNet family and the first version of MobileNet architecture can be used for COVID-19 recognition tasks by medical personnel. We contemplate using convolutional architectures with a few numbers of modifications can create improved customized architectures needed for COVID-19 identification in medical X-ray images. These modifications can include using self-attention modules, concatenation the feature vectors of different architectures, combination of vision transformer family with convolutional architectures, knowledge distillation approach and so forth.

VI. Conclusion

This study began by creating a medical dataset using two publicly accessible repositories. Afterward, we performed the preprocessing procedure, corrected dataset imbalances, and normalized medical chest x-ray images. We then used data augmentation to improve the efficiency of our classifiers. Finally, we evaluated 27 transfer learning-based algorithms using fivefold cross-validation and reported the average results. The results demonstrate that the EfficientNet and MobileNet families of algorithms are viable candidates for our medical computer vision task. Along with low-computational architectures that are well-suited for deployment on mobile devices, two cumbersome models, Inception-ResNetV2 and NasNet-Large, demonstrated excellent results for the given problem. Utilizing Transformer-based architectures and knowledge distillation techniques and developing robust classifiers against adversarial attacks can be considered future research and study areas in the field of medical image analysis.

References

[1] A. Gupta, Anjum, S. Gupta, and R. Katarya, "InstaCovNet-19: A deep learning classification model for the detection of COVID-19 patients using Chest X-ray," Appl. Soft Comput., vol. 99, Feb. 2021, doi: 10.1016/J.ASOC.2020.106859.

[2] C. Ouchicha, O. Ammor, and M. Meknassi, "CVDNet: A novel deep learning architecture for detection of coronavirus (Covid-19) from chest x-ray images," Chaos, Solitons and Fractals, vol. 140, Nov. 2020, doi: 10.1016/J.CHAOS.2020.110245.

[3] P. Saha, M. S. Sadi, and M. M. Islam, "EMCNet: Automated COVID-19 diagnosis from X-ray images using convolutional neural network and ensemble of machine learning classifiers," Informatics Med. Unlocked, vol. 22, Jan. 2021, doi: 10.1016/J.IMU.2020.100505.

[4] M. Z. Islam, M. M. Islam, and A. Asraf, "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," Informatics Med. Unlocked, vol. 20, Jan. 2020, doi: 10.1016/J.IMU.2020.100412.

[5] S. Sheykhivand et al., "Developing an efficient deep neural network for automatic detection of COVID-19 using chest X-ray images," Alexandria Eng. J., vol. 60, no. 3, pp. 2885–2903, Jun. 2021, doi: 10.1016/J.AEJ.2021.01.011.

[6] Y. Kutlu and Y. Camgözlü, "Detection of coronavirus disease (COVID-19) from X-ray images using deep convolutional neural networks," Nat. Eng. Sci., vol. 6, no. 1, pp. 60–74, Jan. 2021, doi: 10.28978/NESCIENCES.868087.

[7] R. C. Joshi et al., "A deep learning-based COVID-19 automatic diagnostic framework using chest X-ray images," Biocybern. Biomed. Eng., vol. 41, no. 1, pp. 239–254, Jan. 2021, doi: 10.1016/J.BBE.2021.01.002.

[8] M. M. Rahaman et al., "Identification of COVID-19 samples from chest X-Ray images using deep learning: A comparison of transfer learning approaches," J. Xray. Sci. Technol., vol. 28, no. 5, pp. 821–839, 2020, doi: 10.3233/XST-200715.

[9] M. Loey, F. Smarandache, and N. E. M. Khalifa, "Within the lack of chest COVID-19 X-ray dataset: A novel detection model based on GAN and deep transfer learning," Symmetry (Basel)., vol. 12, no. 4, Apr. 2020, doi: 10.3390/SYM12040651.

[10] I. Lorencin et al., "Automatic evaluation of the lung condition of COVID-19 patients using X-ray images and convolutional neural networks," J. Pers. Med., vol. 11, no. 1, pp. 1–31, Jan. 2021, doi: 10.3390/JPM11010028.

[11] [R. Jain, M. Gupta, S. Taneja, and D. J. Hemanth, "Deep learning-based detection and analysis of COVID-19 on chest X-ray images," Appl. Intell., vol. 51, no. 3, pp. 1690–1700, Mar. 2021, doi: 10.1007/S10489-020-01902-1.

[12] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. Jamalipour Soufi, "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning," Med. Image Anal., vol. 65, Oct. 2020, doi: 10.1016/J.MEDIA.2020.101794.

[13] S. C. R and R. K. Dubey, "Deep Learning-Based Hybrid Models for Prediction of COVID-19 using Chest X-Ray," Aug. 2020, doi: 10.36227/TECHRXIV.12839204.V1.

[14] J. D. Arias-Londono, J. A. Gomez-Garcia, L. Moro-Velazquez, and J. I. Godino-Llorente, "Artificial Intelligence applied to chest X-Ray images for the automatic detection of COVID-19. A thoughtful evaluation approach," IEEE Access, 2020, doi: 10.1109/ACCESS.2020.3044858.

[15] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," Sci. Rep., vol. 10, no. 1, Dec. 2020, doi: 10.1038/S41598-020-76550-Z.

[16] A. I. Khan, J. L. Shah, and M. M. Bhat, "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images," Comput. Methods Programs Biomed., vol. 196, Nov. 2020, doi: 10.1016/J.CMPB.2020.105581.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Dec. 2016, vol. 2016-December, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[18] A. G. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," Apr. 2017, Accessed: Dec. 15, 2021. [Online]. Available: https://arxiv.org/abs/1704.04861v1.

[19] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., Sep. 2014, Accessed: Jul. 26, 2021. [Online]. Available: https://arxiv.org/abs/1409.1556v6.

[20] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Nov. 2017, vol. 2017-January, pp. 1800–1807, doi: 10.1109/CVPR.2017.195.

[21] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," 36th Int. Conf. Mach. Learn. ICML 2019, vol. 2019-June, pp. 10691–10700, May 2019, Accessed: Jun. 17, 2021. [Online]. Available: http://arxiv.org/abs/1905.11946.

[22] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, vol. 2017-January, pp. 2261–2269, Aug. 2016, Accessed: Jul. 26, 2021. [Online]. Available: https://arxiv.org/abs/1608.06993v5.