



Prediction of Covid-19 using Self-Supervised Learning and CNN

Prashant Sharma^{1*}, Sonal Sharma²

¹Mandsaur University

²Jain University

Corresponding Author: Prashant Sharma prashant.sharma@meu.ed.in

ARTICLE INFO

ABSTRACT

Keywords: COVID-19, VGG-16, Corona Disease (COVID19) is an infectious disease that CT- Scan, RT-PCR, Convolutional Neural Network (CNN)

Received : 3 September

Revised : 19 September

Accepted: 20 October

©2023 Sharma, Sharma: This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/).



is rapidly producing respiratory contaminations and demanding medical attention on a global scale. It has had a profound effect on people lives, society, and the global economy. Due to limitations of RT-PCR (switch record polymerase chain reaction)-based methods for COVID19 diagnosis, various healthcare-based imaging approaches have lately been used. In this paper, we took a freely available image dataset with an extensive number of CT scans which were positive for COVID-19. Then utilizing a least amount of training X-Ray images, we build sample-efficient deep learning approaches that really an extensive testing demonstrated that the suggested Self-Trans method to achieves a lot of cutting-edge benchmarks. The proposed approach detects COVID-19 instances with a reliability rate of 98.7% and COVID-19 infection could be predicted with high precision by CT imaging

INTRODUCTION

In March 2020, the WHO The COVID-19 epidemic, which is affected by SARS-CoV-2, is familiar around the world. It is extremely difficult to fight the coronavirus, which has the potential to cause fatal respiratory diseases with severe discomfort (ARDS). Finding a fundamental component to stop COVID-19 from spreading early on. The most well-known concealing technique for is the polymerase chain reaction (RT-PCR). Nevertheless, it is a challenging procedure, and preliminary studies revealed that it had a poor affectability. Lung lesions related to COVID- 19 have specific morphological characteristics that have been identified using (CT) scans. Computational intelligence techniques have been investigated as a means to automate and assist with detection, albeit the accuracy of COVID-19 treatment by chest scans is mainly dependent on specialists. Prediction is a method of analyzing the future that is based on present, verifiable evidence and routinely planned investigation. Making the framework function without being specifically tailored is the goal of AI research. It is a sub- field of artificial intelligence (AI) that allows generation of acceptable solutions using current data and computations. Machine learning functions somewhat in relation to human learning. It is used in many areas of daily life, including using Netflix, Facebook, e-mail, Amazon, and so on. Because the Scan data utilized in these investigations is not available to the general public, these findings are difficult to duplicate and implement. Also, the correct diagnostic models needed for these tasks must be trained on a significant range of CTs, which is a demanding task. In this work, we examine such two difficulties. The survey algorithms processed for deep learning that is competent of producing high COVID-19 prediction accuracy from CT scans despite the fact that the amount of trained CT images is constrained. To provide a freely searchable dataset with thousands of CT images that are COVID-19 positive. To reduce the possibility of overfitting, we suggest the Self-Trans method, which coordinate descriptive self-supervised learning and transfer learning.

LITERATURE REVIEW

A Gathered datasets are from the same sources, they established the concept as the starting point Lee, K.S. et al [1] Oxford University's visual geometric group (VGG) team developed deep CNN-based models, VGG-19 and VGG-16 Simonyan et. al [2] VGG-19 model and VGG-16 model have 19 and 16 deep convolutional layers, correspondingly, and use a design with very small (3x3x3) convolution operation. While steadily expanding the extent of convolution layer, image categorization rate achieved a pinnacle in ImageNet Challenge 2014 compliance. They developed 12 developmental models for VGG-16 and 19 models using distinct level of improvements. Then conduct a test to compare the efficacy of each model for the COVID-19 categorization. The ROC curve's AUC value was highest while fine-tuning utilizing VGG-16's Conv blocks 4 and 5. Apart from the constructional work developed by Lee K.S. et al. [1], they evaluate the effectiveness combining two recently published models of our proposed schemes developed by Rahimzadeh .M et al. [3] and Das S et al. [4].[3] showed a two aspects of machine learning technique for COVID-19 categorizing using CT scan data. The suggested technique collects CXR components using the

VGG-19 model processed on the datasets of ImageNet, and utilizes regression analysis to categorize CXR images as normal or abnormal. Finally, using XGBoost [5], CXR images defined as abnormalities during the initial stage are reclassified as pneumonia or COVID-19. Rahimzadeh .M et al. [4] proposed various training approaches to assist the network to learn better when the dataset is uneven. The researchers also presented a neural network which connects the Xception [6] and ResNet50V2 [7] networks. For comparison, overall specificity, accuracy, AUC, F1-score, and accuracy of the VGG-16 model with making improvements by Lee .K .S et al. [1], two-stage predictor by Das .S et al. [3], and ResNet50 and Xception model concerted model by Rahimzadeh .M et al. [4] are defined. Moreover, compared the suggested technique to other well-known representations used in other COVID-19 studies [11,12,13,14], such as ResNet-101, ResNet-50 [8], MobileNet [9], and MobileNetV2 [10]. A CNN is one of the most efficient methods for accurately determining the COVID-19 and other respiratory disease can be detected using lung radiography data. Many researches have been managed to identify key contributions to COVID-19 identification [15]-[18]. a revised variant of the ResNet-50 trained CNN model, for example, was utilized in [19] to classify CT scans into 3 groups: bacterial pneumonia, [20] employed COVID-Net, a machine-learning based CNN network, to classify chest X-ray data into 4 groups: Bacterial Infection, Non-COVID, Normal, and COVID-19 positive Viral disease. [21] describes a minimally supervised system for COVID-19 identification and disease placement based on surface masks created by an unsupervised lung classification procedure and a 3D Rnn pre-trained model utilizing 3-D chest CT data. [22] analyzed the performance of cutting-edge neural network models based on learning algorithms using a collection of chest radiography images from individuals with pneumonia, COVID-19 positive cases, and regular occurrences. According to findings, deep learning can give important indicators for identifying COVID-19 patients. Classifiers has been shown to be a dependable strategy for dealing with a small range of training data from certified COVID-19 cases [23]. [24] employs self-supervised learning with contextual distortion to address identification, separation, and clustering problems in radiography. When categorization, the technique was used for scanning plane identification in new-born 2-D ultrasound images, and it demonstrated identification benefits in numerous instances. Yet, we conclude that our proposed method is superior since contextual distortion provides localized features rather than global visual features, which are more beneficial for text categorization.

METHODOLOGY

Firstly, we acquire and create a dataset of chest radiography (CXR) data in this work from numerous publicly available medical sources. There are two categories of chest X-ray (CXR) scans in this database: Covid and Normal. All visuals have a resolution of 299 x 299 pixels and are uploaded in PNG (Portable Network Graphics) format. The visuals were also grayscale with three channels of repeating RGB values.

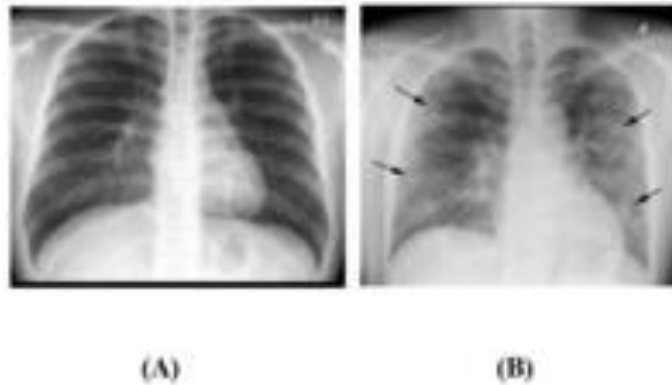


Fig.1 CXR Pictures Showing (A) COVID-19 Negative Case(Normal)
(B) COVID-19 Positive Case.

After that we performed image preprocessing as we know that on a small dataset, DL models typically overfit. As a result, a continuous number of data source is required for rapid learning and better categorization generalization. Data augmentation is a just feasible method of improving model generalization by integrating different variations in the source dataset. In this work, the training data were enhanced using various transformations such as both vertical and horizontal reflections, rotation, and compression. Before allocating the images to CNN for training, they were all adjusted to 224 x 224 x 3.

A. Proposed Block Diagram:

Figure:2 shows a proposed approach architecture. At the component extraction and commotion evacuation stages, the suggested technique provides an effective route. At the testing step, the proposed approach made use of chest-CT filter images and used for pre-processing the data images included more robust element identification, which resulted in a favorable outcome in terms of obtaining an Outstanding characterization. The study shows that the suggested element combining approach would successfully and consistently analyze COVID-19 and support physicians in doing everything it takes to preserve the lives of those affected by COVID-19.

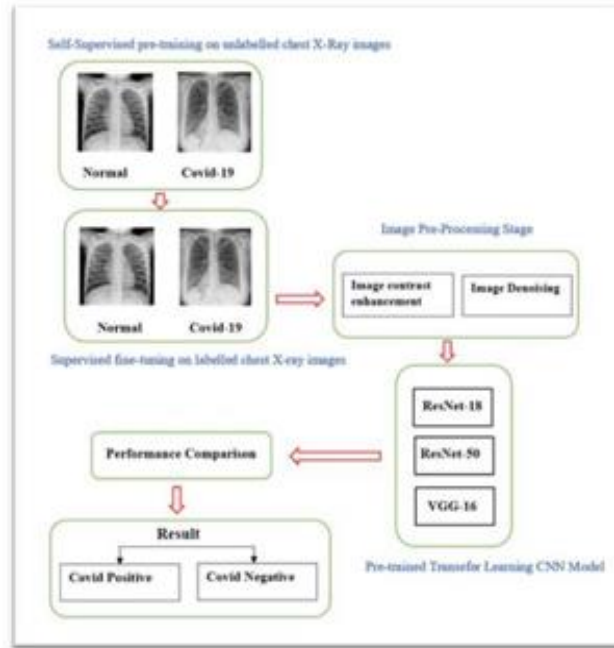


Fig.2 Proposed Architecture

I. Self-Supervised Learning:

Using unlabeled data, self-supervised learning is implemented to build image representation. Self-supervised learning is a cutting-edge approach that utilizes unlabeled data with a predicate goal to collect lexical characteristics for additional visual works such as categorization and to minimize the overfitting and enhance attainment if the information is limited. The image characteristics learned by When used in a transferable manner in other vision tasks, a self-supervised learning technique may be a useful substitute.

II. Self-Supervised Learning with Transfer Learning:

Since while deep learning is highly valued in radiography, there are now two primary concerns with transfer learning in radiography. Pre-trained ImageNet networks, for example, are frequently over specification in subsequent purpose. Second, the original job's database image greatly differs from the target task's dataset image. The image in the ImageNet dataset is an example of natural views of birds, flowers, insects, fish, and so on, but the images in the test dataset are chest CT scans of COVID-19 cases. Furthermore, feature extraction collection includes a significant number of images. The ImageNet collection has 1000 picture categories, although there are fewer images in the medical imaging sector.

III. Pre-trained Convolution Neural Network:

To determine COVID-19 cases from chest radiography data, the following CNN-based pre-trained classifiers were used: Resnet-18 (a), Resnet-50 (b), VGG-16 (c). Researchers from Oxford University proposed the VGG16 architecture, which includes 3 x 3 convolutional layers to increase. To reduce volume size, set the depth and max - pooling layer to zero. The ResNet50 model is based on the deep convolutional neural concept, which overcomes the gradient loss issue in development and improving that as intensity increases, accuracy becomes saturated and reduces rapidly.

IV. Pooling Layer:

In the approach, we used values from their convolutional layers to implement the above-mentioned pretrained models as feature categorization to show inductive learning in identifying COVID-19 infections from lung CT data. After extracting the detecting layers (i.e., classifiers) from the pre-trained Convolution networks and replaced them with suggested classifier, which comprised of a global average pooling layer (GAP). This layer is regarded as a preferable alternative for levelling immediately before the last unit of convolution in order to minimize overfitting by reducing the number of parameter estimates. The GAP layer minimizes the sensory features of a 3-dimensional vector of size by simply combining all hw pixel values from each hw dataset $h \times w \times d$ to $1 \times 1 \times d$.

RESULTS AND DISCUSSION

The data collection is compiled in order to generate a statistical overview of the dataset. The relevant libraries are then uploaded. Disease detection systems are very important in today's technological world since they take less time and are more user intuitive. It is a frequent approach used by clinicians to identify COVID-19 soon on. Researchers may also utilize the concept to extend their future study, particularly for lung illness diagnostics. The classification performance was examined and contrasted, providing positive results and a rapid implementation.

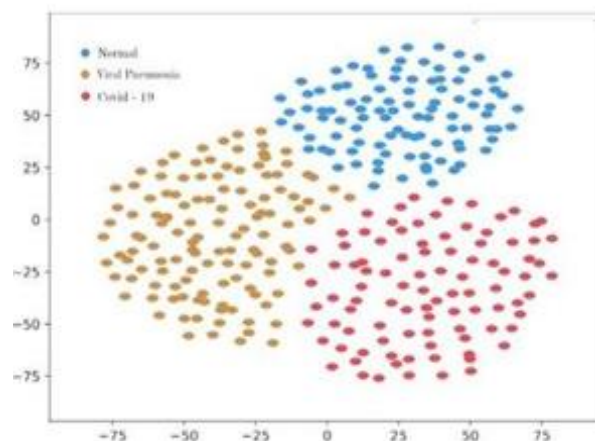


Fig.3 Demonstrates the distribution of CT image characteristics following dimensionality reduction. Each node corresponds to a specific CT scan, the color reflects categorization data, and the text explains the importance of the color.

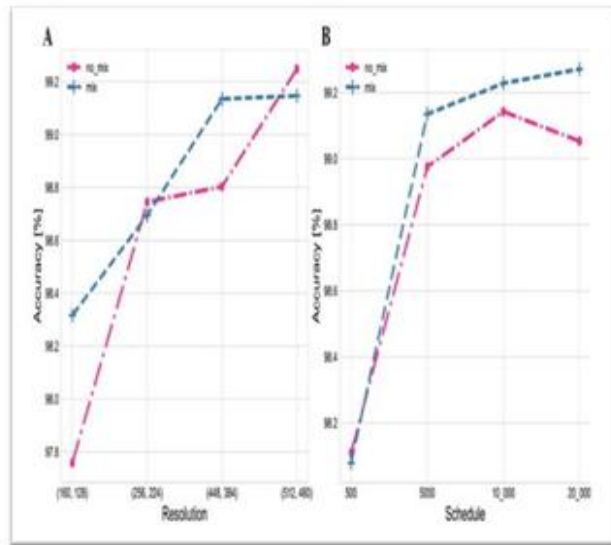


Fig.4 COVID CT-2 precision test with multiple hyperparameters.
(A)Decision. (B) Schedule

CONCLUSIONS AND RECOMMENDATIONS

The prediction of COVID-19 viral disease using chest radiography data is necessary for both healthcare professionals and patients save the testing time and expenses. We used a systematic technique similar to what physicians would use to discover covid-19 in patients. It has the potential to be more precise and faster than the existing RT- PCR method. The proposed approach enables clinicians to quickly and efficiently access information on covid-19 infected patients. Our approach used CT data for COVID-19 testing and examined. The impact of different starting factors on the outcomes demonstrated that the model suggested here, which was pre-trained on ImageNet21k, provided significant generalizability in the CT image context. The proposed model finds COVID-19 occurrences with a 98.7% accuracy. Further explanatory assessments of the designs will be performed in order to decoding the detection phase of COVID-19, identify relevant features in chest radiography data and enhance clinical doctor evaluation. As a result, we can enhance the models in future research.

FURTHER STUDY

This research still has limitations so further research needs to be carried out regarding the topic of Covid-19 Prediction using Self-Supervised Learning and CNN to perfect this research and increase insight for researchers.

REFERENCES

- Lee, K.S.; Kim, J.Y.; Jeon, E.T.; Choi, W.S.; Kim, N.H.; Lee, K.Y. "Evaluation of Scalability and Degree of Fine-Tuning of Deep Convolutional Neural Networks for COVID-19 Screening on Chest X-ray Images Using Explainable Deep-Learning Algorithm." *J. Pers. Med.* 2020.
- Simonyan, K.; Zisserman, A. "Very deep convolutional networks for large-scale image recognition". *arXiv* 2014.
- Das, S.; Roy, S.D.; Malakar, S.; Velásquez, J.D.; Sarkar, R. "Bi-Level Prediction Model for Screening COVID-19 Patients Using Chest X-Ray Images". *Big Data Res.* 2021.
- Rahimzadeh, M.; Attar, A. "A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2." *Inform. Med. Unlocked* 2020.
- Chen, T.; Guestrin, C. "Xgboost: A scalable tree boosting system." In *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13-17 August 2016.*
- Chollet, F. "Xception: Deep learning with depthwise separable convolutions." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21-26 July 2017.*
- He, K.; Zhang, X.; Ren, S.; Sun, J. "Identity mappings in deep residual networks." In *Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8- 16 October 2016.*
- He, K.; Zhang, X.; Ren, S.; Sun, J. "Deep residual learning for image recognition." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27-30 June 2016.*
- Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." *arXiv* 2017.
- 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.*

- Rahaman, M.M.; Li, C.; Yao, Y.; Kulwa, F.; Rahman, M.A.; Wang, Q.; Qi, S.; Kong, F.; Zhu, X.; Zhao, X. "Identification of COVID-19 samples from chest X-ray images using deep learning: A comparison of transfer learning approaches." *J. Ray Sci. Technol.* 2020.
- Rehman, A.; Naz, S.; "Improving coronavirus (COVID-19) diagnosis using deep transfer learning." *medRxiv* 2020.
- Wong, H.Y.F.; Lam, H.Y.S.; Fong, A.H.T.; Leung, S.T.; Chin, T.W.Y.; Lo, C.S.Y.; Lui, M.M.S.; Lee, J.C.Y.; Chiu, K.W.H.; Chung, T.W.H.; et al. "Frequency and distribution of chest radiographic findings in patients positive for COVID-19". *Radiology* 2020.
- L. Srinivasan, Venkatesh, G. Pranay Deepak Reddy, B. Sai Pramath and J. A. Trivedh, "IoT Based Low-Cost Robotic Agent Design for Covid-19 affected people," 2022 International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2022.
- F. Shi et al., "Review of artificial intelligence techniques in imaging data acquisition segmentation and diagnosis for COVID-19", *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 4-15, 2021.
- D. Dong et al., "The role of imaging in the detection and management of COVID-19: A review", *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 16- 29, 2021.
- S. A. Harmon et al., "Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets", *Nature Commun.*, vol. 11, no. 1, pp. 1-7, Dec. 2020.
- Palanikkumar, D, Priya, S. "Brain Storm Optimization Graph Theory (BSOGT) and Energy Resource Aware Virtual Network Mapping (ERVNM) for Medical Image System in Cloud." *J Med Syst* 43, 37 (2019).
- Y. Song et al., "Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images", *IEEE/ACM Trans. Comput Biol. Bioinf.*
- 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, pp. 1-12, Dec. 2020.
- X. Wang et al., "A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT", *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2615-2625, Aug. 2020.
- I. D. Apostolopoulos and T. A. Mpesiana, "COVID-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks", *Phys. Eng. Sci. Med.*, vol. 43, no. 2, pp. 635- 640, Jun. 2020.

Y. Oh, S. Park and J. C. Ye, "Deep learning COVID- 19 features on CXR using limited training data sets", *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2688-2700, Aug. 2020.

Chen, P. Bentley, K. Mori, K. Misawa, M. Fujiwara and D. Rueckert, "Self-supervised learning for medical image analysis using image context restoratio", *Med. Image Anal.*, vol. 58, Dec. 2020