ISBN : 978-93-91535-11-7

CLASSIFICATION OF COVID-19 FOR CHEST X-RAY IMAGES USING DEEP LEARNING

S S Ittannavar[1]
Assistant Professor
Dept. of ECE Engineering
Hirasugar Institute of Technology, Nidasoshi, India
Email id: ishreevijay@gmail.com

B. P. Khot[2]
Assistant Professor
Dept. of ECE Engineering
Hirasugar Institute of Technology, Nidasoshi, India
Email id: pratimak1128@gmail.com

R. H. Havaldar[3]
Professor & Head
Dept. of Biomedical Engineering
K L E's Dr. MSS CET, Belagavi
Belagavi, India,
Email id: raviraj61@gmail.com

Abstract

In this paper, a dataset of X-ray images from patients with common bacterial pneumonia, confirmed Covid-19 disease, and normal incidents, was utilized for the automatic detection of the Corona virus disease. The aim of the study is to evaluate the performance of state-of-the-art Convolutional neural network architectures proposed over the recent years for medical image classification. Specifically, the procedure called Transfer Learning was adopted. With transfer learning, the detection of various abnormalities in small medical image datasets is an achievable target, often yielding remarkable results. The datasets utilized in this experiment are two. Firstly, a collection of CHEST X-ray images including 30 images with confirmed Covid-19 disease and 30 images of normal conditions. Secondly, a dataset including 30 images with confirmed Covid-19 disease and 30 images of normal conditions. The data was collected from the available X-ray images on public medical repositories. The results suggest that Deep Learning with X-ray imaging may extract significant biomarkers related to the Covid-19 disease, while the best accuracy, sensitivity, and specificity obtained is 99.78%, 98.66%, and 96.46% respectively. Since by now, all diagnostic tests show failure rates such as to raise concerns, the probability of incorporating X-rays into the diagnosis of the disease could be assessed by the medical community, based on the findings, while more research to evaluate the X-ray approach from different aspects may be conducted.

***Keywords : Covid-19 - corona virus disease, CNN-Convolutional neural network, CLAHE-Contrast Limited Adaptive Histogram Equalization. SVM-Support Vector Machine, KN-K*-Nearest Neighbor.**

I INTRODUCTION

COVID-19 is an acute resolved disease, but it can also be deadly, with a 2% case fatality rate. Severe disease onset might result in death due to massive alveolar damage and progressive respiratory failure. The early and automatic diagnosis of Covid-19 may be beneficial for countries for timely referral of the patient to quarantine, rapid intubation of serious cases in specialized hospitals, and monitoring of the spread of the disease. Although the diagnosis has become a relatively fast process, the financial issues arising from the cost of diagnostic tests concern both states and patients, especially in countries. In March 2020, there has been an increase in publicly available X-rays from healthy cases, but also from patients suffering from Covid-19. This enables us to study

ISBN : 978-93-91535-11-7

the medical images and identify possible patterns that may lead to the automatic diagnosis of the disease. The development of deep learning applications over the last 5 years seems to have come at the right time. Deep Learning is a combination of machine learning methods mainly focused on the automatic feature extraction and classification from images, while its applications are broadly met is object detection tasks, or in medical image classification tasks. Machine learning and deep learning have become established disciplines in applying artificial intelligence to mine, analyses, and recognize patterns from data. Reclaiming the advances of those fields to the benefit of clinical decision making and computer-aided systems is increasingly becoming nontrivial, as new data emerge.

II LITERATURE SURVEY

The Bayes-SqueezeNet [1] has introduced for detecting the COVID-19 using chest X-ray images. In this the system consists of the augmentation of the raw dataset and model training using the Bayesian optimization concept. The Bayes-Squeeze Net was applied for classifying X-ray images labeled in 3 classes as normal, viral pneumonia, and COVID-19. Using the data augmentation, the net claimed to overcome the problem of imbalanced data obtained from the public databases. As another CNN, the Coro Net [2] was developed for detecting COVID-19 infection from chest X-ray images. This model was based on the pretrained CNN known as the Xception [3]. Coro Net adopted the Xception as base model with a dropout layer and two fully-connected layers added at the end. As a result, Coro Net has 33,969,964 parameters in total out of which 33,969,964 trainable and 54,528 are non-trainable parameters. The net was applied for 3-class classification (COVID-19, pneumonia, and normal) as well as 4-class classification (COVID-19, pneumonia bacterial, pneumonia viral and normal). The Covid GAN [4] was proposed as an auxiliary classifier generative adversarial network based on GAN (generative adversarial network) [5] for the detection of COVID-19. The architecture of the Covid Gan was built on the pretrained VGG-16 [6], which is connected with four custom layers at the end with a global average pooling layer followed by a 64 unit's dense layer and a dropout layer with 0.5 probabilities. The net further utilized the GAN approach for generating synthetic chest X-ray images to improve the classification performance. The Dark Covid Net [7], which was built on the Dark Net model, is another CNN model proposed for COVID-19 detection using chest X-rays. The Dark Covid Net consists of fewer layers and (gradually increased) filters than the original Dark Net. This model was tested for a 2-class classification (COVID-19 and no-findings) and 3-class classification (COVID-19 no-findings, and pneumonia). The work implemented with different ResNet architectures[8], MobileNet-v2 [9], Inception [10], Xception [11], and Inception ResNet-v2 as pretrained CNNs for the detection of COVID-19 from X-ray images. These pretrained CNNs were applied to 2-class and 3-class classification cases using 2 datasets consisting of images of COVID-19, bacterial pneumonia, viral pneumonia, and healthy conditions.

III METHODOLOGY

3.1 Image Dataset:

The COVID-CT dataset contains 360 positive COVID19 cases and 397 negative Chest Computed Tomography images. The positive images were collected from med Rxiv and bio Rxiv. These CT images are in

ISBN : 978-93-91535-11-7

different sizes corresponding to height((maximum=1853, average=491, and minimum=153) and width (maximum=1485,average=383andminimum=124).

As the number of collected images is very small, we further applied Data Augmentation steps such as Rotating all images to 45 degrees, zooming images to 30%, and height shifting with a factor of 0.2 were applied to the dataset to create diversity and will improve the quality of prediction.

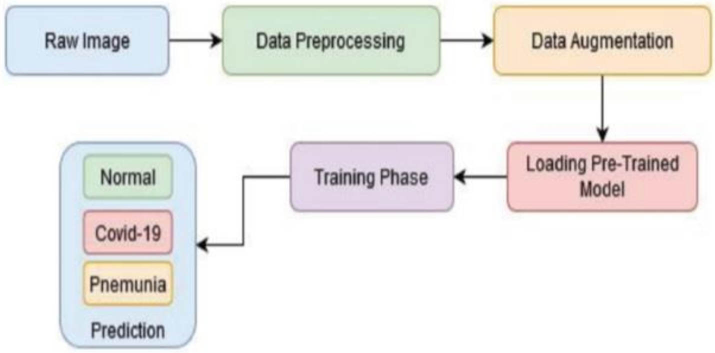

Figure 3.1: Proposed System of Classification of Covid 19 of chest X ray Images.

3.2 Data Preprocessing:

To minimize the effect of sampling bias, histogram equalization is to be applied to images using the CLAHE method. The method both normalizes images and enhances small details, textures and local contrast by first globally normalizing the image histogram followed by application of Contrast Limited Adaptive Histogram Equalization(CLAHE).

3.3 Data Augmentation:

The histogram equalization is applied to enhance the quality of the image without losing the important features of the image. The Weiner filter is used to remove the noises from the image yet preserving fine details and edges of the lungs. The filter size is chosen to be in order to prevent the image from getting over smooth. Weiner filter is typically based on estimation of variance and mean from the local adjacent of individual pixels. As the number of collected images is very small, further we will apply to data augmentation steps such as rotating all images to 45 degrees, zooming images to 30%, and height shifting with a factor of 0.2 were applied to the dataset to create diversity and will improve the quality of prediction. After this, all images were resized to 224 $\times$ 224. This dataset further divided into test, train and validation set with test and validation. The dataset is released publicly with open access for researchers to experiment with methodology and further development at https://github.com/UCSD-AI4H/COVID-CTIV.

3.4 Load Pre trained Model:

CNN models are purposely used for image classification. An image is viewed as an array of pixel which also depends upon the resolution of an image. These CNN models consist series of Convolutional and pooling layers.

ISBN : 978-93-91535-11-7

The data augmentation is achieved using a Convolutional layer. The convolution operation is applied to a region of an image, sampling the values of the pixels in that particular region and converting them into a solitary value. After the data augmentation the images are given to pre-trained CNN such as resnet-50. These CNN models consist series of Convolutional and pooling layer

3.5 Feature Extraction:

After augmentation the regions from the chest images, features are extracted by using LDTP, HOG, and Haralick texture features. The HOG is a global descriptor, which characterize the distribution of intensity gradients or edge directions in the chest images. Additionally, Haralick features (entropy, correlation, angular second moment, contrast, difference entropy, and inverse difference moment) are applied.

3.6 Classification:

Classification refers to a predictive modeling problem where a class label is predicted for an input image. The classification is performed using traditional machine learning classifiers by removing the fully connected layers from the pertained deep learning models. The extracted features were utilized for the final classification using Support Vector Machine (SVM), Decision Tree, Naive Bayes, -Nearest Neighbor (KNN), and Random Forest. In SVM, the input values are plotted in an -dimensional space, and the optimal hyper plane that differentiates the classes is found. In Random Forest, a large number of decision trees are built to operate as an ensemble model where all decision trees predict the class label and eventually the class that gets more votes will be chosen as the predicted label. In Decision Tree, each node acts as a splitting criterion and the branches lead to the final node (leaf node) to provide the output. Naive Bayes is a conditional probability model which used the Bayes theorem for classification. KNN is a non parametric classifier which classifies images based on its-nearest neighbors. The block diagram of proposed system is shown in figure 3.1.

IV RESULTS AND DISCUSSIONS

As the current extracted dataset size is too small, so it seems illogical to train the Residual Network from scratch. Hence results reported are only for ResNet-34 and ResNet-50 trained using transfer learning.. Table 4.1 represents the accuracy and error rate of both the models developed. ResNet - 50 architecture performed better,with an accuracy of 100%. Adding more layers and increasing the number of parameters in Residual Network architecture definitely helps in improving the accuracy of the overall classification.

CNN	Number of Epoches	Accuracy (%)	Error Rate (%)
ResNet-50	10	80	20
	15	83.3	16.7
	20	100	0

Table 4.1 Summarization of Implemented Architecture along with the Accuracy Score and Error Rate

ISBN : 978-93-91535-11-7

V CONCLUSION

This paper has presented an augmented CNN to detect COVID-19 on chest X-Ray images and classify from non-COVID-19 cases. The proposed model need not require extracting the feature manually; it is automated with end-to-end structure. Most of these previous studies have fewer examples for training the deep models. In contrast, the proposed model has used a multi-image augmentation technique driven by first and second order derivative edge operators and this augmentation generates a number of representative edged images. While, CNN is trained with these augmented images, the classification accuracies on an average we will get 87.76% for X-Ray images.

REFERENCES

[1] Yi, P. H., Kim, T. K., & Lin, C. T. (2020). Generalizability of Deep Learning Tuberculosis Classifier to COVID-19 Chest Radiographs: New Tricks for an Old Algorithm? *Journal of thoracic imaging*, *35*(4), W102W104. https://doi.org/10.1097/RTI.0000000000000532.

[2] Yang, W., Sirajuddin, A., Zhang, X. *et al.* The role of imaging in 2019 novel coronavirus pneumonia (COVID-19). *Eur Radiol* ***30,*** *4874–4882 (2020). https://doi.org/10.1007/s00330-020-06827-4.*

[3] Kundu S, Elhalawani H, Gichoya JW, Kahn CE. How might AI and chest imaging help unravel COVID-19's mysteries? Radiology: *Artificial Intelligency. 2020; https://doi.org/10.1148/ryai.2020200053*

[4] Kim TK, Yi PH, Hager GD, Lin CT. Refining dataset curation methods for deep learning-based automated tuberculosis screening. *J Thorac Dis 2020;12(9):5078-5085. doi: 10.21037/jtd.2019.08.34* .

[5] Wong HYF, Lam HYS, Fong AH-T, Leung ST, Chin TWY, et al. Frequency and distribution of chest radiographic fndings in COVID-19 positive patients. *Radiology. 2020. https://doi.org/10.1148/radiol.2020201160*.

[6] Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK (2018) Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLoS Med 15(11):e1002683. https://doi.org/10.1371/journal.pmed.1002683* .

[7] Hurt B, Kligerman S, Hsiao A. Deep learning localization of pneumonia. J Thorac Imaging. 2020;35:W87–9.

[8] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei ImageNet Large Scale Visual Recognition Challenge. *IJCV*, 2015.

[9] Ucar F, Korkmaz D. COVIDiagnosis-net: deepbayes-Squeeze Net based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images. *Med Hypotheses. 2020; 140:109761.*

[10] Khan AI, Shah JL, Bhat MM. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. *Comput Methods Programs Biomed. 2020 Nov;196:105581.*

doi: 10.1016/j.cmpb.2020.105581. Epub 2020 Jun 5. PMID: 32534344; PMCID: PMC7274128.

[11] *Waheed* A, Goyal M, Gupta D, Khanna A, Al- Turjman F, Pinheiro PR., CovidGAN: data augmentation using auxiliary classifer GAN for improved Covid-19 detection. *IEEE Access. 2020; 8:91916–23.*