Machine Learning Based Covid19 Identification from Chest CT Images using GLCM Features

[1]DESETTI VAMSI KRISHNA, [2] DR.S.NARAYANA REDDY

[1]M.Tech Student, [2] Professor
Department of ECE,
S.V. University College of Engineering,
Tirupati, Andhra Pradesh, India.

***Abstract*: The pandemic was announced by the World Health Organization corona virus (COVID-19) universal health dilemma. Any scientific appliance which contributes expeditious detection of corona virus with a huge recognition rate may be excessively fruitful to doctors. In this environment, innovative automation like deep learning, machine learning, image processing and medical image like chest radiography (CXR), computed tomography (CT) has been refined promising solution contrary to COVID-19. Currently, a reverse transcription-polymerase chain reaction (RT-PCR) test has been used to detect the corona virus. Due to the moratorium period is high on results tested and huge false negative estimates, substitute solutions are desired. Thus, an automated machine learning-based algorithm is proposed for the detection of COVID-19. The symptoms seen in these cases were not much different than those seen in case of pneumonia. Earlier the research has been carried out in the field of pneumonia identification and classification through CT- images of chest. The difficulty in identifying Covid19 infection at initial stage is due to high resemblance of its symptoms with the infection caused due to pneumonia. Hence it is trivial to well distinguish cases of corona virus from pneumonia that may help in saving life of patients. The paper uses chest CT images to identify Covid19 infection in lungs using machine learning classifiers and ensembles with Gray-Level Co-occurrence Matrix (GLCM) features. The advocated methodology extracts statistical texture features from CT images by computing a GLCM for each image. The matrix is computed by considering various stride combinations. These GLCM features are used to train the machine learning classifiers and ensembles. The dataset used for evaluating performance of the method is open sourced and can be accessed easily. Proposed method being simple and computationally effective achieves noteworthy performance in terms of Accuracy and Sensitivity. The proposed system developed on MATLAB 2013a version.**

Index Terms: GLCM, Machine Learning, CT Images, COVID19 etc.

I. INTRODUCTION

The COVID-19 (Novel Corona virus) appears as a crown for the microscope to cosmic so named as corona virus. The organizational framework of this study divides the research work in the different sections. The chest radiographs i.e. the chest CT- images of patient infected with the Covid19 exhibits characteristic patterns that are found in CT- images of pneumonia infected patient. Imaging departments have also confirmed that radiological findings of Covid19 on chest CT- images are those very much similar to the pneumonia. Copious amount of research has taken place in identification and classification of pneumonia using chest CT-images. Although Corona virus samples are getting tested widely using Reverse Transcription Polymerase Chain Reaction (RT-PCR), it alone might not offer brake on the global spread of Covid19. Thus, chest CT- images can prove to be a considerate tool for triage of Covid19 infected patients. The ultra-rapid transmission of this virus is creating immense pressure on the global medical and health authorities. In such gigantic provocative situation, there is an urgent need to devise a method that can well distinguish Covid19 suspected patients from Normal or Pneumonia infected patients by making use of chest CT images of concerned individuals. Even after diagnosing chest CT images one finds it difficult to distinguish between pneumonia infected and suspected Covid19 patients. Researchers around the world are attempting to devise some texture based cognitive methods for identification of Covid19. Thus, current research work focuses on extracting texture features from chest CT images using Second-order statistical moments-based techniques. The Harlicka Features (GLCM) are extricated from each chest CT- image to gain some distinguishable insights related to texture of chest CT image. The advocated technique aims at early detection of suspected Covid19 cases by categorizing CT- images into one of the three mentioned classes Covid19, Normal and Pneumonia infected. Moreover, binary classification categorizes images into Covid19 infected or Non-Covid19 infected.
The main contributions of the paper and proposed method are

- Ability to identify Covid19 infection from chest CT images considering high similarity with pneumonia.
- Decision about the GLCM stride combinations giving more efficient Covid19 identification model.
- Performance assessment of Machine Learning classifiers for better Covid19 identification.
- Proposing the best possible ensemble of machine learning classifiers for more efficient Covid19 identification.

The Literature Review is presented in section 2. Further, in section 3 shown fundamentals of underwater imaging is discussed, in section 4 proposed methodology is discussed and In section 5, Simulation Results work is shown. Conclusion and future work are presented by last sections 6.

II. Literature Survey

The Novel Corona virus Pandemic resulted in immense loss of health, wealth, and economy. The medical practitioners or authorities are facing lot of challenges in detection and treatment of patients affected with Covid19. The radiology experts are in progressive search of finding effective and early identification methods for Covid19. The research carried out till date in identification of Covid19 and Pneumonia uses various texture features based cognitive methods. Few of such methods are summarized below.

Nanditha Krishna et al. advocates extraction of all 14 Haralick texture features from chest CT images in [1]. Authors have found that 3 texture features namely variance, sum average and sum variance provide discriminative features to classify CT images into two classes - normal lungs or lungs affected with pneumonia. The results are validated on a dataset obtained from Bangalore based medical college and hospital. The dataset consists of total of 22 images (11 normal lungs and 11 pneumonia affected lungs).

Nitin Singh et al. have proposed a state-of-art technique in [2] that uses dataset consisting of chest CT images of normal people and those infected with pneumonia. Authors have used the combination of wavelet transform method and Gray-Level Co-occurrence Matrix method to extract six time and frequency domain.

Features for detecting Pneumonia. The feature matrix created by fetching statistical GLCM texture features from an image, is used to train algorithms like K-nearest neighbours (KNN) & Support Vector Machine (SVM), achieving better accuracy of 92.6% with weighted KNN model.

Tulin Ozturk et al. have developed a model in [3] that performs two class classification (covid19/Nofindings) as well as three class classification (covid19/No findings/Pneumonia). Authors have used deep learning model, darknet-19 - a classifier model. The inputs are standardized using Batch Normalization method. The performance of this fully automated method is assessed by various parameters of Accuracy, Sensitivity, F-Measure and Precision. Results obtained by using proposed model can be improvised by training the model on larger dataset containing a greater number of covid19 CT images.

In [4] Prabira Kumar Sethy et al. makes use of transfer learning approach for extracting deep features using 13 different pretrained CNN models. The ResNet50 model outperforms other 12 CNN models achieving highest accuracy of 95.33%. The method also compares this approach with other traditional methods of LBP+SVM, HOG+SVM and GLCM+SVM by evaluating performance on these combinations. The author finds out that best classification accuracy is obtained by LBP+SVM followed by GLCM+SVM. The chest CT images are collected from GitHub that consists of in all 381 images. Abhishek Sharma et al. have attempted to automate the diagnosis process of pneumonia infection by considering histogram calculation and OTSU thresholding [5].

Here the resized input chest CT image is histogram equalized and then the abdomen area (region of interest) is cropped. From this region of interest, pneumonia clouds are detected by using OTSU thresholding followed by computing ratio of healthy region to entire lung region.

Abolfazl Zargari Khuzani et al. have proposed a method in [6] which distinguishes a Covid19 patient from pneumonia patient using machine learning classifier. The method extracts global image features from entire CT image without lesion segmentation, which reduces the need of huge training data. The dataset is categorized as images belonging to covid19 class, Normal class, and Pneumonia class.

B.Wang et al.,"AI-assisted CT imaging analysis for COVID-19 screening[7] the novel corona virus pandemic resulted in immense loss of health, wealth, and economy. The medical practitioners or authorities are facing lot of challenges in detection and treatment of patients affected with Covid19. proposed an infection Size Aware Random Forest method for classification of CT images. The feature from an image edge detection of CT images

S.Wang et al., A deep learning algorithm using CT images to screen for Corona Virus disease (COVID-19)[8] in this paper, we try to establish a new tailored deep convolutional neural network (CNN) for segmenting the chest CT images with COVID-19 infections.

III. Proposed Method

The proposed approach shown in figure 1 aims to set up an arrangement that detects COVID-19 among the numerous CT images. The individual detection from images will be tedious and time-consuming. Hence an automated machine learning-based process with rigors pre-processing, segmentation, and feature extraction is needed that assist the doctor to examine the patient.

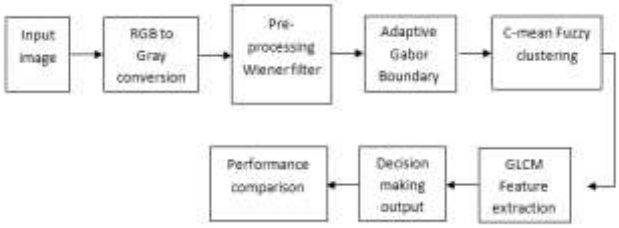

Fig.1 Proposed Block Diagram

Image Acquisition

Image acquisition Our proposed algorithm uses multiple images i.e. Chest Radiography (X-Ray), Computed Tomography (CT scan) images. The dataset downloaded from image processing dataset in Google.

RGB to Gray Converstion

An intuitive way to convert a color image 3D array to a gray scale 2D array is, for each pixel, take the average of the red, green, and blue pixel values to get the gray scale value. This combines the lightness or luminance contributed by each color band into a reasonable gray approximation. Because it is a one layer image from 0-255 whereas the RGB have three different layer image. So that is a reason we prefer grey scale image instead of RGB.

Image Pre-processing

Image pre-processing Images obtained by multiple techniques might consist of distinct buzz that reduces the quality of the image. Consequently, it cannot contribute ample data for image processing. Hence, intensity values are adjusted or color mapping is done for image enhancement. It strengthens the contrast by choosing a dynamic range in the image.

Wiener filtering is utilized to decrease the noise that has corrupted an image and results the same as the first image. The objective is to have least measure of mean square blunder. Wiener filtering explores the earlier information about the noise in an image.

k-means clustering

Fuzzy c-means clustering has can be considered a better algorithm compared to the k-Means algorithm. Unlike the k-Means algorithm where the data points exclusively belong to one cluster, in the case of the fuzzy c-means algorithm, the data point can belong to more than one cluster with a likelihood.

Machine Learning Classifiers with GLCM Feature

The present work comes up with identification of Covid19, putting forward a method that marks the extraction of texture features from jpeg 8-bit input CT image using Gray Level Co-occurrence Matrix (GLCM). The Co-occurrence matrix is created for an entire single input chest CT image. Five different mix of strides i.e. (1), (1, 2), (1, 2, 4, 8), (1, 2, 4, 8, 16, 32) and (1, 2, 4, 8, 16, 32, 64, 128) are considered for computing GLCM. Six statistical features extracted from GLCM are contrast, homogeneity, dissimilarity, correlation, angular second moment and energy. These features are united to constitute a solitary feature vector. Finally, the multiclass classification and two class classification performed using Random Forest classifier. Initially CT images of varied sizes are pre-processed. Size of the Co-occurrence matrix depends upon the maximum number of unique or distinct intensity values of pixels in an image. Greater the number of distinct intensity values, highly accurate will be the extracted textural information. But more number of distinct gray levels increases the size of co-occurrence matrix resulting in increase of computational cost and time. Hence, to reduce the size of GLCM matrix, all the intensity values of image are quantized to 16 levels and these images are further considered for construction of GLCM matrix.

Here decision tree machine learning classifier is used to detect the Covid 19. The flowchart for proposed method shown in figure 2.

Flow Chart

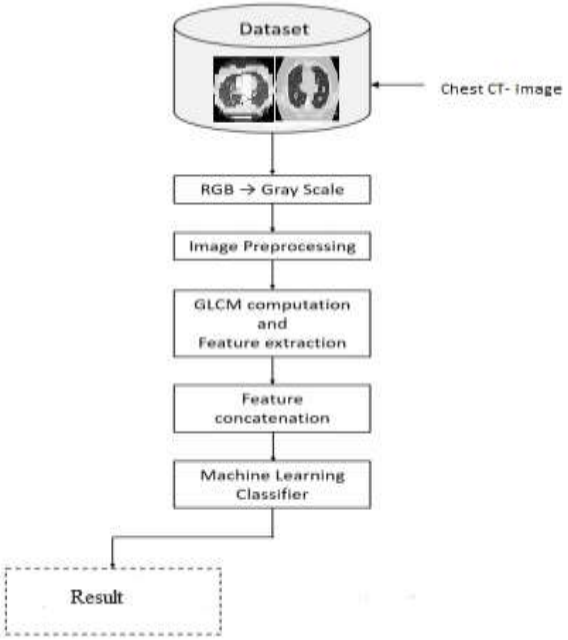

Fig.2 Working Flow chart

IV. SIMULATION RESULTS

Results are simulated using MATLAB 2013a Version.

A. EXISTING METHOD

From figure 3 to figure 10 shows the existing system results. here first we give input took form google net after that it converts gray converted in preprocessing step then it converts into segmentation after that system will show the segmentation image. The segmentation image step then it converts into edge detection image. Then finally shown accuracy and sensitivity.

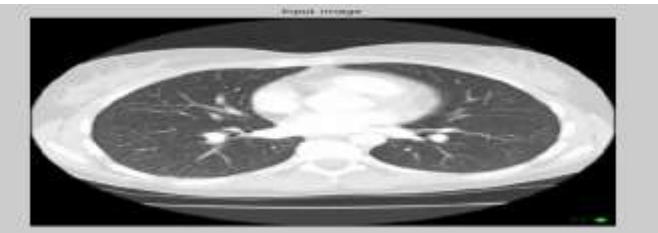

Fig.3: input image

Fig.4: Gray convert image

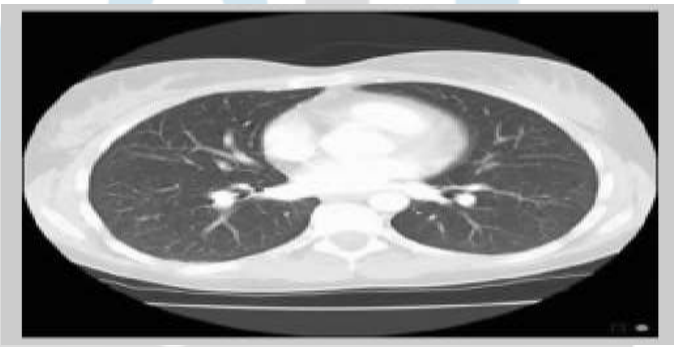

Fig.5: Preprocessing image

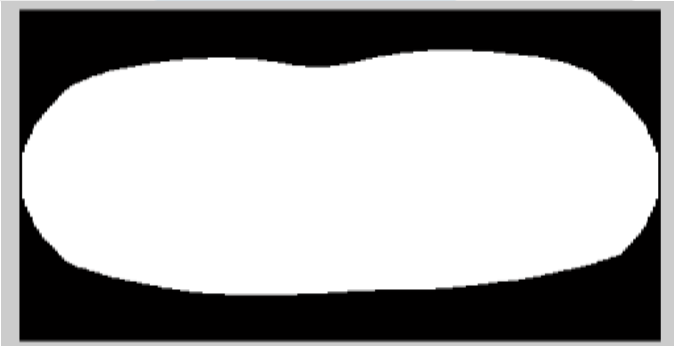

Fig.6: Segmentation image

Fig.7: Edge detection image

Fig.8: Output image

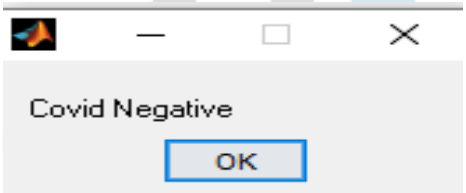

Fig.9: shown covid as negative for selected input image

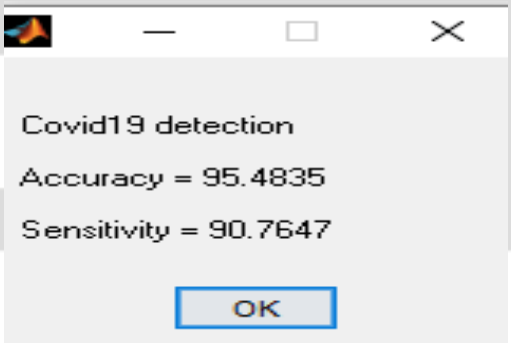

Fig.10: shown Accuracy and Sensitivity

B.PROPOSED SYSTEM

From figure 11 to figure 15 shows the Proposed system results. here first we give input took form googlenet after that it converts gray converted in preprocessing step then it converts into segmentation after that system will show the segmentation image. The segmentation image step then it converts into edge detection image. Then finally shown accuracy and sensitivity

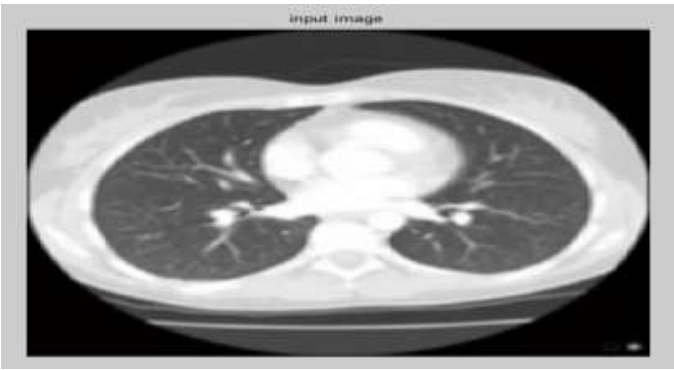

Fig.11: input image

Fig.12: Preprocessing image

Fig.13 Segmentation image

Fig.14: Edge Detection image

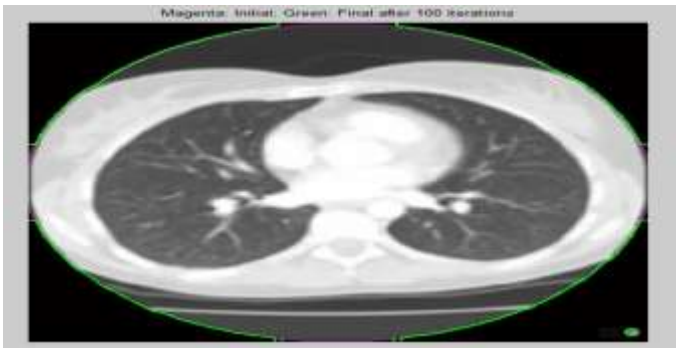

Fig.15: Output image

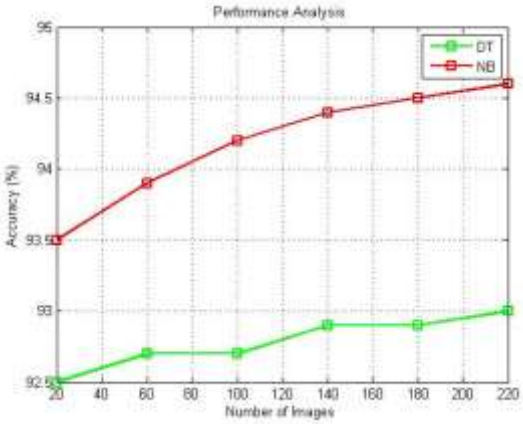

Fig.16: shown performance comparison with accuracy

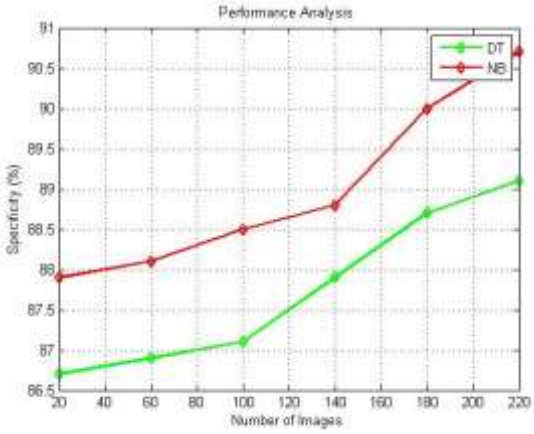

Fig.17: shown performance comparison with accuracy

Figure 16 and 17 shown the accuracy and sensitivity with comparison for both existing and Proposed system.

V. CONCLUSION

The proposed algorithm incorporates firstly pre-processing normalization; then single object image was segmented through fuzzy cmeans clustering; then multiple features statistical, textural, histogram of gradients and discrete wavelet transform were extracted and essential features were selected by GLCM and then Decession Tree, and Navie Bayes classifiers were used to detect the image as normal and COVID positive.

The paper has also explored performance appraise of assorted machine learning classifiers on the best observed GLCM stride combination for Covid-19 identification. Among the experimented two machine learning classifiers (Decision Tree, Naive Bayes); the Decision Tree classifier has shown better ability of Covid19 identification as indicated by all performance metrics.

Future Scope

In future research, Based on best performing individual machine learning classifiers, the proposed method is further experimented with two majority voting-based ensemble combinations of machine learning classifiers as 'Random Forest + Logistic

+ Simple Logistic' and 'Logistic + Simple Logistic + Multilayer Perceptron'. It is observed that the Covid19 identification from chest CT images is efficiently achieved by the ensemble 'Random Forest + Logistic + Simple Logistic'.

REFERENCES

1. N. Deepa, Nanditha Krishna, G. Kumar, "Feature Extraction and Classification of X-Ray Lung Images Using Haralick Texture Features", Smart and Innovative Trends in Next Generation Computing Technologies, 899-907, 2018.
2. Nitin Singh, Rajneesh Sharma, Amit Kukker, "Wavelet Transform Based Pneumonia Classification of Chest X- Ray Images", 2019 International Conference on Computing, Power and Communication Technologies (GUCON), 2019.
3. Tulin Ozturk, et al., "Automated detection of COVID-19 cases using deep neural networks with X-ray images", Computer in Biology and Medicine, Vol. 121, 103792, 2020.
4. Prabira Kumar Sethy, et al., "Detection of Coronavirus Disease (COVID-19) based on Deep Features and Support Vector Machine", International Journal of Mathematical, Engineering and Management Sciences, 5(4):643-651, 2020.
5. Abhishek Sharma, Daniel Raju, Sutapa Ranjan, "Detection of pneumonia clouds in chest CT image processing approach", 2017 Nirma University International Conference on Engineering (NUiCONE), 1-4, 2017.
6. Abolfazl Zargari Khuzani, Morteza Heidari, Ali Shariati, "COVID-Classifier: An efficient machine learning model to assist in the diagnosis of COVID-19 infection in chest CT images", medRxiv 2020.05.09.20096560, 2020.
7. B.Wang et al.,"AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system," Appl. Soft Comput., vol. 98, 2020.
8. S.Wang et al., "A deep learning algorithm using CT images to screen for Corona Virus disease (COVID-19)," MedRxiv, 2020.