

An Overview of Image Processing Techniques for Detecting COVID-19 and Other Infectious Diseases

Raiyan Islam, Shihab Uddin, Md. Shariful Islam, Jamil Mahmud Sakib and Tanvir Ahmed

Abstract—Modern-day medical activities and disease recognition systems are mostly based on the usage of modern technologies. Image processing system is one of the most usable and highly valuable technologies which is used in numerous amount of disease detection process. In this paper, a review will be given based on detecting several infectious and cancerous diseases of different organs in a human body through applying different types of image processing techniques. Image processing system consists of several numbers of image processing techniques which apply to a different category of data and resources. The infectious diseases in a human body possess a certain amount of area in any organ of a human body. Modern medical science of these days is very much advanced that x-ray images, CT or MRI scan images can provide a digital image of a human figure and with the help of these images infections can easily be detected by applying image processing techniques to make sure certain region is affected. A detailed overview will be provided in this review that are the most used image processing techniques to get accurate results on detecting different types of infectious diseases.

Index Terms—COVID-19, COVID-19 Detection, CT Image, Image Processing System, Image Processing Techniques, Infectious Diseases, X-ray Image, MRI Scan

I. INTRODUCTION

In this new era of medical science, the detection of numerous diseases became a lot easier because of modern science and digital technology. Image processing system is one of the most important digital technology which helps to give a proven result of several diseases related to infections of organs in a human body.

Raiyan Islam is a student of American International University-Bangladesh (AIUB). (e-mail: raiyanislam863@gmail.com)

Shihab Uddin is a student of American International University-Bangladesh (AIUB). (e-mail: shahibshihab@gmail.com)

Jamil Mahmud Sakib is a student of American International University-Bangladesh (AIUB). (e-mail: jamilmahmud121@gmail.com)

Md. Shariful Islam is a student of American International University-Bangladesh (AIUB). (e-mail: imdshariful171@gmail.com)

Tanvir Ahmed is with American International University-Bangladesh (AIUB). (e-mail: tanvir.ahmed@aiub.edu)

Infectious diseases can be demonstrated as there is multiple categories of infections that affect the human body and damage a certain area of tissue cells in an organ. Therefore the affected area starts to show abnormal organic activity and day after day it gets spread from cell to cell. At a very early stage, these kinds of infections need to be detected so that necessary steps can be taken in an earlier stage. These days several body parts get infected because of several diseases. In the very last of year 2019, the whole world got affected by Coronavirus contagious disease which later on acknowledged as COVID-19. It damages a little too massive body organ damages and causes mild to severe respiratory illness. Due to this virus, a complicated illness may cause such as Pneumonia. Another major category of body part diseases is Cancer. There are so many categories of cancerous diseases where a part of body cells start to grow at an uncontrollable rate which may form into large tumors and they can start growing into any part of human body even they can or can not be seen from outside of a body part. All of these fatal diseases can get a recovery if they can be detected at a very earlier stage where body effect or damage rate is lower or curable. But without the viewpoint of the damaged region or detection with proper accuracy, necessary steps cannot be taken from medical assistance. At this very point medical science gets merged with digital technology where image processing systems and techniques create major value. The work reviews how image processing system gets applicable in the case of detecting such major diseases.

A. COVID-19 Detection

COVID-19 illness starts by getting affected with coronavirus where a virus cell moves through a human body's respiratory tract and makes the lungs inflamed. This creates a heavy breathing problem and at a certain moment, it leads to pneumonia. To detect pneumonia a Chest CT scan or Chest X-ray images are taken to determine the affected inflamed region of the lungs. To find out such accuracy image processing techniques are applied over the data images of CT or X-ray images. The system is trained by numerous amount of data set of similar category patients and after that, the applied formula can give the proven result to detect if lungs are affected with pneumonia diseases. After the outcome, medical assistance can run over a test and find out the symptoms to make sure if the patient is affected by the virus.

B. Brain, Liver Tumor Detection

The brain and Liver are the most sensitive organ in the human body and there are always higher possibilities of irregular cell growth activity into these important body parts. At some point tissue or cell growth level breaks the boundary line which creates infectious tumors in a certain area. To detect such infected areas with detailed images, MRI or CT scans are taken to examine the infections and to determine if they match with tumor symptoms. In this Magnetic Resonance Imaging and Computed Tomography Scans processes under several image processing techniques where image segmentation, classification and many other approaches are followed which results in an accurate outcome.

C. Lung, Liver, Breast Cancer Detection

Lungs, Liver, Breast these categories of organs faces cancerous diseases because of inner infections and overgrowth of tissue cells. This leads up to a state that it spreads into the body part. Several scan examination gives the output of detailed image view of these irregular behavioral tissue cell regions and from this outcome, image processing techniques are applied by built system application. The system can provide the accuracy output that if the newest input matches with the numerous trained data sets.

II. BACKGROUND RESEARCH

It is known that mankind is researching these several infectious diseases for a long time. Only the COVID-19 is a new disease that came in 2019[1]. To cure the COVID-19, it is important to know whether the patient is a COVID-19 affected patient or not. Thus, several image processing techniques had applied from different types of images to detect COVID-19 patients. When mentioning images, it is about X-ray images, Computed tomography (CT) scan images [15]. Besides this one, other infectious diseases like Brain tumor, Breast Cancer, Liver Cancer, Lung Cancer, Lung tumor are a very old disease which scientists were trying to detect at their earlier stage with the help of image processing techniques, for a long time. As there were some models, method related to image processing techniques which are already been built before COVID-19, so researcher used them for diagnosing COVID-19 and detect them on patient's chest as the virus mainly infect the chest of a patient.

By mentioning the term about previously built algorithm, model, method an observation can be given over the year 2000. In the year 2000, researchers did their research for the detection of Breast Cancer using image processing techniques [13]. Gray-scaled images can be created at that time of old technology. There were mammograms for data analysis and K-nearest neighbor (KNN) and Fuzzy c-means algorithm which were used in breast cancer detection[13]. For these image sources, DDSM (Digital Database for Screening Mammography) was the main place of getting all the collective image data sets.

As CT scan image was mentioned earlier in the paper, the year 2011 can be looked over where it is found that researchers had researched Lung Cancer detection using CT scan images. So back in 2011, there was Lung Image Database (LIDC) where CT scan images were stored. Few of the techniques were applicable for image processing systems at that time. The techniques had multiple names which are (i)Erosion, (ii)Weiner Filter, (iii)Lung Border Extraction Thresholding, (iv)Region Growing Edge Detection, (v)Ridge Detection, (vi)Morphological Operations and (vi)Fitting of Geometrical Models. These are the functions and techniques used in the extraction of lung regions using a segmentation algorithm.[5] At that time, data were analyzed with MATLAB software and an automated CAD system which were used to detect lung cancer.

In 2013, there was Brain Tumor detection-related research which was done with the help of (Magnetic resonance imaging) MRI scan images. The image segmentation algorithm was modified to do the research and at that time when the neural network was getting developed. In this research using neural network techniques, a model was also developed which was named as modified Probabilistic Neural Network model in short PNN and it is based on Learning Vector Quantization. The model was proposed to do the image segmentation process. Edge detection algorithm, MATLAB, grey scaling are some mentionable terms also for the detection process of brain tumor. After two years later, in research(2015), K means clustering technique applied in CT, MRI, USG images were proposed for detection of liver cancer. MATLAB, grey scaling also used in that time. [4]

For these above-mentioned researches, always there was a need for dataset images. So like DDSM, LIDC database, other databases were also growing bigger day by day. In 2017, Lung Image Database Consortium (LIDC) was used in research where the goal is to detect a tumor in the lung from the lung images [6]. Greyscale, noise removal, erosion, feature extraction techniques were applied.

After researching these previously built models, techniques and algorithms, it was realized that in recent years like 2019, 2020, people could work on the previously built technique and improve them a lot. After doing the research, studies found that neural networks, deep learning, image processing are growing and improving day by day. In 2019, several research for lung cancer detection with image processing had been studied for the paper. From the study, it was found that different techniques like clustering technique IPCT(improved profuse clustering technique), trained neural network DITNN(deep learning-based instantaneously trained neural network), convolutional neural network (CNN) were approached for lung cancer detection[8] [10]. So when the COVID-19 virus arrived in 2019, researchers did a lot of research based on the previous study. Different techniques were being applied to get a good percentage result from the image detection.

While discussing the techniques, models, methods, several terms are mention-able. Such as Hybrid Classification Model, Hybrid Selection Approach, MPEG7 histogram filter, Optimized Genetic Algorithm was used for CT images feature analysis and prediction[1]. Then CNN(Convolutional Neural Network) model was a mention-able model which was used on several types of images including X-ray, CT images. [7] [15] [17] [18] [19] Then there were feature extraction techniques, classical image augmentation, Support Vector Machine for classifying COVID-19 from CT and x-ray images[16]. MTL architecture, U-NET architecture, tensor flow, Keras -these are some improved processes for getting a good result from CT scan images for detecting pneumonia from COVID-19 [3]. At first, several X-ray image processing techniques were found for image processing, and later, CT scan images were applied because they could give fast results. As earlier mentioned, this image processing technique needs lots of image data which means data set. So to do the process, experts are improving data sets by collecting CT images, MRI scan images, X-ray images and those data sources can be found online.

III. METHODOLOGY REVIEW

Image processing is a digitization system where images can be processed with a significant amount of data set and from the processing, classification and segmentation a result can be retrieved based on data differentiation, observations, similarities, and statistical possibilities. As most of the fatal infectious organ diseases get imaging diagnosis, multiple image data are easy to get in hand for examination purposes. These images can be categorized based on several medical imagination categories. For the application of image processing techniques, 3 category images can be selected. They are (i) CT scanned images, (ii) MRI images, (iii) X-ray images. The review will be provided based on these 3 category images over which several image processing techniques can be applied for detecting different categories of diseases.

A. Image Processing Techniques for CT Images

There are multiple types of image processing techniques applies on CT images. As for fatal viruses such as COVID-19, different techniques and methods are approached when output needs to be generated from patients' data. In terms of detecting Coronavirus an approach of image processing technique was approached which is a CT scanned image data set was assessed based on the Kaggle benchmark. In this approached method a layout model was declared. In the model, the process begins with the CT Image data collection and data preprocessing. After that, the composing gets done by the Hybrid Selection approach on CT images using MPEG7 Histogram Filter and Optimized Genetic Algorithm. The data Stack go through a Hybrid Classification Model which provides the output result[1]. In another approach, a deep learning technique is used on CT scanned images where CTnet-10 and VGG-16 frameworks were used for Image Augmentation and Fine-Tuning, and later on the result came out by following up Inception V3, DenseNet-169 and VGG-19 category models[2].

Likewise, one more approach is followed for CT imaging analysis using Multitasking Deep Learning-based architecture where images get analyzed and process with MTL architecture based 3 steps. The steps are sequentially

- 1) COVID-19 vs COVID-19 Normal vs Other Infectious Area classification,
- 2) COVID-19 Lesion Segmentation and
- 3) Image Reconstruction. To complete this encoder-decoder is needed for U-NET architecture and these need to be implemented by Tensor flow or Keras [3].

Aside from detecting Corona Virus syndrome by following image processing techniques there is several approaches also get followed for detecting Brain Tumor, Lung Tumor or Cancer, Liver Tumor or Cancer, and many more serious diseases. A method was proposed that liver cancer can be detected with a computer-aided diagnostic system where detection methodology was CT scanned images are used as data set and with this data set K-means clustering was used to segment the images to capture the region of interest where the cancer-relevant tissues are consisted with[4]. A very similar category approach is followed when lung cancer needs to be detected. In that approach, CT scanned lung anatomy images are preprocessed and after that lung region is extracted by an image segmentation process. To examine the outcome accuracy of infection similarities a conditional rule gets applicable on the data set to provide the outcome[5]. Sometimes Lung Tumor needs to be detected in a very early stage so that things don't turn worse after a certain time. To detect Lung Tumor one of the method is the image enhancement technique where the process is followed as Image Greyscale conversion - Normalization - Image Threshold Process

- Noise removal - Feature extraction - Image subtraction. By Image Subtraction the necessary outcome gets retrieved which can be medically examined[6]. A Hybrid bio-inspired algorithm and Convolutional Neural Network method get applied in another process of the image processing system with CT scanned images. In the method, the WOA-APSO algorithm is used on the imagery data. Images get preprocessed at first then get segmented and later on feature gets extracted by using conditional rules and classification[7]. Sometimes Lung Cancer gets detected using deep learning approaches and that also becomes worthy side by side with image processing systems which are detailed herein short. One of the deep learning approaches is Lung Anatomy CT images gets examined by using an Improved Profuse Clustering Technique in short IPCT. Another approach is also based on Deep Learning Neural Network gets instantaneous training. It is called as DITNN for classifying Lung Cancer[8]. It is one of the proven methods when deep learning is used. In a deep learning approach to detect Lung Cancer, CT scanned images formatted as DICOM gets extracted with the central region of interest of the lungs. The extracted region gives similarities of having a Pulmonary Nodule. After that Features on the purpose of extraction and here Deep Residual Network is used for supervised learning.

For the outcome, it gets ensembled for the predictions of various classifiers[9]. Even after approaching all these methods, there is an expert level Lung Cancer detection method that exists where the Deep Convolutional Neural Network approach is used. To detect Pulmonary Nodules a 3-dimensional Convolutional Neural Network(CNN) was designed. It also classifies them into Malignant or Benign diseases-based research where the results are proven in Laboratory orientation and Pathologically[10]. This is applied to the CT image nodule model for the exact outcome of the features.

B. Image Processing Techniques for MRI Images

Magnetic Resonance Imaging (MRI) is an Imaging Technique in the medical sector where it is used in radiology to form the picture of different anatomy of a human body. This MRI helps to find out a very detailed image of several parts so that inner infections can easily be detected without any kind of outer surgery. One of the very sensitive body parts is the brain in the body and an MRI scan is mostly used to find brain tumors. There are many processes and approaches followed to find out brain tumor using MRI scan and using image processing techniques. One of the methods is described as for detecting a brain tumor imaging process goes under 4 individual steps. Based on the MRI Images the steps were used: (i)Pre-Processing, (ii)Filtering, (iii)Segmentation and (iv)Feature Extraction[11]. In an approach detecting Brain Tumor by MRI scanned images Probabilistic Neural Network (PNN) model gets applied. It does the processing with Learning Vector Quantization (LVQ) where images and data gets analyzed with multiple manipulation techniques[12]. Like Brain Tumor, Liver cancer can also be detected in a similar way of using the pragmatic approach of image processing and data mining techniques where K-means clustering technique is used for image segmentation. These are the most followed approach when it is a matter of using imaging techniques to get an output from MRI scanned images. Like other Cancerous disease Breast Cancer can also be detected through mammogram which is also under the MRI subject. To process with mammograms and to detect interesting regions from the affected breast anatomy image processing technique gets applied. The process goes with applying the image segmentation K-nearest neighbor (KNN) algorithm and Fuzzy c-means algorithm on digitized mammograms. Two of the approaches are followed for the image segmentation which is supervised and unsupervised[13].

C. Image Processing Techniques for X-ray Images

X-ray images are widely in medical science to detect most of the inner anatomical problem in a human body and all category image processing techniques are applied for getting the desired outcome. In recent times initial detecting examination of COVID-19 starts with lung X-ray images to find out any type of respiratory problem. Numerous approaches of image processing techniques get applied on X-ray images and most of the valued ones are going to be explained here. At the very first starting with the major Coronavirus detection a technique is followed which is Convolutional Neural Networks method. At first ConvNet Experiment is performed on two sub-categories Covid-19 stage which are (i) COVID-19/Normal, (ii) COVID-19/Pneumonia. With the help of the image preprocessing factor using the APPN approach, the X-ray images are directly sent to ConvNet architecture. In the experiment convolutional network layers were pre-trained system do the observation on resized X-ray images. Three experiments provide the output result which are - ConvNet Experiment, Statistical Measurement Experiment, Transfer Learning Experiment[14]. Aside from regular approaches, deep learning is also get applied to X-ray images to detect multiple anatomy problems.

A Customized Deep Learning method is followed to diagnose COVID-19 by using Chest X-ray images where images get preprocessed with image gray scaling and adaptive Histogram Equalization. Later preprocessed images go into CNN model. After the training and testing the performance metrics gets applied and finally, post-processing gets done using image segmentation by CNN based U-Net model[15]. This provides the focused area as output. Another approach of classifying COVID-19 from X-ray images is using the Shrunk Feature. To process, image augmentation is applied to image data to extract the robust features. After that SMOTE method is applied as well as SAE and PCA architecture gets trained to shrunken the features for the classification purpose[16]. For most of the cases, it is seen that for COVID-19 detection through X-ray imagery data, the deep neural network is used. One of the approach follows as image classification gets done by using ImageNet, SVM, Random Forest, XGBoost and Logistic Regression. Two pipelines of classification get trained. The first approach provides a Binary Classification between COVID-19 cases and Non-COVID-19 cases with the second one being a multi-class classification between COVID-19, pneumonia and normal cases. A pre-trained CNN, ImageNet, is used for the diagnosis task. The approach is a fusion technique between all of the CNNs in three different approaches. The First one being a Mean-Consensus between all of the three approaches. Secondly, avoiding Merely-Averaging of the responses, a Meta-Classifer is used over the top-listed conjugated answers. The last approach was used to extract features from the Penultimate Layer where each of the networks processes with an optimized meta-classifier on the concatenation[17].

For Chest X-ray image classification of COVID-19 cases, another type of CNN model is used which is named DeTraC. To go for this classification approach CXR image sets are divided into training and testing sets with labelled ground-truths which are used as inputs to the system. In the first step of DeTraC a pre-trained CNN model is used as a feature extractor to construct a deep feature space from the CXR images. Principal Component Analysis is applied to the deep feature space for dimension reduction. The reduced feature space of CXR images is used to decompose original classes into several decomposed classes. This concluded the first step of the DeTraC process. In the second step, the final classification layer of an ImageNet pre-trained CNN model is adapted to the decomposed layers which is followed by an adopted Pre-Trained CNN model of fine-tuned parameters. In the final step of the DeTraC process, the predicted labels associated to the decomposed classes are calculated. This is followed by refining the final classification using error-correction criteria. When it is a matter of detecting COVID-19, DeTraC establishes the effectiveness of class decomposition. One of the most challenging issues of detecting COVID-19 cases is dealing with data irregularities where DeTraC can provide a maximum contribution. DeTraC can simplify local data structure with class imbalance by decomposing the class layer. The decomposition gets achieved by investigating class boundaries of the data set and after the adaptation of transfer learning accordingly[18].

Deep Transfer Learning Algorithm can also be applied for the detection process of COVID-19 using chest X-ray images. To apply them at first Features are extracted from pre-trained ImageNet CNNs and leveraged in detection of covid-19 infection. DenseNet 121, NASNetLarge and NASNetMobile are the chosen CNNs that are fine-tuned by pre-loading them with weights trained on ImageNet datasets. Classification layers are not included at the top for ideal feature extraction. A custom classifier gets added on top of the base models which consisted of a pooling layer and a dense fully connected layer. Finally, the convolutional base gets frozen before training the models and that is how the algorithm gets applicable to the model[19].

Another category of deep learning trained model is followed for the detection process of COVID-19 in Chest X-ray images and the model provides a better accuracy rate in providing the output result. It is said to be CheXNet model which is a 121-layer DenseNet based model. To classify the X-ray images by categorizing them into 14 different thoracic disease classes also including pneumonia, the proven model was trained. The frontal-view of Chest X-Ray image which was given can be classified into the following classes sequentially: (i)Normal, (ii)Bacterial Pneumonia, (iii) Viral Pneumonia and (iv)COVID-19 using the proposed system. The model is initialized with pre-trained weights

from CheXNet implementation and the following two stage training are applied: DenseNet's backbone get the weights which are frozen and when the layers are fully connected then the training gets accomplished. 16 sized mini batches is used for the training process, and it was trained for at about 30 epochs. The model which provides lowest validation loss for the training process advances to next stage. After the advancement to next stage, the network gets trained end to end by using the same hyper-parameters in all layers but the network weights are initialized from above at first. For the reason of memory constraints, 8 sized mini batch is used in this stage and it get trained for 10 epochs. The model which provides the lowest validation loss gets the approval for testing[20].

IV. DATA-SET REVIEW

Researches cannot be conducted without proper amount of data set. In this paper there are multiple image processing techniques and approaches for detecting several category of infectious and cancerous diseases are explained with details. The applied techniques are done with a lot of data collection. Here the data set of numerous approaches will be listed with possible sources based on different certainties. On the basis of different medical image category data sets are going to be provided here where what data set is applied on which approach for what purpose is going to be demonstrated.

A. CT scanned imagery Data-Set

1. For a pragmatic approach for detecting liver cancer using image processing and data mining technique, at around 26 clinical patients were examined through CT and USG images. From 26 patients 10 imagery of Liver anatomy were used for the research purpose.

2. On the case of identifying Lung Cancer detection using image processing technique, CT images were collected from NIH/NCI Lung Image Database (LIDC). At around 1000 imagery is passed through the proposed system[21][22].

3. An approach where Improved Profuse Clustering is used for the detection of Lung Cancer by the usage of CT images, deep learning instantaneously trained neural networks is followed. At around 5043 images were analyzed and provide as an input for the analysis process of the lung cancer from different categories and stated patients from 48 individual series. These data were collected from Cancer imaging Archive (CIA)[8]

4. For a Deep Learning approach of Lung Cancer Detection data set was collected from The Lung Image Database Consortium image collection (LIDC-IDRI)[9].

5. For detecting Lung Cancer using the Deep Convolutional Neural Network model, CT images were collected from the Lung Nodule Analysis 2016 Challenge (LUNA16) data set.

This data collection was utilized for training and validating the algorithm[10].

6. On the research purpose of automatic lung tumor detection by using Hybrid bio-inspired algorithm and CNN, at about 120 samples of lung CT images were selected for the experiment of lung tumor detection which contains possible lung tumor states or very early stage lung tumor positioning. They are collected from the NCI Lung Cancer Database Consortium[23].

7. For the screening process of Corona Virus disease (COVID-19) using CT images, total – 51 CT images were collected to examine and testing the system where each of them contains severely Sars and Coronavirus[24].

8. For the classification and segmentation purpose of COVID-19 pneumonia data collection based on CT imaging analysis , 3 data sets from different hospitals including 1396 CT images were collected. One is public available dataset 347 images, 2nd one has 100 COVID-19 CT scans. The 3rd data set came from the Henri Becquerel Cancer Center (HBCC) which is established in France includes a collection of 425 CT scanned images of normal patients and 98 lung cancer patients[25][26].

9. For research on detecting out COVID-19 by using deep learning methods, a collection of 349 COVID-19 positive Computerized Tomography images from 216 individual patients and 463 Non positive-COVID-19 CT images was taken for analysis issue. The data set was All rights reserved. No reuse is allowed without permission[2].

B. MRI scanned imagery Data-Set

1. For a survey of Brain Tumor Detection Using Image Processing Techniques, 42 brain tumor-diagnosed patient's MRI scanned images were used[27][28].

2. For Brain Tumor detection using Neural Network research, a data set which was consisted of 64 MRI-scan Gray-scale images. Each of the images was 220×220 pixels in size[29][30].

3. Research on Breast cancer detection using image processing techniques The data set were collected from - Digital Database for Screening Mammography (DDSM). DDSM images were grey-scaled with 5000x2000 high pixel resolution[31].

C. X-ray imagery Data-Set

1. For research on COVID-19 detection by Chest X-Ray Images using CNN, the data set was – 225 COVID-19 positive patient's chest X-ray images. The X-ray images were collected from 131 male patients and 64 female patients. the average age rate of these people was in between 14 to 58 years[32].

2. On an experimental analysis of a system model based on custom deep learning: COVID-XNet to locate COVID-19, X-ray Images from a healthy patient and patient with COVID

19 cases were considered. For COVID 19 class- BIMCV - COVID-19 data set were collected from Cohen et al. For healthy patient- Padchest dataset also provided by BIMCV[32].

3. By using Shrunken Feature for the Classification of COVID-19 data set: X-ray and CT images, the data set was used which contains 4 individual ARds images, 101 single COVID positive images, 2 of No COVID-19 finding images, 2 pneumocystis-pneumonia images, 11 Sars case images, and 6 streptococcus images[32].

4. Research with Deep Neural Networks usage on COVID-19 detection, a data set consists 13,975 images of x-rays were used which exploded into a much larger data set using data augmentation techniques such as random rotation, zoom, vertical shift and horizontal shift. This helped the research to have more samples of possible anomalies such as image artifacts and noise patterns of different category medical devices connected to the patient, contour layer and volume range of the breast anatomy, differences in size and shape of the lungs due to the patient being an adult or a child[33].

5. For the Classification purpose of X-ray images based on COVID-19 positive patients chest anatomy using DeTraC, a data set of 80 samples contains normal CXR images (ratio: 4020px × 4892px) and CXR images which has 105 and 11 individual data set samples of COVID-19 and SARS cases (ratio: 4248px × 3480px) were used. This number was improved to a total of 1764 samples using Data Augmentation techniques including flipping, translation and rotation from five random different angles[34].

6. On a COVID-19 detection research using Chest X-Ray images experimenting with CNNs Models and Deep Transfer Learning method, a data set of 309 individual COVID-19 positive Chest X-ray images were collected; 236 COVID-19 positive images were obtained from the datasets of Cohen et al and 73 other COVID-19 images were obtained from Kaggle dataset. 2,000 pneumonia and 1,000 healthy Chest X-ray Images were collected from the data set of Kermany et al. All data images were resized to a ratio of 224 x 224 pixels.

V. DATA PROCESSING AND ANALYSIS

The image processing techniques go through multiple steps to get completed and one of the main step is to make it successful is the process with all the data values and data processing. Towards the road of processing medical imagery data to detect multiple category diseases a lot of data need to get processed. In this section an overview of finding out

fatal diseases like detecting COVID-19, Brain Tumor, Lung Tumor, Breast Cancer, Liver Cancer, and many more with the data processing of image processing system will be given. The system will be categorized based on multiple medical imagery data set categories and will provide a detailed explanation of how the process improvises for multiple image processing techniques.

A. Data Processing of CT Images

1. For the screening process of Corona Virus disease (COVID-19) using CT images, the data process system of the proposed model was approached using sub-division models. At first in the Hybrid Selection Model MPEG7 Histogram Filter and Gabor Image Filter were used to train the data set. After that, the CHFS has applied to output the result[1].

2. For the classification and segmentation purpose of COVID-19 pneumonia data collection based on CT imaging analysis experiments were conducted to evaluate the model. By comparing MTL and these other similar U-NET models, MTL was checked whether it is better or not and followed up with best outcome model[3].

3. In a research of A Pragmatic approach for detecting Liver Cancer using Image Processing and Data Mining Techniques with the CT images data set to process with that MATLAB R2011 software was used. As the images were grey scaled that is why for the proficiency the images were converted into RGB. Later, K-means clustering was applied on the RGB converted images to obtain k-means clusters. After that applying Haar Wavelet Transform T_n was computed with the standard deviation. If T_n was in the range then it will result in cancer[4].

4. On identifying Lung Cancer using Image Processing Techniques the data was processed with several steps. At first, the Lung Region gets extracted. Later with image segmentation and analyzing extracted region Diagnostic indicators are applied. It results out with the detected cancer area[5].

5. An approach where Improved Profuse Clustering is used for the detection of Lung Cancer by the usage of CT images and deep learning instantaneously trained neural networks, the process got advanced in a way that is for classification- DITNN method is used and for images conversion to enhancement at first, the edge of the images got detected and later the got segmented applying image segmentation technique. With the help of neural network, the trained features were stored in the database as saved modules for identifying the lung cancer-related non-automated extracted features. Finally, with the help of algorithm, classification is done from weight value, output value comes out from the input node[8].

6. Using a Deep Learning Approach for Lung Cancer

Detection The feature data set gets created at the feature extraction stage and those are fed into several classifiers like XGBoost and Random Forest to get the outcome[9].

7. For detecting Lung Cancer using Deep Convolutional Neural Network model, the trained data set of 4 CT images contain pathologically confirmed malignant diseases get presented. From the targeted nodules value of malignance are automatically circled out for providing the result of probability . If the value is more than 50 percent, it is considered to be a malignant disease and it is resulted out the opposite if those less than 50 percent. Data process and workflow get done based on accuracy, sensitivity, specificity[10].

8. For automatic detection of Lung Tumor a Hybrid bio-inspired algorithm and convolutional neural network is used in research where the sample CT image data set goes under classification process by image classification technique. The performance gets measured through the intelligent lung tumor diagnosis systems which are acquired by comparing different classification and preprocessing algorithms. The evaluation process parameters are used for determining the effectiveness of the model based on accuracy, sensitivity and specificity[7].

9. By using Shrunk Feature for the Classification of COVID-19 data set: X-ray and CT images both were used in the data processing system. At first, the images were processed with a classification technique. After the classification results of raw feature, vectors gave the percentage of performance metrics, which are Sensitivity, specificity, precision, F1-score, balanced accuracy. In that research shows the result using different algorithms, one is shrunken features with sAE algorithm and shrunken features with PCA algorithm[16].

B. Data Processing of MRI Images

1. On a survey on Brain Tumor Detection using Image Processing Techniques data were processed with 4 individual techniques. As the process goes with Pre-Processing: Converting the image into Greyscale mode, Filtering: Using Median Filter for noise removal, Segmentation: Creating pixel separation using and making into one image object using Genetic Algorithm and Feature Extraction: Using Particle Swarm Optimization which is a metaheuristic technique does iterative search on a sample population and gives optimum solution[11].

2. For Brain Tumor Detection using Neural Network research, The data was processed with the proposed model which eradicates the scanned images with image pixel edge detection applying Canny Edge Detection Algorithm. It works on 5 steps:(i) Image Smoothing, (ii) Finding Gradients (iii) Non-Maximum Suppression, (iv) Double Thresholding, (v) Edge Tracking by Hysteresis [12].

3. For research on Breast cancer detection using image processing techniques data were processed with two approaches of image segmentation.(i) Supervised: Train the image pixel data manually at first and then using KNN algorithm to label the remaining pixels.(ii) Unsupervised:In this method FCM algorithm relies on human to label tissue class once the algorithm cluster all the image pixels. After handling both approaches the model combines both results into one datasheet and output the segmented image[13].

C. Data Processing of X-ray Images

1. For research on COVID-19 detection by Chest X-Ray Images using CNN, data got processed with three experiments, (i)ConvNet Exp: On this experiment 4 different network architecture was used with the help of image preprocessing methods. In every structure multiple layer and filters were modified to train the convolutional network structure. (ii) Statistical Measure Experiment: With the data output on image processing in ConvNet informational dataset was measured in this procedure.(iii)Transfer Learning Experiment: The data set of images that provided best output with ConvNet experimental procedure and statistical measurement results, all of them were compared with the pre-trained network models[14].

2. On a Deep Learning system model COVID-XNet for diagnosing and locating COVID-19 in Chest X-ray Images data set got evaluated in two process. They are (i)Quantitative Evaluation - Using the Five-fold cross validation the outcome was achieved by the network where the network was trained and validated with the CNN. They are summarized in a table where these results are categorized as - (i)Sensitivity, (ii)Specificity, (iii)Precision, (iv)F1-score and (v)Balanced accuracy. With the table values a ROC curve for each of the cross-validation folds is also done where reports provides the corresponding AUC values. (ii) Qualitative Evaluation - for finding out the COVID-19 class, the most relevant information comes out based on the prediction performance where the region is highlighted in red which are relevant and regions that are not relevant are presented in dark blue[15].

3. In a research-based Classification and Segmentation technique using Multi-task deep learning-based CT imaging analysis process experiments were conducted to evaluate different models. By comparing MTL, U-NET and other models, MTL was checked that it gives better accuracy rather than other models for both classification and segmentation techniques[3].

4. By using Deep Neural Networks for COVID-19 Diagnosis, the evaluation of CNN architectures was processed by Deep Transfer Learning. For all categories of CNN architectures, pre-trained weights were used and fixed on the ImageNet data set. The newly added classification layer was updated only for specifying the task. The networks were

optimized with cross-entropy loss where an SGD optimizer was used with a learning rate of 0:001. The momentum ratio of the optimizer was 0:9 and weight decay was at the rate of 0:001. Besides, early stopping was implemented, which interrupts the training when the validation-loss stops increasing for several epochs[17].

5. For the Classification purpose of X-ray images based on COVID-19 positive patients chest anatomy using DeTraC process, it starts with class decomposition layer. AlexNet is a pre-trained network that is based on Shallow Learning mode. It is used to extract the exact discriminative features of the 3 original classes. 5 convolutional layer is composed to represent the learned features in AlexNet. There are 3 fully connected one-to-one layers for the classification task. AlexNet generally uses 3×3 max-pooling layers with ReLU activation functions and there are 3 different Kernel Features. The last fully-connected layer was adopted into three classes and the weight parameters were initialised for the specific classification task. To demonstrate the robustness of DeTraC, different ImageNet pre-trained CNN models were used for the transfer learning stage in DeTraC. Performance of DeTraC for different versions of pre-trained models was used for a fair comparison also the pre-trained models were compared in between of both shallow and deep tuning models[18].

6. On research of COVID-19 Detection from Chest X-Ray Images Using CNNs Models with Deep Transfer Learning algorithmic model, data training is done for 10 epochs, with a batch size of 16. ADAM optimization was used with a learning rate of 0.001. The models were trained with categorical cross-entropy. The implementation was conducted in Python 3 using Keras and Tensorflow. Sensitivity and Specificity metrics were used to evaluate the models' performance. Sensitivity = True Positive(TP)/Total positive COVID-19 images.Specificity = True Negative(TN)/Total negative COVID-19 images[19].

7. A DenseNet based CNN Model CheXNet is used to detect COVID-19 using Chest X-Ray. On this analysis focus was mainly the data training process. The training loss due to other classes does not get counted to ensure the actual training loss of the training process. Only a random subset of pneumonia data was considered in individual batches. For avoiding the overwriting on the COVID-19 data the size of this subset was not taken as neither small nor too large so that it does not mask the COVID-19 loss. It is fixed empirically[20].

VI. EXPERIMENTAL EVALUATION

Based on the data set of so many medical imagery categories of different kind of diseases and applying different type of image processing techniques these huge amount of data got processed. Numerous amount of open source data set are collectively gathered from multiple platforms in order to evaluate them for several purposes. One of the major category disease Corona Virus (COVID-19) was selected for

data evaluation. Aside from other disease data also got evaluated by multiple image processing techniques. The imagery data categories were divided into 3 parts because in most of the medical technology data set images of different anatomy is preferred. These 3 categories were CT, MRI and X-ray images. Over these categorized images multiple image processing techniques were easily applicable so that data evaluation got a level advancement. As the data got evaluated, some of the futuristic work scopes got out of the overview. In a few of the evaluations only 1 category of image data sets is being used into the research and analysis. There is very less amount of possible ways got extracted where minimum 2 category images were applicable for the evaluation but not more than two category data can be set as input into any platforms. That is where a few of the obstacles got faced to the evaluation process using several image processing techniques. This paper gives an overview of the data processing system with large number of data set but a survey can improvise more about the techniques and processes which are applicable for image processing systems on different anatomical infectious disease data set. A survey can also help to complete a futuristic evaluation.

VII. CONCLUSION

The paper focuses on image processing techniques and their uses on several infectious diseases like cancer, tumor or detecting symptoms of COVID-19 as early mentioned to identify them at their early stage. The previous technology were discussed in the paper that how the techniques for identifying these infectious diseases are evolving day by day and becoming modern solutions in our daily life so that researcher can use that knowledge for their research and apply in their work. Some specific papers were reviewed to find out approached methods to detect disease and how the data were being analysed and focused for a good accuracy of results. It is reviewed that several data sets were enlarging day by day and nowadays there are a good number of data sets which can be trained for image processing. From the study, it has been realized that most of the result accuracy was above 80 percent and even some of them were above 95 percent [8] [7] [14] [18]. Experimental evaluation based on data sets could tell a little bit realization of those results. The bigger data sets were trained including healthier patients' image and infectious patients' image, the better results with good accuracy can be achieved. So, for future work, it can be said that there are several methods available for training the data for image classification, segmentation, and image processing techniques but always there will be a demand for better data sets with a good available source for detecting such diseases with image processing system.

REFERENCES

[1] Ahmed Abdullah Farid, Gamal Ibrahim Selim and Hatem Awad A. Khater, "A Novel Approach of CT Images Feature Analysis and Prediction to Screen for Corona Virus Disease (COVID-19)" International Journal of Science Engineering Research Volume 11, Issue 3, March-2020, ISSN 2229-5518

[2] Vruddhi Shah, Rinkal Keniya, Akanksha Shridharani, Manav Punjabi, Jainam Shah, Ninad Mehendale, "Diagnosis of COVID-19 using CT scan images and deep learning techniques", <https://doi.org/10.1101/2020.07.11.20151332>

[3] Amine Amyar, Romain Modzelewski, HuaLi, SuRuan, "Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation", Elsevier, Computers in Biology and Medicine Volume 126, November 2020, 104037

[4] Anisha P R, Kishor Kumar Reddy C, Narasimha Prasad L V, "A Pragmatic approach for Detecting Liver Cancer using Image Processing and Data Mining Techniques", SPACES-2015, Dept of ECE, K L UNIVERSITY

[5] Disha Sharma and Gagandeep Jindal, "Identifying Lung Cancer Using Image Processing Techniques", International Conference on Computational Techniques and Artificial Intelligence (ICCTAI'2011)

[6] Kamil Dimiller, Buse UGER, Yoney K. EVER, "Tumor detection on CT lung images using image enhancement", The Online Journal of Science and Technology- January 2017 Volume 7, Issue 1, Page-133

[7] Surbhi Vijh, Prashant Gaurav, Hari Mohan Pandey, "Hybrid bio-inspired algorithm and convolutional neural network for automatic lung tumor detection", Neural Computing Application (2020).

[8] P. Mohamed Shakeel, M.A. Burhanuddin, Mohamad Ishak Desa, "Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks", Measurement, 2019, ISSN: 0263-2241, Vol: 145, Page: 702-712

[9] Siddharth Bhatia, Yash Sinha and Lavika Goel, "Lung Cancer Detection: A Deep Learning Approach", Soft Computing for Problem Solving, 2019, Volume 817, ISBN : 978-981-13-1594-7

[10] Chao Zhang, Xing Sun, Kang Dang, Ke Li, Xiao-wei Guo, Jia Chang, Zong-qiao Yu, Fei-yue Huang, Yun-sheng Wu, Zhu Liang, Zai-yi Liu, Xue-gong Zhang, Xing-lin Gao, Shao-hong Huang, Jie Qin, Wei-neng Feng, Tao Zhou, Yan-bin Zhang, Wei-jun Fang, Ming-fang Zhao, Xue-ning Yang, Qing Zhou, Yi-long Wu, Wen-zhao Zhong, "Toward an Expert Level of Lung Cancer Detection and Classification Using a Deep Convolutional Neural Network", The Oncologist 2019;24:1159-1165

[11] Luxit Kapoor and Sanjeev Thakur "A Survey on Brain Tumor Detection Using Image Processing Techniques", 2017 7th International Conference on Cloud Computing, Data Science Engineering

[12] Pankaj Sapra, Rupinderpal Singh, Shivani Khurana, "Brain Tumor Detection Using Neural Network", International Journal of Science and Modern Engineering (IJISME), ISSN: 2319-6386, Volume-1 Issue-9, August 2013

[13] Tobias Christiaan Cahoon, Melanie A. Suttorf, James e. Bezdek, "Breast Cancer Detection Using Image Processing Techniques" 0-7803-5877-5/00110.00 Q 2000 IEEE journals.sagepub.com/home/jla

[14] Boran Sekeroglu and Ilker Ozsahin, "Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Networks", SLAS Technology 1-13 © The Author(s) 2020, DOI: 10.1177/2472630320958376,A

[15] Lourdes Duran-Lopez, Juan Pedro Dominguez-Morales, Jesús Corral-Jaime, Saturnino Vicente-Diaz, Alejandro Linares-Barranco, "COVID-XNet: A Custom Deep Learning System to Diagnose and Locate COVID-19 in Chest X-ray Images", Appl. Sci. 2020, 10, 5683

[16] Saban Öztürk, Umut Özkaya, Mücahid Barstugan, "Classification of Coronavirus (COVID-19) from X-ray and CT images using shrunken features", August 2020, International Journal of Imaging Systems and Technology

[17] Gabriel Oliveira, Rafael Padilha, André Dorte, Luis Cereda, Luiz Miyazaki, Mauricio Lopes, and Zanoni Dias, "COVID-19 X-ray Image Diagnostic with Deep Neural Networks", 2020, <https://doi.org/10.1101/2020.10.01.20204073>

[18] Asmaa Abbas, Mohammed M. Abdelsamea, Mohamed Medhat Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network", 2020, <https://doi.org/10.1007/s10489-020-01829-7>

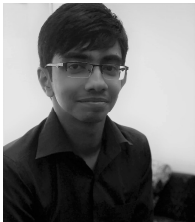
[19] Mohamed Samir Boudrioua, "COVID-19 Detection from Chest X-Ray Images Using CNNs Models: Further Evidence from Deep Transfer Learning", 2020 Available at: <https://ir.library.louisville.edu/jri/vol14/iss1/53>

[20] Arpan Mangal, Surya Kalia, Harish Rajgopal, Krithika Rangarajan, Vinay Namboodiri, Subhashis Banerjee, and Chetan Arora, "CovidAID: COVID-19 Detection Using Chest X-Ray", 2020, arXiv:2004.09803v1

[21] <http://www.singaporecancersociety.org.sg>.

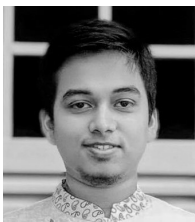
[22] <https://pubmed.ncbi.nlm.nih.gov/21452728/>

- [23] <https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>
- [24] <http://www.oasis-brains.org/Data>
- [25] <http://medicalsegmentation.com/covid19>
- [26] <https://github.com/UCSD-AI4H/COVID-CT>
- [27] <https://en.wikipedia.org/w/index.php?title=Sobel-operator&oldid=750123462>
- [28] <https://en.wikipedia.org/w/index.php?title=Image-segmentation&oldid=753820579>
- [29] <http://www.abta.org/sitefiles/pdflibrary/ABTAfactsandStatistics2010.pdf>
- [30] <http://www.cbtrus.org/2010-NPCR-SEER/CBTRUS/20-WEBREPORT-Final-3-2-10.pdf>
- [31] <https://wiki.cancerimagingarchive.net/display/Public/CBIS-DDSM>
- [32] <https://github.com/ieee8023/covid-chestxray-dataset>
- [33] <https://github.com/lindawang/COVID-Net>
- [34] <https://github.com/asmaa4may/DeTraCCOVID19>



Raiyan Islam was born in Al-Dammam, KSA, in 1996. He is a Bangladeshi citizen and is pursuing his Bachelor's degree from American International University Bangladesh in the Department of Computer Science under Faculty of Science and Technology. He was involved with an education research company (bducation) from 2017 which is also when his undergraduate journey started. He has gained lots of experience as an educational content developer and animator at bducation and is currently dab-

bling in Machine Learning. Currently he is working as a Multimedia Analyst at bducation, Gulshan, Dhaka, Bangladesh. He is also a proficient web-developer with a focus on web deployment. In addition to his career and educational pursuits Raiyan is also an underground football player and is looking forward to becoming an entrepreneur.



Shihab Uddin was born in Dhaka, Bangladesh, in 1997. He is pursuing bachelor of science degree in computer science engineering from American International University-Bangladesh (AIUB) under faculty of science technology department. He started his undergraduate in the year of 2017. Before that he completed his school (2014) and college (2016) from Monipur High School College, Mirpur, Dhaka, Bangladesh. During his undergraduate journey he joined as a general member of ESAB unit face

(engineering students association of bangladesh) in 2020 which is a non-profit volunteer organization. Also in 2020, he joined as a teaching assistant intern in AIUB and completed his internship in the same year. Currently he is a final year undergraduate student in AIUB and doing his thesis to complete the undergraduate program. In future, he is looking forward to receiving a master of science degree in computer science in any reputed university.



Md. Shariful Islam was born in Savar, located near Dhaka city in Bangladesh. Currently pursuing his Bachelor's degree from American International University-Bangladesh (AIUB). From the early days of his university life, he was very passionate about research works and involved with many machine learning-related seminars and workshops. He has also some works on IOT based projects. He was very enthusiastic about computers and artificial intelligence from college days. Currently, he is working

as a web developer in Bangla Puzzle Limited, Shyamoli, Dhaka, and wants to go for higher studies in the automation field.



Jamil Mahmud Sakib was born in Brahmanbaria, Bangladesh, in 1998. He is pursuing his Bachelor degree from American International University-Bangladesh in the Department of Computer Science under Faculty of Science and Technology (FST). He started his undergraduate journey in year 2017. He was an organizing committee student volunteer of ICCA - International Conference on Computing Advancements, 2020 organized by Faculty of Science Technology (FST), American International

University-Bangladesh. Currently he is working as a Product Manager in Virtunus Inc., Middletown, DE, USA. He is a machine learning enthusiast, system designer and data analyzer. He wants to move forward his researches in the field of Medical Data Science and Artificial Neural Network domain.



Tanvir Ahmed Tanvir Ahmed associated with American International University-Bangladesh from 2018 to the Present. He has completed his undergraduate and postgraduate from this university. From 2018-2019 he worked as an Instructor in the Department of computer science, AIUB. Currently, he is serving as a Lecturer in CS at AIUB. He is also in charge of the department's web application. His research interests are computer vision, image processing. Currently, he is working on covid 19 detections from CT or X-ray images. In the future, he is

looking for a Ph.D. opportunity in a reputed university abroad