REPUBLIC OF TURKEY

ALTINBAS UNIVERSITY
Institute of Graduate Studies
Electrical and Computer Engineering

FEATURE SELECTION FOR DIAGNOSE CORONAVIRUS DISEASE BY NEURAL NETWORK AND CALEDONIAN CROW LEARNING ALGORITHM

Mokhalad Abdulameer Kadhim ALSAEEDI

Doctor of Philosophy

Supervisor

Asst. Prof.Dr. Sefer KURNAZ

Istanbul, 2022

FEATURE SELECTION FOR DIAGNOSE CORONAVIRUS DISEASE BY NEURAL NETWORK AND CALEDONIAN CROW LEARNING ALGORITHM

Mokhalad Abdulameer Kadhim ALSAEEDI

Electrical and Computer Engineering

Doctor of Philosophy

ALTINBAŞ UNIVERSITY

2022

The thesis titled FEATURE SELECTİON FOR DIAGNOSE CORONAVIRUS DISEASE BY NEURAL NETWORK AND CALDONIN CROW LEARMIMG ALGORITHIM prepared by MOKHALAD ALSAEEDI and submitted on 20/05/2022 has been **accepted unanimously** for the degree of PhD in Electrical and Computer Engineering

Asst. Prof. Dr. Sefer KURNAZ

Supervisor

Thesis Defense Committee Members:

Asst. Prof. Dr. Sefer KURNAZ	Faculty of Engineering and Natural Sciences, Altinbas University	_____
Assoc. Prof. Yasa EKSIOGLU	Faculty of Engineering and Natural Sciences, Altinbas University	_____
Asst. Prof. Dr. Oguz KARAN	Faculty of Engineering and Natural Sciences, Altinbas University	_____
Asst. Prof. Dr. Mesut RAZBONYALI OZUK	Faculty of Engineering and Architecture, Maltepe University	_____
Asst. Prof. Dr. Zeynap ALTAN	Faculty of Engineering and Architecture, Beykent University	_____

I hereby declare that this dissertation meets all format and submission requirements of a PhD dissertation.

Submission date of the thesis to Institute of Graduate Studies: __ /__/__

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Mokhalad Abdulameer Kadhim ALSAEEDI

Signature

DEDICATION

I sincerely thank my great supervisor Dr. Sefer KURNAZ for helping me conduct the research and prepare the thesis. This thesis is wholeheartedly dedicated to my beloved family, who has been my source of inspiration.

ABSTRACT

FEATURE SELECTION FOR DIAGNOSE CORONAVIRUS DISEASE BY NEURAL NETWORK AND CALEDONIAN CROW LEARNING ALGORITHM

Abdulameer K. Alsaeedi, Mokhalad

Ph.D., Electrical and Computer Engineering,Altınbaş University,

Supervisor: Dr. Sefer KURNAZ

Date: May/2022

Pages: 70

The New Caledonian crow learning algorithm (NCCLA) was used for introducing feature selection approaches in this work. In the suggested algorithms, the crow learning algorithm selects the best features linked to COVID-19 disease in the first step. COVID-19 features selection utilizing artificial neural network (ANN) as training input Experiments at a Brazilian hospital using the COVID-19 disease dataset reveal that crow learning algorithm iteratively decreases feature selection objective function. Reduce quantity of features and reduce error of the classification of the infected patients as healthy by lowering feature selection function. The sensitivity, accuracy, F1, and precision of the suggested technique for diagnosing COVID-19 patients are 94.15%, 94.31%, 94.27%, and 94.38%, respectively, according to experimental data. The suggested technique is more accurate compared with CNN, ANN, CNNRNN, CNNLSTM, RNN, and LSTM to identify COVID-19 patients.

Keywords: COVID-19, Machine Learning (ML), Feature Selection, New Caledonian (NC) row Learning Algorithm (NCCLA)

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

NCCLA: New Caledonian Crow Learning Algorithm

GA : Genetic Algorithm

PSO : Particle Swarm Optimization

GWO : Grey Wolf Optimization

WHO : World Health Organization

SARS : Severe Acute Respiratory Syndrome

MERS : Middle East Respiratory Syndrome

1. INTRODUCTION

1.1 BACKGROUND

Medical records and data are now a powerful resource for uncovering the hidden patterns in diagnosing a wide range of disorders. Machine learning (ML) and Data mining (DM) technologies can be used for extracting hidden knowledge from data. One of the most important applications of ML is in diseases' pattern recognition and classification [1]. Epidemics constitute the most threat to humans currently, and the coronavirus, also referred to as COVID-19, is the latest instance example [2].

COVID-19 disease could be diagnosed using CT-scans of lungs and image processing techniques. One of the difficulties with such study type is that disease doesn't always affect lungs and could be hidden in images of the CT scans of the condition. The difficulty in interpreting a CT scan of the lungs is that just one lung image could be used for diagnosing COVID-19 disease. Those techniques are costly and in the majority of cases unavailable, and they could just be tested in hospitals or medical facilities. Individuals are exposed to pollution in medical centers when they are placed in such environments for testing COVID-19. The principal features regarding the COVID-19, as well as additional clinical indications and other properties linked with virus as well as the person's response to it, are missed by diagnostic procedures.

Coronavirus produces severe acute respiratory syndrome and results in respiratory syndrome in humans. In 2019, the Corona-virus and its pathogen made their first appearance in Wuhan, China [3]. COVID-19 disease is now a global pandemic and poses a threat to human life because of its broad incidence.

Hundreds of thousands of people have died as a result of the coronavirus, and the WHO expects that millions will die as a result of the epidemic [4]. The coronavirus is not just fatal to humans. The Corona-virus has crippled global economy and rendered various people jobless worldwide. Corona-virus causes severe unemployment and poverty worldwide. Moreover, more individuals are sliding into poverty [5][6]. This pandemic had a harmful impact upon every facet of human life. Individuals could no longer appear in public as comfortably as they could prior to COVID-19 outbreak. This epidemic has an impact on both production and demand. Globally,

the disease has terrible economic implications. Because of the virus's devastating consequences on the global economy, human health, and worldwide poverty, several nations are funding coronavirus-related scientific research [7][8]. Many researches were proposed so far for diagnosing COVID-19 disease around the world.

The majority of COVID-19 disease research relies on analysis of the CT scan images as well as the image processing approaches for diagnosing the disease as well as its severity. Computed tomography scans of the lungs are used as training input in image processing techniques, and after that a ML approach is taught for identifying the damaged area [9]. A few of the useful techniques for assessing lung images in diagnosing COVID-19 include artificial neural networks (ANNs) [10], support vector machines (SVM) [11], deep learning (DL) [12], and random forest (RF) [13]. Using CT scans of lungs in order to detect COVID-19 has several benefits, including the ability for diagnosing the disease early on, the ability of creating 3-D image of the lungs, and an ability to pinpoint the damaged areas. The lack of easy access to procedures to diagnose COVID-19 with the use of lung CT scan is a significant barrier. Image processing procedures are not often available, and the usage of small amounts of ionizing radiation could have serious health consequences. Using X-rays to examine lungs is risky for the pregnant women and cancer patients [14]. For decades, ML algorithms were used in medical applications like therapy recommendations, disease diagnosis, and medication prescription. Multilayer neural networks, for example, are now utilized to identify images and evaluate patient information.

In spite of the fact that a great number of research papers in lung imaging were created to identify COVID-19, only a few were created depending on patient records and information. For diagnosing COVID-19, different approaches were created, one of which is the analysis of data provided via medical centers using ML algorithms. Until now, a considerable number of COVID-19-related data sets were uploaded to reliable databases. Each dataset contains multiple traits that can be used for the diagnosis of COVID-19 in people, although not all of these aspects are essential to diagnose COVID-19 disease. A few variables, like gender and age, have been shown to be critical in detecting COVID-19 disease in research papers. A learning approach like neural network (NN) can be used for diagnosing an individual with cod, yet just if the major traits are taken into account during training. The Caledonian crow learning method had been utilized in order to offer a feature selection approach in this work [15].

The goal of employing the crow learning algorithm is to learn in ANNs just the key elements of COVID-19 disease and ignore the less significant features. To help with this research, we created a binary version of the Caledonian crow learning method and ANN for recognizing COVID-19. This article's ultimate purpose is to save patients and develop automated way for identifying corona-virus.

1.2 RESEARCH APPROACH

Increasing the number of machine learning inputs can reduce its accuracy. As a result, a feature selection approach depending on crow learning algorithm had been applied in this work. The relevant features associated to Quaid 19 disease are picked by crow learning approach and employed for learning in the ANN in the suggested approach.

1.3 RESEARCH MOTIVATION

In this study, we will present an approach for diagnosing the Covid-19 pandemic disease using an ANN and NCCLA. For the dataset we will use the kaggle. We will present a binary version of Caledonian Crow Learning to feature selection. We will combine the bird learning and artificial neural network to detect corona. We will save patients and provide an automated approach to diagnose Corona disease.

1.4 RESEARCH SIGNIFICANCE AND CONTRIBUTIONS

For diagnosing COVID-19, a variety of approaches were created, one of which is analysis of data that has been provided by the medical centers through the use of ML algorithms.

Till now, a considerable number of data sets about corona disease were published in respected databases like github or kaggle.

Each data set has several features for diagnosing this disease in the subjects, however, not all features of corona disease are inevitably important.

Others are important because of their age and gender.

Here is an approach of feature selection, which is based upon Caledonian Crow Learning that has been proposed first in 2020 to use only important features of Corona-virus for learning in the ANN and less important features.

An important way to determine how much a person has corona is through using a learning approach, like the NNs, however, only in the case where important features are taken under consideration for the training.

The significant innovations of this thesis are as follows:

Use living learning to detect Corona disease

Presenting a binary version of Caledonian Crow Learning

The combination of bird learning and artificial neural network to detect corona

Save patients and provide automated approach for the diagnosis of Corona-virus

1.5 ORGANIZATION OF THE THESIS

This article is divided to various sections. The problem has been introduced in the first section. The second section of the research background focuses upon researches in COVID-19 field. Caledonian crow learning, which had been utilized in feature selection, is introduced in third section. The suggested approach for corona-virus diagnosis has been formulated in fourth section utilizing binary version of Caledonian crow learning. Implementations and experiments, as well as the suggested approach and its comparison to existing approaches, are discussed in the fifth sections. The research findings and future ideas for developing a suggested algorithm in diagnosing corona-virus have been described in the sixth and final section.

2. CORONAVIRUSES, LITERATURE REVIEW AND ARTIFICIAL NEURAL NETWORK

2.1 CORONAVIRUSES

Coronaviruses belong to a big viral family. Coronaviruses have been linked to serious lung illnesses. The Corona-viruses represent the main cause of various colds, yet 3 of them are particularly dangerous [16]. Transmission between animals and humans was common in early stages of the disease's progression, yet it has since evolved into human-to-human infection. Corona-virus represents a rapidly spreading disease which has put humanity in an unusual dilemma. Former examples include Middle East respiratory sickness (MERS) and severe acute respiratory syndrome (SARS). COVID-19 refers to the last 19 cases of the virus that spread widely in 2019 [17]. In 1918, the H1N1 virus had triggered the outbreak of the Spanish flu, which resulted in the mortality of no less than 50 million people. COVID-19, such as the 1918 outbreak, is regarded as the deadliest epidemic in modern history. The most recent example is the 2009 swine flu epidemic. Figure 2.1 depicts the historical trend of many worldwide epidemics between (1918 and 2019), depending on data from the WHO.

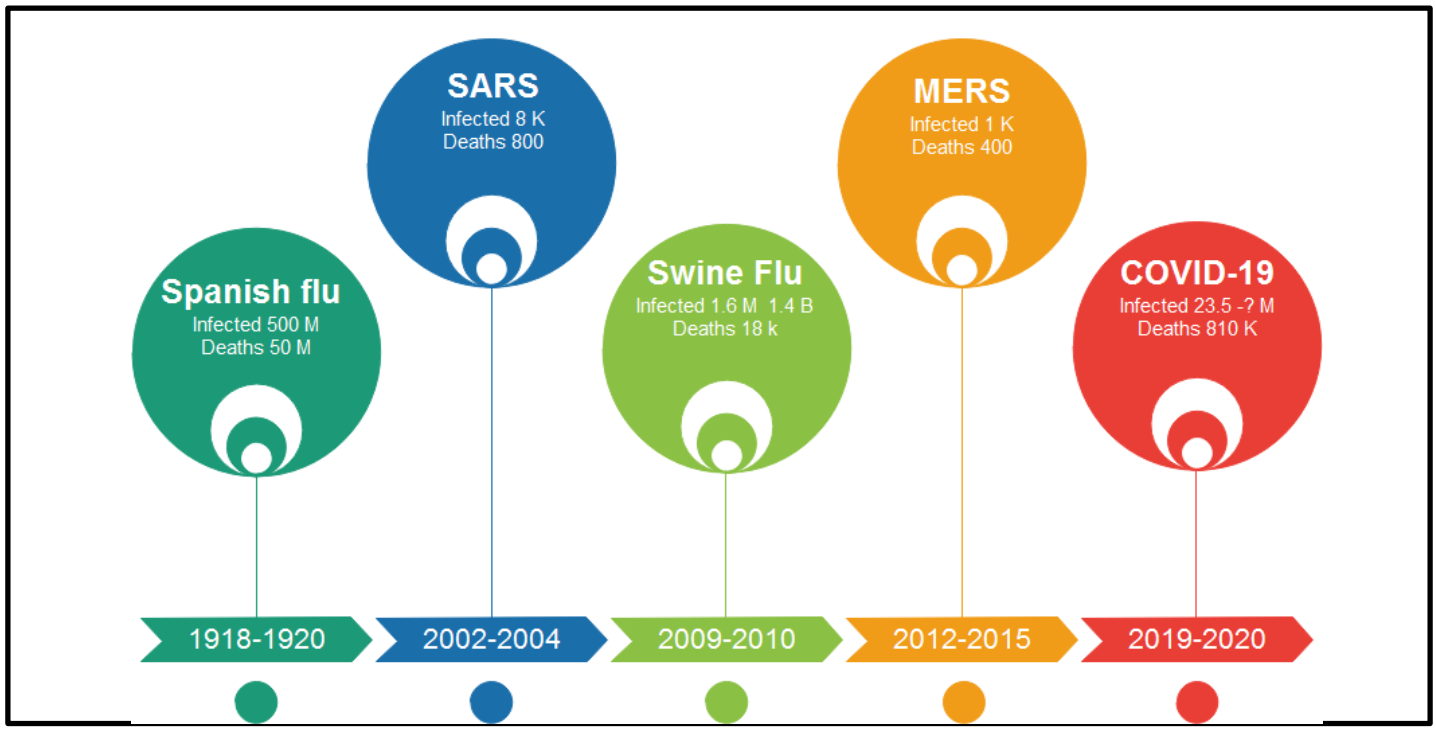

Figure 2.1: Trends of epidemics of the past century and COVID-19 [18]

From 1918 to 1920, approximately 500 million people were infected with Spanish flu, according on analyses of epidemic trends from previous century and corona-virus. According to studies, nearly 50 million individuals have died as a result of the disease. From 2002 to 2004, the next SARS pandemic struck, infecting 8,000 individuals and resulting in the mortality of 800. Between 200 and 2012, the next swine flu outbreak broke out, infecting between (1.6 and 1.4) billion individuals and causing mortality in roughly 18,000 people [18]:

MERS was an epidemic in the Middle East between the years of 2012 and 2015, causing mortality of 400 people, while COVID-19 that had emerged in 2019, infected over 23 individuals and caused the death of over 810,000. COVID-19 has caused numerous mortalities in the short time after it was discovered and diagnosed. Since the Coronavirus epidemic in December 2019, the disease has affected over 23.5 million individuals. Based on reports by WHO, this outbreak will not cease and will continue, with huge numbers of people predicted to be infected and dead due to the disease. According to researches, the virus infects a huge number of people, and Quaid 19 disease could harm vital organs. COVID-19 disease could be diagnosed by gathering data from patients. Figure 2.2 depicts the percentage of each data source which may be analyzed [19][20].

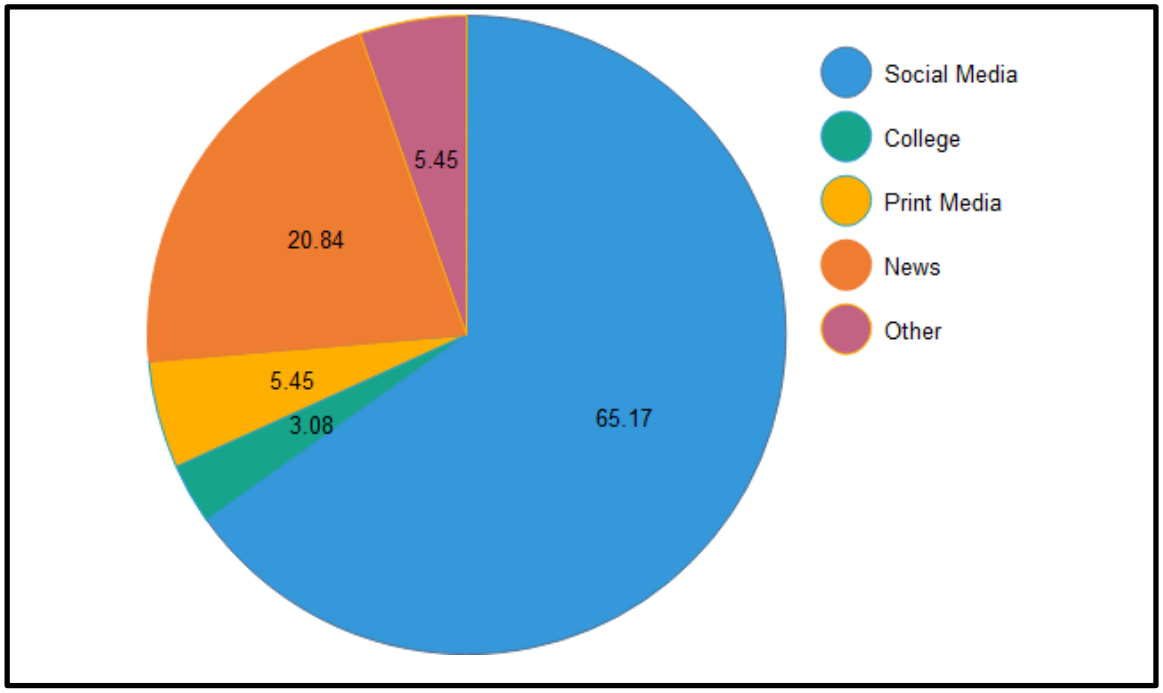

Figure 2.2: Sources of data collection of COVID-19

According to chart analysis, around 65.17% of COVID-19 data is released on social media. According to studies, 20.84% of COVID-19 data has been linked as well to patient information in medical centers' records. This information could be combined with other information to uncover hidden COVID-19 patterns. The first stage in treating any disease is diagnosing it, and various research papers were undertaken with the use of such approach for diagnosing Quaid disease, with a focus on DM and ML.

2.2 LITERATURE REVIEW

The associated COVID-19 disease dataset, as well as DM approaches like text mining, are investigated in [21]. The use of DM approaches to analyze the data of scientific researches on mortality regarding the first corona-virus wave was recommended in the work [22]. They use DM to attempt to uncover hidden COVID-19 patterns with the use of immunity, mortality, and vaccination data. The impact of several medicines on COVID-19 was examined with the use of DM approaches in the work [23]. The findings of their studies reveal that the drug combination of Azithromycin and Hydroxychloroquine, while effective in treating COVID-19 disease, can also cause digestive infections and thin illness in people. According to their findings, such medication combinations should be used with caution in individuals with immune system diseases, CVD, or acquired fat problems. A demonstrated method for the prediction of the crown pestilences using a half breed model of ANN has been offered in the study [24]. The COVID disease that had spread to over two-hundred countries and was declared a pandemic via the WHO, could be a challenging task. With limited information, predicting COVID-19 infection could be complicated. To predict the COVID-19 pestilence, researchers hope to propose a cross breed model that includes test mode assessment and ANN. Time-arrangement data between Jan. 22nd, 2020, and May 18th, 2020, was used in this investigation. Time was organized using exploratory mode parsing to construct sub-flags and misshape genuine data, and the consolidated information was prepared using ANN. The results of their studies suggest that the suggested approach works better than traditional measurable examination.

A DL technique for the modelling and forecasting of COVID-19 in 5 Indian states has been presented in the paper [25]. DL is employed in this work to develop comparative analytical model

regarding COVID-19 outbreak using incremental accuracy of ANNs. In the case when new datasets have been obtained from ever-changing education data, the suggested model intelligently adjusts to new real-time facts. They evaluated and examined DL methods to predict coronavirus infection in [26]. As a result of the rising numbers of the patients on a daily basis, which is a high rate in various nations, lab analyses are time-consuming, and findings and treatment restrictions emerge. A clinical choice model with prescient computations is required due to the limitations of present indicative frameworks. By differentiating illnesses, prescient computations may be able to reduce the strain on medical service systems. They evaluate clinical forecast models using DL and lab data in this investigation. Their suggested approach was validated by 10 approval approaches and 18 lab discoveries from a total of 600 patients. The diagnoses of corona-virus patients has a sensitivity of 99.42% and accuracy of 88.66%, according to experimental studies. A feature selection approach combined with classification algorithms was provided in [27] for diagnoses of COVID-19 with utilizing computed tomography scans. A CNN has been used to determine the image's main features in the first step. Second, the suggested approach employs the whale optimization algorithm for performing feature selection, and lastly, for diagnosing coronary COVID-19 disease, it employs a suggested classification depending on SVM voting, NNs, nearest neighbors, and decision trees. Experimentations have shown that the suggested method is at least as effective in property selection as particle optimization algorithm. In [28], a DL-based recommendation system for COVID-19 disease was suggested. Physician diagnostic volume is reduced by a system that advises diagnosis. To identify COVID-19 fast, the suggested technique examines a huge volume of CT-Scan images. Their findings suggest that DL models based on the CNN artificial neural network may be able to detect illness. In 2019, irregular backwoods indicator has been introduced with order clinical data using highlight positions in [29]. The authors of this paper have conducted large-scale testing on 10 benchmark datasets and found promising results. They provided extremely precise indications of ten different infections. According to study [30], a whale advancement computation that has been based upon the theory of choppiness will be available in 2020 for better help the precision of the machine in the clinical applications. The suggested disease detection model has been compared to numerous competitive SVM models depending on other algorithms of optimization, such as BFO, PSO and GA.

Jainy et. al. [31], principal component analysis (PCA) is utilized for the reduction of dimensionality of the feature space and ANN is used for classification.

In [32], they proposed a strategy to upgrade the characterization execution. To start with, the increased tumor locale by means of picture expansion is utilized as the ROI rather than the first tumor district since tumor encompassing tissues can likewise offer significant hints for tumor types. Second, the expanded tumor area is parted into progressively fine ring-structure sub-regions. The adequacy of this technique was assessed on an enormous dataset with three component extraction strategies, specifically, force histogram [33].

Thejaswini P et. al. [34], presented the method for the MRI image for segmentation. They used the Adaptive Regularized Kernel-based Fuzzy C-Means Clustering. For brain tumor segmentation and classification, the SVM and ANN combined and used for feature extraction. Their suggested approach was tested and evaluated on 94 images and they obtained 98%, 78%, and 91.4% for Sensitivity, Specificity, and Accuracy respectively. A truthful and fully automatic system based on preprocessing that does not take huge time is represented in [35]. For extraction of features from MRI brain images, the deep transfer learning approach has been utilized. For improving the performance, a confirmed classifier model was implemented on the feature matrix. The GoogleNet is used in the proposed model and a 98% accuracy is obtained. The deep convolutional neural network-based multi-grade is provided in [36] for the classification of brain tumor. In this method, in multi-grade tumor classification for sidestepping lack of data problem the data augmentation techniques are used. In [37], for improving the accuracy of classification, they adopted and incorporated the CapsNets. Also for the real set data for magnetic resonance imaging system, the over-fitting problem of CapsNets is investigated. They developed a visualization model for the output of the CapsNet to obtain the best features for improving the accuracy. The deep learning method with pre-trained VGG-16 CNNs combination, for classification of 3 brain tumor types that contain the glioma tumor, meningioma and pituitary is presented in [38]. The GLCM that contains the contrast and energy image and also the input image has been utilized for the feature extraction. The accuracy rate that has been obtained was 96.5%. Navid Ghassemi et. al. [39] introduced an unsupervised pretraining technique to defeat overfitting utilizing generative adversarial networks. The unsupervised pretraining permits utilizing comparable unlabeled data-sets. This technique may be utilized for preparing complex networks on little data-sets. Simulation results have been acquired on brain tumor classified dataset.

Some researches were provided for diagnosing COVID 19 disease in the world, most of which utilize the CT scan image analysis and attempt to diagnose coronary heart disease in individuals with image processing methods. In this method, lung images of individuals are considered as educational input and ML technique is taught to diagnose the affected area [9], including methods for analyzing lung images to diagnose COVID disease. Used include ANN [10], SVM [11], DL [12], RF [13], Stacked-autoencoder-based model [40], Deep Convolutional Autoencoder [41], CNN with Autoencoder [42].

Abdu Gumaei et. al. [43], provided a hybrid feature extraction approach depending on the regularized extreme learning machine. They have utilized the brain MRI images for the evaluating of their proposed method. For the preprocessing and enhancement of the contrast the Min-Max used. In their study 94.233% is obtained from the experiments.

Kaplan et. al. presented a method based upon modified local binary pattern for feature extraction from MRI brain images [44]. In their method at first the filtering has been used in order to normalize and smooth the image, and then use the different modified LBP method that called nLBP and aLBP that the nLBP method was obtained in high accuracy 94.56%. For classification they used the ML such as ANN, RF, K-NN and decision tree.

The liver and brain tumor disease classification based on the CNN, discrete wavelet transform and LSTM is proposed in [45]. In their work the dataset be used from Firat university that contain 56 benign and 56 malignant images. The accuracy that they obtained for the liver tumor and brain tumor was 98.60% and 99.10% respectively.

Swati et. al. [46], used the transfer learning and fine-tuning for MRI classification. The used the pretrained CNN model and the new fine tuning scenario that it was depended on the transfer learning. They achived 94.82% from 5 fold cross validation number.

Brain tumor detection in MRIs using a hybrid technique depending on CNNs and neural autoregressive distribution estimation [47]. A significant issue in the use of lungs' CT scan to diagnose coronary heart disease is that this approach is not often available and, at the same time, it cannot be utilized in large numbers because of the use of ionizing radiation. Using X-rays to examine lungs is risky for pregnant women and cancer patients [14]. For decades, ML algorithms were widely employed in medical applications like therapy recommendations, disease diagnosis, and medication prescription. Multilayer neural networks, for example, are a subset of ML algorithms with a distinctive structure of ANNs which are utilized for classifying images or

analyzing patient information. In spite of the fact that many researches in lung image processing were established for diagnosing COVID disease, only a few were created for diagnosing coronary heart disease depending on patient records and information. Many approaches were created for diagnosing COVID 19, one of them is the analysis of data provided by medical centers using ML approaches. A huge number of data sets about coronary heart disease were given in reliable databases so far. Each one of the datasets has some traits that can be used to diagnose the condition in individuals, however not all of the variables found in coronary heart disease are necessary. The gender and age of individuals are also crucial due to their attack. A significant approach for determining the extent to which an individual has has a coronary artery is to use learning approach like NN, provided that just significant features of education are taken into account. Here, a feature selection technique depending on the Ant Colony Optimization algorithm [48] is utilized to learn in an ANN just the significant elements of coronary artery disease and ignore the less relevant features. Our role in this study is using living things' learning for the diagnosis of the emerging coronary heart disease, provide a binary version of Ant colony learning algorithm, combine ant learning and ANNs for diagnosing COVID 19 dataset, and present an automated approach for the diagnosis of COVID 19. They introduced the COVID-Net, a DNN designed for COVID-19 detection with the use of an X-ray imaging system, in [49]. In addition, they unveiled a new open-access benchmark which contained 13975 CXR images from 13870 patient cases. [50] describes a three-player knowledge transfer and concentration system that includes a pre-trained appearing network extracting the CXR imaging features from a large set of the lung disease CXR images. The data transfer from AP to RF networks in their study. Parnian Afshar et al. [51] provided an approach for COVID-19 case detection from X-rays depending on capsule network. They concluded that their approach is more effective compared with CNN, with sensitivity, accuracy, area under the curve (AUC), and specificity scores of 90%, 95.7%, 0.97%, and 95.8%. The DL is based on the DarkNet model, which was utilized in [52] to classify COVID-19 binary and multi-class cases. The classification model is utilized in this model for the YOLO real-time object detection system. For binary classes and multi-class cases, the classification accuracy rate was 98.08% and 87.02%, respectively. [53] uses an ANN to diagnose COVID-19 in a quick and accurate approach. For the identification of COVID-19 cases from non-COVID-19 cases, they utilized and evaluated 10 CNNs. The sensitivity, AUC, accuracy and specificity of ResNet-101 were 100%, 0.994%, 99.51%, and 99.02%, respectively. The sensitivity, AUC, accuracy, and

specificity for the Xception model have been 98.04%, 0.994%, and 99.02%, and 100%, respectively. Abdu Gumaei et. al. [43], had presented hybrid feature extraction approach that is based upon the regularized extreme learning machine. They used brain MRI images for the evaluating of their proposed method. For the preprocessing and enhancement of the contrast, the Min-Max was used. In their study, 94.233% is obtained from the experiments.

Kaplan et. al. presented a method based upon modified LBP for feature extraction from MRI brain images [44]. In their method, at first, the filtering is used to normalize and smooth the image, and then use the different modified LBP methods called nLBP and αLBP that the nLBP method was obtained in high accuracy 94.56%. For classification, they used ML such as ANN, RF, K-NN, and decision tree.

The liver and brain tumor disease classification based on the CNN, discrete wavelet transform, and LSTM is proposed in [45]. In their work, the dataset is used from Firat university that contains 56 benign and 56 malignant images. The accuracy that they obtained for the liver tumor and brain tumor was 98.60% and 99.10% respectively.

Swati et. al. [46], used transfer learning and fine-tuning for MRI classification. They have utilized the pre-trained CNN model and the new fine-tuning scenario that was dependent on transfer learning. They achieved 94.82% from 5-fold cross-validation number.

The hybrid method based on the CNN and neural autoregressive distribution estimation used to brain tumor detection in MRI images [47].

One of the interesting behaviors in nature for finding food, which has a high intelligence nature, is the behavior of ants. Ants have a clever way of finding the food that can be used to find optimal and to reduce their output error in finding the shortest way. In this thesis, the behavior of ants is first introduced and a modified ant colony optimization for feature selection approach is introduced based upon it. In the second phase, the deep learning and autoencoder-based method are used for classifying the COVID-19 disease and brain tumors data [48].

The ant colony optimization algorithm has been suggested in 1990 by Marco Dorigo as an innovative method for solving hybrid optimization problems [48]. This algorithm is derived from actual behavior of ants to find food in the shortest path [54]. Every one of the ants leaves a chemical material that called pheromone in the path of finding food, which other ants choose the shortest path using the previous pheromones. This algorithm is very useful for solving non-deterministic

polynomial (NP) problems and is used in problems such as itinerant vendor, scheduling problem, vehicle routing problem, etc. [55].

a) To solve any NP problem using the ant colony algorithm, the following must be specified:

b) First, you have to turn the problem into a graph, including nodes and edges.

c) See distance nodes (η) are raised and specified.

d) A possible solution is created according to the problem.

e) The pheromone update rule is used to determine the effective edges in achieving the best answer.

f) The probabilistic transition rule is used to find the next node [56].

Several different implementations of ant colony algorithms such as Ant System, Max-Min Ant System, and Ant Colony System have been proposed, main difference of these methods is in the formula of pheromone update [56].

2.3 ARTIFICIAL NEURAL NETWORK

2.3.1 Background

The ANN is a numerical representation of the NNs in individual's cerebrums, which are made up of a massive number of nerve cells called neurons that are linked together. The average number of the neurons in a human brain is between 10 and 11. The correspondences between neurons are established through fleeting motivations of the electrical signals that are generated by cell's mass, which is referred to as a layer. This information is delivered as electrochemical intersections, referred to as neurotransmitters, from one neuron to the next, in which such intersections are located on branches that connect them to the body of the cell, which is referred to as the dendrites, which carry the signals from a large number of neurons to the nerve cell's body, in which qualities are processed with the aim of selecting the neuron's result, whether or not to create an electrical sign, also referred to as "Fire," for neurons that are related to the result of such neurons, referred to as the axon. A signal passed from one neuron to the next could have one of two effects on the

receiving neuron: excitatory or inhibitory. The inhibitory effect prevents the getting neuron from terminating, whereas the excitatory effect causes it to fire. Regardless, the conductivity regarding the electrochemical intersections that transport various information sources into the cell body limits their influence [57][58]. Figure 2.3 shows an adapted kind of a neuron's basic components.

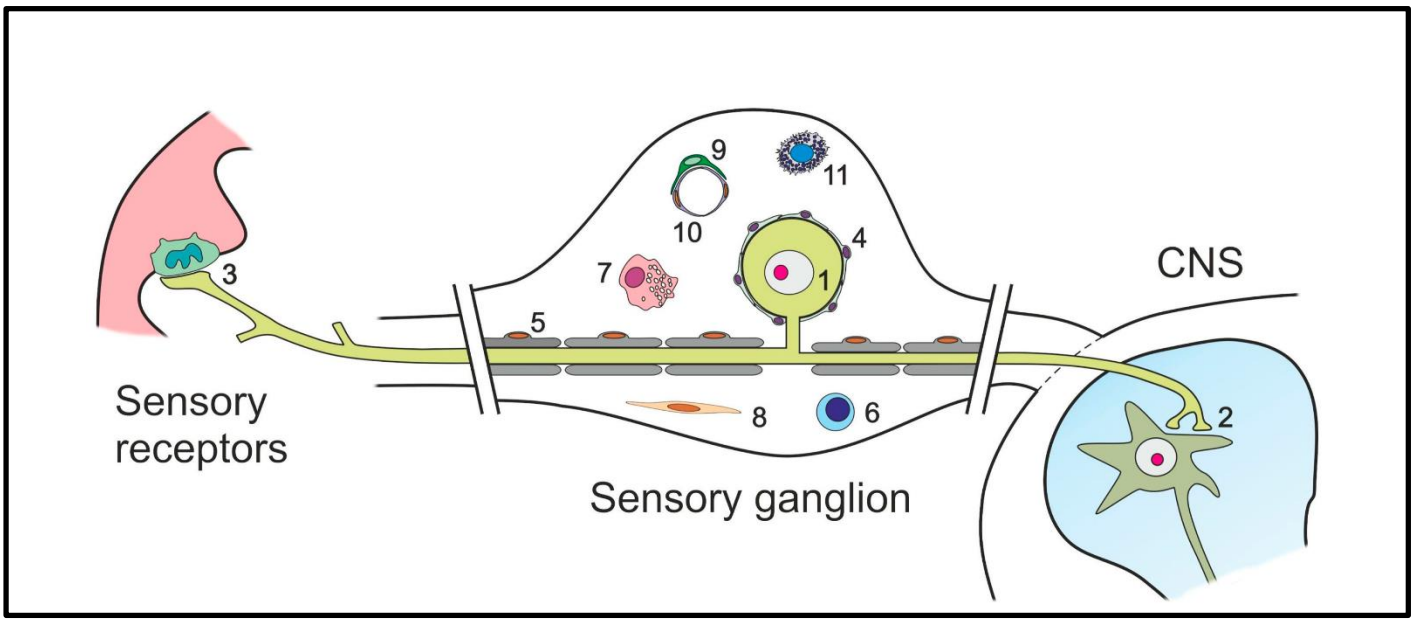

Figure 2. 3: The Essential Components of the neuron’s [59]

The effect of contributions to the neuron is altered in ANNs using mathematical properties referred to as loads. Prior to being conveyed to neurons, each information is duplicated through the comparing weight. Then, a neuron aggregates such weighted characteristics in spite of a predisposition worth and passes results of summation via actuation work in order to determine neuron's outcome. In addition, the bias values present the neuron with more flexibility when it comes to changing the output value depending on the features it needs to detect. In the case when no bias is utilized, for example, the neuron's output is 0 whenever the inputs are 0, because the weights are multiplied by such inputs. As a result, the presence of bias value enables neurons to evaluate needed output depending upon the feature [60]. Equation 2.1 shows the synopsis S of inputs and the bias of neuron j.

$$s_j = \sum_i x_i \, w_{ij} + b_j \qquad (2.1)$$

xi represent contributions of that neuron, wi are loads in relation to those contributions, and bj represents the neurons' bias. The visual layout of the way that a neuron grips the contributions to request registering result is shown in fig. 2.4.

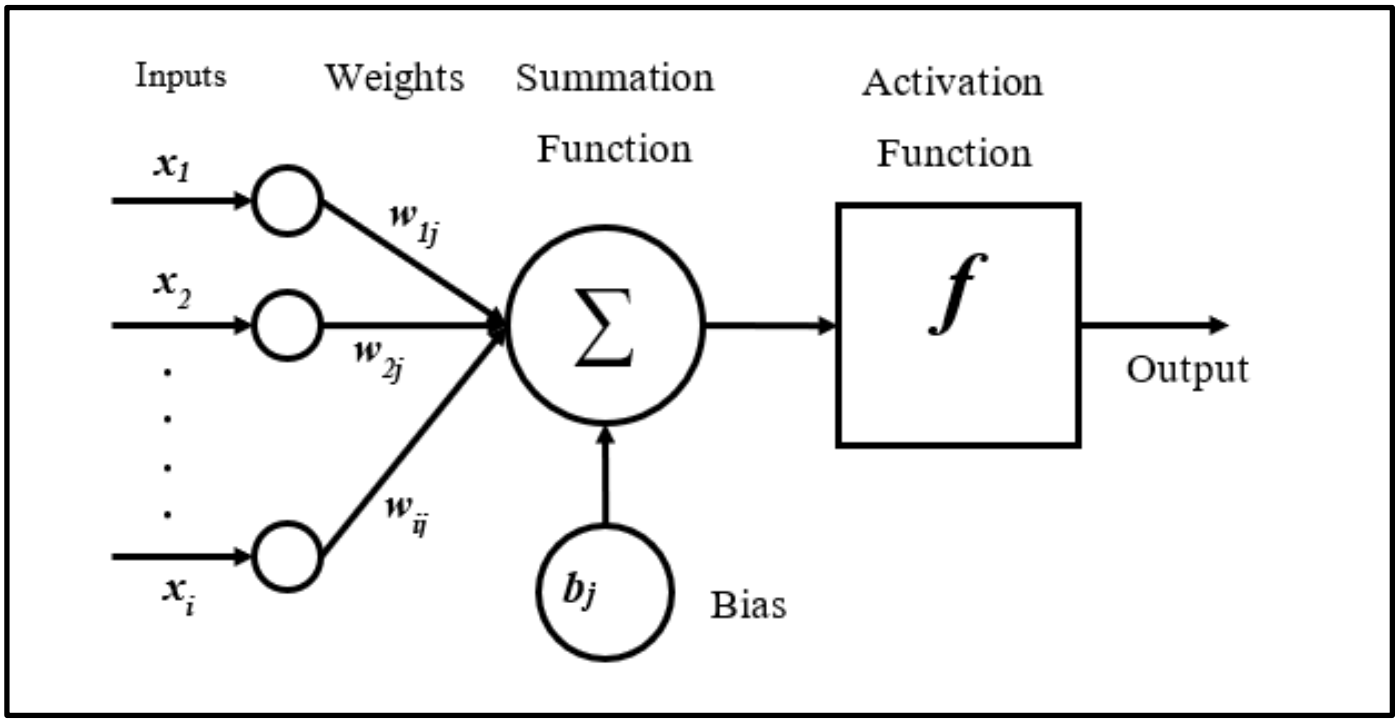

Figure2. 4: Visual representation of Neuron's Calculation operations [60]

By converting the summation's results into an actuation effort, the neuron gains the ability to set nonlinear limits for independent direction. If an actuation work is eliminated from the neuron's calculation, the only possible limit that a neuron may utilize to divide tuples in a data-set into classes is a straight limit that limits the neuron's capability to make more exact forecasts. Furthermore, neurons identified further in NN could be able to create more perplexing limits for every one of the classes, which can improve precision regarding NN's predictions. Furthermore, through modifying the areas of every piece of mind-boggling limit, created by linking neuron limits before neuron, the use of inclination values inside each neuron can aid in the development

of these complicated limits. The Hyperbolic Tangent (TanH), Sigmoid, and Rectified Linear Unit (ReLU) [61] are three often used initiation capacities, as shown in Figure 2.5. NNs having ReLU initiation capacities have shown to be far superior to other enactment capacities in terms of execution [62][63]. The outcome can be calculated from the contributions through detecting the required elements due to the non-linearity of the calculations. However, because a NN might take different courses to come at a certain result, and thus could be the result of a single element of a mix of a variety of the features, such networks are used as single-direction capacities for producing hash values which could be used to represent data, though hash esteem cannot be used for recovering first information [64][65][66].

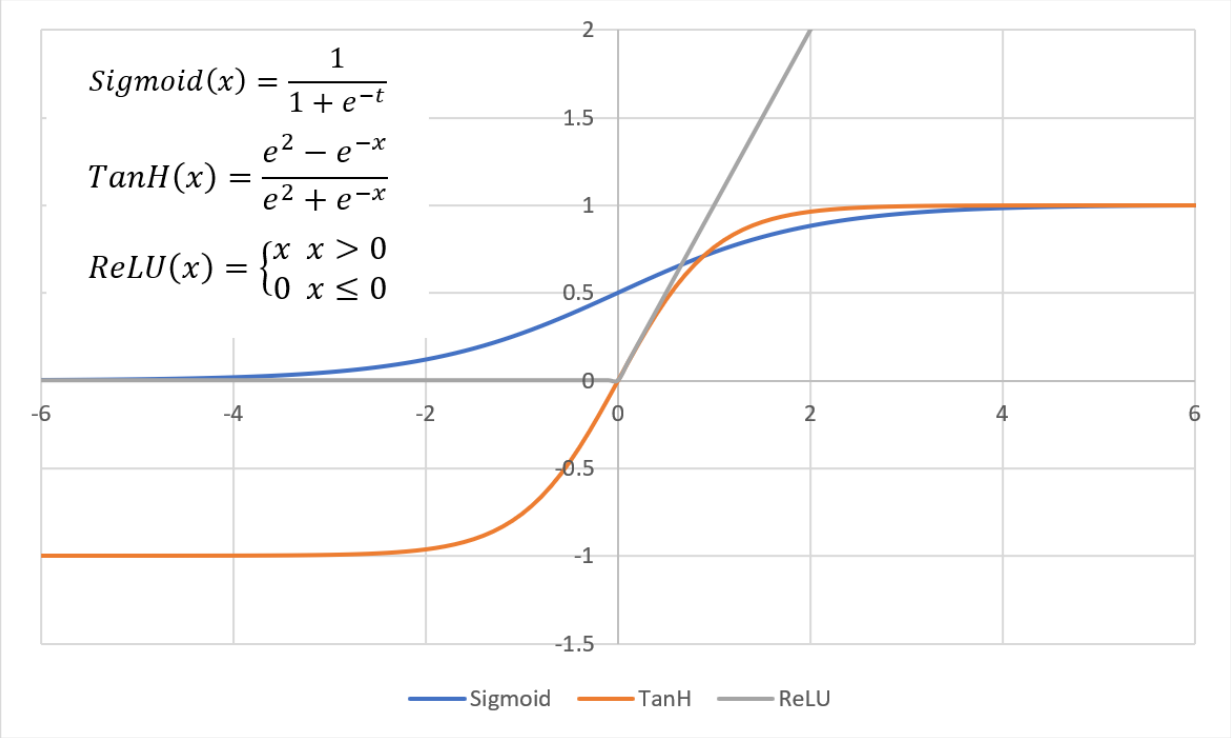

Figure 2. 5: Neurons' Activation Function [67][62]

2.3.2 The Feed Forward Neural Networks (FFNN)

The neurons in FFNN are disseminated in layers, with the results of neurons in each layer being linked to contributions to neurons in subsequent layers. The yield, information, and secret layers are the 3 layer types in the network. The information layers are the main layer with which the contributions from the outside world are associated, and the number of neurons in this layer corresponds to number of contributions that have been made by NN, whereas the number of neurons in result layer corresponds to number of results that are required by the framework. As a result, the number of neurons in such layers is limited by amount of execution that is required in order to complete the assignment for which an NN is being performed. Furthermore, in all cases, using only information and result layers is insufficient to complete the required task, due to the fact that it limits the amount and complexity of elements that the NN could recognize for the purpose of providing correct forecasts. [68][69]. Figure 2.6 shows an example of entirely associated feedforward deep neuronal architecture.

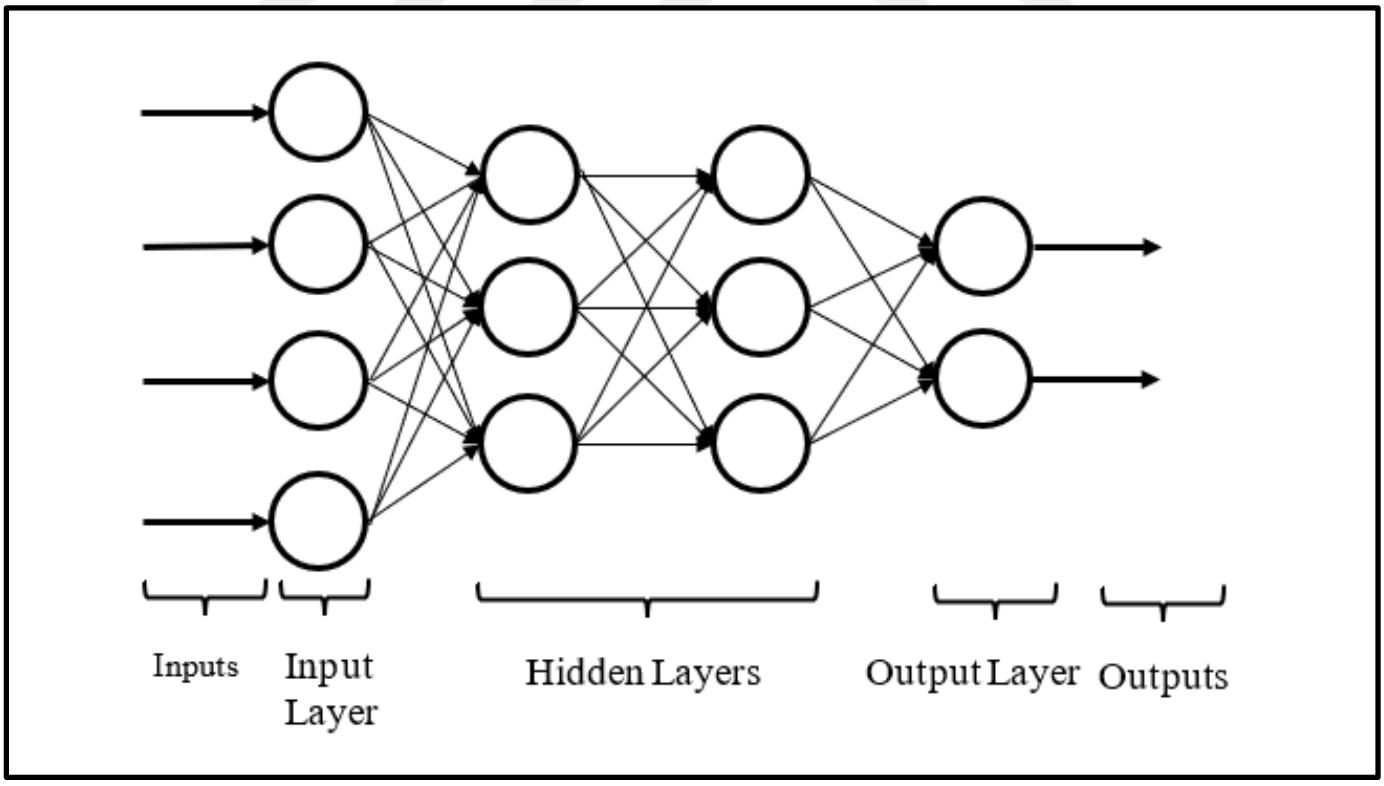

Figure 2. 6: Sample Feed-Forward Artificial Neural Network.

The majority of layers are inserted between the info and result layers when working on the presentation of NNs. This layer is known as a covered layer because it is invisible to the outside world, unlike the information and result layers which are linked with framework's data sources and results. Although the number of neurons in the hidden layers isn't limited by outside factors, increasing the number of neurons in a hidden layer or number of stowed away the layer increases the difficulty of the computations that are required in order to predict a class for information, which resulted in increasing time that is required for that prediction. While increasing the number of layers, increasing the number of neurons in a secret layer results in increasing the number of the highlights which can be identified in that layer, i.e., blends that fire the neurons in hidden layer. For instance, as the profundity of the network grows, so does the complexity of the element that could be identified in more profound layers. As a result, networks with more than one stowed away are capable of detecting more mind-boggling highlights, which are referred to as profound NNs [70][71].

The feature of overfitting, where expectations are dependent on explicit elements in the NN, makes such forecasts very limited to such elements, is one of the main challenges investigated with profound NNs. As a result, any additional data sources that could fit into that class but don't fire the neurons associated with such elements are likely to be wrongly classified. To overcome this problem, a predetermined proportion of neurons in the hidden layer are dropped indiscriminately during each preparation stage cycle, forcing the NN to find alternative ways to make the same forecast and reducing the reliance on explicit elements. This approach is called as Dropout, and it revealed a significant improvement in NN forecasts [72][73], as seen in Fig. 2.7.

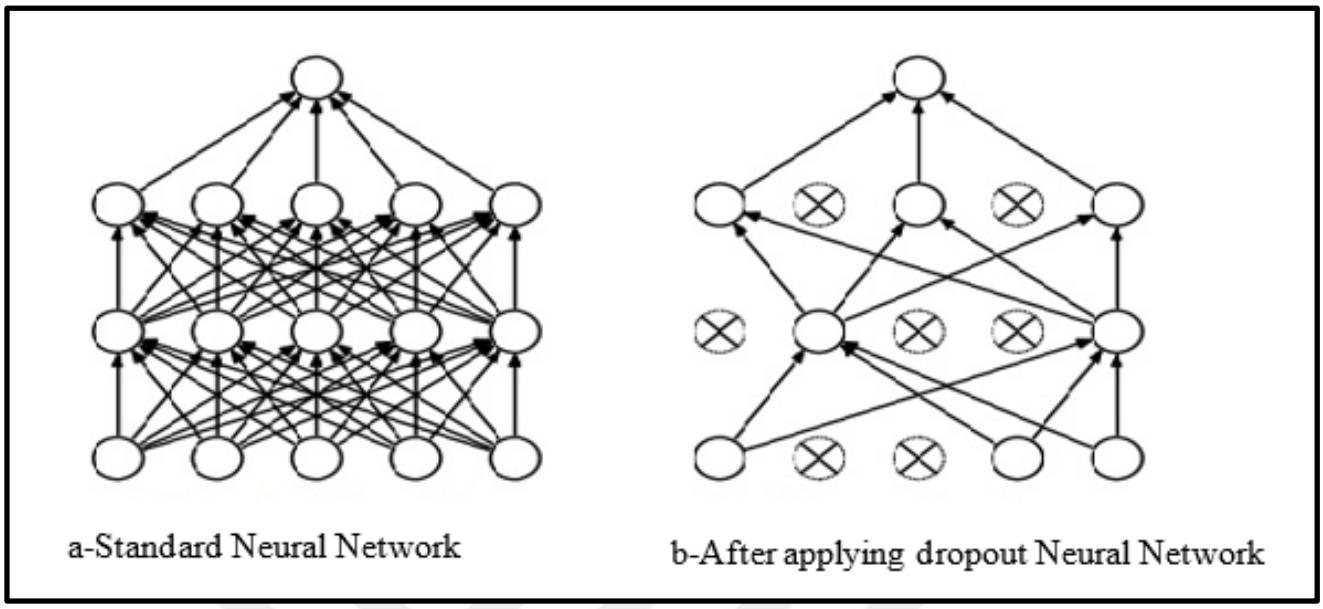

Figure 2. 7: Dropout Neural Net Model [72].

The forecasts that are provided by am NN are results of calculations that have been carried out in every one of the neurons, from the information layer to result layer, which are based mostly on the inclinations and loads of the NN, notwithstanding the information upsides of the information layers. The preparation stage entails adjusting the inclinations and loads of the NN for working on the precision of the forecasts provided by such networks, as the information values are behindhand the management of such networks. Using the Sum of Squared Errors (SSE) approach shown in Eq. 2.2, this update is completed through the estimation of the difference between necessary and anticipated quality, which is referred to as expense.

$$C = \sum_i \frac{1}{2}(y_i - \hat{y}_i)^2 \qquad (2.2)$$

where y represents the anticipated value and ŷ, represents the actual obligatory incentive for i yields in result layer [74]. The predispositions and loads are refreshed by back-spreading across NN and updating their qualities based on slope drop computation in light of drop calculation. This computation deals with fractional first subsidiary regarding determined blunder in the previous layer for each weight and predisposition. The characteristics calculated from each subsidiary are after that deducted from the first value of that inclination or weight. The refreshed attributes ensure that the blunder in the result layer is reduced since the worth reported from the halfway subsidiary

is positive in the case when the effect of that boundary creates the blunder. Furthermore, assuming the subsidiary's value is negative, extending the worth increases the blunder to the point where, when combined with deduction in the update condition, the new worth increases, lowering the error rate and further developing NN's expectations [75].To handle such contributions to a NN in previous and before NNs, two-layered sources of information, such as pictures, were converted into one-layered vectors. This modification results in a data deficiency in the two-layered information, in which highlights in one of the aspects are extended based on that aspect's size. For example, in the case where pixels in an image are straightened equally and the columns of a picture are ordered close to one another, the elements created by vertical pixels are lost since distance between 2 adjoining pixels in an upward direction equals the picture's width. To overcome this problem, NNs are used to find highlights in two-layered data sources, allowing the neighborhood highlights to be found regardless of the direction of such elements in the data [76].

In the visual arrangement of humans, NNs are used for repeating the inferotemporal way. For distinguishing highlights, such networks use two-layered channels that are tangled throughout the entire image, with each channel identifying a specific component for each layer. The recognized elements are also two-layered since the channels are 2-layered. This approach allows the separation of the highlights from 2-layered contribution without affecting input's state [77].The channels in the layer are actually loads that are conveyed in 2-layered clusters and refreshed throughout dataset preparation, resulting in channels which have the ability of recognizing different highlights based on the sources of information and the assignment required by the NN, which is the primary difference from traditional PC vision approaches that detect explicit types of elements, such as corners and edges. Those are tangled throughout the entire picture, with the steps indicating the progression size of the development toward each path. Another exhibit is created by combining the load results with information values, with the size of new cluster that has been determined by the steps that have been defined for comparing layer [78].Pooling layers that are divided to max-pooling and normal pooling layers, are one more important type of layer in the NNs. However, max-pooling layers have demonstrated a strong commitment to overfitting avoidance by reducing the size of one channel's results before transmitting the qualities to the next layer [79]. Max-pooling layers have two-layered channels that are tangled as a result of the layer's contribution. Nonetheless, after the channels, max-pooling layers are used to create another cluster

that has the most severe features identified in each maximum pooling channel. The steps that regulate the progression size that such channels move per convolution should be specified similarly to the size of the channels in these layers [80]. Figure 2.8 shows an illustration of a maximum pooling layer with (2x2) channel size and (2x2) steps.

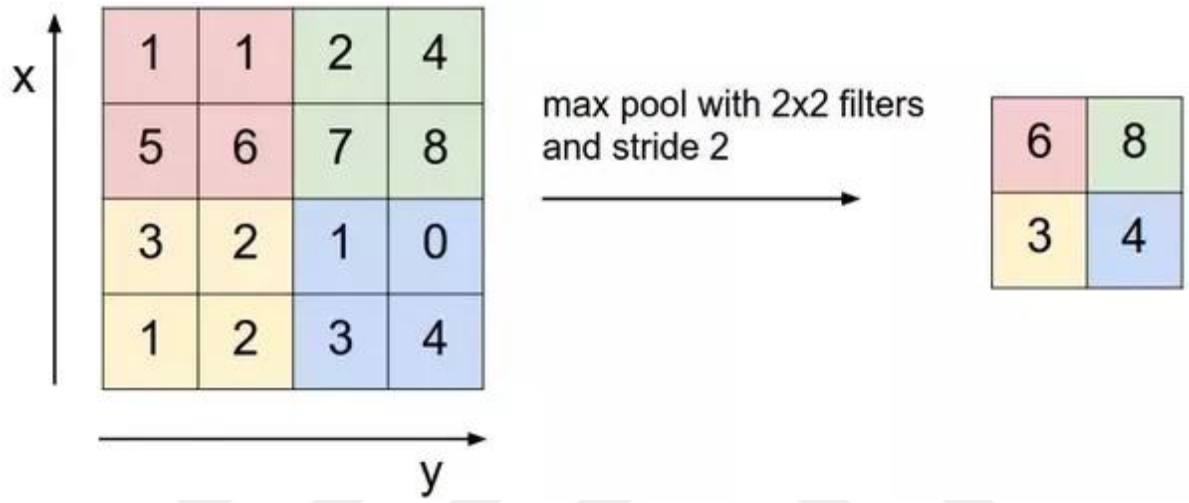

Figure2. 8: A example of the Output of a MaxPooling

The output of final layer, which could be a channel or a maximum pooling layer, is leveled to a single vector and completely that is associated with FFNN. In light of diverse blends of the channels in the layers, the neurons in the entirely linked layers may now distinguish advanced two-layered neighborhood highlights in the information. It's also possible to handle a three-layered contribution to NN, in which data is divided into layers of two-layered exhibits. The NN after that creates many channel layouts for each layer, with more profound levels gaining the ability to combine highlights from many layers [81].

2.3.3 Training Artificial Neural Networks

Like people's cerebrums, where geography of the organic NN and conductivities of the neuro-transmitters characterize the choices that have been made by the mind, the ANNs are additionally dependent upon the dispersion of the neurons and the loads among them for the purpose of settling on the required option. Two indistinguishable NNs may be used in entirely different tasks

where different decisions are produced with the use of different load values amongst their neurons. The type of impact, the result of the neuron in the past layer over neuron in the following one, in addition to meaning of that impact on the result of neuron in the subsequent layer, is determined by the worth of a load between two neurons [82].

Backpropagation is important in the ubiquity of NNs because it allows the network's exhibition to be substantially improved when this process has been utilized for the purpose of updating the value of the network's loads. For the purpose of refreshing loads w of an ANN, the back-propagation requires 3 characteristics, as can be seen from Eq. 2.3, which are the pace of progression of network's result, concerning misfortune being refreshed $\frac{\partial O}{\partial w}$, the mistake E between result of the network and the one that is really required from it and rate of learning L [83].

$$\widehat{w} = w - \frac{\partial O}{\partial w} \times E \times L \qquad (2.3)$$

In spite of the type of error work that is used by an NN, like the cross-entropy and Mean Squared Error (MSE) capacities, such capacities compute one value that addresses the contrast between NN's result using current loading values and the qualities required from the network. The NN's result is acquired by using the forward pass of the NN to handle a cluster of the testing inputs from the preparation data-set, whilst the real results have been obtained directly from the preparation data-set or through handling the sources of information using pre-defined capacity. In backpropagation, the determined mistake esteem is after that used. However, because large blunder values could result in large delta values, a learning rate has been utilized for controlling delta values in lower ranges for loading refreshes. This delta values control ensures the avoidance of the detonating weight values, allowing the loads values which cause the basic mistake to be identified [84].Through ascertaining pace of the progression of result blunder, concerning weight values, 3 possible characteristics may be produced [59], namely: A positive worth, demonstrating the fact that increasing one's weight self-esteem increases the error. To reduce the contrast between the NN and the necessary results, the weight esteem must be reduced by the work out delta esteem in this way.A negative worth, which is an indication of the fact that the error is lessened through the

increase of weight's worth. To reduce the blunder worth and generate more accurate findings, the current weight esteem must be enlarged by the chosen delta esteem.A zero worth, which shows that no changes are needed to current weight value.

The upsides of the loads in NN could be refreshed based on such potential qualities and using the recipe shown in Equation 2.3 to reduce the distinction between the NN's expectations and the actual result required to complete the ANN's assignment. However, in order to reduce the delta esteem employed for refreshing weight values, the optimal NN exhibition, which has been formed through limiting blunder via refreshing load values, working out optimal load values necessitates a number of emphases, such as ages [85].

2.3.4 Transfer Learning

The NN preparation is dependent upon misfortune that has been determined between qualities that are yielded by the neural network, using its present predispositions and loads, and results that are required for this information. Nevertheless, those boundaries cannot be quickly refreshed, to attempt at not missing world-wide least misfortune.in addition to that, as update begins from the yield layer to information layer, in other words, back-propagation, updates to boundaries of layers that are nearer to information layer are not significant. Which is why, the preparation of profound neural network necessitates massive assets and preparing tests for the purpose of accomplishing necessary presentations [86][87].

Layers that are closer to info layer are accountable for the identification of the crude high-lights, which will most likely be divided between different applications that NNs are used for. For instance, in a case where a Convolutional NN is required for perceiving face high-lights [63] or distinctive mark includes, channels in layer which is closer to information layer could identify the comparative elements, instance upward, level and inclining lines. Blending those elements in after layer results in characterizing complicated high-lights which are more engaged with the accomplishment of necessary assignments. Anyway, throughout preparation, layers that are closer

nearer to result layer get quicker refreshes. As a result of that, the strategy of the preparation may be sped up essentially through starting with the lately perceived crude high-lights [88][89].

In move learning, NN preparation does not start with no preparations, for instance, loads and inclinations aren't set in a random manner. In addition to that, NN construction, identified by number of the layers, neurons quantity in every one of the layers and the way through which layers are related to each other, could be differing, starting with one application then onto the following one, concerning the number of the elements that are to be recognized at every intricacy degree and the intricacy of elements that are to be identified to arrive at expected results. After that, a pre-trained neural network could be deployed for preparation, in spite of whether NN is prepared for comparative application, instead of indistinguishable application.

Preparation uses crude highlights that neural network had definitely knew, just as making sure that pre-owned constructions are reasonable for applications. Thereby, move learning had resulted in permitting the critical decrease is preparing time at the same time as providing required presentations [90][91].

3. MATERIAL AND METHOD

3.1 NEW CALEDONIAN CROW LEARNING ALGORITHM

Crow Learning Algorithm can be defined as a calculation that has been first proposed in 2020. The crow that takes in computation isn't exactly the same as crow search computation that has been proposed in 2016. In the approach of the crow learning, crows' learning process has been shown on the guardians as well as other relatives. A crow can learn how to hunt worms in a tree trunk from other members of family.

A representation of crows' intelligent behavior for the hunting has been depicted in Fig. 3.1. Every one of the crows is learning to find optimum solution (i.e. worm). Such meta-heuristic computation has high capacity for looking through complex. Crow calculations may be deployed in the highlight determinations as multi-dimensional problem.

Figure 3. 1: Crow learning mechanism for hunting [15]

According to Figure 3.2, a crow, or solution to some task could be learning from other solutions [15]:

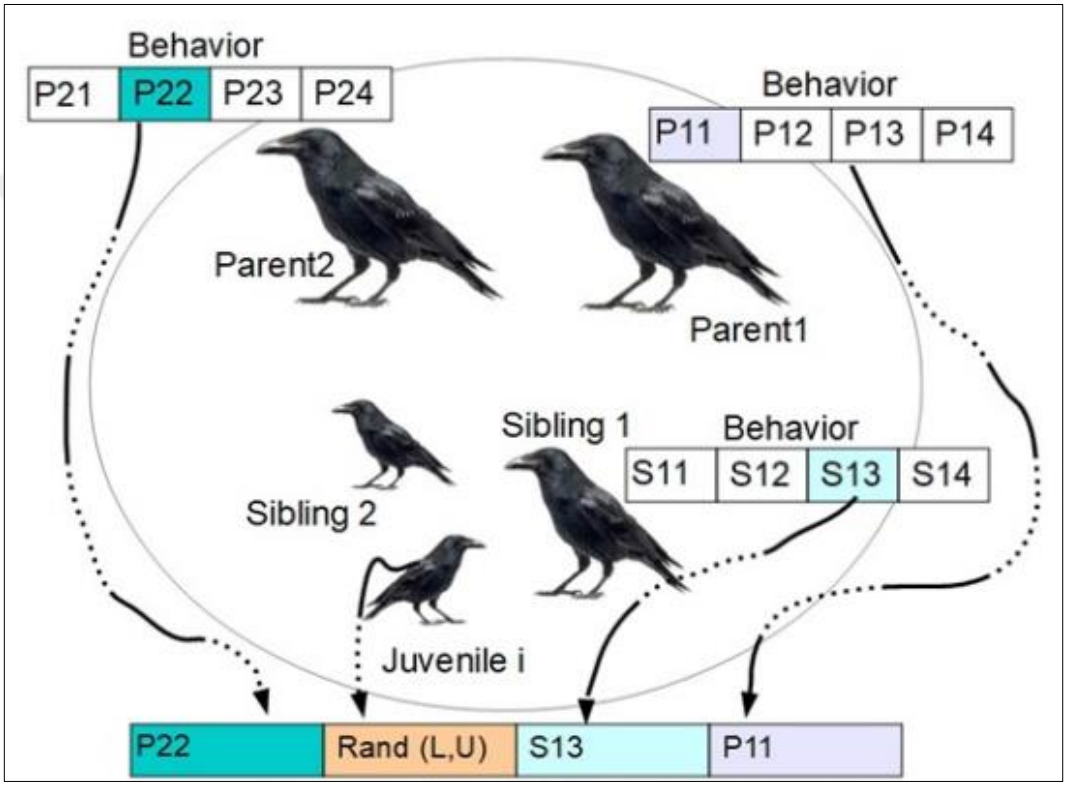

Figure 3. 2: Updating and learning mechanisms in the Crow algorithm [26]

As it has been represented by this figure, parents 1 and 2 have been represented as P1 & P2, respectively. Highlight No2 is parent No2 or P22. Highlight No1 is parent No1 or P11. High-light No3 is its sibling or No1 vector with S13. A crow utilizes the 3 characteristics that have been listed, and its 2nd feature performs the random generation of many other features. Rand (L, U) has been utilized for the purpose of creating random characteristic, representing a random number that ranges from L to U.

3.2 INITIAL POPULATION

Crow algorithm necessitates parameters like the size of the population or n, likelihood of learning or Rpprob, vertical learning likelihood or VSLprob, social learning probability or Slprob, parental choice probability or P1-prob, and trial and error probability or TaEprob and maximal number of the iteration initializations. In some families such as F, several crows signify population. Every crow's position or solution to the problem has been represented as X_i. The members of the crow population and their families have been placed within a matrix like Equation (3.1) [15]:

$$F = \begin{bmatrix} (X_{11}, X_{12}, X_{13}, \dots, X_{1d}) => X_1 \\ (X_{21}, X_{22}, X_{23}, \dots, X_{2d}) => X_2 \\ (X_{31}, X_{32}, X_{33}, \dots, X_{3d}) => X_3 \\ \vdots \\ (X_{n1}, X_{n2}, X_{n3}, \dots, X_{nd}) => X_n \end{bmatrix} \quad (3.1)$$

In this equation, X_{ij} represents solution of its ith then jth solutions. In 1st iteration, a random population of solutions like Equation (3.2) is created [15]:

$$X_{ij} = L + (U - L) \times U(0,1) \quad (3.2)$$

U (0,1) represents random vector which is between 0 and 1 with an even distribution. U & L represent upper and lower bounds of every one of the solutions, respectively. objective function of this problem may be utilized for the evaluation of any crow or feature vector. The competence of the crows has been defined on the basis of the experience and learning in relation matrix (3.3) [15]:

$$Fittness = \begin{bmatrix} Fittness(X_{11}, X_{12}, X_{13}, \dots, X_{1d}) \\ Fittness(X_{21}, X_{22}, X_{23}, \dots, X_{2d}) \\ Fittness(X_{31}, X_{32}, X_{33}, \dots, X_{3d}) \\ \vdots \\ Fittness(X_{n1}, X_{n2}, X_{n3}, \dots, X_{nd}) \end{bmatrix} \quad (3.3)$$

Population members may be ordered based on their merits. After the sorting of population, 2 worthy members, e.g. X1 & X2, are taken under consideration as the 2 solutions, and the rest n-2 crows are taken as the young crows, which are learning and experiencing. In the phase of the learning, each one of the individuals or crows in a family enhances their conduct with the use of

a variety of the learning systems. The crows can have social and individual types of learning. In the phase of the learning stage, each one of the young adults in a family with little experience makes an attempt of socially or separately refreshing their conduct qualities. Every one of the teens must utilize behavior of the rest of the crows for the purpose of updating their social-behavioral features.

3.3 HORIZONTAL AND VERTICAL LEARNING

In crow learning algorithm, in the case where juvenile crow makes the decision of learning socially based on the SLprob probability, it can make the decision of learning from its parents or the bigger and more skilled siblings. In this algorithm, a likelihood of the learning from parents equals the VSLprob and probability of the learning from vector and sibling equals -1 VSLprob. Larger VSLprob amounts provides the crows with the ability of learning more than the parents. Increased 1- VSLprob results in making the crow learn more than its siblings. The modelling of the learning behaviors from the siblings or parents in eq. (3.4) is expressed below [15]:

$$\begin{aligned} &X_{ij}(t) = X_{kj}(t-1) \\ &for\ Vertical - Learning: k = 1\ or\ 2 \\ &for\ Horizontal - Learning: 3 \leq k \leq i-1 \end{aligned} \tag{3.4}$$

In the formula above, t represents new iteration number. t-1 represents current algorithm's iteration. k represents number one crow in its population. In the case where the learning has been based upon the parental behavior, then k equals 1 or 2. In this algorithm, feature that is utilized in vector may be utilized in the phase of learning. Learning could be chose from a sibling crow, whose number equals k = 3 to i-1, and may be randomly chosen based on Eq. (3.5) [15]:

$$k = 3 + [rand \times (i-3)]\ \&\ i \geq 3 \tag{3.5}$$

3.4 INDIVIDUAL LEARNING

In crow learning algorithm, in the case of inexperienced or juvenile crow decides learning individually, it can perform a random update to its behavioral features with the use of Eq. (2). Crows don't take part in the social learning with (1- SLprob) probability. TaEprob has an identical probability of individual learning for every one of the crows. The likelihood of the individual-

learning for every one of the crows is the same (1- SLprob). High TaEprob values provide the global search agents with the ability of doing this algorithm in a more random manner around global search and search space [15].

3.5 REINFORCEMENT LEARNING OF CROWS JUVENILES

After completion of group or individual learning phases, some behaviors and properties that have been learned by the adolescent members from the parents could be including rewards. The likelihood of receiving prize is the same as Rpprob's. Based on Equation (3.6), an individual receives prize [15]:

$$X_{ij}(t) = X_{ij}(t) \pm RW \tag{3.6}$$

RW denotes reward that has been computed with the use of Equation (3.7) and Equation (3.8), where any characteristic, j, Xi adolescent behavior with a repetition t may be decreased or increased by the RW [15]:

$$\alpha = |X_{ij}(t) - X_{ij}(t-1)| \tag{3.7}$$

$$\beta = X_{ij}(t-1) \times \exp(-lf \times r \times t \times mean(j)) \tag{3.8}$$

For the calculation of β, average of some characteristic j, like the average (j), has been utilized. The r value represents a random number that ranges from 0 and 1. Value of lf represents learning factor value that has been computed based on Equation (3.9) [15]:

$$lf = lf_{min} + ((lf_{max} - lf_{min})/Max\_t) \times t \tag{3.9}$$

The value of lf_{min} is minimal factor of learning. lf_{max} represents maximum factor of learning. Max_t value is maximal iterations' number. t represents number of current iteration of crow learning process. The factor of learning in this algorithm has been increased. The value of the factor of learning has been increased from lf_{min} in 1st iteration to lf_{max} in final iteration. The coefficient of the reward is computed based on Equation (3.10) [15]:

$$RW = \begin{cases} \beta - \alpha & if \quad i < n/2 \\ r1 \times ((r2 \times \beta) - \alpha) & otherwise \end{cases} \tag{3.10}$$

3.6 REINFORCEMENT LEARNING OF THE CROW PARENTS

Following the strengthening of the adolescent learning, X1 & X2 parents reward their behaviours with the increase in their experience and knowledge based upon their earlier experiences. Via phase of learning reinforcement, X1 and X2 update some of their behavioral properties, like Equation (3.11) [15]:

$$X_{ij}(t) = \begin{cases} X_{ij}(t-1) - [X_{1j}(t-1) + \exp\left(r1 \times mean(j) - X_{ij}(t-1)\right)] & i = 1 \\ X_{ij}(t-1) - [r2 \times X_{1j}(t-1) + \exp\left(r1 \times mean(j) - X_{ij}(t-1)\right)] & i = 2 \end{cases} \quad (3.11)$$

In the equation above, a parent crow or i = 1 utilizes 1st equation. Mother crow utilizes 2nd equation. $X_{ij}(t-1)$ represents optimal solution. The r1 & r2 values are 2 uniform random numbers that range between 0 and 1, respectively.

3.7 THE SUGGESTED APPROACH

The issue in the approach that has been proposed is the issue of the classification of patients who have COVID-19 as healthy subjects. The main objective is reducing error of the classification of COVID-19 patients as healthy individuals. In the present chapter, the suggested approach for COVID-19 diagnosis is presented and formulated with the use of feature selection through crow algorithm and ANN.

3.8 PROPOSED FRAMEWORK

The proposed approach's framework for the feature selection and COVID-19 diagnosis utilizing multi-layer ANN with crow learning process had been illustrated in Figure 3.3. In this framework, the data are gathered initially from the medical centers followed by the pre-processing. Data in suggested approach include patient information as well as their features that have been obtained. The data and the samples that have been associated with COVID-19 have been apportioned to 2 classes, which are: training and testing with 70% to 30% ratio. A component vector of the highlights associated to patients with corona-virus has been considered as subject from crow

learning computation. Underlying population of the highlight vectors, which are 0s & 1s, have been deployed for applying the informational collections and diminishing counterfeit neural organization's information measurements. At every one of the stages, a feature vector is chosen as crow, and the two key factors below are computed for the evaluation of every one of the feature vectors*:*

a) COVID-19 patient classification error
b) Number of the chosen features

Any feature vector reducing the 2 error indices and number of features that have been chosen has been referred to as the optimum feature vector in disease diagnosis. The minimization of those 2 aims goals enables crow to be considered as a parent in addition to training the rest of the crows. The suggested approach's steps can be summarized below*:*

Several data that are associated with COVID-19 collected and placed in dataset or available data is utilized.

Data that is related to the patients data is preprocessed. For the purpose of increasing the accuracy of learning, the phase of normalization is utilized in the pre-processing.

Training samples, representing approximately 70% of total samples of the patient, are utilized in the processes of machine training and learning, and the remaining samples are utilized as testing samples for the evaluation of suggested approach in diagnosing COVID-19.

Utilizing binary version of crow learning approach, significant features have been chosen from amongst disease-associated features.

The features that were chosen are utilized for ML in ANNs.

In the phase of feature selection, the feature vector has been chosen as crow learning algorithm member, and that feature vector could have 0 and 1 components, indicating non-feature selection and feature selection that are associated with the Quaid 19 disease, respectively.

Numerous feature vectors are considered arbitrarily as crow learning algorithm members.

Through the implementation of crow approach, significant features that are associated with the illness have been identified and this optimum feature vector has been utilized for the learning of the ANN.

The Crow Learning Algorithm utilizes the individual, social, and enhanced mechanisms of learning for the purpose of updating the feature vectors.

The feature vectors undergo conversion into the binary at every one of the steps with the use of functions of conversion.

Updates are carried out on the feature vectors and ultimately, optimum feature vector of the parent and child types is specified in every one of the iterations.

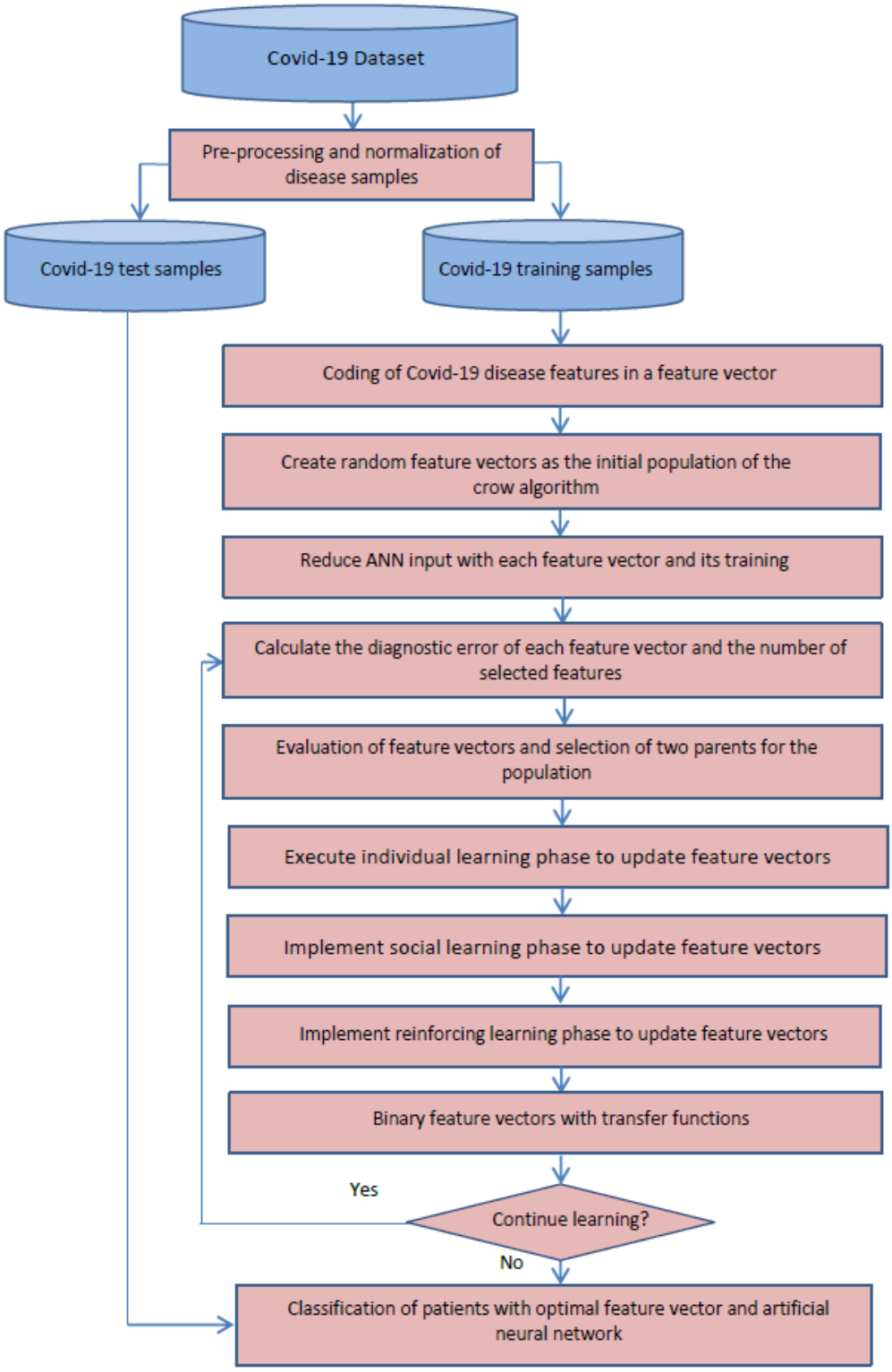

Figure 3. 3: Flow-chart of suggested feature selection algorithm and COVID-19 patient classification

Utilizing optimum feature vector, input filter has been created for ANN for reduction of sample size as for learning in ANN.

NN has been utilized by utilizing optimum feature vector for the classification of the test samples that are related to COVID-19.

3.9 PREPROCESSING COVID-19 PATIENTS

In the present work, corona-virus data from [26] was utilized. Information that is associated with the patients who have COVID-19 has been utilized in the dataset for the evaluation of the suggested approach. This data can be preprocessed as normalization for the more accurate classification of the patients and users. For normalizing, Eq. (3.1) has been utilized, and every one of the columns of separate dataset has been normalized with the following equation:

$$T' = a + \frac{T - min(T)}{max(T) - min(T)}(b - a) \quad (3.1)$$

The abnormalized and normalized values of the feature that is related to the users and patients of corona-virus equals T and T', correspondingly. The maximal and minimal non-normalized property values can be represented respectively by max(T) and min(T). In the majority of the researches, a normal limit has been considered equal to [0,1] which is why, the criterion above may be expressed in the form of as eq. (3.2):

$$T' = \frac{T - min(T)}{max(T) - min(T)} \quad (3.2)$$

3.10 CODING OF COVID-19 FEATURE VECTORS

In the suggested approach, every one of the solutions is characterized and encoded with feature vector. Every feature component is associated with diagnosing the binary disease and it equals 0

and 1. Every one of the feature vectors or crow algorithm members is identified and calculated based on Eq. (3.3):

$$F_i = \ll F_i^1, F_i^2, \dots, F_i^D \gg \quad (3.3)$$

Here, F_i represents binary vector of different COVID-19 features. D represents number of the features that are utilized for this disease. In initial step, an attempt has been performed for the creation of a random feature vectors' population as initial crow algorithm population, which is based on Equation (3.4):

$$Pop = \{F_1, F_2, \dots, F_N\} \quad (3.4)$$

Pop represents feature vectors' initial population or crow algorithm's initial population. F_i represents feature vector. N represents size of population of feature vectors that are related to corona-virus.

3.11 OBJECTIVE FUNCTION FOR THE FEATURE SELECTION

The chosen features may be taken as inputs to multi-layer ANN. The generated error amount may be utilized as one of the criteria for the evaluation of feature vectors. Any feature vector which could lead to less error for the prediction and classification of corona-virus patients and number of the chosen traits is minimum is then more optimum. The suitable objective function based on Eq. (3.5) needs minimization of corona-virus patients' mean classification error and minimization of the number of the chosen characteristics:

$$Cost(F^i) = \begin{cases} \min mse = \frac{1}{m} \cdot \sum_{i=1}^{m} (\bar{Y}_i - Y_i)^2 \\ \min FS = \frac{F}{D} \end{cases} \quad (3.5)$$

In the objective function, Y_i equals actual class value of the sample concerning being sick or otherwise. $\bar{Y}_i$ represents predicted sample value concerning corona-virus disease, m represents number of the testing or training samples. F represents number of the features that have been chosen and D represents total number of the possible COVID-19 features. The function of

evaluation has 2 goals and several researches attempts at presenting uni-modal function which can be easier to solve. Objective function may be expressed as eq3.6:

$$Cost = \alpha \times \frac{1}{m}\sum_{i=1}^{m}(\bar{Y}_i - Y_i)^2 + \beta \times FS \quad (3.6)$$

Here, α & β represent 2 random numbers between 0 and 1, and any of the property vectors which minimizes it is more optimum.

3.12 UPDATING FEATURE VECTORS

Based on diagram in Fig. 3.4, every one of the feature vectors is updated with the use of Crow learning approach, and those steps can be summarized below:

The parameters of implementation of ANN like the number of the layers, number of feature vectors, number of the hidden neurons in every one of the layers, and maximal number of the crow algorithm iterations have been initialized.

Data that is associated with COVID-19 are pre-processed then normalized.

A random and initial population of the feature vectors have been generated for the detection of COVID-19 disease as crow learning algorithm population.

The NN input is reduced by every feature vector. Multi-layer ANN is training with every one of the feature vectors.

Every one of the attribute vectors utilizes a mean value of the errors of corona-virus diagnoses and number of the features that have been chosen for assessment.

The 2 optimum feature vectors in a population have been considered to be the 2 population parents and the remaining feature vectors are young crows.

The phases of individual, reinforcement and social learning processes are carried out for updating feature vectors at every one of the stages.

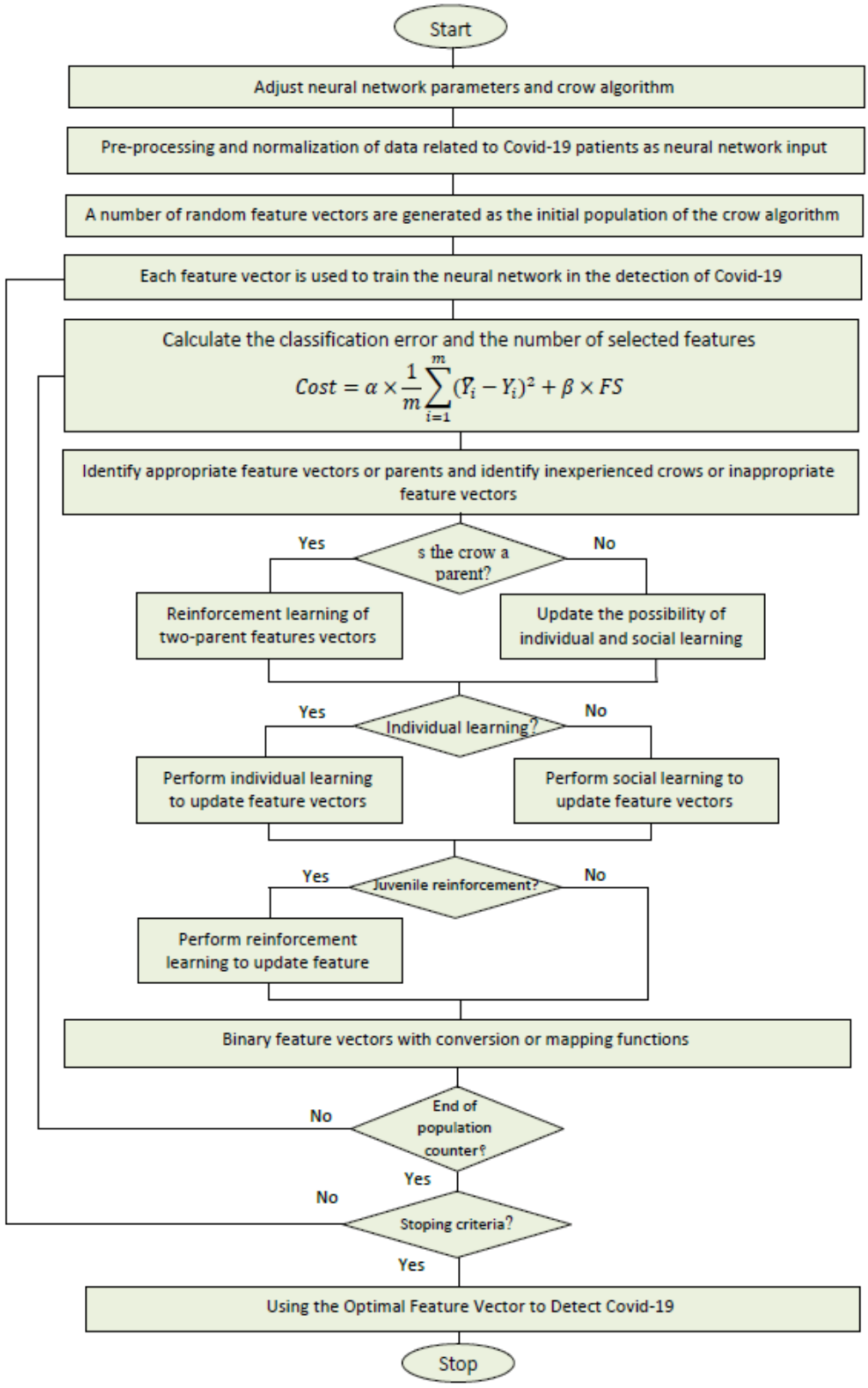

Figure 3. 4: Flow-chart of suggested approach for the detection of COVID-19 disease

Feature vectors are re-binary at every one of the steps utilizing functions of conversion, like V and S.

In every one of the iterations, all of the attribute vectors have been updated.

In every one of the iterations, a unit is added to the proposed algorithm's iteration counter and the algorithm is repeated under the condition that this counter is less than maximal iteration value.

In final iteration, optimum feature vector has been utilized for the training of multi-layer ANN in diagnosing patients with COVID-19.

3.13 BINARIZATION OF THE FEATURE VECTORS

After application of crow learning equations for updating feature vectors, the values that have been updated may be binary again. Binary representation returns feature vector values to 0 and one so that it may be utilized for selecting a feature. In order to choose feature and binary, 2 functions of S and Vmay be utilized , correspondingly (3.7) & (3.8):

$$T\left(F_i^k(t)\right) = \frac{1}{1+e^{-F_i^k(t)}} \tag{3.7}$$

$$t\left(F_i^k(t)\right) = \left|\frac{2}{\pi}\arctan\left(\frac{2}{\pi}F_i^k(t)\right)\right| \tag{3.8}$$

The reason behind the use of those 2 functions is that the range of every one of them is from 0 to 1. Those functions could map the non-binary values from 0 to 1 and after that, binary property vectors based on Eq. (3.9) for Gaussian function and based on Eq. (3.10) for V:

$$F_i = \begin{cases} 1 & sig \geq rand \\ 0 & otherwise \end{cases} \tag{3.9}$$

$$F_i = \begin{cases} 1 & rand \leq T(F_i) \\ 0 & rand > T(F_i) \end{cases} \tag{3.10}$$

In the diagram in Figure 3.5, the S and V functions of conversion have been plotted, and roles of those 2 functions is the binarization of feature vectors:

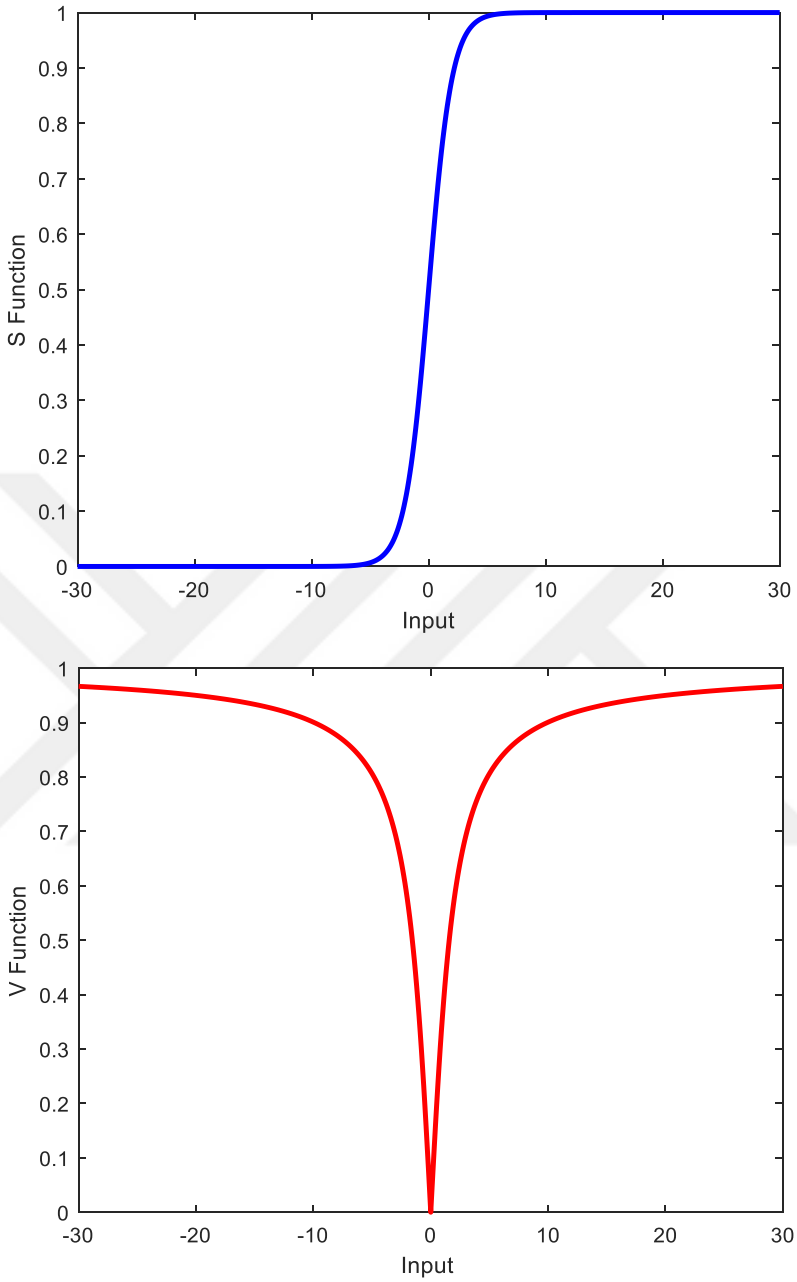

Figure 3. 5: Left to right S and V conversion functions, respectively

4. EXPERIMENTAL RESULTS AND DISCUSSION

4.1 RESULTS

In this chapter, implementation and results of suggested method's analysis in diagnosing COVID19 will be discussed. The implementations and experimentations have been are conducted on Matlab program v2019.

4.2 DATA SET

The data-set that has been used in the present study includes lab results of the patients in a São Paulo hospital in Brazil [92]. The data-set includes samples that have been obtained from SARS-CoV-2 patients in the first months of 2020. This dataset includes data that has been gathered from 5,644 patients. Approximately 10% of subjects and samples that have been infected by COVID-19, and of those 10%, 2.50% to 6.50% needed being admitted to an intensive care in the hospital. There has not been any gender information in this data-set. Based on this study, 19 significant and basic features were utilized for the diagnosis of COVID-19. Feature #20 represents the type of the output, indicating the type of label the subject is virus infection.

4.3 CRITERIA OF EVALUATION

The criteria of sensitivity, accuracy, F1-score and precision have been utilized for the evaluation of the suggested approach and comparing it to approaches of COVID-19 diagnosis. The accuracy percentage in best and worst cases have been respectively 0 and 100, and the proximity to the value of 100 has been an indication of proper accuracy of the algorithm of classification. Eqs. (4.1), (4.2), (4.3), & (4.4) showed the calculations of the accuracy, sensitivity, precision, and F1 score, respectively [93][94][95]:

$$Accuracy = \frac{(TP+TN)}{(TP+FP+TN+FN)} \quad (4.1)$$

$$Recall = \frac{TP}{(TP+FN)} \quad (4.2)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (4.3)$$

$$F1 - Score = 2.\frac{Precision.Recall}{(Precision+Recall)} \quad (4.4)$$

TP index represents the subjects that had the disease diagnosed correctly. TN represents the healthy individuals who have been correctly diagnosed as healthy. FP represents the individuals who have been misidentified as patients. FN represents the individuals that were mistakenly diagnosed as healthy.

4.4 PARAMETERS

Parameters that have been listed in Table4.1 have been utilized for the implementation of the suggested approach. Some parameters have been set on the basis of crow algorithm. NN -related parameters have been estimated through the trial and error. A 2-layer ANN that includes 20 neurons in every one of the layers has been utilized for implementing it. In the implementations, 70% of samples were specified for the training and the remaining 30% have been set aside for the testing. The size of the population was 20 and number of the iterations of crow learning method has been 50. Every one of the tests has been repeated for 30 times and mean value of the indicators like error and accuracy has been computed and compared to similar approaches in COVID-19 diagnosis.

Table 4.1: Proposed Method Parameters

Parameters	Values
Number of the of crows in the family	20
Max. number of the Iterations	50
Probability of Reinforcement	0.90
Probability of Social Learning	0.990

Table 4.2: Proposed Method Parameters " Table continued "

Probability of Vertical Learning	0.990
Probability of 1st Parent Selection	0.950
Probability of Trial and Error	0.30
Lf-min	0.00050
Lf-max	0.020
ANN Layers	2
Neurons of the Hidden Layers	20

4.5 SAMPLE OUTPUTS

For the analysis and evaluation of the suggested approach, the objective function value of the selection of the feature and accuracy of the classification of patients from healthy persons may be shown in output based on crow algorithm repetition. Figure 4.1 illustrates objective function of selection of feature and value of accuracy of classification index of suggested method for the diagnosis of patients who have been infected with COVID-19 in MATLAB:

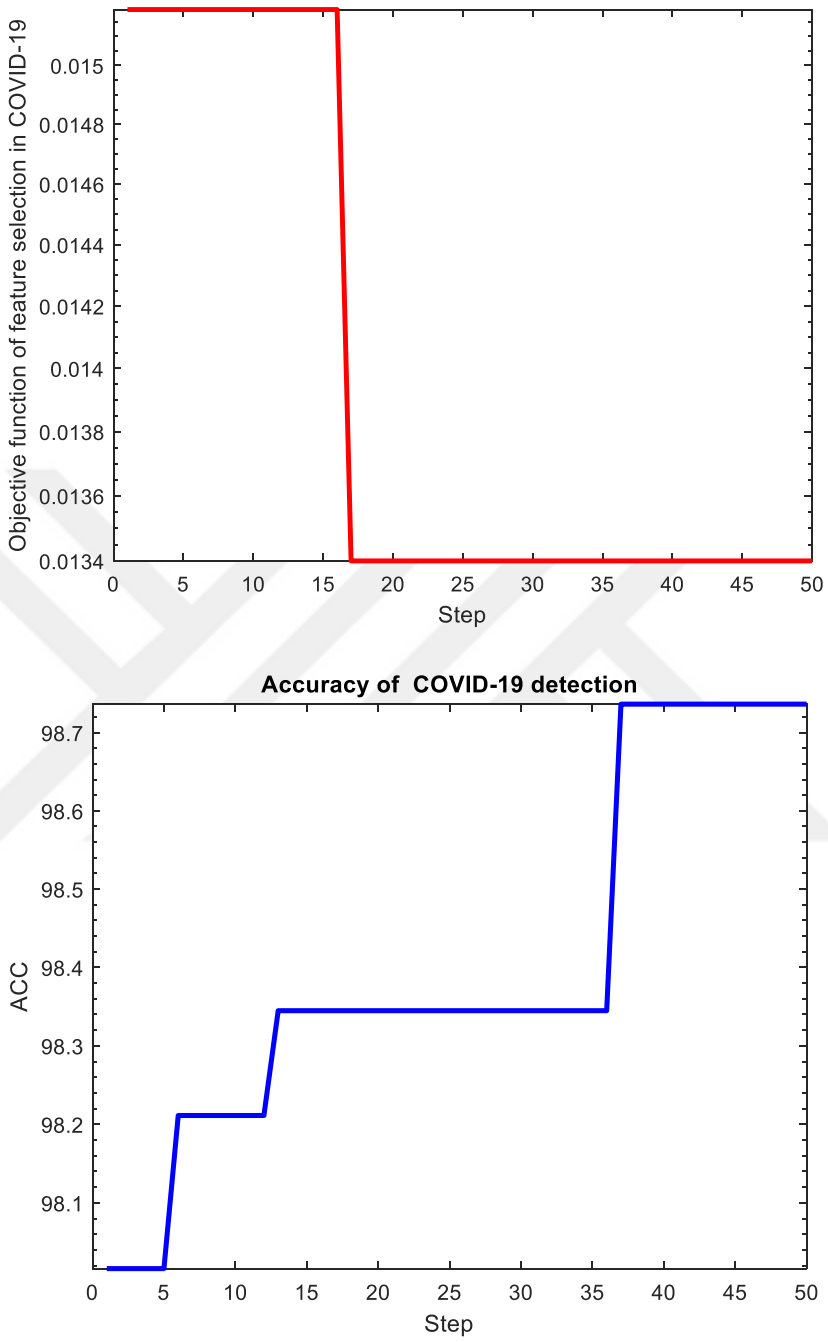

Figure 4. 1: From left to right, objective function of selection of features and accuracy of COVID-19 diagnosis

The objective function's value of the property selection has had a descending tendency concerning crow learning algorithm's iteration. The objective function of the feature selection has 2 aims, which are: the reduction of the classification errors as well as reduction of the number of the chosen

attributes. Through the reduction of objective function, the feature selection in the general cases, those two goals have been reduced as well. The decreased feature function, feature selection in diagnosing corona-virus is due to diagnostic error as well as number of the features is reducing. The number of the features that have been chosen was reduced as well by implementing crow learning method. The accuracy of classification of people who have been infected by corona-virus and the healthy individuals was increased according to the frequency, because of the two factors of the NN training and feature selection.

4.6 RESULTS

For the implementation of the suggested approach, total of 18 lab results from 600 patients have been taken under consideration for the prediction of COVID-19 infections. All of the samples have been obtained from lab results of the patients in a hospital in Brazil. The experimentations have shown that the suggested approach's accuracy for the diagnoses of COVID-19 patients has been 94.31% and its precision, sensitivity, and F1-score have been respectively 94.38%, 94.15%, and 94.27%. The sensitivity, accuracy, F1-score and precision of the multilayer ANN without crow learning approach have been computed as well to be 88.24%, 88.67%, 88.44% and 87.96%, respectively. comparison of suggested approach with the feature selection and multi-layer ANN with no feature selection has been depicted in Figure 4.2 [26]:

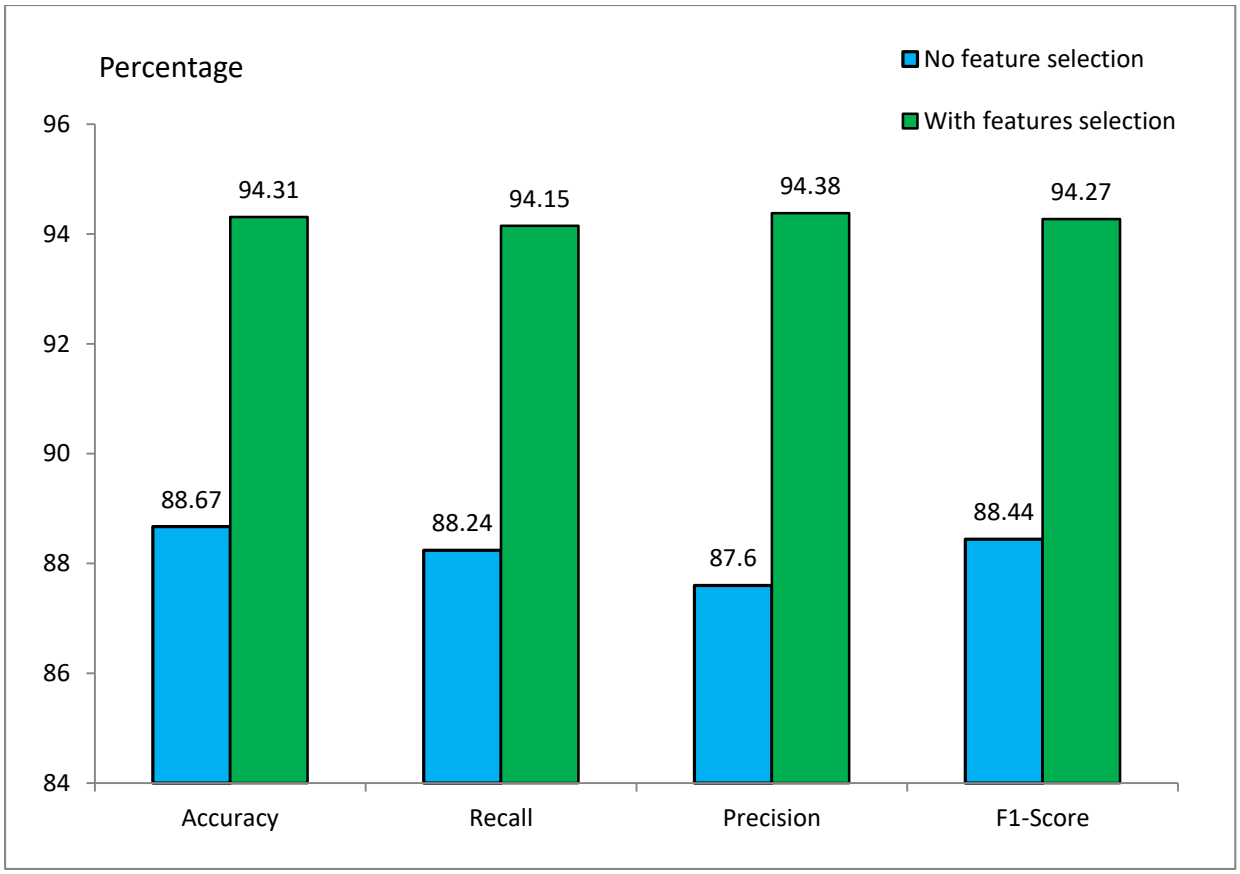

Figure 4. 2: Comparisons of the sensitivity, accuracy, F1-score and precision of suggested approach with the approach with no feature selection

The experimentations have shown that COVID-19 patient classification accuracy in the suggested approach is 94.31%. The experimentations have shown that in the case of not using feature selection, its accuracy has been 88.67%. Utilizing the feature selection using crow algorithm could result in increasing the multi-layer ANN accuracy by approximately 5.64%. The suggested approach accomplished better performance in comparison with ANN in the 4 indices without feature selection. The suggested approach had enhanced over 5.91% in the index of sensitivity, approximately 5.83% in F1-score, and approximately 6.78% in precision index in comparison with the approach without feature selection. The suggested approach has been compared to 6 DL approaches as the classifiers. Results have been compared in Table4.2, on the basis of sensitivity, accuracy, F1-score and precision.

Table4. 3: Comparison of accuracy of suggested approach with other algorithms

Methods	Accuracy	F1-Score	Precision	Recall
ANN	0.86900	0.87130	0.87130	**0.87130**
CNN	0.87350	0.88560	0.88470	**0.88670**
CNNRNN	0.86240	0.87550	0.87550	**0.87550**
CNNLSTM	0.92300	0.93000	0.92350	**0.93680**
RNN	0.84000	0.84270	0.84280	**0.84270**
LSTM	0.90340	0.89970	0.89970	**0.89980**
Proposed	0.94310	0.94270	0.94380	**0.94150**

Figure 4.3 presents comparison of accuracy, F1-score and sensitivity of CNN, ANN, CNNLSTM, LSTM, CNNRNN, RNN learning approaches, and the suggested approach in COVID-19.

The comparisons have shown that the suggested method showed higher efficiency in every one of indicators compared to DL approaches. The proposed method's analysis in diagnosing COVID-19 infection has shown that it's of a higher accuracy in terms of sensitivity, accuracy, precision, and F1 when compared to the CNN, ANN, CNNRNN, CNNLSTM, RNN and LSTM algorithms. The suggested approach has higher values concerning percentage in those indicators in the classification of the patients from the healthy subjects. In the index of accuracy, the suggested approach had a 94.31% accuracy. In the 2nd place, in terms of the index of accuracy is CNN-LSTM approach with 92.3% accuracy. The least efficient performance in the index of accuracy for diagnosing COVID-19 is associated with RNN approach, an approximately 84% accuracy.

In F1-score, the suggested approach's accuracy has been superior to other approaches. In 2nd rank in terms of the F1-score is CNN-LSTM approach and its value has been 93%. The least efficient performance in F1-score has been associated with the RNN, where the value has been 84.27%.

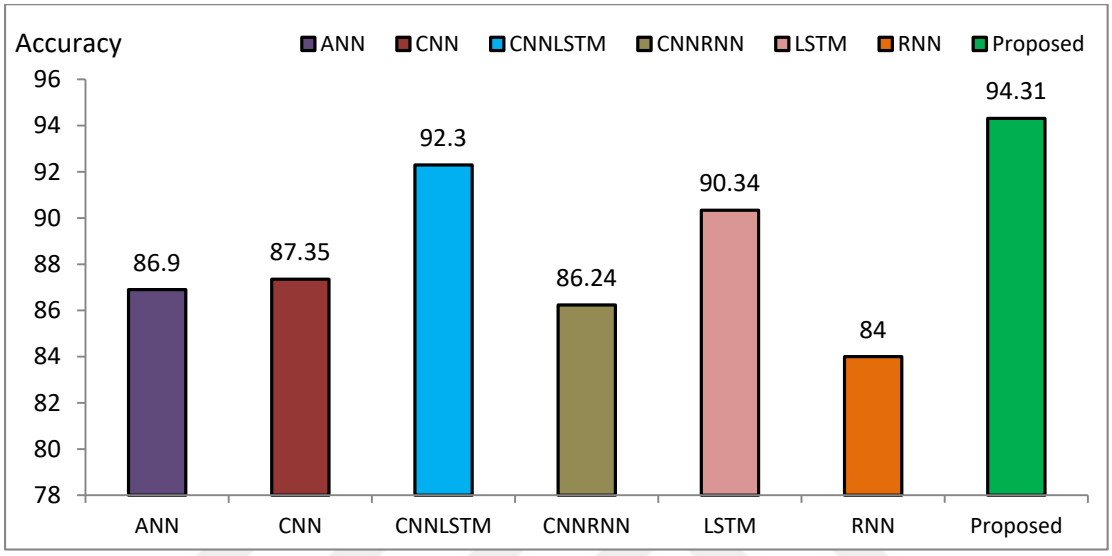

Figure 4. 3: Comparisons of the accuracy of suggested approaches with learning algorithms

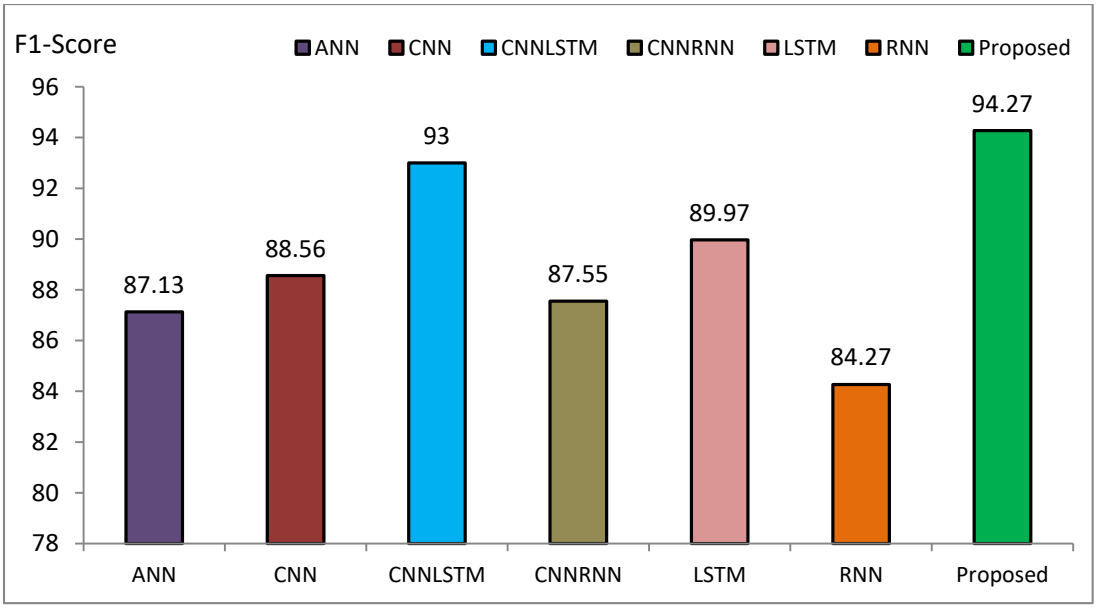

Figure 4. 4: Comparisons of the F1-score of suggested approaches with learning algorithms

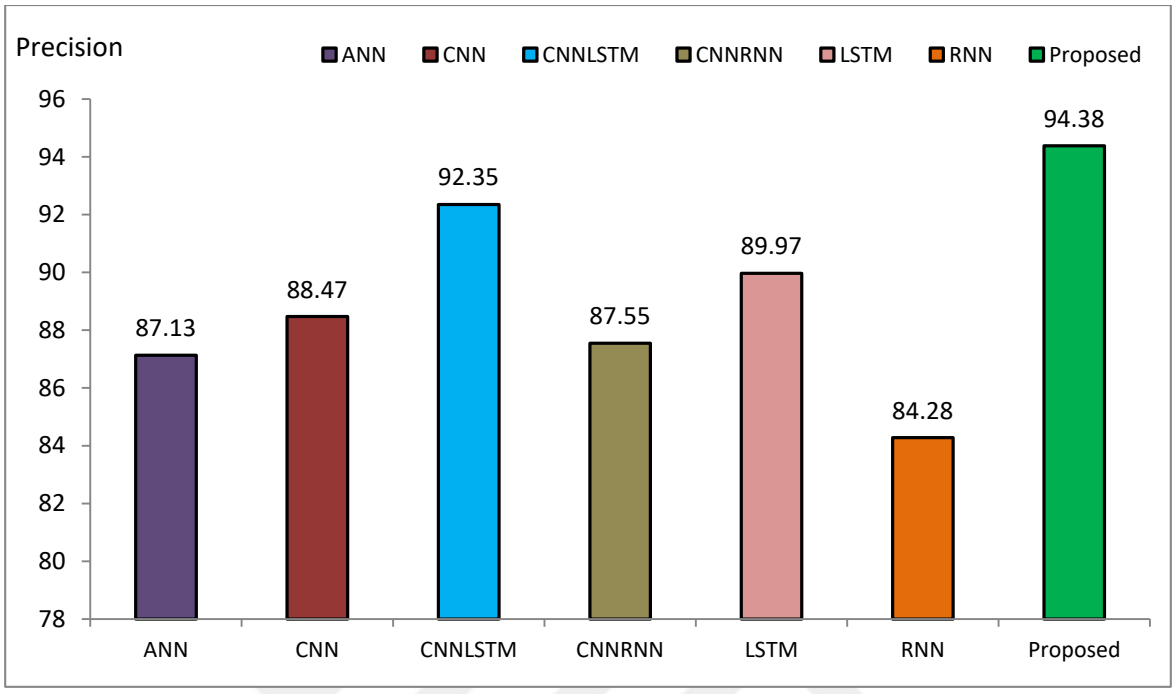

Figure 4. 5: Comparisons of the precision of suggested approaches with learning algorithms

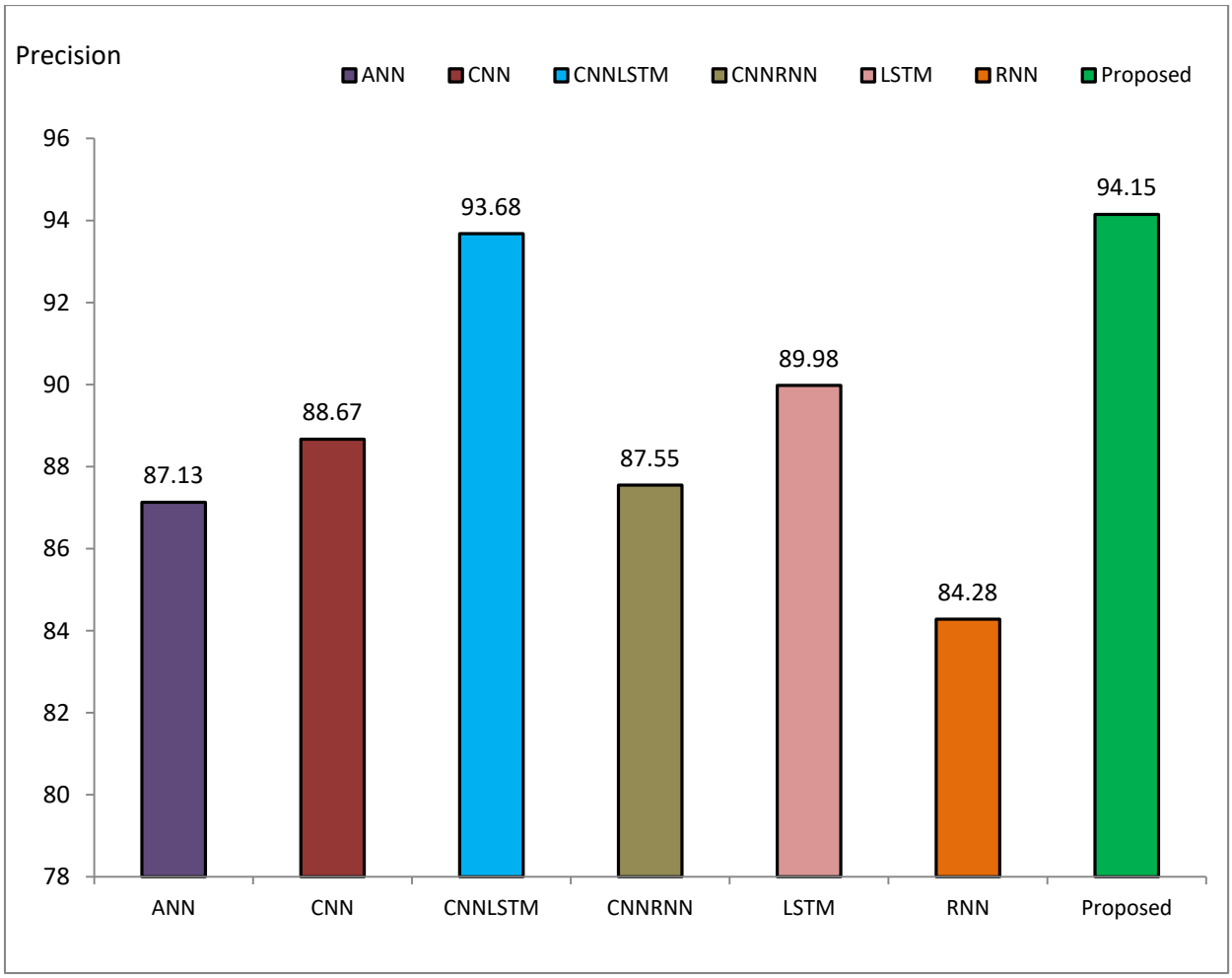

Figure 4. 6: Comparisons of the precision of suggested approaches with learning algorithms

In Figure 4.4, the index of the diagnostic accuracy of corona-virus patients in the suggested approach has been compared to 2 researches in Sao Paulo and Wenzhou hospitals in Brazil with 2 learning approaches, which are: CNNLSTM and SVM. The suggested method's accuracy has been 94.31%. The accuracy of patient diagnosis with the SVM in a Chinese hospital has been equal to 80%.

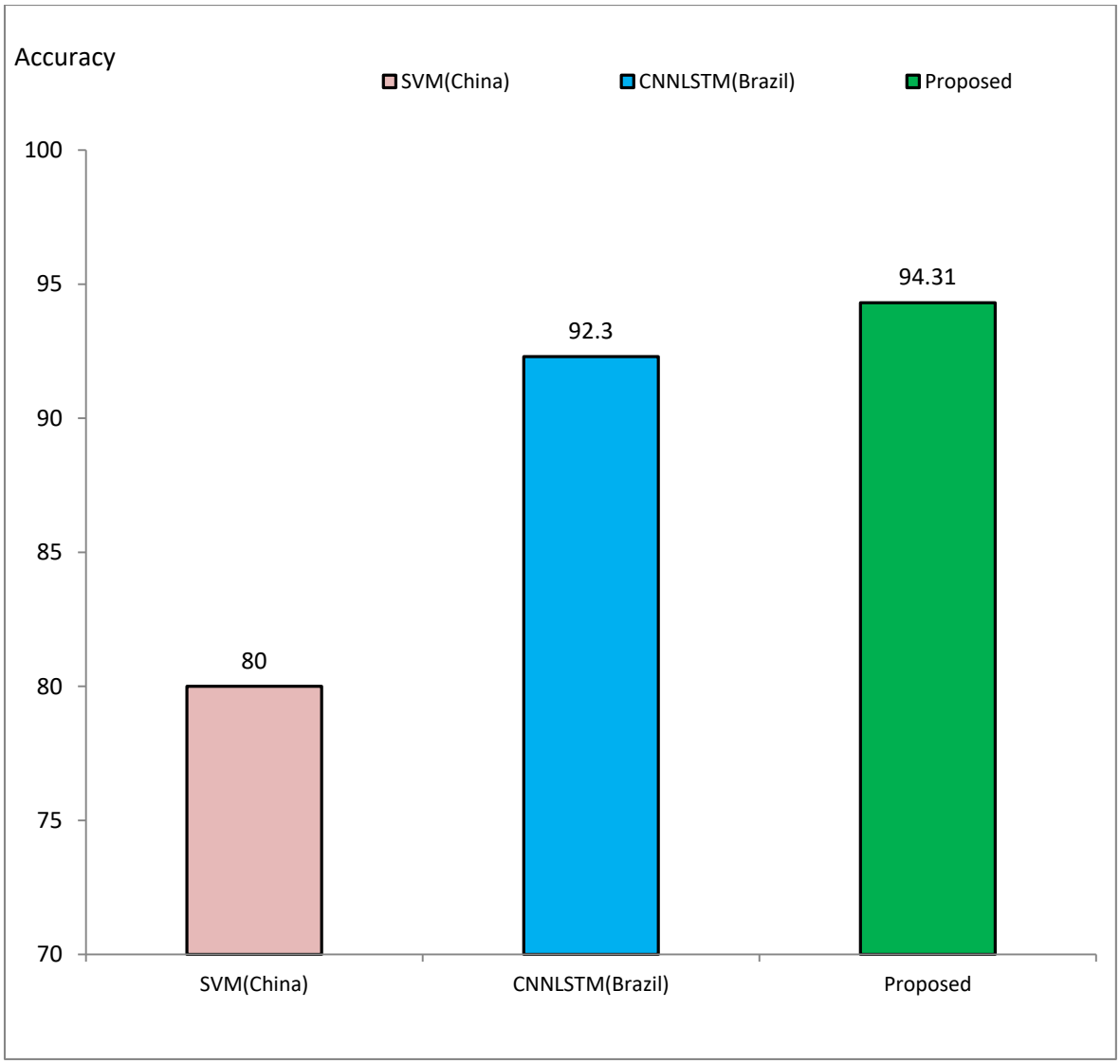

Figure 4. 7: Comparisons of accuracy of COVID-19 diagnosis in suggested approach with the ones that had been studied in Brazil and China

Concerning the accuracy of performance, the suggested approach is more efficient compared to other algorithms, like the CNN, CNNRNN, CNNLSTM, RNN and LSTM, then it's in 2nd place of the CNN-LSTM approach. The least efficient performance has been associated with RNN network with an 84.28% accuracy. In the index of sensitivity, the suggested approach had the optimal performance, followed by the CNNLSTM. For the purpose of more efficiently evaluating the suggested approach, the research results have been compared to 2 researches in hospitals of

Cangnan People's Hospital and Wenzhou Central Hospital in Wenzhou, China, and Sao Paulo, Brazil, concerning the index of accuracy. The experimentations have shown that CNN-LSTM learning diagnosis accuracy in the Brazilian hospital has been 92.3%. The suggested approach has been of a higher accuracy in the diagnosis of COVID-19 from 2 researches that have been conducted in Brazil and China based upon the data of the patients.

5. CONCLUSIONS AND FUTURE WORK

5.1 CONCLUSIONS

In the present work, multi-layer NN and feature selection have been utilized for increasing COVID-19 diagnosis effectiveness. In the suggested approach , a group of the characteristics that are associated with COVID-19 patients has been characterized as input to AI network, and important characteristics have been chosen by crow learning algorithm then utilized for the learning. The suggested approach for the diagnosis of COVID-19 has 2 basic layers. In 1st layer, the process of feature selection is carried out with crow algorithm's binary version. In 2nd layer, the patients have been classified into healthy persons utilizing multi-layer ANN and optimum feature vector. The suggested approach has been utilized as tool of classification in diagnosing COVID-19. The implementations in this work have been carried out on the data that has been gathered at the Sao Paulo Hospital with the use of MATLAB program. The experimentations have shown that the suggested approach results in reducing the function of feature selection in diagnosing COVID-19 through the replication of crow learning approach. Objective function reduction results from feature selection through crow learning algorithm. Through the characteristic of feature selection, amount of COVID-19 patients' diagnostic error had reduced based on crow learning algorithm iteration as well as the accuracy amount is an increment. The experimentations have shown that the suggested approach is more efficient in index of sensitivity, accuracy, accuracy compared to COVID-19 diagnosis methods. The feature selection makes ML to focus more upon the essential characteristics, resulting in the reduction of risks of the classification of infected individuals from the healthy ones. The experimentations have shown that CNNLSTM approach has high accuracy for the diagnosis of corona-virus patients and is in 2nd place in comparison with the suggested approach. In feature work, a data fusion of the computed tomography scans with patient records will be utilized for accurate classification of COVID-19. For increasing the suggested approach's accuracy in the detection of corona-virus, a combination of swarm intelligence (SI) algorithms in feature selection and DL phases based upon the CNN-LSTM will be utilized.

5.2 FUTURE WORK

For Future work we can work on other meta-heuristic optimization method such as gray wolf optimization method.

REFERENCES

[1] S. Lalmuanawma, J. Hussain, and L. Chhakchhuak, "Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review," *Chaos, Solitons & Fractals*, vol. 139, p. 110059, 2020.

[2] F. Stephany, N. Stoehr, P. Darius, L. Neuhäuser, O. Teutloff, and F. Braesemann, "The CoRisk-Index: A data-mining approach to identify industry-specific risk assessments related to COVID-19 in real-time," *arXiv Prepr. arXiv2003.12432*, 2020.

[3] W. H. Organization, "Coronavirus disease 2019 (COVID-19): situation report, 73," 2020.

[4] A. S. Albahri *et al.*, "Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): a systematic review," *J. Med. Syst.*, vol. 44, pp. 1–11, 2020.

[5] N. Tang, D. Li, X. Wang, and Z. Sun, "Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia," *J. Thromb. Haemost.*, vol. 18, no. 4, pp. 844–847, 2020.

[6] N. Fernandes, "Economic effects of coronavirus outbreak (COVID-19) on the world economy," *Available SSRN 3557504*, 2020.

[7] X. Xu *et al.*, "A deep learning system to screen novel coronavirus disease 2019 pneumonia," *Engineering*, vol. 6, no. 10, pp. 1122–1129, 2020.

[8] M. Yadav, M. Perumal, and M. Srinivas, "Analysis on novel coronavirus (COVID-19) using machine learning methods," *Chaos, Solitons & Fractals*, vol. 139, p. 110050, 2020.

[9] U. Özkaya, Ş. Öztürk, and M. Barstugan, "Coronavirus (covid-19) classification using deep features fusion and ranking technique," in *Big Data Analytics and Artificial Intelligence Against COVID-19: Innovation Vision and Approach*, Springer, 2020, pp. 281–295.

[10] A. Mollalo, K. M. Rivera, and B. Vahedi, "Artificial neural network modeling of novel coronavirus (COVID-19) incidence rates across the continental United States," *Int. J. Environ. Res. Public Health*, vol. 17, no. 12, p. 4204, 2020.

[11] P. K. Sethy and S. K. Behera, “Detection of coronavirus disease (covid-19) based on deep features,” 2020.

[12] O. Gozes, M. Frid-Adar, N. Sagie, H. Zhang, W. Ji, and H. Greenspan, “Coronavirus detection and analysis on chest ct with deep learning,” *arXiv Prepr. arXiv2004.02640*, 2020.

[13] C. M. Yeşilkanat, “Spatio-temporal estimation of the daily cases of COVID-19 in worldwide using random forest machine learning algorithm,” *Chaos, Solitons & Fractals*, vol. 140, p. 110210, 2020.

[14] D. Dong *et al.*, “The role of imaging in the detection and management of COVID-19: a review,” *IEEE Rev. Biomed. Eng.*, 2020.

[15] W. Al-Sorori and A. M. Mohsen, “New Caledonian crow learning algorithm: A new metaheuristic algorithm for solving continuous optimization problems,” *Appl. Soft Comput.*, p. 106325, 2020.

[16] M. M. Ahsan, K. D. Gupta, M. M. Islam, S. Sen, M. Rahman, and M. S. Hossain, “Study of different deep learning approach with explainable ai for screening patients with COVID-19 symptoms: Using ct scan and chest x-ray image dataset,” *arXiv Prepr. arXiv2007.12525*, 2020.

[17] A. Ulhaq, J. Born, A. Khan, D. P. S. Gomes, S. Chakraborty, and M. Paul, “Covid-19 control by computer vision approaches: A survey,” *IEEE Access*, vol. 8, pp. 179437–179456, 2020.

[18] D. Debone, M. V Da Costa, and S. G. E. K. Miraglia, “90 days of COVID-19 social distancing and its impacts on air quality and health in Sao Paulo, Brazil,” *Sustainability*, vol. 12, no. 18, p. 7440, 2020.

[19] K. H. Gohel, P. B. Patel, P. M. Shah, J. R. Patel, N. Pandit, and A. Raut, “Knowledge and perceptions about COVID-19 among the medical and allied health science students in India: an online cross-sectional survey,” *Clin. Epidemiol. Glob. Heal.*, vol. 9, pp. 104–109, 2021.

[20] D. Cucinotta and M. Vanelli, “WHO declares COVID-19 a pandemic,” *Acta Bio Medica Atenei Parm.*, vol. 91, no. 1, p. 157, 2020.

[21] M. Roche, “COVID-19 and Media datasets: Period-and location-specific textual data mining,” *Data Br.*, vol. 33, p. 106356, 2020.

[22] P. Radanliev, D. De Roure, and R. Walton, “Data mining and analysis of scientific research data records on Covid-19 mortality, immunity, and vaccine development-In the first wave of the Covid-19 pandemic,” *Diabetes Metab. Syndr. Clin. Res. Rev.*, vol. 14, no. 5, pp. 1121–1132, 2020.

[23] K. Baralić *et al.*, “Safety assessment of drug combinations used in COVID-19 treatment: in silico toxicogenomic data-mining approach,” *Toxicol. Appl. Pharmacol.*, vol. 406, p. 115237, 2020.

[24] N. Hasan, “A methodological approach for predicting COVID-19 epidemic using EEMD-ANN hybrid model,” *Internet of things*, vol. 11, p. 100228, 2020.

[25] J. Farooq and M. A. Bazaz, “A deep learning algorithm for modeling and forecasting of COVID-19 in five worst affected states of India,” *Alexandria Eng. J.*, vol. 60, no. 1, pp. 587–596, 2021.

[26] T. B. Alakus and I. Turkoglu, “Comparison of deep learning approaches to predict COVID-19 infection,” *Chaos, Solitons & Fractals*, vol. 140, p. 110120, 2020.

[27] E.-S. M. El-Kenawy, A. Ibrahim, S. Mirjalili, M. M. Eid, and S. E. Hussein, “Novel feature selection and voting classifier algorithms for COVID-19 classification in CT images,” *IEEE Access*, vol. 8, pp. 179317–179335, 2020.

[28] R. Sethi, M. Mehrotra, and D. Sethi, “Deep learning based diagnosis recommendation for COVID-19 using chest x-rays images,” in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2020, pp. 1–4.

[29] M. Z. Alam, M. S. Rahman, and M. S. Rahman, “A Random Forest based predictor for medical data classification using feature ranking,” *Informatics Med. Unlocked*, vol. 15, p. 100180, 2019.

[30] M. Wang and H. Chen, “Chaotic multi-swarm whale optimizer boosted support vector machine for medical diagnosis,” *Appl. Soft Comput.*, vol. 88, p. 105946, 2020.

[31] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, "Segmentation, feature extraction, and multiclass brain tumor classification," *J. Digit. Imaging*, vol. 26, no. 6, pp. 1141–1150, 2013.

[32] J. Cheng *et al.*, "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PLoS One*, vol. 10, no. 10, p. e0140381, 2015.

[33] M. R. Ismael and I. Abdel-Qader, "Brain tumor classification via statistical features and back-propagation neural network," in *2018 IEEE international conference on electro/information technology (EIT)*, 2018, pp. 252–257.

[34] P. Thejaswini, M. B. Bhat, and M. K. Prakash, "Detection and classification of tumour in brain MRI," *Int. J. Eng. Manufact.(IJEM)*, vol. 9, no. 1, pp. 11–20, 2019.

[35] S. Deepak and P. M. Ameer, "Brain tumor classification using deep CNN features via transfer learning," *Comput. Biol. Med.*, vol. 111, p. 103345, 2019.

[36] M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, and S. W. Baik, "Multi-grade brain tumor classification using deep CNN with extensive data augmentation," *J. Comput. Sci.*, vol. 30, pp. 174–182, 2019.

[37] P. Afshar, A. Mohammadi, and K. N. Plataniotis, "Brain tumor type classification via capsule networks," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 3129–3133.

[38] O. N. Belaid and M. Loudini, "Classification of Brain Tumor by Combination of Pre-Trained VGG16 CNN," *J. Inf. Technol. Manag.*, vol. 12, no. 2, pp. 13–25, 2020.

[39] N. Ghassemi, A. Shoeibi, and M. Rouhani, "Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images," *Biomed. Signal Process. Control*, vol. 57, p. 101678, 2020.

[40] D. Li, Z. Fu, and J. Xu, "Stacked-autoencoder-based model for COVID-19 diagnosis on CT images," *Appl. Intell.*, vol. 51, no. 5, pp. 2805–2817, 2021.

[41] F. Khoshbakhtian, A. B. Ashraf, and S. S. Khan, "Covidomaly: A deep convolutional

autoencoder approach for detecting early cases of covid-19," *arXiv Prepr. arXiv2010.02814*, 2020.

[42] F. Khozeimeh *et al.*, "CNN AE: Convolution Neural Network combined with Autoencoder approach to detect survival chance of COVID 19 patients," *arXiv Prepr. arXiv2104.08954*, 2021.

[43] A. Gumaei, M. M. Hassan, M. R. Hassan, A. Alelaiwi, and G. Fortino, "A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification," *IEEE Access*, vol. 7, pp. 36266–36273, 2019.

[44] K. Kaplan, Y. Kaya, M. Kuncan, and H. M. Ertunç, "Brain tumor classification using modified local binary patterns (LBP) feature extraction methods," *Med. Hypotheses*, vol. 139, p. 109696, 2020.

[45] H. Kutlu and E. Avcı, "A novel method for classifying liver and brain tumors using convolutional neural networks, discrete wavelet transform and long short-term memory networks," *Sensors*, vol. 19, no. 9, p. 1992, 2019.

[46] Z. N. K. Swati *et al.*, "Brain tumor classification for MR images using transfer learning and fine-tuning," *Comput. Med. Imaging Graph.*, vol. 75, pp. 34–46, 2019.

[47] R. Hashemzehi, S. J. S. Mahdavi, M. Kheirabadi, and S. R. Kamel, "Detection of brain tumors from MRI images base on deep learning using hybrid model CNN and NADE," *Biocybern. Biomed. Eng.*, vol. 40, no. 3, pp. 1225–1232, 2020.

[48] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE Comput. Intell. Mag.*, vol. 1, no. 4, pp. 28–39, 2006.

[49] 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, 2020.

[50] X. Li and D. Zhu, "Covid-xpert: An ai powered population screening of covid-19 cases using chest radiography images," *arXiv Prepr. arXiv2004.03042*, 2020.

[51] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K. N. Plataniotis, and A. Mohammadi, “Covid-caps: A capsule network-based framework for identification of covid-19 cases from x-ray images,” *Pattern Recognit. Lett.*, vol. 138, pp. 638–643, 2020.

[52] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, “Automated detection of COVID-19 cases using deep neural networks with X-ray images,” *Comput. Biol. Med.*, vol. 121, p. 103792, 2020.

[53] A. A. Ardakani, A. R. Kanafi, U. R. Acharya, N. Khadem, and A. Mohammadi, “Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks,” *Comput. Biol. Med.*, vol. 121, p. 103795, 2020.

[54] M. Dorigo and K. Socha, “An introduction to ant colony optimization,” in *Handbook of Approximation Algorithms and Metaheuristics, Second Edition*, Chapman and Hall/CRC, 2018, pp. 395–408.

[55] M. Dorigo and C. Blum, “Ant colony optimization theory: A survey,” *Theor. Comput. Sci.*, vol. 344, no. 2–3, pp. 243–278, 2005.

[56] H. R. Kanan, K. Faez, and S. M. Taheri, “Feature selection using ant colony optimization (ACO): a new method and comparative study in the application of face recognition system,” in *Industrial conference on data mining*, 2007, pp. 63–76.

[57] B. Kröse, B. Krose, P. van der Smagt, and P. Smagt, “An introduction to neural networks,” 1993.

[58] K. Gurney, *An introduction to neural networks*. CRC press, 2018.

[59] J. Fernández-Montoya, C. Avendaño, and P. Negredo, “The glutamatergic system in primary somatosensory neurons and its involvement in sensory input-dependent plasticity,” *Int. J. Mol. Sci.*, vol. 19, no. 1, p. 69, 2018.

[60] S. K. Esser, R. Appuswamy, P. Merolla, J. V Arthur, and D. S. Modha, “Backpropagation for energy-efficient neuromorphic computing,” *Adv. Neural Inf. Process. Syst.*, vol. 28, pp. 1117–1125, 2015.

[61] L. Wan, M. Zeiler, S. Zhang, Y. Le Cun, and R. Fergus, "Regularization of neural networks using dropconnect," in *International conference on machine learning*, 2013, pp. 1058–1066.

[62] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," *arXiv Prepr. arXiv1505.00853*, 2015.

[63] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2016.

[64] M. Turčaník and M. Javurek, "Hash function generation by neural network," in *2016 New Trends in Signal Processing (NTSP)*, 2016, pp. 1–5.

[65] M. Turčaník, "Hash function generation based on neural networks and chaotic maps," in *2017 Communication and Information Technologies (KIT)*, 2017, pp. 1–5.

[66] N. Abdoun, S. El Assad, R. Assaf, O. Déforges, M. Khalil, and S. Belghith, "Design and implementation of robust Keyed Hash functions based on Chaotic Neural Network," 2018.

[67] B. Karlik and A. V. Olgac, "Performance analysis of various activation functions in generalized MLP architectures of neural networks," *Int. J. Artif. Intell. Expert Syst.*, vol. 1, no. 4, pp. 111–122, 2011.

[68] M. Sheinfeld, S. Levinson, and I. Orion, "Highly accurate prediction of specific activity using deep learning," *Appl. Radiat. Isot.*, vol. 130, pp. 115–120, 2017.

[69] A. A. M. Nureddin, J. Rahebi, and A. Ab-BelKhair, "Power Management Controller for Microgrid Integration of Hybrid PV/Fuel Cell System Based on Artificial Deep Neural Network," *Int. J. Photoenergy*, vol. 2020, 2020.

[70] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*, vol. 1, no. 2. MIT press Cambridge, 2016.

[71] S. Ahmed, M. Frikha, T. D. H. Hussein, and J. Rahebi, "Optimum Feature Selection with Particle Swarm Optimization to Face Recognition System Using Gabor Wavelet Transform

and Deep Learning," *Biomed Res. Int.*, vol. 2021, 2021.

[72] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[73] A. Ab-BelKhair, J. Rahebi, and A. Abdulhamed Mohamed Nureddin, "A Study of Deep Neural Network Controller-Based Power Quality Improvement of Hybrid PV/Wind Systems by Using Smart Inverter," *Int. J. Photoenergy*, vol. 2020, 2020.

[74] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv Prepr. arXiv1412.6980*, 2014.

[75] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 249–256.

[76] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[77] H.-C. Shin *et al.*, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.

[78] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[79] L. Tian and Z. Mu, "Ear recognition based on deep convolutional network," in *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2016, pp. 437–441.

[80] A. Giusti, D. C. Cireşan, J. Masci, L. M. Gambardella, and J. Schmidhuber, "Fast image scanning with deep max-pooling convolutional neural networks," in *2013 IEEE International Conference on Image Processing*, 2013, pp. 4034–4038.

[81] K. R. Konda, “Unsupervised Relational Feature Learning for Vision.” Univ.-Bibliothek Frankfurt am Main, 2016.

[82] I. K. Sethi and A. K. Jain, *Artificial neural networks and statistical pattern recognition: old and new connections*. Elsevier, 2014.

[83] R. Hecht-Nielsen, “Theory of the backpropagation neural network,” in *Neural networks for perception*, Elsevier, 1992, pp. 65–93.

[84] R. Miikkulainen *et al.*, “Evolving deep neural networks,” in *Artificial intelligence in the age of neural networks and brain computing*, Elsevier, 2019, pp. 293–312.

[85] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural networks*, vol. 61, pp. 85–117, 2015.

[86] L. Torrey and J. Shavlik, “Transfer learning,” in *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*, IGI global, 2010, pp. 242–264.

[87] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, “A survey on deep transfer learning,” in *International conference on artificial neural networks*, 2018, pp. 270–279.

[88] H. Ravishankar *et al.*, “Understanding the mechanisms of deep transfer learning for medical images,” in *Deep learning and data labeling for medical applications*, Springer, 2016, pp. 188–196.

[89] X. Yin, X. Yu, K. Sohn, X. Liu, and M. Chandraker, “Feature transfer learning for face recognition with under-represented data,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 5704–5713.

[90] K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and A. Agrawal, “Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection,” *Constr. Build. Mater.*, vol. 157, pp. 322–330, 2017.

[91] Y. Yu, H. Lin, J. Meng, X. Wei, H. Guo, and Z. Zhao, “Deep transfer learning for modality classification of medical images,” *Information*, vol. 8, no. 3, p. 91, 2017.

[92] P. Schwab, A. D. Schütte, B. Dietz, and S. Bauer, “predcovid-19: A systematic study of clinical predictive models for coronavirus disease 2019,” *arXiv Prepr. arXiv2005.08302*, vol. 76, 2020.

[93] A. T. H. Al-Rahlawee and J. Rahebi, “Multilevel thresholding of images with improved Otsu thresholding by black widow optimization algorithm,” *Multimed. Tools Appl.*, 2021.

[94] S. M. B. K. Albargathe, E. Kamberli, F. Kandemirli, and J. Rahebi, “Blood vessel segmentation and extraction using H-minima method based on image processing techniques,” *Multimed. Tools Appl.*, vol. 80, no. 2, pp. 2565–2582, 2021.

[95] A. A. I. Mohamed, M. M. Ali, K. Nusrat, J. Rahebi, A. Sayiner, and F. Kandemirli, “Melanoma skin cancer segmentation with image region growing based on fuzzy clustering mean,” *Int. J. Eng. Innov. Res.*, vol. 6, no. 2, p. 91C95, 2017.