Check for
updates

Neighborhood Rough Neural Network Approach for COVID-19 Image Classification

S. Nivetha[1] · H. Hannah Inbarani[1]

Abstract

The rapid spread of the new Coronavirus, COVID-19, causes serious symptoms in humans and can lead to fatality. A COVID-19 infected person can experience a dry cough, muscle pain, headache, fever, sore throat, and mild to moderate respiratory illness, according to a clinical report. A chest X-ray (also known as radiography) or a chest CT scan are more effective imaging techniques for diagnosing lung cancer. Computed Tomography (CT) scan images allow for fast and precise COVID-19 screening. In this paper, a novel hybridized approach based on the Neighborhood Rough Set Classification method (NRSC) and Backpropagation Neural Network (BPN) is proposed to classify COVID and NON-COVID images. The proposed novel classification algorithm is compared with other existing benchmark approaches such as Neighborhood Rough Set, Backpropagation Neural Network, Decision Tree, Random Forest Classifier, Naive Bayes Classifier, K- Nearest Neighbor, and Support Vector Machine. Various classification accuracy measures are used to assess the efficacy of the classification algorithms.

Keywords CT scan images · Neighborhood rough set · Neighborhood Rough Neural Network (NRNN) · COVID-19

1 Introduction

One of the most serious problems of the twenty-first century is SARS-CoV-2 and the related COVID-19 disease. As of 12th October 2020, the total number of worldwide cases of Coronavirus is 136,660,311. Of these, 2,950,152 people were deaths and 109,907,478 were recovered. The number of active patients is 23,802,681 [1]. As a result of increases in COVID cases, Computer-Aided Detection Diagnosis (CADD) must be used to assist radiologists in making the diagnosis process to alleviate a huge number of COVID patients. COVID-19

✉ S. Nivetha
nivethas@periyaruniversity.ac.in

H. Hannah Inbarani
hhinba@periyaruniversity.ac.in

[1] Department of Computer Science, Periyar University, Salem, Tamil Nadu, India

can be diagnosed using three different methods: blood tests, X-rays, and Computed Tomography (CT) scans [2, 3]. CT technology, CT applications, and innovations are helping to improve image quality and reduce acquisition time [4]. CT technology is essential in diagnostic medicine, image-guided intervention, and the evaluation of therapeutic and surgical outcomes [5].

CT is a useful tool for early screening and diagnosis of COVID-19, it can segment texture features, shape, size, density, and texture. CT scans, in comparison to traditional X-rays, provide a much more accurate description of the patient's condition. This accurate information can be used to decide whether a medical condition exists, as well as the nature and precise location of the problem.

Rough Set theory (RST) is a method for reducing dimensionality while retaining the semantics of the features [6]. The weakness of RST to deal with real-valued data is its central problem. Before the implementation of RST, methods of discretizing the data were used to solve this issue. However, since such approaches can lead to information loss, several RST extensions have been developed to remedy this inability to operate on real-valued domains. The Neighborhood Rough Set Classification (NRSC) can operate effectively with real-valued data, resulting in minimal information loss. The Neighborhood Rough Neural Network (NRNN) which is made up of Neighborhood Rough Neuron and decision-making neuron is used in this study for COVID and NON-COVID image classification and it has achieved a high accuracy in the test of fault classification and the benchmark classifiers such as Support Vector Machine (SVM), Decision Tree Classifier (DTC), Random Forest Classifier (RFC), Back Propagation Neural Network (BPN), Neighborhood Rough Set (NRS) are used for comparative analysis.

2 Research Motivation

Medical imaging has enhanced our understanding of the integral and efficient nature of human life systems, and it is now widely used for the exploration, mediation, and handling of clinical issues. The ongoing COVID-19 pandemic is wreaking havoc on people's health, as well as their social and economic lives. Reviving public life while mitigating the risk of infection is a vital and difficult challenge. Extracting relevant features from CT-COVID and CT-NON COVID images and classify the images is the ultimate goal of our research. The rules are derived from NRSC and distinguish as COVID and NON-COVID. The proposed solution for COVID-19 follows a methodology that uses Neighborhood Rough Neural Network and Back Propagation techniques. This methodology includes segmentation, feature extraction, and classification of COVID or NON-COVID. To enhance the performance of the proposed approach, Back Propagation Neural Network is employed in the NRSC. This leads to updated decision tables are generated. From the experimental results, it is deduced that the overall classification accuracy of the proposed NRNN is 98%, 97%, 100%, 100%, for GLCM 0°, GLCM 45°, GLCM 90°, GLCM 135° datasets respectively. The proposed approach correctly distinguishes between COVID and NON-COVID compared to the existing approaches.

3 Research Contribution

The motivation of the study is to propose an effective classification of COVID and NON-COVID CT Scan Images using Neighborhood Rough Set hybridized with Backpropagation

technique. Machine learning is used to achieve effective classification results with limited computing resources. In this work, 349 CT scan COVID images and 397 CT NON-COVID images are used to prove the effectiveness of the proposed method. The experimental results are shown to measure the correctness in terms of its accuracy, precision, and recall values against Decision Tree, Naive Bayes Classifier, Random Forest Classifier, K-Nearest Neighbor Classifier, Support Vector Machine, and Neighborhood Rough Set Classification.

4 Research Objectives

Rough Set theory can be used for classifying data that is imprecise or noisy in order to determine structural relations. The neural network has the ability to handle complex and Non-Linear relationship in the data and provides support for parallel processing. Rough Set Theory (RST) and neural networks are the two key models for intelligent information processing that have recently emerged. The Rough Set theory and neural networks are the two techniques which have a higher adaptive ability to work with erroneous and partial knowledge. To apply Rough Set, data must be discretized, due to the discretization procedure in Rough Set, there are many limitations such as it may result in data loss. This difficulty is solved using Rough Set that is based on the neighborhood relation. Because Neighborhood Rough Set avoids the discretization procedure, data loss is minimized. It is capable of dealing with both continuous and discrete data. Backpropagation Neural Network technique helps to improve the accuracy while updating the weights to minimize loss and optimizes the performance of the network. By fine-tuning the weights, the error rate is reduced, and the model's generalization increases, the model becomes more accurate and consistent. As a result, for classification of COVID-19 images, a hybridization of Neighborhood Rough Set and Back Propagation approach is proposed in this work.

5 Related Work

Many studies have used Machine learning and deep learning as a classification model for COVID-19 using Computed Tomography (CT) images as described in related work. In [7], the author has developed a deep learning-based framework for detecting COVID-19 pneumonia on high-resolution CT. 46,096 anonymous images from 106 admitted patients in Renmin Hospital of Wuhan University, including 51 patients with laboratory-confirmed COVID-19 pneumonia and 55 control patients with other diseases, were retrospectively collected for model creation and validation. In an internal retrospective dataset, the model had a per-patient accuracy of 95.24% and a per-image accuracy of 98.85%. It had a 96% accuracy in an external dataset. In [8], the author has implemented an AI system that analyses CT images automatically and calculates the risk of infection to diagnose COVID-19 pneumonia quickly. The proposed AI system prioritizes each CT image due to the possibility of infection, allowing physicians to confirm and segregate infected patients in real-time. In this work, the author has adapted models previously built for different diagnosis, such as ResNet-50 and 3D Unet + + and have built a training-inference pipeline for fast testing of a selection of candidate models. They developed a three-stage annotation and quality control pipeline for segmentation and classification. Furthermore, the technology identified all lesion locations automatically for faster assessment. The application for proposed system was installed in 16 hospitals so far, and it's currently performing over 1,300 screenings per day. In this

research dataset, 1,136 training cases (723 COVID-19 positives) from five hospitals were used which included several pulmonary diseases. The proposed system achieves a sensitivity of 0.974 and a specificity of 0.922. In [9] the COVID-19-infected patients are classified as infected (+ ve) or not infected (-ve) using Convolutional Neural Networks (CNN) and Multi-Objective Differential Evolution (MODE). The COVID-19 is classified using the data of chest CT images in this paper. Various training and testing dataset ratios are considered for experimental purposes, including 20,80%, 30:70%, 40:60%, 50:50%, 60:40%, 70:30%, 80:20%, and 90:10%, respectively. In terms of Precision, F-Measure, Sensitivity, Specificity, and Kappa Statistics, the proposed model outperforms competitive models such as ANN, ANFIS, and CNN by 1.9789%, 2.0928%, 1.8262%, 1.6827%, and 1.9276%. In [10], a new explainable deep learning method is presented for COVID-19 detection using CT scan images. A publicly accessible SARS-CoV-2 CT scan dataset with 1252 CT scans that are positive for SARS-CoV-2 infection (COVID-19) and 1230 CT scans for patients who are not infected with SARS-CoV-2. The data was gathered from actual patients in hospitals in Sao Paulo, Brazil. In this work, the authors have used an explainable Deep Learning approach (xDNN) and were able to achieve a very promising F1 score of 97.31%. The xDNN classifier generated the best results in terms of accuracy. In [11], a total of 305 COVID-19 (11%), 872 CAP (33%) and 1,498 non-pneumonia (56%) tests were included in the study. In 251 patients, 305 COVID-19 CT tests were obtained. Community-Acquired Pneumonia (CAP) data included 872 CT scans from 869 patients. The non-pneumonia data consisted of 1,498 CT scans taken on 1,475 people. To extract features from the COVID-19, a novel self-supervised learning method (i.e. Rubik's cube Pro) was suggested. The Rubik's cube Pro is found to outperform the traditional Rubik's cube, with a + 1.8 percent increase in F1 score.

In [12], chest CT data from multiple centers and scanners were collected and they proposed a weakly supervised deep learning system for fully automatic detection of COVID-19 con-taminated regions. The Cancer Imaging Archive (TCIA) Public Access provides access to the records. COVID-19, CAP, and NP patients were included in this study, which included 150 3D volumetric chest CT exams. The proposed model achieves a mean ACC of 96.2%, PRC of 97.3%, SEN of 94.5%, SPE of 95.3%, and AUC of 0.970, indicating that it is relatively good at separating COVID-19 cases from NP cases, as indicated by the results. In [13], a novel Joint Classification and Segmentation (JCS) method is proposed for performing COVID-19 chest CT diagnosis in real-time and with explanations. The JCS algorithm was trained using a COVID-19 Classification and Segmentation (COVID-CS) dataset that included 144,167 chest CT images from 400 COVID-19 patients and 350 uninfected cases. A total of 3,855 chest CT images from 200 patients have been annotated with fine-grained information. On the classification test set, it achieves an average sensitivity of 95.0%, a precision of 93.0%, and a Dice score of 78.5% on the segmentation test set of the COVID-CS dataset. In [14], the authors have used chest CT images to classify COVID-19 using Adaptive Feature Selection driven Deep Forest (AFS-DF). They extract location-specific features from CT images. Then, to learn high-level representations of these features with the relatively small-scale data, they use a deep forest model to learn high-level representations of the features. With 1495 COVID-19 patients and 1027 community-acquired pneumonia patients, they proposed AFSDF on the COVID-19 dataset Community-Acquired Pneumonia (CAP). The accuracy, sensitivity, specificity, and precision (ACC, SEN, SPE, SPE, SPE, SPE, SPE, SPE, and SPE were used to find the effectiveness of the proposed system. The proposed system yielded AUCs of 91.79%, 93.05%, and 89.95%. 96.35% and 96.35%, respectively using pre-trained deep learning architectures as an automated platform for COVID-19 identification and diagnosis in chest CT. In this work, to identify patients as COVID infected or not, a DenseNet201-based Deep Transfer Learning (DTL) is proposed, i.e., COVID-19 (+) or COVID (-). The

proposed model extracts features from the ImageNet dataset using its own trained weights and a convolutional neural network. The SARS-CoV-2 CT scan dataset from Kaggle (www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset) was used. The dataset contains 2492 CT scans, 1262 of which are positive for SARS-CoV-2 infection (COVID-19 (+) and 1230 of which are negative for SARS-CoV-2 infection (COVID-19 (-). The proposed model classifies chest CT-scans as having 99.82%, 96.25%, and 97.4% training, testing, and validation accuracy, respectively. The proposed model achieves a 97% accuracy rate. The accuracy of VGG-16 and Resnet152V2 is, however, 96% and 95%, respectively [15]. In [16] provides a clinical Computer-Aided Diagnosis method that uses CT features to distinguish COVID-19 pneumonia patients from NON-COVID-19 pneumonia patients. A total of 612 patients (306 COVID-19 pneumonia and 306 NON-COVID-19 pneumonia) were enrolled in the study. a clinical Computer-Aided Diagnosis (CAD) method that uses CT features to automatically distinguish COVID-19 pneumonia patients from those that do not have the virus. To assess the pattern, location, and distribution of lesions in both classes, twenty radiological features were extracted from CT images. All of the important CT features were fed into five different classifiers: Decision Tree (DT), K-Nearest Neighbor (KNN), Naive Bayes (NB), and Support Vector Machine (SVM) to form an ensemble classifier. Using the ensemble COVIDiag model classifier, the proposed CAD scheme achieved a sensitivity, precision, and accuracy of 0.965, 93.54%, 90.32%, and 91.94%, respectively. In [17] an ensemble deep learning model for the novel COVID-19 was developed. Dataset was collected from 2500 CT images of a lung tumor and 2500 regular lung images from Ningxia Medical University's general hospital in China. The gathering of the dataset from Previous journals, authoritative media accounts, and public databases provided 2933 lung CT images of COVID-19 patients. The results showed that our EDL COVID (Ensemble Deep Learning) model outperformed a single individual classifier in terms of overall classification precision, with the fastest detection speed of 342.92 s and an accuracy of 97%; the ensemble accuracy can thus reach 99.05%. In [18], a deep learning-based CAD (Computer-Aided Diagnosis) method is proposed to distinguish COVID-19 infection from other atypical and viral pneumonia diseases. A total of 1020 CT slices were collected from 108 patients with COVID-19 (COVID-19 group) and 86 patients with other atypical and viral pneumonia diseases (NON-COVID-19 group). The ten well-known convolutional neural networks evaluated to discriminate COVID-19 infection from NON-COVID-19 infection was AlexNet, VGG-16, VGG-19, SqueezeNet, GoogleNet, MobileNet-V2, ResNet-18, ResNet-50, ResNet-101, and Xception. ResNet-101 and Xception had the highest results of all the networks. With an AUC of 0.994 (sensitivity), ResNet-101 was able to differentiate COVID-19 cases from NON-COVID-19 cases. In [19], the Neighborhood Rough Set classification algorithm (NRSC) results in a wide variety of biomedical recognition and classification abilities. The proposed Novel Rough Set Classification algorithm is comparable to a multi-layered algorithm in terms of experimental analysis. The obtained results revealed that the proposed algorithm was successful. Using NRSC and DWT achieve a classification accuracy of 99.32%. The results showed that the proposed method performed well. The NRSC classification system was much superior to the previous one. The Neighborhood Rough Set-based Classification (NRSC) algorithm is used to classify five medical data sets in the paper. Six performance tests such as Precision, Recall, F-Measure, Folkes-Mallows Index, Kulczynski Index, Rand index are used to validate the classification algorithm's efficiency. The proposed NRSC produces the best results compare to other existing benchmark classification algorithm [20].

6 Dataset Description

This study uses the publicly accessible COVID-19 CT scan dataset, which may be found at https://github.com/UCSD-AI4H/COVID-CT [21]. Radiologists have commended this CT image-based radiography dataset. This COVID-CT dataset consists of 349 positive COVID-19 cases and 397 negative Chest Computed Tomography images. 153, 491, and 1853 are the minimum, average, and maximum heights of the images in the data set respectively. 124, 383, and 1485 are the minimum, average, and maximum widths of the images in the data set respectively. To maintain consistency in the dataset for studies, all images were transformed to Portable Network Graphics (.png) format. Both positive and negative images were also resized to 256*256. Table 1 describes the COVID – 19 dataset distribution.

6.1 Dataset Distribution

7 Pre-processing

The most common approach is to eliminate the noise from the image and then add various filters to it. The median filter is a non-linear algorithm that is used to minimize the noise in images. It is one of the most effective strategies for removing background noise. The Median filter is used for COVID pre-processing in this study. The image is enhanced and the noise is reduced without losing the important features of COVID images for diagnosis during the pre-processing phase.

8 Segmentation

Segmentation is a technique for converting images into a more understandable form. It simplifies image processing [22]. Many images processing functions, such as medical image analysis, object recognition, tumor detection, satellite imagery, and so on, require segmentation. In this Research, Otsu Thresholding is applied for the images, it is used to perform automatic thresholding in image processing. It selects an effective threshold for dividing image pixels into different regions and isolating artifacts from their background based on their level distribution. Morphological operations are needed to obtain specific parts of an image and also aid in the removal of unnecessary parts. Erosion is a binary image processing technique that "thins" or "shrinks" objects. After Segmenting the images, Erosion is performed for extracting the specific parts of the image.

9 Feature Extraction

The method of detecting and describing global or local features in medical images is referred to as Feature Extraction. Feature extraction is the process of extracting the properties of objects in images. The Gray-Level Cooccurence Matrix (GLCM) is one of the most widely used metrics (GLCM). GLCM is a statistical texture analysis method that uses second-order statistics. It investigates the spatial relationship between pixels and determines how often a particular combination of pixels appears in an image in a given direction and distance d. The number of rows and columns in a GLCM matrix equals the number of grey levels, G, in the image. The GLCM determines how often a pixel with the value i of grey level corresponds

Table 1 Dataset distribution

Dataset	Date of publication	Modality	Total images	No. of patients	Open-source	Class	File type	Data repository
COVID -19 (Positive)	March 28,2020	Chest CT scan	349	216	Yes	COVID	RAR FILE	https://github.com/UCSD-AI4H/COVID-CT
NON-COVID-19 (Negative)	March 28, 2020	Chest CT scan	397	15	Yes	NON-COVID	RAR File	https://github.com/UCSD-AI4H/COVID-CT

(A) Sample Test Pixel Values for the Image   **(B)** Eight Directions for the Pixel.
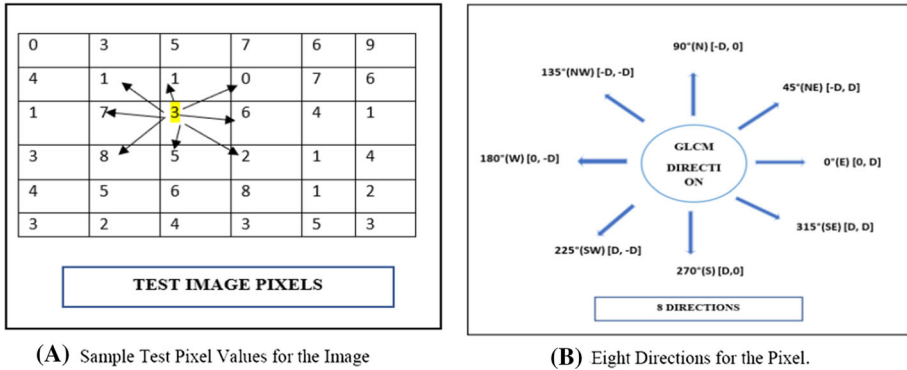
Fig. 1 a Sample test pixel values for the image, **b** eight directions for the pixel

horizontally, vertically, or diagonally to the neighboring pixels with the value j of grey level (grayscale intensity or Tone). Every pixel has eight neighbors, giving the co-occurrence grey levels of 0° (East), 45° (North East), 90° (North), 135° (North West), 180° (West), 225° (South West), 270° (South), or 315° (South East) the possibility of occurrence. Figure 1a shows an example of a sample test image pixel. For example, the 3rd row pixel with grey level 3 as a reference pixel has eight neighbors along the 0° (East), 45° (North East), 90° (North), 135° (North West), and 180° (North West) directions (West), 225° (South West), 270° (South), and 315° (South East), the co-occurrence grey levels are (3,6), (3,0), (3,1), (3,7), (3,8), (3,5), (3,5) and (3,2) [23]. Figure 1b depicts the eight possible directions forms from the sample test image pixel.

In terms of the GLCM criteria, the co-occurring pairings obtained by selecting 0° is consistent with the results obtained by selecting 180°.This principle equally extends to 45°, 90°, and 135° [23]. Due to this symmetric property of GLCM, Horizontal GLCM which is addition of East and West ($\theta = 0°$ and 180°), Vertical GLCM which is sum of North and South ($\theta = 90°$ and 270°), and Diagonal GLCM is the North East and South West ($\theta = 45°$ and 225°). Otherwise, it can be referred as North West and South East ($\theta = 135°$ and 315°). The Diagonal GLCM is the same for both diagonals. As a result, there are four options for determining the value of θ. As a result, GLCM directions are calculated as Horizontal $= 0°$(East), Vertical $= 90°$(North), Diagonal $= 45°$ (North East), and Diagonal top left to bottom right $= 135$(North West) [23]. From this reason, GLCM was calculated for the dimensions of 0°, 45°, 90°, and 135° to extract the features from the COVID images and NON-COVID images. In texture-based features, for the dimensions 0°, 45°, 90°, and 135°, the Grey Level Co-occurrence Matrix (GLCM) was computed. For each segmented image, a total of 6 features were extracted, which consists of Contrast, Dissimilarity, Homogeneity, Energy, Correlation, and ASM.

10 Neighborhood Rough Set

The Neighborhood Rough Set Classification (NRSC) replaces the standard Rough Set model's equivalent approximation with a neighborhood relation that supports both continuous and discrete datasets. The main NRS concepts were used in this study [24–28]. The entire NRS model is based on an θ—neighborhood relation, which replaces the equivalence relation

in standard Rough Sets with a neighborhood relation based on distance metric functions (Euclidean distance). The Euclidean distance metric is defined as follows:

$$f(x_i, x_j) = \sum_{k=1}^{n} \sqrt{(x_i + x_j)} \tag{1}$$

Definition 1 To the non empty set, universe $U = \{x_1, x_2, x_3, \ldots, x_n\}$ metric function satisfies,

1. Non- Negative $f(x_i, x_j) \geq 0, if x_i = x_j, then f(x_i, x_j) = 0$
2. Symmetry $f(x_i, x_j) = f(x_j, x_i)$
3. Triangle Inequality $f(x_i, x_j) \leq f(x_i, x_k) + f(x_j, x_k)$

From the above notation f is defined as metric function of universe U, and $< U, f >$ is the neighborhood relation [24–28].

Definition 2 Assuming $< U, f >$ as the neighborhood relation, $\forall x_i \in U, \theta \geq 0$, having.

$$\theta(x_i) = \{x | f(x_i, x) \leq \theta, x \epsilon U\}$$

Then, $\theta(x_i)$ is the θ neighborhood relation set of x_i[13].

10.1 Neighborhood Decision System

A decision system can be defined as $< U, (A \cup D) >$. U is a universe that is a set of records, A and D which is respectively the conditional and decision attribute set [13], which will create a neighborhood relation based on the condition. In Neighborhood Decision System, θ is a basic decision system and is referred to as NDS $= < U, A \cup D, \theta >$.

Definition 3 In $NDS = < U, A \cup B, \theta >$, B is subset of A (B $\subseteq$ A), for arbitrary X $\subseteq$ U, two sets of records, called lower and upper approximations of X in terms of relationship between N with respect to B,

$$\underline{N_B X} = \{x_i | \theta_B(x_i) \subseteq X, x_i \in U)\} \tag{2}$$

$$\overline{N_B X} = \{x_i | \theta_B(x_i) \cap X \neq \varnothing, x_i \in U\} \tag{3}$$

here $\theta_B(x_i)$ is calculated as follows,

$$\theta_B(x_i) = \{x | f(B(x_i), B(x)) \leq \theta, x \epsilon U\} \tag{4}$$

and B (x) is a function to extract the sub vector from x, only the whose attributes that are contained in the attribute set B will be chosen. That is B $(x) = \{a (x) | \forall aB\}$ The boundary region of the decision D with respect to attributes B is defined as,

$$BNR(D) = \overline{N_B X} - \underline{N_B X} \tag{5}$$

The neighborhood lower approximation is the union lower approximation of each D decision classes. The neighborhood lower approximation of decision is also called as positive region [20, 24–28].
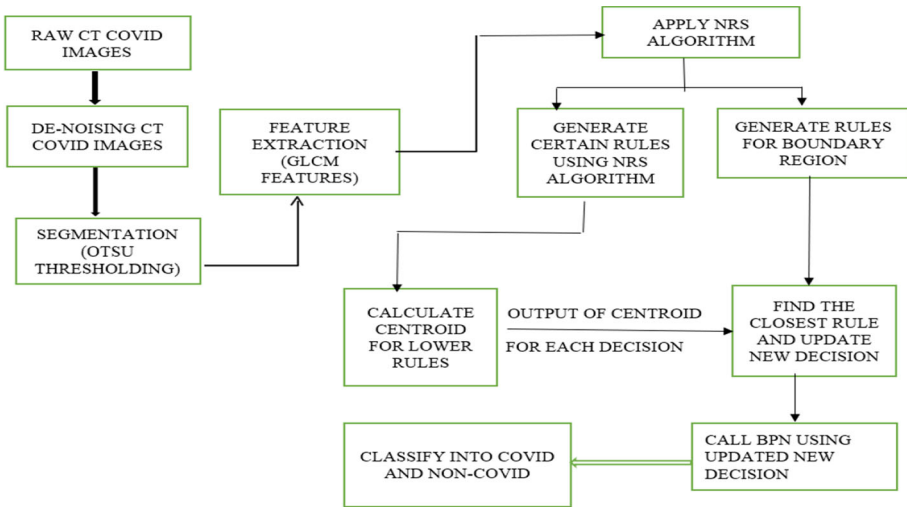
Fig. 2 Block diagram of the proposed method

11 Proposed Methodology

In this research, the neighborhood roughest based backpropagation algorithm is proposed to make decision rules to efficiently classify the COVID and NON-COVID images. There are three phases in this proposed methodology. In the First Stage, the Collecting COVID-19 CT dataset used in this research (https://github.com/UCSD-AI4H/COVID-CT) is available to the public [21]. 349 CT images from COVID patients and 397 CT images from NON-COVID patients are included in the dataset, GLCM features are extracted. The complete methodology adopted in this work for the classification of COVID is exposed in Fig. 2. A dataset is divided into two parts: a training set and a testing set. The effectiveness of a classifier is determined by providing it with a testing set. The suggested algorithm in Figs. 3 and 4 shows the procedure and outcomes. Back Propagation Neural Network is used in the third phase.

12 Backpropagation Neural Network

The core of neural net training is Back Propagation. It's a technique for fine-tuning the weights of a neural network using the error rate from the previous epoch. By fine-tuning the weights, it helps to reduce error rates and improve the model's generalization, making it more accurate. The Main Layers in the Backpropagation neural network are Input Layer, Hidden Layer, Output Layer. For Example, the input layer is given x, Weights w is used to model the input. The output is calculated by each hidden layer, and data is ready at the output layer. The error is the difference between the real and desired output. Return to the hidden layers and tweak the weights to reduce the error in future runs.

INPUT: The set of conditional attributes, D = {d$_1$,d$_2$,..., d$_n$} is the decision attribute.

OUTPUT: Set of Rules for each class.

Step 1: Define the Neighborhood relation for the conditional attributes using,

$$[\theta]_C \Leftarrow C$$

$$\theta_C(x_i) \Leftarrow [x]_C$$

Where $\forall \theta_C(x_i)$ is referred as

$$\forall \theta_C(x_i) = \{x \mid f(C(x_i), B(x)) \leq \theta, x\epsilon U\}$$

Step 2: Develop the Equivalence relation for the decision attribute using the equation,

$$[x]_D \Leftarrow D$$

Step 3: Construct the Neighborhood Rough Set lower approximation space for each class.

$$\underline{N_B X} = \{x_i \mid \theta_C(x_i) \subseteq X, x_i \in U)\}$$

Step 4: Construct the Neighborhood Rough Set upper approximation space for each class.

$$\overline{N_B X} = \{x_i \mid \theta_C(x_i) \cap X \neq \emptyset, x_i \in U\}$$

Step 5: Find the boundary region of the data set by using the Neighborhood Rough Set boundary region.

$$\mathbf{BND_C(X)} = \overline{N_B X} - \underline{N_B X}$$

Step 6: Generate certain rules for each class separately using a Neighborhood Rough Set based on lower approximation.

Step 7: Calculate centroid for Lower Rules using Lower approximation.

$$CEN_D(x) \Leftarrow \underline{N_B X}$$

Step 8: Compute distance between Boundary Rules and centroid for lower approximation rules for each class.

$$\forall BND_C(X)$$

$$\forall CEN_D(X)$$

$$DIS_{ij} \Leftarrow DIS(BND_C(D), CEN_D(X))$$

Where,

$$DIS_{ij} = \sum_{k=1}^{n} \sqrt{(X_{ik} - X_{jk})^2}$$

Step 9: Update minimum distance to the matching closest rule by using the equation

$$ND \Leftarrow MIN(DIS_{ij}, CEN_D(X))$$

Step 10: Update the new decision Feature,

$$D \Leftarrow ND$$

Step 11: Call BPN (C, D)

Fig. 3 Neighborhood rough neural network (NRNN) classification algorithm

INPUT: The set of conditional attributes, D = {d_1, d_2…, d_n} is the decision attribute.

OUTPUT: Decision

Step 1: For each weight W in a network, W← A random Weights, b← bias

Step 2: For each training set, apply Forward Propagation which helps to evaluate the predicted output against expected output.

Step 3: Apply Backward Propagation, for all weights from the hidden layer to the output layer. Minimize the cost function and adjust the network's weights and biases.

Step 4: Backward pass continues for all weights from the input layer to the hidden layer.

Step 5: Update Network Weights.

Step 6: Predict the class Label.

Fig. 4 Back Propagation Neural Network algorithm for Neighborhood Rough Neural Network (NRNN) Classification

12.1 Proposed Neighborhood Rough Neural Network Algorithm (NRNN)

There are three significant reasons for the given interpretation of the algorithm that has been proposed. The distance metric function is one, the value of the –neighborhood relation is the other, and a neural network is a powerful tool for classification. The shape of neighborhoods is defined by the Euclidean distance function, and the value of the neighborhood relation is used to regulate the flow of the algorithm.

12.2 Backpropagation Neural Network Algorithm-Call BPN (C, D)

The proposed model has been improved in the following ways:

1. Rough Set can be applied to discrete-valued attributes. As a result, continuous-valued features must be discretized before being used. Typically, pixels in images are viewed as continuous values. Hence to classify continuous numerical Image data, *Neighborhood Rough Set* is used in this work which constructs Lower and Upper approximation based on Neighborhood relation.
2. The boundary values are regarded to be uncertain, and so *inference decision-making* has been used in this work to handle the uncertainty in decision making.
3. New *updated decision feature table* is constructed in this work through updation of the minimum distance to the closest rule formed using Neighborhood Rough Set which is used as input to the Back Propagation Neural Network (BPN) which helps to improve the net classification ability to respond correctly to the training and test datasets.

13 Result and Analysis

The Proposed algorithm was executed using a windows system with Intel(R) Core (TM) i5-10210U CPU @ 1.60 GHz 2.11 GHz machine with 8 GB RAM. NRSC algorithm was implemented using the Anaconda Software. In this research, Novel Neighborhood Rough

Table 2 Features and extracted rules using NRSC

GLCM features	Total features extracted	Number of COVID rules	Number of NON-COVID rules	Boundary rules
GLCM 0°	6	302	360	84
GLCM 45°	6	319	351	76
GLCM 90°	6	292	344	110
GLCM 135°	6	309	308	129

Set Based Neural Network was employed to choose the most relevant features. Six different textual features were applied to COVID-CT datasets and GLCM-0°, GLCM-45°, GLCM-90°, and GLCM-135° were considered. They are six features in each feature set. By Using the NRSC algorithm, lower approximation rules are generated for COVID and NON-COVID images. Table 2 contains a thorough description of the datasets, the number of features retrieved, lower approximation rules for COVID AND NON-COVID, and boundary rules.

13.1 Evaluation of NRNN

Decision Tree (DT) is a tree structure that looks like a flowchart, with each internal node denoting a test on an attribute, each branch denoting a test result, and each leaf node holding a class label [29]. The root node is the topmost node in a tree. The idea of a decision tree is to identify an optimal partitioning of the space of possible observations, mostly by recursive splits [30]. Lepetit et al. [31] were the first to introduce *Random Forests (RF)*. A random tree is one that is generated at random from a set of possible trees, each with K random features at each node [32]. In a Random Forest, each decision split's features are chosen at random. By selecting characteristics at random, the correlation between trees is reduced, resulting in improved prediction power and efficiency. The *Naive Bayes (NB) approach* is commonly referred to as Bayesian classifier because it is based on the Bayes Theorem [33]. Bayesian classifiers are based on the idea that because the feature values in a class are similar, a class can possibly predict the features of its members. When an agent recognizes a class, predicting feature values is simple. If, on the other hand, the class is unknown and the primary source of information is some of the feature values, the Bayes rule can be used to forecast the matching class. *The K-nearest Neighbor (KNN)* is often known as the KNN classification algorithm [34] and is one of the most widely used classification algorithms with the best classification effect. The main concept is that when adding new data of unknown category to be categorized, the category of the data to be classified must be established based on the category of other samples. Vapnik was the first to propose the *Support Vector Machine* (SVM) [35]. SVM is a supervised Machine learning algorithm that can be used to solve problems like classification and regression. SVR is a Support Vector Machine-based non parametric regression approach [36]. To recognize the pattern between predictor and predicted values, the SVR uses the structural risk minimization concept. The GPR model is a probabilistic kernel model. Gaussian processes can be viewed of as a Bayesian variant of SVM algorithms [36]. In this research, SVM is used for the classification of the CT-Scan images into COVID and NON-COVID images. The parameter settings used by the classification approaches are depicted in Table 3.

Table 3 Parameter settings for classifiers

Number of Hidden Layers:8 Classes: {COVID and NONCOVID} Iterations = 100 Activation function

Classifier	Parameter Setting	Test Size
Random Forest Classifier	Cross Validation = 3, Scoring = "ROC_AUC"	Test Set Size = 30% Training Set Size = 70%
Naive Bayes Classifier	Gaussian Naive Bayes	
K-Nearest Neighbor Classifier	K = 5	
Support Vector Machine	Kernel = Linear	
Back Propagation Neural Network	Number of Inputs: 6 Number of Hidden Layers:8 Classes: {COVID and NONCOVID} Iterations = 100 Activation function = ReLu	

The Features are employed as input for the classifiers, which are divided into two distinct datasets, training and testing datasets (70% of samples as training data and 30% of samples as testing data). A more accurate system is produced by prediction algorithms that learn from the training set. This system is used to predict the appearance of new objects. A confusion matrix is used to test the classifier in Machine learning. The number of correct and incorrect predictions is shown in a confusion matrix. To represent counts based on expected and actual values, confusion matrices are used. The output "TN" stands for True Negative, and it shows the number of negative instances that have been correctly classified. Similarly, "TP" stands for True Positive, which refers to the number of positive examples that have been correctly classified. In comparison to other classifiers, Table 4 shows the confusion matrix entries GLCM 0°, GLCM-45°, GLCM-90°, GLCM-135°.

In this section, the classification accuracy of both the NRSC (Neighborhood Rough Set Classification) and Back Propagation Neural Network (BPN) classifiers are tested and compared with the proposed hybrid approach for COVID-19 Image dataset. The performance of these two approaches is almost similar and there is a little marginal increase in the results. Hence Hybridization of Neighborhood Rough Set with Neural Networks is proposed in this work which combines advantages of NRSC and BPN. Neighborhood Rough Set very well handles the noise and uncertainty in images along with Back propagation which is a standard method for classification. In BPN, the gradients of the loss or error function are calculated, and existing parameters are updated in response to the gradients. As a result, the combination of a Neighborhood Rough Set and Backpropagation Neural Network works better than other approaches for classification of COVID and NON-COVID images. The efficacy of the proposed NRNN algorithm is compared to the standard benchmark classifiers such as Neighborhood Rough Set Classification, Backpropagation Neural Network, Decision Tree Classification, Random Forest Classification, Naive Bayes, K-Nearest Neighbor and Support Vector Machine.

According to the interpretation of the data, the proposed algorithm achieved a correctly classified as COVID and NON-COVID when compared to the results of the different classifiers. The confusion matrix of GLCM-90 and GLCM-135° using various classifiers such as Decision Tree, Random Forest Classifier, Naive Bayes Classifier, K-Nearest Neighbor Classifier, Support Vector Machine, Back Propagation Network, Neighborhood Rough Set Classification, and Proposed NRNN.

Table 4 Confusion matrix and various validation measures for GLCM- 0°, 45°, 90°, 135°

GLCM features	Classification algorithm	Desired output	Output result for confusion matrix		Precision	Recall	F1-score	Support
			CO	N-CO				
GLCM 0°	NRNN	CO	99	3	0.98	0.97	0.98	102
		N–CO	2	120	0.98	0.98	0.98	122
	NRSC	CO	96	6	0.93	0.94	0.94	102
		N–CO	7	115	0.95	0.94	0.95	122
	BPN	CO	100	0	0.83	0.91	0.87	100
		N–CO	13	111	0.92	0.85	0.89	124
	DTREE	CO	92	11	0.91	0.89	0.90	103
		N–CO	9	112	0.91	0.93	0.92	121
	RFC	CO	46	51	0.92	0.92	0.92	106
		N–CO	61	66	0.93	0.92	0.93	118
	NBC	CO	79	18	0.83	0.78	0.80	98
		N–CO	13	114	0.83	0.87	0.85	126
	KNN	CO	79	18	0.87	0.93	0.90	101
		N–CO	13	114	0.94	0.89	0.89	123
	SVM	CO	79	18	0.86	0.81	0.84	97
		N–CO	13	114	0.86	0.90	0.88	127
GLCM 45°	NRNN	CO	100	3	0.97	0.97	0.97	103
		N–CO	2	119	0.98	0.97	0.97	121
	NRSC	CO	84	20	0.87	0.92	0.86	104
		N–CO	30	90	0.81	0.87	0.84	120
	BPN	CO	91	9	0.83	0.91	0.87	100
		N–CO	18	106	0.92	0.85	0.89	124
	DTREE	CO	74	21	0.70	0.78	0.74	95
		N–CO	19	95	0.82	0.76	0.79	129
	RFC	CO	87	19	0.85	0.82	0.84	106
		N–CO	15	103	0.84	0.87	0.86	118
	NBC	CO	66	40	0.82	0.62	0.71	106
		N–CO	14	104	0.72	0.88	0.79	118
	KNN	CO	90	10	0.72	0.90	0.80	100
		N–CO	35	89	0.90	0.72	0.80	128
	SVM	CO	62	35	0.70	0.64	0.67	97
		N–CO	27	100	0.74	0.79	0.76	127
GLCM 90°	NRNN	CO	100	0	1.00	1.00	1.00	228
		N–CO	0	124	1.00	1.00	1.00	294
	NRSC	CO	110	9	0.98	0.92	0.95	119
		N–CO	2	103	0.92	0.98	0.95	105
	BPN	CO	82	28	0.94	0.90	0.88	110

Table 4 (continued)

GLCM features	Classification algorithm	Desired output	Output result for confusion matrix		Precision	Recall	F1-score	Support
			CO	N-CO				
		N–CO	13	101	0.88	0.88	0.88	134
	DTREE	CO	88	18	0.82	0.83	0.83	106
		N–CO	19	99	0.85	0.84	0.84	118
	RFC	CO	91	15	0.88	0.86	0.87	106
		N–CO	12	106	0.88	0.90	0.89	118
	NBC	CO	53	44	0.72	0.55	0.62	97
		N–CO	21	106	0.71	0.83	0.77	127
	KNN	CO	87	9	0.85	0.91	0.88	96
		N–CO	15	113	0.93	0.88	0.90	128
	SVM	CO	52	45	0.74	0.54	0.62	97
		N–CO	18	109	0.71	0.86	0.78	127
GLCM 135°	NRNN	CO	94	0	1.00	1.00	1.00	94
		N–CO	0	130	1.00	1.00	1.00	130
	NRSC	CO	104	13	0.95	0.89	0.92	117
		N–CO	6	101	0.89	0.94	0.91	107
	BPN	CO	100	0	0.98	1.00	0.99	100
		N–CO	2	122	1.00	0.98	0.84	124
	DTREE	CO	91	13	0.83	0.88	0.85	104
		N–CO	19	101	0.89	0.84	0.86	120
	RFC	CO	96	10	0.91	0.91	0.91	106
		N–CO	10	108	0.92	0.92	0.92	118
	NBC	CO	65	34	0.76	0.66	0.70	99
		N–CO	21	104	0.75	0.83	0.79	125
	KNN	CO	98	4	0.88	0.96	0.92	102
		N–CO	14	108	0.96	0.89	0.92	122
	SVM	CO	68	29	0.77	0.70	0.74	97
		N–CO	20	107	0.79	0.84	0.81	127

13.2 Performance Analysis of Various Metrics

The proposed Novel NRNN is tested using performance measures extracted from the confusion matrix, such as Precision, Recall ,F1-score, Support and G-Mean. The proposed NRNN algorithm is evaluated using a variety of classification evaluation metrics. Each metric's systematic interpretation is provided in Table 5. [37–45].

In general, Precision, Recall, F-measure, and higher values indicate better efficacy of the system Table 6 shows the performance comparison of the proposed NRNN and other benchmark classification algorithms and the overall classification accuracies of NRNN, NRS, BPN, Decision Tree, Random Forest Classifier, Naive Bayes Classifier, KNN Classifier, Support

Table 5 Various evaluation metrics

S. no	Validation measure	Formula
1	Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
2	Sensitivity	$\frac{TP}{TP+FP}$
3	Specificity	$\frac{TN}{TN+FP}$
4	Error rate	$\frac{FP+FN}{TP+TN+FP+FN}$
5	Mattews correlation coefficient(MCC)	$\frac{(TP*TN)-(FP*FN)}{\sqrt{(TP+FP)*(TP+FN)*(TN+FP)*(TN+FN)}}$
6	Lift	$\frac{(TP/(TP+FP))}{((TP+FN)/(TP+TN+FP+FN))}$
7	Youden's index	Sensitivity + Specificity − 1
8	Balanced classification rate	$\frac{1}{2}$ (Sensitivity + Specificity)
9	Balanced error rate	1 − BCR

*TN true positive; TP true negative; FP false positive; FN false negative

Vector Machine. For GLCM- 0° findings showed that the Naive Bayes Classifier had the lowest classification accuracy, although the proposed NRSC system had a 0.98% higher classification accuracy than the other methods. NRSC and BPN are closely following all performance indexes. For GLCM-45° the proposed NRNN algorithm is more effective than the Neighborhood Rough Set Classification. The accuracy of the proposed NRNN algorithm is 0.97%, while that of the NRSC, BPN, DTC, RFC, NBC, KNN, SVM algorithms is 0.86%,0.88%, 0.80, 0.77% 0.5,0.84%,0.72% respectively. For GLCM 90° and GLCM 135°, output of the NRNN classifier is found to have 100% accuracy and no error rate than other classifiers. Furthermore, the proposed novel NRNN method classifies 100% correctly for COVID and NON-COVID Classification. According to the analysis of the data, the proposed NRNN achieved a prediction accuracy of 100%, classifies correctly all values when compared to the results of other algorithms with zero error rate.

For the GLCM 0° and GLCM 45°, the proposed NRNN classification produces good results. For the GLCM 90° the proposed method shows 100% accuracy for various measures than the NRSC, BPN, Decision Tree, Random Forest, Naive Bayes, KNN, and SVM. It also demonstrates how effective the proposed NRSC algorithm is in comparison to other classification algorithms. It outperforms other classifier algorithms in terms of precision, recall, and F1-score. By comparing GLCM features on various angles, 90° and 135° features show 100% more accuracy than other classifiers. Figure 5 reveals the prediction accuracy of the DTC, RFC, NBC, KNN, SVM, BPN, NRSC, and the proposed NRNN for each GLCM dataset. The proposed algorithm is found to have a higher prediction accuracy value. The highest prediction accuracy value, 100% is achieved by using the GLCM 90° and GLCM 135° datasets. The lowest prediction accuracy score, 77% is produced by Naive Bayes Classifier.

Figure 6 reveals the classification of error rate for various classifiers such as DTC, RFC, NBC, KNN, SVM, BPN, NRS, and the proposed NBRNN for each GLCM dataset. The proposed algorithm is found to have the best error rate of 0% value. The error rate of 0% is achieved by using the GLCM 90° and GLCM 135° datasets. The worst error rate, 0.33% is produced by Naive Bayes Classifier and BPN too.

Table 6 Evaluation of various performance measures

| GLCM performance features measures		Classification algorithm							
		NRNN	NRSC	BPN	DTC	RFC	NBC	KNN	SVM
GLCM 0°	Accuracy	0.98	0.94	0.94	0.88	0.92	0.77	0.94	0.86
	Sensitivity	0.98	0.93	0.88	0.87	0.91	0.84	0.95	0.85
	Specificity	0.97	0.95	1.0	0.89	0.92	0.73	0.92	0.86
	Error rate	0.02	0.05	0.05	0.11	0.08	0.22	0.06	0.13
	MCC	0.95	0.88	0.87	0.76	0.84	0.55	0.87	0.71
	Lift	2.15	2.04	1.96	1.94	1.93	1.79	2.06	1.98
	Youden index	2.95	2.88	2.85	2.76	2.83	2.57	2.87	2.72
	BCR	0.97	0.94	0.94	0.88	0.91	0.78	0.93	0.86
	BER	0.02	0.05	0.05	0.11	0.08	0.21	0.06	0.13
GLCM 45°	Accuracy	0.97	0.86	0.88	0.80	0.77	0.75	0.84	0.72
	Sensitivity	0.97	0.83	0.83	0.80	0.70	0.78	0.79	0.69
	Specificity	0.97	0.83	0.79	0.79	0.80	0.72	0.81	0.74
	Error rate	0.05	0.13	0.12	0.20	0.23	0.28	0.15	0.27
	MCC	0.94	0.70	0.70	0.62	0.53	0.44	0.69	0.43
	LIFT	2.13	1.87	1.87	1.64	1.66	1.64	1.68	1.60
	Youden index	2.93	2.75	2.75	2.59	2.52	2.72	2.43	2.43
	BCR	0.94	0.87	0.87	0.79	0.76	0.86	0.71	0.71
	BER	0.03	0.12	0.12	0.20	0.23	0.13	0.15	0.28
GLCM 90°	Accuracy	1.00	0.95	0.90	0.82	0.86	0.67	0.91	0.72
	Sensitivity	1.00	0.98	0.85	0.84	0.87	0.65	0.89	0.74
	Specificity	1.00	0.93	0.90	0.79	0.84	0.66	0.92	0.70
	Error rate	0.00	0.04	0.33	0.18	0.14	0.33	0.08	0.28
	MCC	1.00	0.90	0.53	0.74	0.73	0.73	0.85	0.42
	Lift	2.28	1.84	1.79	1.77	1.85	1.48	1.85	1.71
	Youden index	3.0	2.90	2.55	2.64	2.71	2.32	2.82	2.45
	BCR	1.0	0.95	0.79	0.82	0.85	0.66	0.91	0.72
	BER	0.00	0.04	0.21	0.17	0.14	0.33	0.08	0.27
GLCM 135°	Accuracy	1.00	0.92	0.99	0.88	0.91	0.76	0.92	0.78
	Sensitivity	1.00	0.94	0.98	0.81	0.89	0.81	0.88	0.77
	SPECIFICITY	1.00	0.88	1.0	0.87	0.92	0.72	0.96	0.78
	Error rate	0.0	0.08	0.99	0.15	0.08	0.24	0.07	0.21
	MCC	1.00	0.83	0.98	0.72	0.81	0.40	0.80	0.55
	Lift	1.00	1.81	2.19	1.76	1.88	1.66	1.95	1.78
	Youden index	3.0	2.83	2.98	2.69	2.82	2.53	2.84	2.55
	BCR	1.00	0.91	0.99	0.84	0.91	0.76	0.92	0.77
	BER	0.0	0.08	0.009	0.15	0.08	0.23	0.07	0.22
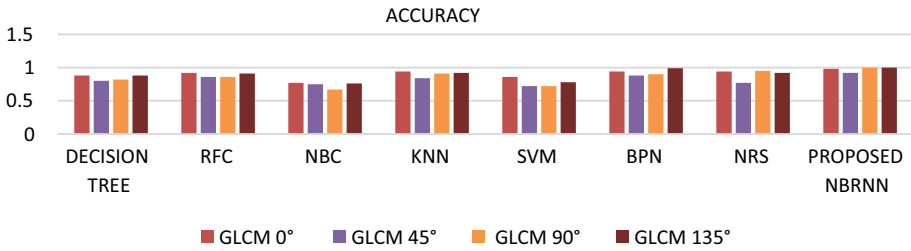
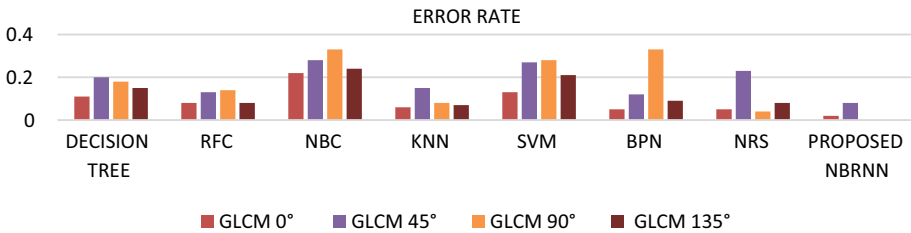**Fig. 5** Comparison of accuracy using various classifiers

Fig. 6 Comparison of error rate using various classifiers

13.3 Graphical Performance Assessment for Proposed NRNN

Receiver Operating Characteristic (ROC) curve is a graph that shows the various true positive rate cutoffs about the false positive rate. It's important to look at the success of different classifiers. ROC graphs are often used in the fields of decision-making and Machine learning. ROC curves can be used in this study to improve the superiority of the proposed NRNN. The algorithm decision-making rules appear in the top left corner of the ROC space, indicating that the model accurately forecasts the class. The diagonal line represents the method of predicting a class at random. Any classifier that appears in the ROC graph's bottom right corner performs worse than the others. The data on the ROC graph's far left side is becoming increasingly relevant. The ROC curve comparison of the proposed NRNN and decision-making algorithms is shown in Figs. 7,8,9 and 10 for GLCM-0°, GLCM-45°, GLCM-90°, and GLCM-135°.The proposed Novel NRNN is evaluated graphically using the ROC curve. The proposed NRNN outperformed the other current classification algorithms on COVID datasets. This refers to that proposed NRNN correctly classify the COVID and NON-COVID.
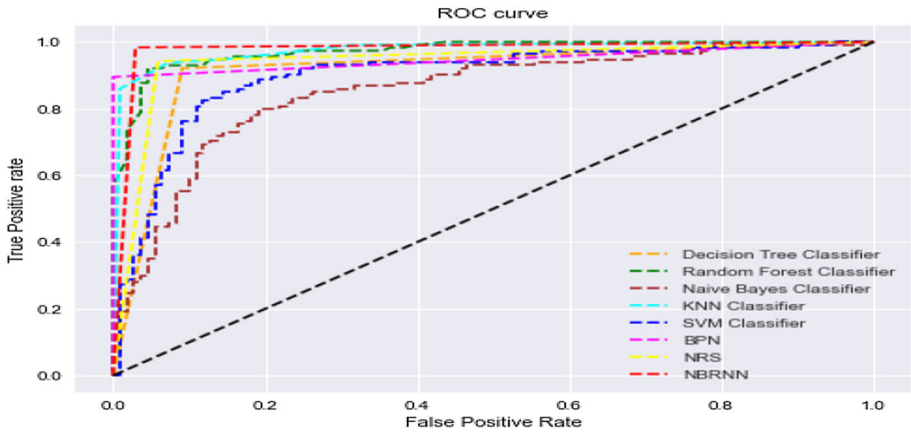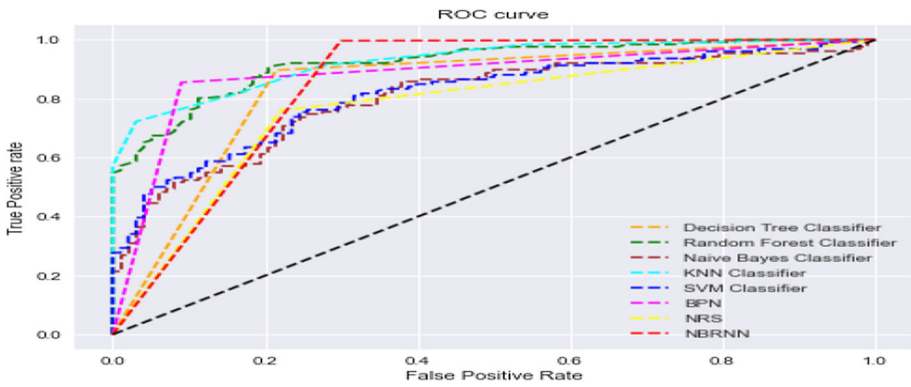
Fig. 7 ROC curve analysis proposed NRNN—GLCM 0°

Fig. 8 ROC curve analysis proposed NRNN—GLCM 45°
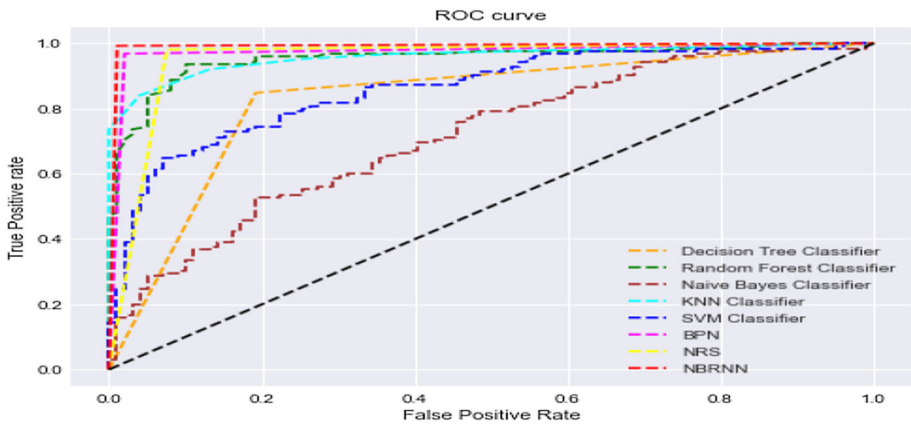

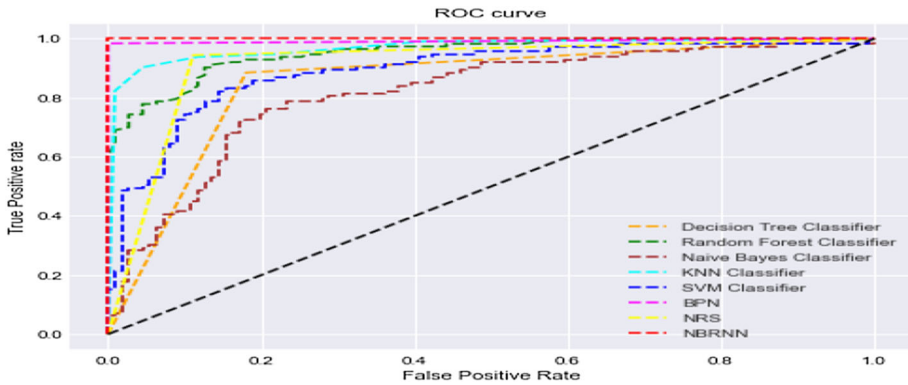**Fig. 9** ROC curve analysis proposed NRNN—GLCM 90°

Fig. 10 ROC curve analysis proposed NRNN—GLCM 135°

14 Conclusion

In this paper, a novel Neighborhood Rough Neural Network (NRNN) is proposed to classify the COVID and NON-COVID images. In this paper, a novel Neighborhood Rough Neural Network (NRNN) is proposed, which is an effective grouping of De-noising, feature extraction, segmentation, and classification methods. The Median Filter has been applied for De-noising and GLCM feature extraction of textual features for COVID and NON-COVID images. In the experiment analysis, accuracy was 0.98%, 0.92%, 1.00%, 1.00% for the proposed NRNN respectively. The obtained result shows that the proposed approach outperforms other comparative classification algorithms in terms of accuracy. As compared to other classification schemes, the proposed novel approach achieves a significant improvement in terms of Precision, Recall, F1-score, Sensitivity, Specificity, Bit Error Rate, Mathews Coefficient Correlation, Fowlkes Mallow Score, Lift, Youden Index, Balanced Classification Rate, and Bit Error Rate. The proposed Novel NRNN algorithm effectively classifies COVID and NON-COVID images, according to the experiential findings. The proposed system's success in accurately diagnosing the COVID and NON-COVID disease is accurately displayed using the ROC curve study. In the future, COVID-19 X-ray images will be incorporated with the Neighborhood Rough Set to classify images.

References

1. COVID-19 Coronavirus Pandemic. worldometer: https://www.worldometers.info/coronavirus/, 2020
2. Fan L, Li D, Xue H, Zhang L, Liu Z, Zhang B, Jin Z (2020) Progress and prospect on imaging diagnosis of COVID-19. Chin J Acad Radiol 3(1):4–13
3. Hu Z, Song C, Xu C, Jin G, Chen Y, Xu X, Shen H (2020) Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. Sci China Life Sci 63(5):706–711
4. Samei E, Pelc NJ (2019) Computed tomography: approaches, applications, and operations. Springer, Berlin

5. Zonneveld FW (1987) Computed tomography of the temporal bone and orbit. J Comput Assist Tomogr 12(3):540

6. Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. J Mach Learn Res 3:1157–1182

7. Zheng, C., Deng, X., Fu, Q., Zhou, Q., Feng, J., Ma, H., & Wang, X. (2020). Deep learning-based detection for COVID-19 from chest CT using weak label. MedRxiv

8. Jin, S., Wang, B., Xu, H., Luo, C., Wei, L., Zhao, W., & Xu, W. (2020). AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system in four weeks. MedRxiv

9. Singh D, Kumar V, Kaur M (2020) Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based convolutional neural networks. Eur J Clin Microbiol Infect Dis 39(7):1379–1389

10. Angelov, P., & Almeida Soares, E. (2020). SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification. medRxiv

11. Li Y, Wei D, Chen J, Cao S, Zhou H, Zhu Y, Zheng Y (2020) Efficient and effective training of COVID-19 classification networks with self-supervised dual-track learning to rank. IEEE J Biomed Health Inform 24(10):2787–2797

12. Hu S, Gao Y, Niu Z, Jiang Y, Li L, Xiao X, Yang G (2020) Weakly supervised deep learning for covid-19 infection detection and classification from ct images. IEEE Access 8:118869–118883

13. Wu YH, Gao SH, Mei J, Xu J, Fan DP, Zhang RG, Cheng MM (2021) Jcs: an explainable covid-19 diagnosis system by joint classification and segmentation. IEEE Trans Image Process 30:3113–3126

14. Sun L, Mo Z, Yan F, Xia L, Shan F, Ding Z, Shen D (2020) Adaptive feature selection guided deep forest for covid-19 classification with chest ct. IEEE J Biomed Health Inform 24(10):2798–2805

15. Jaiswal A, Gianchandani N, Singh D, Kumar V, Kaur M (2020) Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2020.1788642

16. Ardakani AA, Acharya UR, Habibollahi S, Mohammadi A (2021) COVIDiag: a clinical CAD system to diagnose COVID-19 pneumonia based on CT findings. Eur Radiol 31(1):121–130

17. Zhou T, Lu H, Yang Z, Qiu S, Huo B, Dong Y (2021) The ensemble deep learning model for novel COVID-19 on CT images. Appl Soft Comput 98:106885

18. Ardakani AA, Kanafi AR, Acharya UR, Khadem N, Mohammadi A (2020) 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 121:103795

19. Kumar SU, Inbarani HH (2017) Neighborhood rough set based ECG signal classification for diagnosis of cardiac diseases. Soft Comput 21(16):4721–4733

20. Inbarani HH (2015) A novel neighborhood rough set based classification approach for medical diagnosis. Procedia Comput Sci 47:351–359

21. https://github.com/UCSD-AI4H/COVID-CT

22. El-Baz A, Jiang X, Suri JS (2016) Biomedical image segmentation: advances and trends. CRC Press, London

23. Zubair AR, Alo OA (2019) Grey level co-occurrence matrix (GLCM) based second order statistics for image texture analysis. Int J Sci Eng Investig 8(93):64–73

24. Hu Q, Yu D, Liu J, Wu C (2008) Neighborhood Rough set based heterogeneous feature subset selection. Inf Sci 178(18):3577–3594

25. Hu Q, Yu D, Xie Z (2008) Neighborhood classifiers. Exp Syst Appl 34(2):866–876

26. Du Y, Hu Q, Zhu P, Ma P (2011) Rule learning for classification based on neighborhood covering reduction. Inf Sci 181(24):5457–5467

27. Lin, T. Y. (2001). Granulation and nearest neighborhoods: rough set approach. In: Granular computing (pp 125–142). Physica-verlag: Heidelberg

28. Yong L, Wenliang H, Yunliang J, Zhiyong Z (2014) Quick attribute reduct algorithm for neighborhood Rough Set model. Inf Sci 271:65–81

29. Wu X, Kumar V, Quinlan JR, Ghosh J, Yang Q, Motoda H, Steinberg D (2008) Top 10 algorithms in data mining. Knowl Inf Syst 14(1):1–37

30. Patel BN, Prajapati SG, Lakhtaria KI (2012) Efficient classification of data using decision tree. Bonfring Int J Data Min 2(1):06–12

31. Lepetit V, Fua P (2006) Keypoint recognition using randomized trees. IEEE Trans Pattern Anal Mach Intell 28(9):1465–1479

32. Ali J, Khan R, Ahmad N, Maqsood I (2012) Random forests and decision trees. Int J Comput Sci Issues (IJCSI) 9(5):272

33. Stuart A (1994) Kendall's advanced theory of statistics. Distribution theory, vol 1. Wiley, London

34. Huang X (2018) An improved KNN algorithm and its application in real-time car-sharing prediction. Dalian University of Technology, Daian, China
35. Vapnik VN (1995) The nature of statistical learning theory. Springer, New York
36. Thapa S, Zhao Z, Li B, Lu L, Fu D, Shi X, Qi H (2020) Snowmelt-driven streamflow prediction using machine learning techniques (LSTM, NARX, GPR, and SVR). Water 12(6):1734
37. Bekkar, M., Djemaa, H. K., &Alitouche, T. A. (2013). Evaluation measures for models assessment over imbalanced data sets. J Inf Eng Appl, 3(10)
38. Sokolova M, Lapalme G (2009) A systematic analysis of performance measures for classification tasks. Inf Process Manage 45(4):427–437
39. Demsar J (2006) Statistical comparisons of classifiers over multiple data sets. J Mach Learn Res 7:1–30
40. Ganesan J, Inbarani HH, Azar AT, Polat K (2017) Tolerance rough set firefly-based quick reduct. Neural Comput Appl 28(10):2995–3008
41. Sayed GI, Hassanien AE, Azar AT (2019) Feature selection via a novel chaotic crow search algorithm. Neural Comput Appl 31(1):171–188
42. Inbarani HH, Kumar SU, Azar AT, Hassanien AE (2018) Hybrid rough-bijective soft set classification system. Neural Comput Appl 29(8):67–78
43. Kumar SS, Inbarani HH, Azar AT, Polat K (2017) Covering-based rough set classification system. Neural Comput Appl 28(10):2879–2888
44. He K, Cao X, Shi Y, Nie D, Gao Y, Shen D (2018) Pelvic organ segmentation using distinctive curve guided fully convolutional networks. IEEE Trans Med Imaging 38(2):585–595
45. Xiao Y, Zijie Z (2020) Infrared image extraction algorithm based on adaptive growth immune field. Neural Process Lett 51(3):2575–2587