Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
An intelligent deep convolutional network based COVID-19 detection from chest X-rays

Mohammad M. Alshahrni, Mostafa A. Ahmad, Monir Abdullah, Nadir Omer, Muzzamil Aziz

College of Computing and Information Technology, University of Bisha, Bisha, Saudi Arabia
Gesellschaft für wissenschaftliche Datenverarbeitung mbH GWDG, Göttingen, Germany

Received 24 July 2022; revised 30 August 2022; accepted 7 September 2022
Available online 14 September 2022

Abstract Coronavirus disease-2019 (COVID-19) seems to be a fast spreading contagious illness that affects both humans and animals. This catastrophic deadly virus has an impact on people's daily lives, their wellbeing, and a nation’s economy. According to a clinical research of COVID-19 affected patients, these individuals have been most commonly infected with a lung illness after coming into touch with the virus. A chest X-ray (also known as radiography) or a chest CT scan seems to be more efficient imaging techniques for detecting lung issues. Nonetheless, when compared to a chest CT, a significant chest X-ray remains a less expensive procedure. Thus, in this research, a novel Deep convolution neural network algorithm is presented to detect the COVID-19 from X-ray image. Moreover, to enhance diagnostics sensitivity and reduce error rate, a hybrid Two-step-AS clustering approach with Ensemble Bootstrap aggregating training and Multiple NN methods used. In addition, TSEBANN model has been employed to explore the qualification procedure effects. The proposed algorithm was trained before and after classification while compared to traditional Convolutional Neural Network (CNN). After, the process of pre-processing and feature extraction, the CNN strategy was adopted as an identification approach to categorize the information depending on Chest X-ray recognition. These examples were then classified using the CNN classification technique. The testing was conducted on the COVID-19 X-ray dataset, and the cross-validation approach was used to determine the model’s validity. The result indicated that a CNN system classification has attained an accuracy of 98.062 %.

1. Introduction

Coronavirus, a huge group of infections that might affect people and cause significant sickness, is a major risk to human-kind. COVID-19 has compelled individuals throughout the nation to use surgical masks, which resulted in slowing the...
nation’s economic development. COVID-19 is already increasing quickly over practically each of the nation’s regions. It has already impacted a significant proportion of individuals, and it is growing every day. Although humanity is acquainted with several coronaviruses, the current outbreak is indeed the result of a unique wildlife sickness [1]. The phrase “new” is attributed to the idea because people are still to be infected by the disease, and “anthropogenic” indicates the disease’s transfer from species to species. Individuals appear to lack an intrinsic measure of assurance due to the uniqueness of the bacteria, which contributes to inhibiting the growth and effects of other infections. Infections, especially those that are recently identified, have the ability to spread locally or globally. Throughout this section, influenza is characterized as an epidemic of every contagious disease resulting in a significant rise in illness and death across a vast particular area, whereas an epidemic is described as a disease that usually develops. The Severe Acute Respiratory Syndrome (SARS) virus lately afflicted 8,095 people and killed over 775, while the chickenpox outbreak attacked many across the nation. The very first table contains the much more dangerous diseases of the previous decades, along with breakouts and pandemics associated with them. The virus infecting COVID-19 (the condition’s formal term) is SARS-CoV-2, which belongs to a large group of germs that frequently potentially cause breathing problems. The SARS-CoV infection began spreading in 2002, when it was discovered in Guangdong, China. In the Middle East, a new coronavirus epidemic (MERS-CoV) began in 2012, resulting in a severe respiratory ailment known as Middle East Respiratory Syndrome [2].

858 people died from the infection, and another 2494 people were infected. Coronavirus is currently fighting the world, and means have started growing in December 2019, though it is still up for contention. COVID-19 is a contagious disease that spreads primarily through indirect or direct contact with the infected humans, breathing particles produced whenever a human resource is supplied or disease has indeed happened, or droplets in the environment (airborne transmission) [3]. A sign of infection, a high temperature, and breathing problems are major markers that can lead to additional internal bleeding or serious respiratory issues, and even death in some cases. The ability to combat the infection is limited due to its easy transmission as well as its frequency of just evolving significantly and persistently amongst humans. To put it differently, there seems to be a finite number of researchers and clinicians, and once the outbreak hits the peak of transmission, there might not be any sufficient capability to fully contain it. As a consequence, there is a clear need for technologies that might also help such professionals do their respective duties more efficiently. It might entail the creation of client management apps, analytics, statistics visualizations, machine learning techniques, and other AI-based solutions [4].

It has been stated that in order to combat the development of coronavirus disease, mandatory patient challenging and quick treatment symptoms for infected individuals are critical. The Reverse Transcription Polymerase Chain Response (RTPCR) test on pulmonary samples is indeed the best quality level screening approach utilized for assessing COVID-19 individuals. Although it is the most commonly used technique for COVID-19 detection, it is a laborious, perplexing, persistent, and inefficient approach with a positive rate of only 64%. Symptomatology assessment, epidemiological information, positive radiography imaging (chest radiograph (CXR)/computed tomography (CT)), and positive microbiological tests are some of the additional tools used for identifying COVID-19. Bronchopneumonia with dyspnea, fever, cough, and pulmonary insufficiency with acute respiratory distress syndrome (ARDS) are among the diagnostic features of significant coronavirus exposure [5]. Another important symptomatic tool for coronavirus is the quickly available radiographic scanning. In the preliminary phase, many COVID-19 cases have similar radiography features, including ground-glass opacities, multifocal and reciprocal with an outlying or posterior dispersion mainly in the inferior projection, and pulmonary consolidation in the later phase. Although normal chest X-ray scans could assist in the early recognition of probable occurrences, the imaging of assorted virus-related infected patients is comparable to and combined with those of other infectious and parasitic lung ailments [6]. As a result, radiology has difficulty in distinguishing COVID-19 from those other viruses’ pneumonia. In the current situation, where treatment centers are overburdened and work continuously, the adverse reactions of coronavirus, which are comparable to viral pneumonia, could result in an erroneous diagnosis. As a consequence, an inaccurate diagnosis could mean to a non-COVID viral infection existence mistakenly recorded as extremely suspect of consuming COVID-19, leading in behavior deferment and the threat of presenting to optimistic coronavirus individuals.

Several investigators were focusing on machine learning techniques in order to recognize hidden connections and extract important information from the raw data. Sentiment classification, rainfall forecasting, network intrusion detection, software fault prognostication, and healthcare and health data gathering are among the study disciplines where machine learning has played a significant role. Furthermore, many studies have been focusing on utilizing machine learning systems to identify COVID-19 victims through examining patterns in CBC testing results, and some of these studies are presented below. Utilizing full blood lab tests, researchers constructed a machine learning technique [7]. Classification, dimensionality reduction, regression, clustering, fault diagnosis, and incentive maximization are all common ML methods. The ML models are implemented on labeled sets of data in the SL paradigms, meaning that they will have a ground-truth result (discrete or continuous) for each intake. In UL, on the other hand, there is no ground-truth output, and the methods generally try to find trends in the data. The algorithm tries to increase the reward function, rendering it better suited to consecutive decision-making problems. Regression and categorization are features of supervised methods; unsupervised machine learning features clustering algorithms and dimension reductions; and reinforcement learning (RL) features categorization and management [8].

Artificial intelligence (AI) is progressively important for the development of new knowledge and techniques. With rapid decision-making, enhanced image processing, and priority-based therapy, AI applications and services are altering and lifesaving. AI algorithms that ultimately aid the decision-making framework may lessen the differences in the analysis of a significant quantity of information in the form of therapeutic data and laboratory findings [8]. Additionally, AI technologies have revolutionized testing and diagnosis procedure. Artificial intelligence-based smartphone software has already been created to gather and analyze patient data distinctly,
An intelligent deep convolutional network based COVID-19 detection from chest X-rays

artificial intelligence (AI), focusing on machine learning and deep learning, could help in the fight with COVID-19 in numerous ways. Machine learning can be utilized by academics and medics to examine big quantities of information in order to calculate the development of COVID-19, serve as an early warning system for possible outbreaks, and detect underserved communities. In a broad range of image application domains, particularly picture separation, machine learning has shown impressive effectiveness in image analysis and image classification. Classification techniques and imaging sequences and SIFT are utilized for identifying major characteristics from photos, which would then be used in models trained with classifying systems such as SVM. In contrast to handmade parameters, the DNN-based approach performs exceptionally well at categorizing photos based on retrieved properties.

Several efforts are exerted to classify lung X-ray images in the coronavirus clinical study or the healthy occurrence classification utilizing machine learning approaches, according to ML parameters. Deep learning methods were used in several of these efforts. For example, scientists created a convolutional network model for identifying COVID-19 from chest X-ray pictures on the fly. Using the MobileNet framework, classification performance is stated to be 96.78%. The study employed the following algorithms. Inception-ResNetV2 and InceptionV3 have claimed estimation accuracy of 95% and 82%, correspondingly. Orthogonal moment and its derivatives have lately proven to be particularly useful in a range of data acquisition and data processing techniques. The use of image instant to extract information from images has indeed been demonstrated to be useful in a multitude of conditions. Combining orthogonal quaternion harmonic transition events with picture display and attribute selection ideal condition, for instance, is successful in classifying color galaxy photos.

The goal of the research was to progress and design an AI-based customizable scanning analytical model. The scientists used a variety of databases and assays to track the progression of COVID-19 sickness in every individual across periods. It was achieved through the use of the 3D measurement technique, which resulted in the production of the “Corona Score.” According to the results of the experiment, 157 people with ubiquitous CTs have been included in the study. Utilizing pre-existing methodologies, the proposed A component examines clusters or sections of focused opacification and instance proportions in 3 components. Subsystem B has indeed again created two-dimensional assessment of every layer enabling powdered glasses penetration segregation and containment, as well as larger diffused opacities. In contrast, a case-by-case therapy has indeed been proposed to address the increased transparency caused by diseases. Because it concentrates on the proposed detection operations, existing lung disease testing approaches are ineffective in identifying progressively widespread GG opacification. In, a network known as the COVID-Net is offered as an instance of a hybrid computer design methodology. A publicly available benchmarking database has indeed been created from this database, dubbed “COVIDx,” with manufacturing being the production of training and validation of the system, COVID-Net, even more effectively.

To train and analyze COVIDx for training and analyzing the COVID-Net suggested in this work, a total of 13,976 chest X-ray images from about 13,875 patient cases were employed. Furthermore, a research conducted a survey of virologists, communicable diseases experts, and politicians who will benefit from the investigation of medical and technology components of the COVID-19 pandemic. The report also studied how various technology methods are being used in the background of COVID-19, as well as a variety of artificial intelligence techniques which have been proposed to help with the COVID-19 global epidemic, starting from initial prognosis through the use of image diagnosing to designs that could also clarify the condition’s propagation and anticipate different potential epidemic locations. Supervised learning, a popular area of artificial intelligence (AI), facilitates end-to-end designs to be constructed in terms of generating expected outcomes using model parameters without a need to individually feature extracted and has recently intrigued the attention of the health industry as a vital resource for doctors. Research employed X-ray pictures to classify features gathered from multiple convolutional network models, followed by an SVM classifier to
Coronavirus is a contagious agent that has recently spread across the world. It is a global pandemic that primarily causes shortness of breath. If not diagnosed and treated early, it poses significant hazards including blood coagulation, cardiac arrest, and, in the majority of cases, mortality. The aim of the present study is to identify COVID-19 in its initial stages to reduce its number of fatalities. Early detection of the coronavirus reduces economic losses due to its significant influence on labor and manufacturing. For the first period, COVID-19 has been detected utilizing blood work that is both simple to obtain and inexpensive. COVID-19 was detected and have its own dataset that researchers created. The dataset consists of 134 examples of various males and females, ages, and positive and negative instances. Blood tests are included in the database (CBC, ALT, D-Dimer, CRP, S-Ferritin, and LDH). Doctors could use this blood research to test whether or not a patient is infected with COVID-19. The method employed a machine learning approach, including Random Forest (RF), Support Vector Machine (SVM), and Naive Bayes classifications. The accuracy of RF was 76%, SVM was 88%, and Naive Bayes classifier was 85%. The model used the classifier model because it had the highest accuracy. As per the American Centers for Disease Control and Preventives (CDC), PCR has a reliability of 80% and is less expensive than blood and urine tests, and researchers accomplished this by using the approach [16].

2.1. Problem statement

A massive proportion of reported incidents must be evaluated for proper quarantine and therapy to limit the transmission of COVID-19. Infective advancement screening is the finest quality grade, although it is inefficient and results in several incorrect negatives. To combat the disease, a quick and accurate variety of applications are essential. Because of COVID-19 radiography alterations in X-ray images, researchers intended to develop a deep learning method for removing the graphic aspects of COVID-19 in order to provide a therapeutic assessment before the pathologic testing, thus conserving vital time for disease prevention [17]. A machine learning procedure is utilized for identifying the chest X-ray pictures in this research. Because precision has been the most important component in this problem, the DCNN performance can be enhanced by taking a larger collection of descriptions for training the neural network and doubling the value of repetitions. Google’s Tensor Flow is a huge machine teaching method, and Google’s convolution layer framework is Inception V3. Hence to overcome this issue, the detection of COVID-19 founded on chest X-ray image is developed by means of deep convolution layer [18].

3. Methodology

3.1. Dataset

Chest X-ray images of 342 COVID-19 samples were selected from Dr. Joseph Cohen’s Freeware GitHub Repository for this research. This resource contains chest X-ray/computed tomography (CT) pictures of individuals with ARDS, bronchitis, Middle East respiratory (MERS), COVID-19, and severe acute respiratory disorder (SARS). From the “ChestX-ray8” database, 2800 typical (healthy) chest X-ray pictures were chosen. Furthermore, 2773 bacteria and 1495 virus pneumonia chest X-ray pictures from Kaggle’s "Chest X-ray Images (Pneumonia)" database were employed. The experiments used chest X-ray pictures from three input generated sets of data [19]. Table 1 displays the distribution of photos by class in the generated databases. Within the training sample, the data augmentation strategy was adopted with a step function of 1/256, shear variation of 0.2, zoom region of 0.2, and horizontal flip permitted. In the data sources, all pictures were rescaled to 224 x 224-pixel resolution [20].

Deep learning is a type of pattern recognition influenced by the structure of the human brain. Deep neural networks have managed to show remarkable results in biological image recognition, as well as many other fields, in latest dates. It is anticipated that significant data could be gleaned by combining deep learning techniques into patient records. Deep learning algorithms have been widely utilized to segment, categorize, and identify lesions in medical information. Through the use of deep neural networks, picture and signaling information from medical data modes including magnetic resonance (MRI), X-ray, and computed tomography (CT) were analyzed [21]. As a consequence of these investigations, ailments including skin cancer, diabetes, brain tumors, and breast cancer can be detected and diagnosed in research with simplicity. Convolutional neural networks (CNNs) are a kind of deep learning model utilized for object recognition. For CNN to function, the images provided as inputs should be predictable by the computer and translated into a detailed integration [22].

3.2. Data sampling

Researchers intended to use exclusively genuine X-ray, ultrasonography, and computed tomography information in this project, and they will not explore creating and using simulated data at this time. For the model experiments, researchers also utilized a rather equal database of images, with imbalances resolved through the use of computed class training values [23]. Researchers constructed a master database for all tests using the source information listed in Table 2. A sample photo from the X-ray COVID-19 collection was minimally processed to remove improperly identified projections. Again, for multiple methods, all of the other COVID-19 photos would be included. Non-COVID-19 photos also were gently vetted to eliminate incorrectly labeled projects and images occupied by obstructive surgical equipment. This provided us with acceptable X-ray, ultrasonography, and computed tomography image examples. Non-COVID-19 photos too were gently vetted to eliminate mislabeled estimates and images occupied by obstructive medical instruments. This provided us with acceptable X-ray, ultrasonography, and CT scan imaging sets [24]. Because the resulting sample source was too limited for a deep tool for learning, researchers utilized several data preprocessing transitions during the experimental procedure, such as horizontal flip and rotation, and vertical shift and rotation, to boost the amount and complexity of the sample group, as shown in Table 2.
Table 1  Image dataset.

| Database            | Bacterial pneumonia | COVID-19 | Viral | Normal |
|---------------------|---------------------|----------|-------|--------|
| No. 1               | *                   | 342      | *     | 2800   |
| No. 2               | *                   | 342      | 1495  | *      |
| No. 3               | 2773                | 342      | *     | *      |

Table 2  Sample dataset.

| Assortment            | No. of data | Representative | Note                                           |
|-----------------------|-------------|----------------|------------------------------------------------|
| Data collection based on COVID-19 | 116 (COVID-19) | Size, quality, contrast, and brightness | This article’s only supply of public information on corona PA X-ray pictures. |
| Chest X-ray image (NIH) | 323 (pneumonia) 60,362 (NF) | The COVID-19 dataset has comparable intra-dataset consistency. The photos are all 1024 by 1024 pixels in length. | The COVID-19 image collection of data is practically highly comparable. In this investigation, it has been used. |
| COVID-CT              | 350 (COVID-19) 398 (Non-COVID) | Size, quality, contrast, and brightness | The only publicly available COVID-19 CT images are often used in this research. |
| POCOVID-Net Data      | 655 (COVID-19) 278 (pneumonia) 173 (NF) | Size, quality, contrast, and brightness | This research used only access to the public COVID-19 ultrasound pictures available. |

3.3. Fractal technique for Feature extraction

Extraction of features reduces the amount of material needed to adequately describe a huge quantity of data. One of the primary issues when undertaking a sophisticated analysis of the data is the quantity of different quantity. A huge number of parameters necessitates a lot of storage and computer resources or a classification technique that services the instructive sample and generalizes to new circumstances. Removal of features is an extensive word that refers to approaches for striking the composed set of factors to tackle high-precision challenges [27]. The goal of image examination is to make an exclusive way to showcase the basic properties of pictures. A grey surface matrix is generated in the fractal methods used to achieve statistical parameters. The picture characteristics of the probability density function of the measured substances are derived in descriptive statistics from the intensity of light of the specified sites compared to one another in the visual [28]. The amount of strength opinions (pixels) in each combination affects the statistic.

The fractal approach is employed to excerpt the characters in the research. Feature selection was utilized to lower the dimensionality and pick the most important characteristics that may adequately differentiate the dissimilar categories to cope with excessive input information. To generate the eigenvectors form of the picture and lower the dimensionality, the fractal algorithm was used in conjunction with covariance examination. The input for the fractal method is the same dimension, and one picture is mentioned to as two-dimensional matrices and a single vector [29]. Grayscale photographs with a certain quality are essential. By resizing matrices, every piece of information is transformed into a column vector, and the pictures are retrieved from the matrices of size $X \times Y$, where $Y$ is the no of pixels for each picture and $X$ is the number of photos. To regulate the standard deviation of each actual picture, the average image should be determined. The correlation coefficient is then computed, and the correlation matrix’s eigenvectors and eigenvalues are formed. The fractal computation strategy is that $X$ represents the number of instructional pictures, $B_i$ represents the average of the pictures, and $c_i$ represents each picture in the $A_i$ vectors. There are originally $M$ pictures, each of which has the $Y \times Y$ dimensions. Every picture could be presented in an N-dimensional field (Eq. (1)), and average procedures could be performed in Eq. (2).

$$M = Y \times Y \times X$$  \hspace{1cm} (1)

$$B_i = \frac{1}{X} \sum_{x=1}^{X} A_s$$  \hspace{1cm} (2)
The fractal technique calculates determining the standard error to be a critical issue, which would be likewise determined using Eq. (3) and the covariance matrices from Eq (4).

\[ V = \frac{1}{X} \sum_{i=1}^{X} A_i - B_i \]  

\[ \text{Covariance} = VV^T \]  

In the above Eq. (3) \( M = [\text{Var}_1, \text{Var}_2, \ldots, \text{Var}_n] \) and Covariance = \( Y^2 \times Y^2 \) and M = \( Y^2 \times X \). As a result, Covariance equals a huge number. The eigenvalues of Covariance may now be found using Eq (5).

\[ C_i = AD_i \]  

3.4. Feature extraction process

- **Statistical Feature**
  
  When it is examined the X-ray photos more closely, it can be seen that the outstanding texturing and probabilistic correlations are also most probably the much more conspicuous clear representation within. In the latest days, a growing number of scholars are begun to use texture and statistic qualities to simulate classifying challenges, and the trend is projected to continue. The success of this work stems from its accessibility, as compared to computer engineering, which is typically a time-consuming procedure that necessitates a thorough knowledge of issue classifications and methodologies that allow for hand layout descriptions. That procedure isn’t necessary [30]. Despite unmanufactured descriptions getting some advantages, that’s worth noting because handmade descriptions offer unique qualities which can be useful for a classification task. For example, throughout this circumstance, the benefits of employing handmade characteristics outweigh the drawbacks since those techniques are much more effective since they often identify the element that defines a problem in a somewhat more definite way [31]. Whenever handcrafted elements are employed instead of unpolished components, a more accurate description of the relying primarily upon the manufactured properties of the photos is available. Despite the fact that attempts have turned to using the other two categories in extracting features, that wasn’t always the case. As a consequence, it can be evaluated the two separately before combining the data of many two studied groups to arrive at an outcome. In this way, it might take advantage of the complementary description techniques, as demonstrated in and to protect individuals against trying the same thing during performing a multiclass classification. In Table 3, characteristics are retrieved and the derivation is discussed. The article specifically mentions the descriptors that were used to characterize the performance. To achieve outstanding performance and effectiveness purposes or, most specifically, in computer-aided diagnostic activities, particular texture descriptors have been used. Here are some instances of statistical characteristics [32].

While \( z_i \) is the distribution of the area level of intensity, \( p(z_i) \) represents a randomized indicator of brightness factors. The standard error of various levels of intensity is represented by the letter \( L \).

The GLCM approach was utilized to create the Material characteristics grouping. Research developed the GLCM notion, which is a connection among pixels in matrices that are commonly used in texturing research. Whenever multiple pixels were placed side by side, the distances among them and the angles across their individual axis are used to define the connection. As a result, the GLCM variables are the satellite’s length and degrees [33]. The GLCM algorithms measure the timing at which a combination of cells having distinct characteristics and in a specific spatial connection appears in an image, hence quantifying its material. The GLCM creates the matrices of matched images having different ideals in a certain spatial relationship, out of which statistics measurements are extracted. As stated previously, statistical assessments of textured filtering functionalities and spatial relationships of image pixels proved determined to be inadequate for information delivery regarding form in feature extraction. These second-order characteristics are used to build the GLCM features and functionality. The total median of measures of similarity among adjacent pixels in different aspects (homogeneity, uniformity, etc.) can be used to determine the reflection. Among the major factors influencing the GLCM’s differentiating capacities is pixel spacing [34]. The relationship among image pixels (i.e., short-term neighborhood connectivity) is depicted whenever position 1 is accounted for. However, changes in the Euclidean distance correlate to variation in the number of related pixels.

- **GLCM Features**

  GLCM includes functionality that identifies the attribute connection amongst pixels in the textured image, as described in statistics and structure texturing method. The variables to be examined are influenced by several formulas for obtaining attributes from the co-occurrence matrix. It identified the four more relevant Haralick texture components for more analysis of the characteristics of the X-ray imagery collections, including such correlations, uniformity, power, and brightness. Table 4 lists many methods for calculating GLCM Texture Features.

3.5. Preprocessing

The proposed dataset is augmented during pre-processing using an augmentation approach dubbed “Image Data Generator.” The image data source generates additional data by making little adjustments to the image, such as brightening, magnification, horizontally flipping, and moving. COVID Image database was processed using Image Data Converter, which included pre-processing (such as edge sharpness, luminance, and so on) because the database was smaller than the other subcategories in the database and the following Table 5 shows the dataset of augmentation images after augmentation [35].

3.6. Deep convolutional neural network

Convolutional neural networks (CNN), which teach multiple levels efficiently, are amongst the most important deep learning methods. To reduce preparation, CNN employs a range of multi-layered views. A Convolutional system is made up of three main layers: the convolutional layer, the max - pooling, the fully-connected layers, and several layers upon layer
that fulfill various purposes. There are two levels of training for each architectural style: forward and back-propagating [36]. The input goes from the concealed layer to the output nodes in the forward step, and the data only change forward. The input image transfers the information to the hidden layer, which subsequently generates the output. When a neural system is established in backpropagation, the input image is input to the system in the first step. When the output is achieved, the network error value can be determined, and this number, together with the cost function diagrams, is returned to the system to update the weights in the network as seen in Fig. 1. The above are the many sorts of hidden sublayers found in CNN.

### 3.6.1. Convolution neural network

Convolutional System are equipped to accommodate multidimensional data including pictures and data over the period. The feature extraction process and weight computation throughout the training stage are examples of this. A convolutional operation is used to determine the composition of these systems, which is essential for tackling complicated work [37].

The basis of the convolution network is the convolution layer, which produces a three-dimensional stack of the neuron as outputs. In simple terms, each layer produces a three-dimensional mass. The CNN network convolves the contribution picture as well as the central feature maps in such regions using a different kernel. There are three key advantages to using a pooling layer:

- Each feature map’s scale factor approach reduces the dimensionality dramatically.
- The local connection establishes a link between adjacent pixels.
- It makes changes in the object’s location more stable.

Table 3 Extracted Features and their Formulation.

| Feature No | Feature Name | Formulation | Description |
|------------|--------------|-------------|-------------|
| 1          | Mean         | \( m = \sum_{i=0}^{L-1} z_i * p(z_i) \) | The Mean of a set of adjacent pixels in a picture is a metric indicating the average brightness of the pixels. |
| 2          | Standard Deviation (SD) | \( \text{STD} = \sum_{i=0}^{L-1} (z_i - m)^2 * p(z_i) \) | The SD is a measure of how widely spaced out a set of data is. |
| 3          | Skewness     | \( \text{Skewness} = \sum_{i=0}^{L-1} (z_i - m)^3 * p(z_i) \) | The skewness, or more precisely, the absence of symmetry, is a metric for measuring symmetrical relationships. In this case, the distribution or data set is symmetric if it is identical to both the right as well as left in the center. The skewness of a regular distribution is 0, and the skewness of a few symmetrical samples must be close to zero. Skewness numbers that are negative indicate data that is skewed to the left, while positive values suggest data that is skewed to the right. |
| 4          | Contrast     | \( \text{Contrast} = \sum_{i=0}^{L-1} \sqrt{(z_i - m)^2 * p(z_i)} \) | The difference in brightness and/or color that differentiates the object is referred to as the contrast. Contrasting is defined as the gap in color intensity among a focus as well as other entities in the very same range of view in real-world visual processing. |
| 5          | Kurtosis     | \( \text{Kurtosis} = \sum_{i=0}^{L-1} (z_i - m)^4 * p(z_i) \) | Kurtosis is a statistic that indicates whether data are peaks or flat when compared to a normal distribution. In other words, datasets with a high kurtosis seem to have a distinct, near-normal peak. Measures the changes in relative intensity within a given area. |
| 6          | Smoothness   | \( \text{Smoothness} = 1 - \frac{1}{\text{STD}} \) | The SD is a measure of how widely spaced out a set of data is. |

### 3.6.2. Contrast

Contrast is a metric that indicates the gap in color intensity among a focus as well as other entities in the very same range of view in real-world visual processing.

### 3.6.4. Energy

Energy is a metric for the changes in relative intensity within a given area.

### 3.6.5. Smoothness

Smoothness is a statistic that indicates whether data are peaks or flat when compared to a normal distribution. In other words, datasets with a high kurtosis seem to have a distinct, near-normal peak. Measures the changes in relative intensity within a given area.

Table 4 Formulas for the GLCM Texture Features.

| Feature No | Feature Name | Formulation |
|------------|--------------|-------------|
| 1          | Correlation  | \( \sum_{i,j=0}^{N-1} \frac{p_{ij} \times (i-j)^2}{\sigma(i)} \) | Measures the changes in relative intensity within a given area. |
| 2          | Homogeneity  | \( \sum_{i,j=0}^{N-1} \frac{p_{ij} \times (i-j)^2}{\sigma(i)^3} \) |
| 3          | Energy       | \( \sum_{i,j=0}^{N-1} \frac{p_{ij} \times \ln(p_{ij})}{(i-j)^2} \) |
| 4          | Contrast     | \( \sum_{i,j=0}^{N-1} p_{ij} \times (i-j)^2 \) |

Table 5 Total No of Images after Augmentation.

| Augmentation | Augmentation image after Covid-19 | CXR Normal image | Descriptions of Viral and Pneumonia |
|--------------|-----------------------------------|-----------------|------------------------------------|
| 206          | 1362                              | 1352            | 1246                               |
health information basis on the results of the Smyth investigation utilizing information about the study. The proposed strategy was using a Two-step grouping technique for heterogeneity unbounded nonlinear functions to accommodate numerical or categorical inputs: firstly, a joint mix of probability model and Gaussian, and secondly, a combined mixture of probability model and Gaussian. Then, to deal with category input order, a hidden, Markov network was combined with a variety of different models. The method was pushed to the limits on a proper machine, and the results suggest that it is successful in identifying health care services for huge families [38].

Deneshkumar et al [48] using 2 different grouping techniques as well as other data analysis methodologies, proposed a method for finding anomalies and evaluating the effect size in diabetic patients. The investigators aimed to find connections and patterns amongst an enormous volume of healthcare data in order to get new treatment insights [39]. A novel approach for revealing natural clusters (or groups) within an overall framework for collecting data is linked to the Two-step-AS clustering algorithm Two-step-AS Cluster is an exploratory technique for uncovering natural groups (or clusters) within a data gathering which would otherwise go undiscovered. Thus, an approach makes it through a technique that had several favorable qualities that differentiate it from ordinary grouping techniques.

1. Working either with the classified or dependent variable. If the parameters are believed to be independent of each other, a combined multinomial-normal distribution can be employed for discrete and numerical data.
2. The clustering algorithm is chosen for you dynamically. That used an optimization approach that analyses levels of model-choice criteria throughout alternative grouping methods, it is necessary to detect the optimal parameters immediately.
3. In Scaling, the Two-step-AS approach creates a Cluster Feature (CF) tree that summarizes the elements in each group in evaluating large data sets.

Retailers and consumer items companies, for instance, frequently employ grouping techniques to analyze information on their customers’ buying habits, sexuality, aging, level of income, as well as other factors. Such companies’ branding and marketing production approaches are geared to every individual customer group in order to increase sales and build brand loyalty. Pre-clustering is conducted by screening every dataset and assessing if that could be allocated to one of the pre-existing groupings or if it has to be assigned to a clustering process depending on the separation requirement [40]. Just after the pre-clustering process, segmentation is done. Researchers used the Two-step-AS technique of log-likelihood separation in combination with the log-likelihood similarity in the study.

The pre-clustering procedure begins with the creation of a database table called the CF (cluster feature) tree, which contains the group centers as well as many other necessary details. Stages are used to organize nodes, with every degree comprising a certain number of members of every node. A leaf entry is indeed the cluster’s final sub-cluster. Every piece of information is found by climbing the CF tree iteratively, starting at the parent node and working straight to the local child node. Whenever the process comes to a leaf node, it looks for the leaf entry which is nearest to it. If a recording is inside a particular range of the nearest leaf entry, it is appended to the leaf entry. If the information is inside a predetermined radius, the CF tree is modified. If it is not the case, a specific value again for the root of the tree is generated [41]. Whether there is enough room in the leaf node for a further value, the leaf is divided into 2 values, every-one of which is assigned to one of the two leaves. As depicted in the figure, the farthest pairing operates as seeds, and the remaining numbers are dispersed depending on the proximity criterion. The clustering process might assemble the sub-clusters into the proper number of clusters using the sub-clusters produced out from pre-cluster phases as inputs (including the noises if the alternate phase has been used). Because the collection of subs is substantially smaller than that of the collection of original information,
standard classification methods might well be useful in this case. It uses an agglomerative hierarchical technique that estimates the number of groups to use in a specific circumstance dynamically. The difference between the two of relevance is determined by the log-likelihood relationship. The log-likelihood value can be used to establish a connection for both categorical and continuous data. When two clusters are combined into the same, the separation separating them decreases due to a decrease in the standardized residuals of the likelihood function, indicating that they have been getting closer. The probability measurement provides a probabilistic model for the parameter’s values of a society. Continuous variables have a distribution layer, whereas measures of variability have a multivariate regression allocation. Every parameter is supposed to be impartial to each other. As demonstrated in Eqn. (6), the clustering separation among Euclidean (x) and log-likelihood (y) [42].

\[ z(x, y) = \xi_k^2 + \xi_j^2 - \xi_{k,j}^2 \]  
(6)

Where,

\[ \xi_i = -N_i \left( \sum_{k=1}^{K} \frac{1}{2} \log \left( \tilde{\sigma}_k^2 + \tilde{\sigma}_{ik}^2 \right) + \sum_{k=1}^{K} \frac{1}{N_i} \tilde{E}_{ik} \right) \]

\[ \tilde{E}_{ik} = -\sum_{k=1}^{K} \frac{N_{iklj}}{N_i} \log \frac{N_{iklj}}{N_i} \]

The characteristics of such formulas are as follows:

- \( K^q \) signifies the scope categorization amount of input data characteristics, \( K^q \) is the symbolic number classification of inputs characteristics, \( L_k \) is the typical amount for the kth representational characteristics, \( N_i \) is the number of incidents in group \( v \), \( N_{iklj} \) is the number of incidents in group \( v \) – that is equivalent to the \( l \)th type of the \( k \)th symbolic feature – \( \tilde{\sigma}_k^2 \) is the plausible variability of the \( k \)th global attributes including every incidence, \( \tilde{\sigma}_{ik}^2 \) is the plausible variability [43].

If \( \tilde{\sigma}_k^2 \) is omitted in the equation for \( \xi_k^2 \) and \( \tilde{\sigma}_{ik}^2 \) is omitted in the equation form, the coolness among groups \( i \) and \( j \) would’ve been correctly equivalent to a decrease in log probability again when the groups are joined [44]. Researcher could be avoiding the problem caused by \( \tilde{\sigma}_k^2 = 0 \) by introducing the component, which would make the presence of the natural logarithm ambiguous. The method is separated into two parts, each of that is used to dynamically specify the number of clusters. Following, the indication Schwarz’s Bayesian Information Criterion (BIC) is generated for all groupings amount from a given Eqn. (7); the indication is then used to establish an approximate solution for the group’s amount in the sec.

\[ BIC(J) = -2 \sum_{j=1}^{J} \xi_j + m_j \log (N) \]  
(7)

\[ m_j = J \left\{ 2K^q + \sum_{k=1}^{K^q} (L_k - 1) \right\} \]  
(8)

\[ f'_{mk} = \left\{ \begin{array}{ll} \rho \cdot \text{binom} \left( N, \frac{1}{2} k \cdot \text{binom} \left( N - \sum_{l=1}^{k-1} f'_{ml}, \frac{l}{2} \right) \right) & \text{oth} \\ \rho \cdot \text{binom} \left( N, \frac{1}{2} k \cdot \text{binom} \left( N - \sum_{l=1}^{k-1} f'_{ml}, \frac{l}{2} \right) \right) & \text{oth} \end{array} \right. \]  
(9)

In regards to some aspects, Two-step Clustering marked a transition similarity measure, and it has numerous benefits. In Eqn. (8), two stages of grouping proportion are derived for starters, grouping parameters could be continuous or categorical, allowing for a greater choice of applicability. Secondly, the Two-Step Clustering approach consumes less storage and performs calculations more quickly [7]. Thirdly, characteristics are employed as a grouping separation indicator, and the information could be “automatically” restructured with both the greatest number of nodes at the very same moment. As a result, the Two-step Clustering method is implemented and investigated as a way to combine the Multiple Neural Network and Ensemble Bootstrap aggregating pattern recognition. Fig. 2 shows the MNN-Ensemble Bagging Learning technique [45].

To have a more accurate interpretation, the hybrids modules integrate the Ensemble Bootstrap aggregating classifier with the Multiple NN. The modules use a hybrid mixed approach to integrate recommendations from independent classifications specifically related while categorizing multiple kinds of X-ray Covid-19 information in an effort to expand classification performance. The hybrid algorithm divides information for training and evaluating methods into two sub-groups using tenfold cross-validation, including one that contains Ensemble Bootstrap aggregating learning while the other doesn’t. The maximum possible performance is achieved by combining the results of the Ensemble Bootstrap aggregating supervised learning with both the Multiple NN classifiers. Every-one of the training/testing trials uses X-ray Covid-19 database features as input parameters, that are subsequently supplied into the Multiple NN. The target field then becomes a property of a specific class (Non-COVID, COVID-19, or Pneumonia) [46]. The outcomes of the Multiple NN classifiers with Ensemble Bootstrap aggregating training result show a performance improvement over preceding Multiple NN methodologies when categorizing the information with Ensemble Bootstrap aggregating learning output.

But on the other hand, a mixture of winner-takes-all and weighted-most-majority methodologies relates to the use of the Ensemble Bootstrap aggregating learning algorithm in conjunction with other treatments. The purpose of this combi-
nation process is to minimize the amount of overdiagnosis while also ensuring the quality of COVID-19 identification in normal individuals. The overall modeling with moderate value projections could benefit from the minority version with significantly greater price estimations, as compared to the minority model having minuscule value predictions, under the modified rule of the majority. In contrast to the prediction’s factors of the single product with the maximum priority, the forecast factors among most predictions with sufficiently large numbers are destroyed in the succeeding procedure [47]. The embedded component is built during the cognitive stage of the Ensemble Bootstrap aggregating procedure. Every Multiple NN receives classification results from the information with each cross-validation fold in the database. To achieve the entire presented with a range, this forecast should be integrated into the general forecasting model.

By averaging the values out from source information and then aggregating these duplicates altogether, bootstrap aggregation (bagging) produces replicas of the training information. It has the advantage of producing training examples that are the same length as the entire dataset. The technique is continued over \( k = 1, \ldots, K \) and \( m = 1, \ldots, M \) iterations to yield frequency values, as shown in equation (10):

\[
{f}_{mk} = \binom{N}{k} \sum_{t=1}^{K} \binom{N - \sum_{i=1}^{t-1} f_{mk} + \frac{t}{C_{1}^{2}}} f_{mk} \quad \text{otherwise}
\]

While \( f_{mk} \) is the simulation incidence frequency for the \( k^{th} \) record of the \( m^{th} \) random subset, and \( f_{mk} \) is the appearance frequency attributed to the \( k^{th} \) record of the \( m^{th} \) random subset.

The training database has a total of \( K \) unique records. The overall dataset is \( N \). After that, a model is created for every duplicate. An ensembles report is the result of combining these algorithms. The network method creates new recordings using one of the ways indicated above; the choices available depend on the target’s measurement levels. Equation (11) is used to evaluate the reliability of the prediction classifier associated with each ensemble’s methodology and base models:

\[
\frac{1}{N} \sum_{k=1}^{K} f_{mk} (y_k - \hat{y}_k)
\]

Where \( \hat{y} \) is the target rate for \( k^{th} \) record \( \hat{y}_k \) is the modal category for the target values.

The main aspect of CNNs is the automatic feature extraction concept. The generated source data is frequently sent to the Extraction Network function, which then sends the retrieved characteristics to the Classifier Networking [48]. In the feature extraction methods, different convolutional and pooling layer pairings are being used. To accomplish a data input convolution procedure, a sequence of convolution filters is processed. The pooling layer is utilized as a density estimation layer that defines the thresholds. Several variables should be modified during backpropagation, which effectively reduces interactions inside the neural network’s architecture [35].

### 3.6.2. Activation function

Activation functions are typically utilized in neural network models to acquire necessary outcomes from systems engineering. Various activation rates could be employed for neural network models, with the sigmoid and nonlinear activation functions being the most essential. The sigmoid function takes the input of any value among + and - and returns a value in the range of 0 to 1. The resulting value of the nonlinear activation operator is between 1 and -1. This functionality reduces the utilization of these two roles on CNN systems [49]. Since the matrix pictures have distinctive features, image information is removed, and the device remains unworkable for most circumstances following applying these methods. The rectified linear unit ReLU function is among the most modern training algorithm to be developed. The activation function \( g \) is performed on all components in ReLU. Its objective is to give the system nonlinear behavior. This method inserts all of the image’s squares and sets all negative signs to zero. The goal of utilizing ReLU is to construct a nonlinear section of the deep convolutional neural network as well as its nonlinear development (convolution is a straightforward procedure that involves doubling and adding the components) [50].

#### 3.6.3. Max pooling

Max-pooling has a variety of consequences in neural networks. The utilization of max pooling enables the system to recognize the item with relatively minimal image modifications initially. Secondly, it provides the system with extra visual space to recognize characteristics. In the CNN, pooling is utilized to recalculate the feature throughout the removing sample job, allowing us to go deeper into the system [51]. The spatial storage media performance deteriorates as researchers approach the end of each phase and would like to lower the sample. Therefore, to preserve this data, research must begin pooling our resources. Max and Average pooling are the two more prevalent types of pooling [52].

#### 3.6.4. Conceptualization

Based on image analysis, several strategies for detecting the diagnosis of the disease have indeed been presented. Its primary goal is to develop a methodology for analyzing COVID-19 victims using Computed tomography pictures of their lungs [53]. Fig. 3 depicts a conceptual design of the system techniques. In terms of the theoretical architecture, researchers perform three studies for COVID-19 image diagnosis and classification. The first step is to use the fractal approach to retrieve feature representations. Then research used Convolutional system methods to categorize the photos of the victims. The design process of a convolutional neural network (CNN) to categorize treatment founded on Computed tomography pictures is the standard option. Lastly, the network can be trained to segment affected tissues in lung Image data [54-56].

### 4. Experimental design and dataset

The chapter explained how the proposed Multiple NN was examined, but also how the experiment settings were achieved utilizing the Ensemble Bootstrap aggregating classification and the Multiple NN methodologies. In the computing effectiveness chapter, the influence of the suggested approach on its own correctness and processing phases is discussed. All of the experiments were done in the MATLAB environment with the SPSS IBM modularity resources, and the findings were given. The current X-ray database is enhanced by having crossing balancing Covid-19 class photographs as supplemen-
tary, as per the proposed approach. The experiment aims to demonstrate the negative impact of unbalanced patterns on the raw dataset’s effectiveness. It’s worth noting that the two-step-AS-AS-MNN has been tweaked to perform a specific training operation with the greatest parameter estimation accessible [57]. Inside this paper, a Covid-19 diagnosis prediction strategy is provided that combines a hybrid Two-step-AS clustering approach with Ensemble Bootstrap aggregating training and Multiple NN methods to increase diagnostics sensitivity and decrease misinterpretation error even while enhancing the accuracy of classification.

As a result, a novel approach that blends unsupervised and supervised training methods to create a hybrid approach of training is developed. On the X-ray chest image feature extraction process utilizing the Multiple NN classification structure, a competent investigation was carried out on the Ensemble Bootstrap aggregating learning and Two-step-AS grouping database structure. Whenever the clusters’ findings have been used as parameters to the classification algorithm, they are forecasted to use Multiple NN classifications as forecasts for positive Covid-19 cases, pneumonia, and Covid-19 cases that had not been discovered [27]. The protocol was applied consistently. To explore the effects of the qualification procedure, the TSEBANN model has been used. Given the multitude of occurrences, the X-ray chest dataset has indeed been connected to many incidents. The samples are divided into ten sections for testing and training the TSEBANN approach, which was performed using the tenfold cross-validation method.

Multiple situations have been used in tandem for the identification and classification of COVID-19 in X-ray images. To begin, the TSEBANN approach was used to classify the X-ray pictures into three categories: COVID-19, No-Finding, and Pneumonia. In addition, classifications are trained in the TSEBANN framework: COVID-19 classes and No-Findings classes. The TSEBANN algorithm was used to train the COVID-19 sections. For challenges involving ternary and binary classifications, the suggested model’s outputs are validated using a tenfold cross-validation technique. It’s employed in the training records; 90 % of the X-ray images are tested four times, dependent on the rest of the database, which is part of the pre-processing stage. In the original study sample of subjects, 43 women and 82 males are determined to be positive. Not many of the individuals in this database get a detailed health background. As per the information given, the mean lifespan of the 26 COVID-19 participants is around 55 years old. [58] provided a chest X-ray 8 picture library that could be used for healthy and pneumonia imaging.

To verify that the outcomes are matched, randomized photos out of this group of 500 no-finds and 500 pneumonia frontal chest X-rays have been used as a reference. Every group comprised 375 cases, with 125 COVID19 cases classified as class 1, 125 non-COVID19 people with the disease classified as class 2, and 125 pneumonia instances classified as class 3. The methodology recognized the created groups once the information was matched, which were then were using to detect every subgroup independently via diagnostics clustering investigations. The TSEBANN training classification was used in investigations with tenfold cross-validation to evaluate the correlations factor of the X-ray current diagnostic classification. The new balanced information is divided into ten elements in aggregate [59–61]. Every element represented 10 % of the entire dataset, providing for the conversion of every database set into the testing dataset. A total of 9 experiments have been used for training and one series of experiments was used for assessment throughout every round. The Chest X-ray dataset was grouped utilizing the Two-step-AS approach based on non-COVID-19 factors like pneumonia and COVID19 traits of each kind.
5. Results analysis and discussion

The effectiveness of the NN classification technique can be improved using the Two-step Cluster (TSC) approach. It’s because discretizing continuous characteristics improves their performance. Machine learning uses the discretization approach to handle continuous values; discretization of global attributes can make information simply compact and efficient induction methods of learning, as well as simplify and improve the efficiency of the algorithm. TSC is a grouping technique applied with the goal of handling exceedingly large samples. It is a classification algorithm included with the statistics application SPSS. With the Two-step-AS cluster, you can manage both categorical and continuous information. The Two-step-AS model specifications are shown in Table 6.

Whenever processed information, or one which has been characterized, is input into the computer, TSC is being used to construct a classification model from the collected information. The class label is comprised of two labels that are joined together, cluster-1 and cluster-2. Following the establishment of every classification model, the prior probabilities of every class label are calculated for the NN computation, which is necessary. The TSC method generates a minimum and maximum variety of general clusters of 2 and 15, respectively, as shown in Table 7. As a Feature Importance recommendation, the TSC methods use the Bayesian Information Criterion (BIC) Procedure and the Natural Logarithm measurement as a Euclidean Distance.

Quality is a metric for group separation and cohesion. Cluster-1 has 985 entries with a perfection rating of 0.89 and a maximum significance rating of 1.00, while cluster-2 has 416 entries with a quality rating of −0.25 and a standard significance rating of 1.00. The total efficiency of the model is = 0.76. (-1 to 0.2 Poor | 0.2 to 0.5 Fair | 0.5 to 1 Excellent).

The significance, on the other hand, is a measurement of clusters cohesiveness that ranges from 0 to 0.2 Poor, 0.2 to 0.6 Fair, and 0.6 to 1 Good. The outlier’s intensity is a metric that determines how distinct an outlier group seems to be from the rest of the groupings. The outlier analysis was based on the records’ upper and lower similarity.

The approach incorporates an additional stage in the construction of the CF tree that allows the resolution of odd numbers to be included (outliers). Outliers are recordings that don’t seem to belong in some of the groups where they’re discovered. In SPSS, leaf elements are drawn from different if the entire number of records in the leaf is within a predefined proportion (by default, 25 %) of the overall volume of data in the clustered CF tree’s greatest leaf entry. The approach searches for and eliminates any possibly unexpected values before generating the CF tree. Following the CF tree’s rebuilding, the method assesses if those variables could be absorbed within the tree without increasing its size. Lastly, outliers are statistics which does not fall into some of the previous systems. If the CF tree grows to be greater than the maximum length allowed, it is recreated to use the present CF tree, with the threshold length extended to fit the larger tree. The modified CF tree is shorter and allows input records to be added. With centering modes for feature inputs, Table 6 illustrates the relevance of an input to a specific cluster. Fig. 4 depicts the relevance of an input to a certain cluster using centers methods for features inputs.

With centering options for features inputs, Fig. 5 illustrates the relevance of the information to a specific cluster. The values of standard deviation cluster center produced by the essential features entries (feature7, feature4, feature3, and feature1) are (0.61, 0.41, 0.32, and −0.16) for the higher cases, and (0.61, 0.41, 0.32, and −0.16) for the lower cases, as shown in Table 8. (-0.120, −0.89, −0.75, and −0.23).

Among clusters 1 and 2, the main characteristics obtained were 1.0, 0.87, 0.78, and 0.46 sample variance. A COVID-19 dataset was gathered with the purpose of learning the information in order to conduct an experimental analysis. The investigators used a tenfold cross-validation technique for both training and validation of the database in the study, as discussed previously. To test the enhanced consequences of the hybrid technique, cross-dataset testing was performed where the MNN and MBANN classifications are utilised both without and with Two-step-AS clustering discoveries. Obtained from equation (12), the outcomes of the cross-validation technique were computed to yield the relevant diagnosis:

| Table 6 TwoStep-AS-AS grouping method Conditions. |
|-----------------------------------------------|
| **TwoStep-AS Clustering Specifications**      |
| Minimum Number of Regular Clusters: 2         |
| Maximum Number of Regular Clusters: 15        |
| Feature Importance Method: Information Criterion |
| Information Criterion Bayesian Information Criterion (BIC) Log Likelihood |
| Distance Measure Final Model: Number of Regular Clusters: 2 |
| Number of Outlier Clusters: 87                |
| Important Feature Inputs: Feature No-3, Feature No-4, Feature No-7, Feature No-10 |

| Table 7 TwoStep-AS Grouping Method Excellence. |
|-----------------------------------------------|
| **Cluster Quality**                           |
| Cluster Number of Records:                     |
| Goodness: Importance                           |
| Cluster-1: 985                                |
| 0.89: 1.00                                    |
| Cluster-2: 416                                |
| −0.25: 1.00                                   |
\[ \text{Acc} = \frac{(\text{True Negative} + \text{True Positive})}{(\text{True Negative} + \text{False Positive}) + (\text{True Positive} + \text{False Negative})} \times 100 \]  

The percentage of COVID-19 executables accurately classified is denoted by True Positive (TP). The number of COVID-19 executables wrongly classified is referred to as False Positive (FP). The number of non-COVID-19 and pneumonia executables mistakenly categorized is referred to as True Negative (TN).
AS is introduced to the database as a feature even during the testing procedures on the database. The result of the Two-step-AS in the lack of grouping, as well as the outcomes of the training/grouping and with grouping to use the Two-step-AS technique the Ensemble Multiple NN classification techniques without Table 9, enhances diagnostic performance by a rate of 99.055 percent, which is impressive. Table 9 illustrates the results of the Ensemble Multiple NN classification techniques without grouping and with grouping to use the Two-step-AS technique in the lack of grouping, as well as the outcomes of the training/testing procedures on the database. The result of the Two-step-AS is introduced to the database as a feature even during the process of consolidation, enabling every occurrence in the database to be labeled with a clustering identifier.

As per the statistics, cross-validation was conducted ten times, and the median classification performance attained by using Multiple NN without grouping is 88.73 percent for training and 84.492 percent for testing. Figs. 6 and 7 illustrate the greatest with worst performances of the COVID-19 reliability Clustering technique based on MN.

Without using grouping, the training and testing of high-performance results achieved in fold number 5 used to have an accuracy ratio of 98.51 and 98.11 percent, correspondingly. Lower training/testing scoring predictive performance utilising grouping was achieved in folds 10 and 1, correspondingly, with 86.69 and 80.82 percent with a recognition rate of 99.055 percent and a testing accuracy of 97.949 percent.

Folds 1 and 4 have the highest train and test rating predictions outcomes utilizing grouping, with 99.3 and 97.949 percent, correspondingly. Predictions outcomes utilizing grouping were reached in fold 10th position with 98.49 and 98.28 percent, correspondingly, despite the low training/testing ratings. The results show that whenever the Two-step-AS cluster approach is utilised, performance can be enhanced. The outcomes of the Multiple NN with grouping are enhanced whenever the outcome is merged with the Two-Step-AS Clustering method, and the efficiency of breast cancer detection is improved. In addition, Table 10 reflects the effectiveness of the Ensemble Multiple Neural Network algorithm on the X-ray COVID-19 database. The identical scenario was replicated with Multiple Neural Networks alone and Ensemble Bootstrap aggregating classification. Figs. 6 and 7 show the resulting effectiveness in both the training and testing, without the use of Two-step-AS grouping.

As per the data, the median classification performance attained by using Ensemble Multiple NN without grouping for the training and testing trials is 92.204 percent for the training and 87.889 percent for the testing. Table 10 also summarizes the output of the Ensemble Multiple NN classifiers with grouping utilizing the Two-step-AS approach, with a precision of 99.115 percent and a testing accuracy of 98.062 percent. Without using grouping, the training process of high-

| Table 9 According to the MNN Methodology, the effectiveness of the X-ray COVID-19 Database. |
|----------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| **Fold No** | **COVID-19 Diagnosis Results With TwoStep-AS Clustering Approach %** | **Training Accuracy** | **Training Error** | **Testing Accuracy** | **Testing Error** | **COVID-19 Diagnosis Results Without TwoStep-AS Clustering Approach** | **Training Accuracy** | **Training Error** | **Testing Accuracy** | **Testing Error** |
| 1st Fold | 99.3 | 0.7 | 97.26 | 2.74 | 87.53 | 12.47 | 80.82 | 19.18 |
| 2nd Fold | 98.89 | 1.11 | 98.64 | 1.36 | 86.99 | 13.01 | 85.71 | 14.29 |
| 3rd Fold | 99.29 | 0.71 | 97.36 | 2.64 | 87.82 | 12.18 | 83.7 | 16.3 |
| 4th Fold | 99.24 | 0.76 | 98.71 | 1.29 | 88.75 | 11.25 | 84.79 | 15.21 |
| 5th Fold | 99.01 | 0.99 | 98.21 | 1.79 | 98.51 | 1.49 | 98.11 | 1.89 |
| 6th Fold | 99.04 | 0.96 | 98.49 | 1.51 | 88.24 | 11.76 | 85.31 | 14.69 |
| 7th Fold | 99.07 | 0.93 | 97.55 | 2.45 | 87.64 | 12.36 | 85.89 | 14.11 |
| 8th Fold | 99.09 | 0.91 | 97.42 | 2.58 | 87.16 | 12.84 | 84.84 | 15.16 |
| 9th Fold | 99.13 | 0.87 | 97.57 | 2.42 | 88.01 | 11.99 | 86.98 | 13.02 |
| 10th Fold | 98.49 | 1.51 | 98.28 | 2.72 | 86.69 | 13.31 | 68.77 | 13.23 |
| **Average** | 99.055 | 1.045 | 97.949 | 2.15 | 88.734 | 11.266 | 84.492 | 13.708 |
performance results generated in fold number 5 used to have an accuracy ratio of 99.26 and 97.17 percent, correspondingly. In contrast, with 89.98 and 82.19 percent, the lowest train and test scoring predicting outcomes utilizing the grouping were produced in fold number 1.

With 99.37 and 98.64 percent, correspondingly, high training and testing ratings predictive outcomes are produced util-

Table 10 Performance on the X-ray COVID-19 Dataset Based on Ensemble MNN Algorithm.

| Fold No | COVID-19 Diagnosis Results With TwoStep-AS Clustering Approach % | COVID-19 Diagnosis Results Without TwoStep-AS Clustering Approach % |
|---------|---------------------------------------------------------------|---------------------------------------------------------------|
|         | Training Accuracy | Training Error | Testing Accuracy | Testing Error | Training Accuracy | Training Error | Testing Accuracy | Testing Error |
| 1st Fold| 99.3              | 0.7            | 97.26           | 2.74          | 89.98            | 10.02          | 82.19           | 17.81         |
| 2nd Fold| 98.89             | 1.11           | 98.64           | 1.36          | 91.13            | 8.87           | 88.44           | 11.56         |
| 3rd Fold| 99.37             | 0.63           | 98.24           | 1.76          | 91.12            | 8.88           | 87.22           | 12.78         |
| 4th Fold| 99.33             | 0.67           | 98.38           | 1.62          | 91.52            | 8.48           | 87.06           | 12.94         |
| 5th Fold| 99.01             | 0.99           | 98.21           | 1.79          | 99.26            | 0.74           | 97.17           | 2.83          |
| 6th Fold| 99.04             | 0.96           | 98.49           | 1.51          | 90.74            | 9.26           | 87.04           | 12.96         |
| 7th Fold| 99.27             | 0.73           | 98.16           | 1.84          | 91.07            | 8.93           | 87.53           | 12.47         |
| 8th Fold| 99.32             | 0.68           | 98.39           | 1.61          | 91.82            | 8.18           | 87.42           | 12.58         |
| 9th Fold| 99.13             | 0.87           | 97.57           | 2.43          | 92.26            | 7.74           | 87.27           | 12.73         |
| 10th Fold| 98.49             | 1.51           | 97.28           | 2.72          | 93.14            | 6.86           | 87.55           | 12.45         |
| Average | 99.115            | 0.885          | 98.062          | 1.938         | 92.204           | 7.796          | 87.889          | 12.111        |
ising grouping in folds 3 and 2. Forecasting outcomes utilising grouping were achieved in folds 10 and 1 with 98.49 and 97.26 percent, correspondingly, despite poor training and testing ratings. The results show that whenever the Two-step-AS clustering technique is utilized, the performance is improved. The outcomes of the Multiple NN with grouping were enhanced whenever the outcome is merged with the Two-step-AS clustering approach, and the efficiency of breast cancer detection is improved.

The T-test methodology was used in this investigation to establish statistically significant between both the findings of the first investigation utilizing Multiple NN and the findings of the second experiment that use the Ensemble Multiple NN methods, as well as to illustrate the benefits of someone using the Two-step-AS methodology. The T-test revealed a statistical distinction between the two parameters with a low significance level (typically less than 0.05). This condition is underscored in the assessment instruments generated analysis of the gathered observations in Tables 11 and 12, which demonstrate that the significance level (0.000211 and 0.000008) is significant. Whenever paired with the Two-step-AS clustering technique, the TSEBANN yields statistical significance.

| Table 11 | statistically significant outcomes for multiple neural networks using the T-test. |
|----------|----------------------------------------------------------------------------------|
| prior to and after the development, variations in reliability resulted in 50 percent, and 70 %. | Avg | Standard Deviation | Std. Error Mean | Std. Error Mean | 95 % CI | df | Significant amount |
| With Clustering VS Without Clustering | 13.4570 | 7.13236 | 2.25545 | 8.35482 | 18.55918 | 9.966 | 9 | 0.000211 |

| Table 12 | statistically significant outcomes for ensemble multiple neural networks using the T-test. |
|----------|----------------------------------------------------------------------------------|
| prior to and after the development, variations in reliability resulted in 50 percent, and 70 %. | Avg | Standard Deviation | Std. Error Mean | Std. Error Mean | 95 % CI | df | Significant amount |
| With Clustering VS Without Clustering | 10.1730 | 3.52945 | 1.11611 | 7.64818 | 12.69782 | 9.115 | 9 | 0.000008 |

**Comparison between the COVID-19 Diagnosis Methods**

![Comparison between the COVID-19 Diagnosis Methods](image)

**Fig. 8** The TSEBANN and the Present COVID-19 Classifiers are compared.
6. Conclusion

The outbreak of coronavirus disease (COVID-19) is causing a public health emergency. Among the most essential and difficult aspects of computer vision that is crucial for identifying the disease, the current research focuses on improving a test for detecting COVID-19 in pneumonia individuals. Clustering methods may be found in a wide range of fields that utilize large datasets to find hidden connections. Traditional approaches, on the other hand, are often unable to effectively function on these collections although most findings received from the real world comprise both mathematical and category properties. It is shown that the two-step-AS methodology, which really is simple to get and immediately estimates the optimal number of clusters, would be utilized for solving this problem. During the first phase of the TSEBNN, reported symptoms are classified as pneumonia, COVID-19, or healthy instances. Moreover, because COVID-19 is produced by viruses, all of the cases tested are partitioned into three groups utilizing TSEBNN during the two stages of prognosis: positive COVID-19, pneumonia, and negative COVID-19 (normal). The purpose of TSEBNN method is to provide a rapid, methodical, and accurate computer-aided approach for describing COVID-19 instances to patients coming to institutions and undergoing initial evaluation with a chest X-ray scan. To demonstrate the efficiency of the proposed technique, comprehensive examinations were performed, utilizing both the training and testing processes, as well as a tenfold cross-validation technique to demonstrate the usefulness of the TSEBNN methodology. A series of research were also conducted to demonstrate the effectiveness of the TSEBNN in recognizing COVID-19 occurrences if contrasted with certain other COVID-19 approaches addressed in the ongoing study. The proposed methodology would be further developed in the future in order to make it predictable for other types of COVID-19 data, including CT-scan and analytical information. Optimization strategies based on machine learning and data mining techniques would be employed in tandem to improve the effectiveness of the prediction strategy in COVID-19 sickness.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number (UB-57-1442).

References

[1] R. Sujath, J.M. Chatterjee, A.E. Hassnian, A machine learning forecasting model for COVID-19 pandemic in India, Stoch. Environ. Res. Risk Assess 34 (7) (Jul. 2020) 959–972, https://doi.org/10.1007/s00477-020-01827-8.
[2] N. Poonawala-Lohani, P. Riddle, M. Adnan, and J. Wicker, “A Novel Approach for Time Series Forecasting of Influenza-like Illness Using a Regression Chain Method,” in BioComputing 2022, Kohala Coast, Hawaii, USA, Dec. 2021, pp. 301–312. doi:10.1142/9789811250477_0028.
[3] C. Zhang, L. Qin, K. Li, Q. Wang, Y. Zhao, B. Xu, L. Liang, Y. Dai, Y. Feng, J. Sun, X. Li, Z. Hu, H. Xiang, T. Dong, R. Jin, Y. Zhang, A novel scoring system for prediction of disease severity in COVID-19, Front. Cell. Infect. Microbiol. 10 (2020).
[4] R. Majhi, R. Thangada, R.P. Sugasi, N. Kumar, Analysis and prediction of covid-19 trajectory: A machine learning approach. J. Public Affairs 21 (4) (2021) Nov, https://doi.org/10.1002/ppa.2537.
[5] Z. Malki, E.-S. Atlam, A.E. Hassnien, G. Dagnew, M.A. Elhosseini, I. Gad, Association between weather data and COVID-19 pandemic predicting mortality rate: Machine learning approaches, Chaos, Solitons Fractals 138 (2020) 110137.
[6] V. Chaurasia, S. Pal, Application of machine learning time series analysis for prediction COVID-19 pandemic, Res. Biomed. Eng. 38 (1) (2022) 35–47.
[7] N.S. Punn, S.K. Sonbhadra, S. Agarwal, COVID-19 epidemic analysis using machine learning and deep learning algorithms, MedRxiv (2020).
[8] F. Rustam, A.A. Reshi, A. Mehmoond, S. Ullah, B.-W. On, W. Aslam, G.S. Choi, COVID-19 Future Forecasting Using Supervised Machine Learning Models, IEEE Access 8 (2020) 101489–101499.
[9] S.F. Ardabili, A. Mosavi, P. Ghamisi, F. Ferdinand, A.R. Varkonyi-Koczy, U. Reuter, T. Rabczuk, P.M. Atkinson, Covid-19 outbreak prediction with machine learning, Algorithms 13 (10) (2020) 249.
[10] S. Sengupta, S. Mugde, G. Sharma, Covid-19 Pandemic Data Analysis and Forecasting using Machine Learning Algorithms, Public and Global Health, preprint (2020), https://doi.org/10.1101/2020.06.25.20140004.
[11] K. Santosh, COVID-19 prediction models and unexploited data, J. Med. Syst. 44 (9) (2020) 1–4.
[12] C. Long et al, Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT?, Eur J. Radiol. 126 (2020) 108961.
[13] M.S.K. Inan, F.I. Alam, R. Hasan, Deep integrated pipeline of segmentation guided classification of breast cancer from ultrasound images, Biomed. Signal Process. Control 75 (2022) 103553.
[14] P. Zimmermann, N. Curtis, Coronavirus infections in children including COVID-19: an overview of the epidemiology, clinical features, diagnosis, treatment and prevention options in children, Pediatr. Infect. Dis. J. 39 (5) (2020) 355.
[15] K. McIntosh, M. Hirsch, A. Bloom, Coronavirus disease 2019 (COVID-19): Epidemiology, virology, and prevention, Lancet. Infect. Dis 1 (2020) 2019–2020.
[16] W. Kong, P.P. Agarwal, Chest imaging appearance of COVID-19 infection, Radiology: Cardiothoracic Imaging 2 (1) (2020) e200028.
[17] A. Bernheim et al, Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection, Radiology (2020) 200463.
[18] J. Rasheed et al, A survey on artificial intelligence approaches in supporting frontline workers and decision makers for the COVID-19 pandemic, Chaos, Solitons Fractals 141 (2020) 110337.
[19] Y.E. Almalki, A. Qayyum, M. Irfan, N. Haider, A. Glowacz, F. Varkonyi-Koczy, U. Reuter, T. Rabczuk, P.M. Atkinson, Covid-19 prediction of COVID-19 trajectory: A machine learning approach, Front. Cell. Infect. Microbiol. 10 (2020).
images using ensemble learning with convolutional neural network, Pattern Anal. 24 (3) (2021) 1111–1124.

[21] A. Narin, C. Kaya, Z. Pamuk, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, Pattern Anal. 24 (3) (2021) 1207–1220.

[22] S. Asif, Y. Wenhui, H. Jin, S. Jinhai, “Classification of COVID-19 from Chest X-ray images using Deep Convolutional Neural Network.” in: 2020 IEEE 6th International Conference on Computer and Communications (ICCC), Chengdu, China, Dec. 2020, pp. 426–433. doi: 10.1109/ICCC51575.2020.9344870.

[23] M.S. Satu et al, Convolutional Neural Network Model to Detect COVID-19 Patients Utilizing Chest X-ray Images, Infectious Diseases (except HIV/AIDS), preprint (2020), https://doi.org/10.1101/2020.06.07.20124594.

[24] A.S. Kwekha-Rashid, H.N. Abduljabbar, B. Ahlayani, Coronavirus disease (COVID-19) cases analysis using machine-learning applications, Applied Nanoscience 2021 1–13.

[25] M. J. Awani, M. W. Imitiaz, M. Usama, A. Rehman, N. Ayesha, and H. M. F. Shehzad, “Covid-19 Detection by using Deep learning-based Custom Convolution Neural Network (CNN).” in: 2021 International Conference on Innovative Computing (ICIC), Lahore, Pakistan, Nov. 2021, pp. 1–7. doi: 10.1109/ICIC53490.2021.9693071.

[26] A. Akhtar, S. Akhtar, B. Bakhtawar, A.A. Kashif, N. Aziz, M. S. Javeid, COVID-19 Detection from CBC using Machine Learning Techniques, Int. J. TIM 1 (2) (2021) 65–78, https://doi.org/10.54849/ijtim.v1i2.22.

[27] N.-A.-A. Alam, M. Ahsan, M.A. Based, J. Haider, M. Kowalski, COVID-19 Detection from Chest X-ray Images Using Feature Fusion and Deep Learning, Sensors 21 (4) (2021) 1480, https://doi.org/10.3390/s21041480.

[28] M.J. Hurry, S. Chakraborty, M. Paul, A. Ulhaq, B. Pradhan, M. Saha, N. Shukla, COVID-19 Detection Through Transfer Learning Using Multimodal Imaging Data, IEEE Access 8 (2020) 149808–149824.

[29] F.M. Salman, S.S. Abu-Naser, E. Alajrani, B.S. Abu-Nasser, B. A. M. Ashqar, “COVID-19 Detection using, Artificial Intelligence” 4 (3) (2020) 8.

[30] S. Hassantabar, M. Ahmadi, A. Sharifi, Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches, Chaos, Solitons Fractals 140 (Nov. 2020), https://doi.org/10.1016/j.chaos.2020.109170.

[31] A. Borkowski, Using Artificial Intelligence for COVID-19 Chest X-ray Diagnosis, Federal Practitioner 37 (9) (2020) Sep, https://doi.org/10.12788/fp.0045.

[32] T. Agrawal, P. Choudhary, FocusCovid: automated COVID-19 detection using deep learning with chest X-ray images, Evolving Systems 13 (4) (2020) 519–533.

[33] S. Varela-Santos, P. Melin, A new approach for classifying coronavirus COVID-19 based on its manifestation on chest X-rays, Pattern Recogn. 121 (2022) 108242.

[34] S. Bakheet, A. Al-Hamadi, Automatic detection of COVID-19 using pruned GLCM-Based texture features and LDCRF classification, Comput. Biol. Med. 137 (2021) 104781.

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

[36] M. Abbaspour Onari, S. Yousefi, M. Rabieepour, A. Alizadeh, M. Jahangoshai Rezaee, A medical decision support system for predicting the severity level of COVID-19, Complex Intell. Syst. 7 (4) (2021) 2037–2051.

[37] M.I. Uddin, S.A.A. Shah, M.A. Al-Khasawneh, A Novel Deep Convolutional Neural Network Model to Monitor People following Guidelines to Avoid COVID-19, Journal of Sensors 2020 (Jul. 2020) 1–15, https://doi.org/10.1155/2020/8856801.

[38] A. Davoudi, M. Ahmadi, A. Sharifi, R. Hassantabar, N. Najafi, A. Tayebi, H.A. Kasgari, F. Ahmadi, M. Rabiee, Q. Zou, Studying the effect of taking statins before infection in the severity reduction of COVID-19 with machine learning, Biomed Res. Int. 2021 (2021) 1–12.

[39] A. Abbas, M.M. Abdelsamea, M.M. Gaber, Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network, Appl Intell 51 (2) (Feb. 2021) 854–864, https://doi.org/10.1007/s10489-020-01829-7.

[40] R. Kafieh, R. Arian, N. Saeedizadeh, Z. Amini, N.D. Serej, S. Minaeec, S.K. Yadav, A. Vaezi, N. Rezaee, S. Haghjoo Javanmard, K. Bluiss, COVID-19 in Iran: Forecasting Pandemic Using Deep Learning, Comput. Math. Methods Med. 2021 (2021) 1–16.

[41] M.D.K. Hasan, S. Ahmed, Z.M.E. Abdullah, M. Monirujaman Khan, D. Anand, A. Singh, M. AlZain, M. Masud, K. Sun, Deep Learning Approaches for Detecting Pneumonia in COVID-19 Patients by Analyzing Chest X-Ray Images, Mathematical Problems in Engineering 2021 (2021) 1–8.

[42] I. Ozsahin, B. Sekeroglu, M.S. Musa, M.T. Mustapha, D. Uzun Ozsahin, Review on Diagnosis of COVID-19 from Chest CT Images Using Artificial Intelligence, Comput. Math. Methods Med. 2020 (2020) 1–10.

[43] R.G. Babukarthik, V.A.K. Adiga, G. Sambasivam, D. Chandramohan, J. Amudhavath, Prediction of COVID-19 Using Genetic Deep Learning Convolutional Neural Network (GDCNN), IEEE Access 8 (2020) 177647–177666, https://doi.org/10.1109/ACCESS.2020.3025164.

[44] X. Zhang, X. Lu, Q. Shi, X.-Q. Xu, H.-C. Leung, L.N. Harris, J. D. Iglehart, A. Miron, J.S. Liu, W.H. Wong, Recursive SVM feature selection and sample classification for mass-spectrometry and microarray data, BMC Bioinf. 7 (1) (2006).

[45] M. Singh, S. Bansal, S. Ahuja, R.K. Dubey, B.K. Panigrahi, N. Dey, Transfer learning–based ensemble support vector machine model for automated COVID-19 detection using lung computerized tomography scan data, Med. Biol. Eng. Comp. 59 (4) (2021) 825–839.

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

[47] V. Guarrasi, N.C. D’Amico, R. Sicilia, E. Cordelli, P. Soda, Pareto optimization of deep networks for COVID-19 diagnosis from chest X-rays, Pattern Recogn. 121 (2022) 108242.

[48] A.S. Ashour et al, Ensemble-based bag of features for automated classification of normal and COVID-19 CXR images, Biomed. Signal Process. Control 68 (2021) 102656.

[49] M. Jahangoshai Rezaee, A medical decision support system for the diagnosis of COVID-19 respiratory diseases, Multimodal Biomedical Imaging XVI 11634 (2021) 116340P.

[50] H.A. Rothan, S.N. Byrareddy, The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak, J. Autoimmun. 109 (2020) 102433.

[51] A. Boukari, F. Boukari, Transfer and deep learning techniques for the diagnosis of COVID-19 respiratory diseases, Multimodal Biomedical Imaging XVI 11634 (2021) 116340P.

[52] E. Ong, M.U. Wong, A. Huffman, Y. He, COVID-19 coronavirus vaccine design using reverse vaccinology and machine learning, Front. Immunol. 11 (2020) 1581.

[53] M. Abbaspour Onari, S. Yousefi, M. Rabieepour, A. Alizadeh, M. Jahangoshai Rezaee, A medical decision support system for predicting the severity level of COVID-19, Complex Intell. Syst. 7 (4) (2021) 2037–2051.

[54] M.I. Uddin, S.A.A. Shah, M.A. Al-Khasawneh, A Novel Deep Convolutional Neural Network Model to Monitor People following Guidelines to Avoid COVID-19, Journal of Sensors 2020 (Jul. 2020) 1–15, https://doi.org/10.1155/2020/8856801.
An intelligent deep convolutional network based COVID-19 detection from chest X-rays

19 pandemic.” Information Sciences Letters, 11, no. 2 (2022): 537-548.

[56] A.I. Taloba, M.A. Fouly, T. Soliman, Developing an Efficient Secure Query Processing Algorithm on Encrypted Databases using Data Compression, Information Sciences Letters 12 (1) (2023) 1–8.

[57] C.E. von Schacky, N.J. Wilhelm, V.S. Schäfer, Y. Leonhardt, F. G. Gassert, S.C. Foreman, F.T. Gassert, M. Jung, P.M. Jungmann, M.F. Russe, C. Mogler, C. Knebel, R. von Eisenhart-Rothe, M.R. Makowski, K. Woertler, R. Burgkart, A.S. Gersing, Multitask deep learning for segmentation and classification of primary bone tumors on radiographs, Radiology 301 (2) (2021) 398–406.

[58] B. Zheng, Y. Cai, F. Zeng, M. Lin, J. Zheng, W. Chen, G. Qin, Y.i. Guo, S. Yu, An interpretable model-based prediction of severity and crucial factors in patients with COVID-19, Biomed Res. Int. 2021 (2021) 1–9.

[59] B. Gallo Marin et al., “Predictors of COVID-19 severity: A literature review,” Rev Med Virol, vol. 31, no. 1, pp. 1–10, Jan. 2021, doi: 10.1002/rmv.2146.

[60] O.R. Shahin, R.M. Abd El-Aziz, A.I. Taloba, Detection and classification of Covid-19 in CT-lungs screening using machine learning techniques, Journal of Interdisciplinary Mathematics 25 (3) (2022) 791–813.

[61] O.R. Shahin, H.H. Alshammari, A.I. Taloba, R.M.A. El-Aziz, Machine Learning Approach for Autonomous Detection and Classification of COVID-19 Virus, Comput. Electr. Eng. 101 (2022) 108055.