Deep learning models/techniques for COVID-19 detection: a survey

Kumari Archana[1], Amandeep Kaur[1], Yonis Gulzar[2]*, Yasir Hamid[3], Mohammad Shuaib Mir[2] and Arjumand Bano Soomro[2,4]

[1]Department of Computer Science, Chitkara University, Chandigarh, Punjab, India, [2]Department of Management Information Systems, College of Business Administration, King Faisal University, Al-Ahsa, Saudi Arabia, [3]Information Security and Engineering Technology, Abu Dhabi Polytechnic College, Abu Dhabi, United Arab Emirates, [4]Department of Software Engineering, Faculty of Engineering and Technology, University of Sindh, Jamshoro, Pakistan

The early detection and preliminary diagnosis of COVID-19 play a crucial role in effectively managing the pandemic. Radiographic images have emerged as valuable tool in achieving this objective. Deep learning techniques, a subset of artificial intelligence, have been extensively employed for the processing and analysis of these radiographic images. Notably, their ability to identify and detect patterns within radiographic images can be extended beyond COVID-19 and can be applied to recognize patterns associated with other pandemics or diseases. This paper seeks to provide an overview of the deep learning techniques developed for detection of corona-virus (COVID-19) based on radiological data (X-Ray and CT images). It also sheds some information on the methods utilized for feature extraction and data preprocessing in this field. The purpose of this study is to make it easier for researchers to comprehend various deep learning techniques that are used to detect COVID-19 and to introduce or ensemble those approaches to prevent the spread of corona virus in future.

KEYWORDS

artificial intelligence, deep learning, COVID-19, pandemic, classification

1 Introduction

An infection of the respiratory system called COVID-19 mainly affects the lungs. One of the most crucial measures in identifying and managing COVID-19 cases is getting a chest X-ray. COVID-19, a unique illness, is extremely contagious and has spread quickly over the world [1]. The pandemic was deemed a global public health emergency by the World Health Organization (WHO) on January 30, 2020, after it had spread to 216 nations. The World Health Organization designated this novel corona virus-associated acute respiratory illness Corona Virus disease-19 (COVID-19) on 11th February in 2020. Our study's objective is to find out various deep learning techniques used to identify COVID-19 from chest radiographic pictures.

The sample normal and corona virus pictures are shown in Figures 1, 2, respectively. As can be observed, it is exceedingly challenging to distinguish between the two with the naked eye [1].

This paper exclusively focuses on deep learning techniques for corona-virus identification. This paper reviews all the approaches mentioned in the literature in the hope of assisting researchers in the development of improved corona-virus detection techniques. This paper discusses methodology, datasets and common criteria for evaluation and comparison of all methods and future directions.

The following is a summary of this review's major contributions:

(i) To carefully review the most recent systems for diagnosis of COVID-19 based on deep learning using CT and X-ray medical imaging samples.

(ii) To present the assessed works and the pertinent information in a clear, succinct, and understandable way by taking into account some crucial components like the data utilized for experiments, the feature extraction methods, different deep learning based classification methods and the performance assessment metrics used.

(iii) To draw attention to and talk about the difficult elements of the current deep learning-based COVID-19 diagnosing systems.

(iv) To outline potential future research trajectories for the advancement of effective and trustworthy COVID-19 detection systems.

The remaining paper reads as follows: Section 2 contains the background of study and Section 3 includes the related work done in the field of Deep Learning for COVID-19 detection. The different methods created for COVID-19 diagnosis from both CT and X-ray samples employing pre-trained models with deep transfer learning are also explained. Section 4 explains the evaluation matrices required to measure the performance of models and the summary of studied models along with their performance measures like accuracy, sensitivity, specificity and F-score. Section 5 lists some research gaps from the studied literature, Section 6 informs about some limitations of the study and Section 7 involves concluding observations and future scope.

2 Background

According to reports, as of December 6, 2020, the United States, India, and France, respectively, had 15,318,189, 9,703,908, and 2,295,908 infected individuals, and 290,136, 140,994, and 55,521 individuals had passed away in these nations [2]. The risk of death from this condition has been reduced with the use of various medications. It should be mentioned that there is no effective medication available anywhere in the world to treat COVID-19.

The patients who are in the acute phase typically receive these prescriptions and use them. A supply chain network (SCN) that can keep track of drug inventories and manage communication between supply chain participants. Fariba Goodarzian et al. have fulfilled this purpose. They proposed a mathematical model MILP i.e., Mixed Integer Linear Programming model. They designed a sustainable-resilience network for COVID-19 pandemic. This network was divided into five levels and a stochastic chance constraint programming method was also used to deal with uncertain parameters. Three hybridization based meta-heuristic algorithms were developed. The authors main motive was to analyze the impact of COVID-19 on environment [3].

The chest radiographs (X-Ray, CT scan) are mainly used to detect the occurrence of COVID-19 infection inside human body.

Figure 3 shows the general procedure to detect COVID-19 from chest images using deep learning.

Data pre-processing is the first step which comes under data preparation and includes procedures like noise reduction, resizing, augmentation, and so forth. The data is divided into training, validation, and testing sets for the experiment during the data partitioning stage. For data partitioning, the cross-validation technique is typically used. A specific model is created using training data, and its performance is assessed using validation data and test data. The feature extraction and classification process is a key stage in the deep learning-based COVID-19 diagnostic. At this point, the deep learning technique automatically extracts the feature by repeatedly completing a number of operations, and classification is then completed using the different class labels like healthy, normal infection, lung opacity, COVID-19 positive etc. Last but not least, the constructed system is evaluated using metrics such as accuracy, sensitivity, specificity, precision, F1-score, and others [4].

2.1 Image datasets

Although the data may take on numerous forms, imaging methods like CT and X-ray samples are taken for COVID-19 diagnosis.

Some of the publicly available datasets based on chest radiographs are.

2.2 Image augmentation

The X-ray imaging (X-ray) technology has been around for a long time and is used in a wide range of security systems at
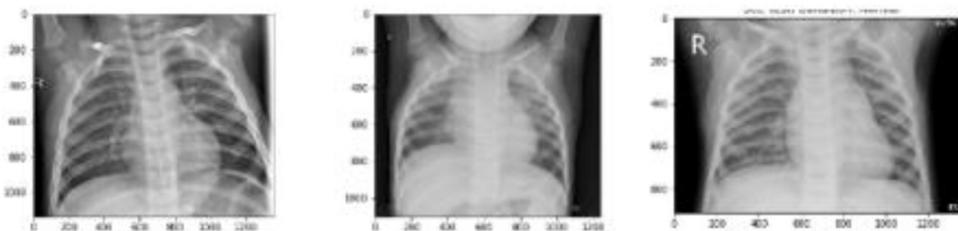
FIGURE 1
Sample images without COVID-19 infection.

FIGURE 2
Sample images with COVID-19 infection.
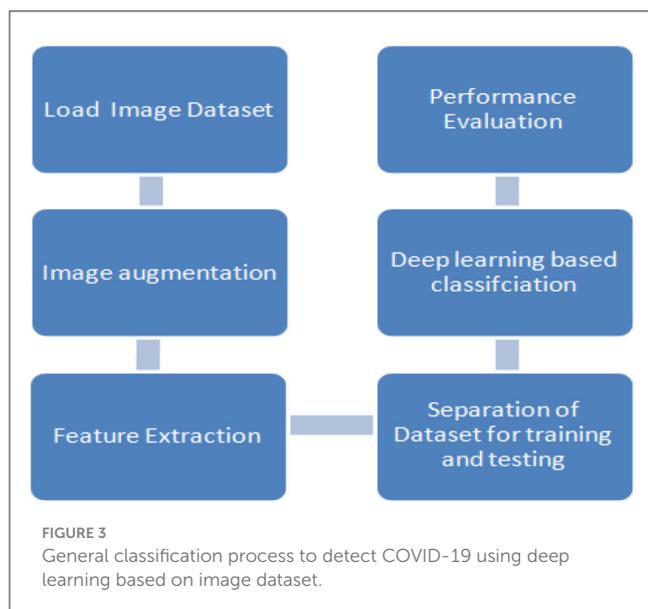
FIGURE 3
General classification process to detect COVID-19 using deep learning based on image dataset.

ports, borders and in specific facilities [5]. X-ray imaging provides a method for non-destructive detection of hidden threats by analyzing X-ray pictures [6]. As illustrated in Figure 4, we divide the method into two primary categories: image synthesis and image transformation [7].

2.2.1 Image transformation

An image transformation can be employed to change an image from one representation to another. Observing an image in alternative domains like frequency or Hough space allows for the detection of characteristics that might be less straightforward to discern in the spatial domain. The primary aim is to prepare X-ray images for use in real-world security systems. This means the altered images should be consistent with what would be encountered in actual security scenarios.

2.2.2 Image synthesis

The conventional approach to generate new images from positive samples includes modifying their characteristics through various manipulation methods. A contemporary technique known as image synthesis combines threat-related data from positive samples with data from negative samples to create new images.

Generative Adversarial Network (GAN) is the most widely used image augmentation technique. GANs create distinctive images that mirror the feature distribution of the input data by using random noise from a latent space [7, 8].

2.3 Feature extraction

The objective of feature extraction is to enrich training data by introducing new attributes, thus enhancing the effectiveness of machine learning algorithms. Feature extraction achieves this by generating novel features from the existing ones and subsequently removing the original features, effectively reducing the feature count in a dataset. Machine learning methods, such as deep learning, can be employed to detect features within images. This approach utilizes a multi-layered neural network, designed to mimic the functioning of the human brain. Each layer in an image-processing pipeline is capable of extracting one or more distinct features. To expedite analysis, processing is often parallelized. Text feature extraction, which gathers textual information, serves as a fundamental technique for capturing the content of a text message and forms the foundation for numerous text processing tasks [8]. The fundamental building blocks of these features are referred to as "text features" [9]. In the process of feature extraction, irrelevant or redundant features are removed. Feature extraction, as a data preprocessing technique for learning algorithms, can significantly enhance algorithm accuracy and reduce processing time.

3 Related work

When Deep Learning models are used to chest X-ray (CXR) pictures, it has been successful in helping researchers diagnose pulmonary disorders such as COVID-19 and pneumonia.

Shelke et al. established a deep learning based automated COVID-19 Screening Chest Xray Classification that can further classify mild, medium, and severe COVID-19 [10]. Hammoudi et al. created a hierarchical classification of COVID-19 from pneumonia viral classification using a comparison research for performance evaluation, and their average accuracy above 84% [11].

Keidar et al. achieved 90.3% accuracy with augmentation and normalization of CXR images utilizing ResNet50, ResNet152, and vgg16 for COVID-19 classification [12].

Alam et al. [13] conducted feature extraction by training a Convolutional Neural Network (CNN) with extracted features

FIGURE 4
Taxonomy for image augmentation [7].

FIGURE 5
Measurement of performance for different feature extraction model.
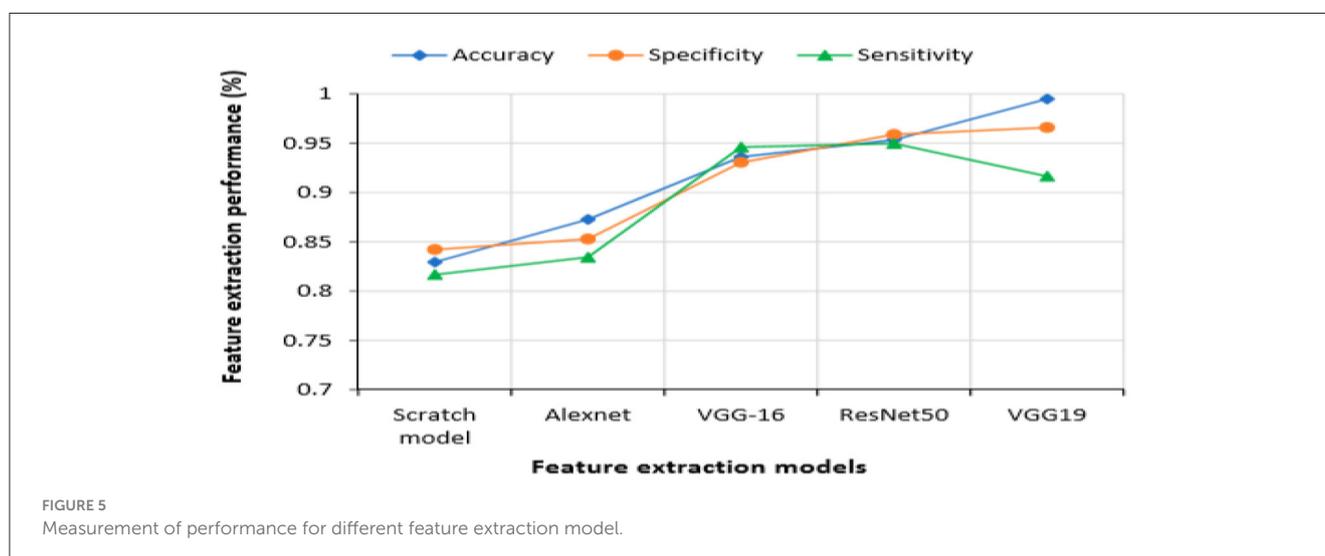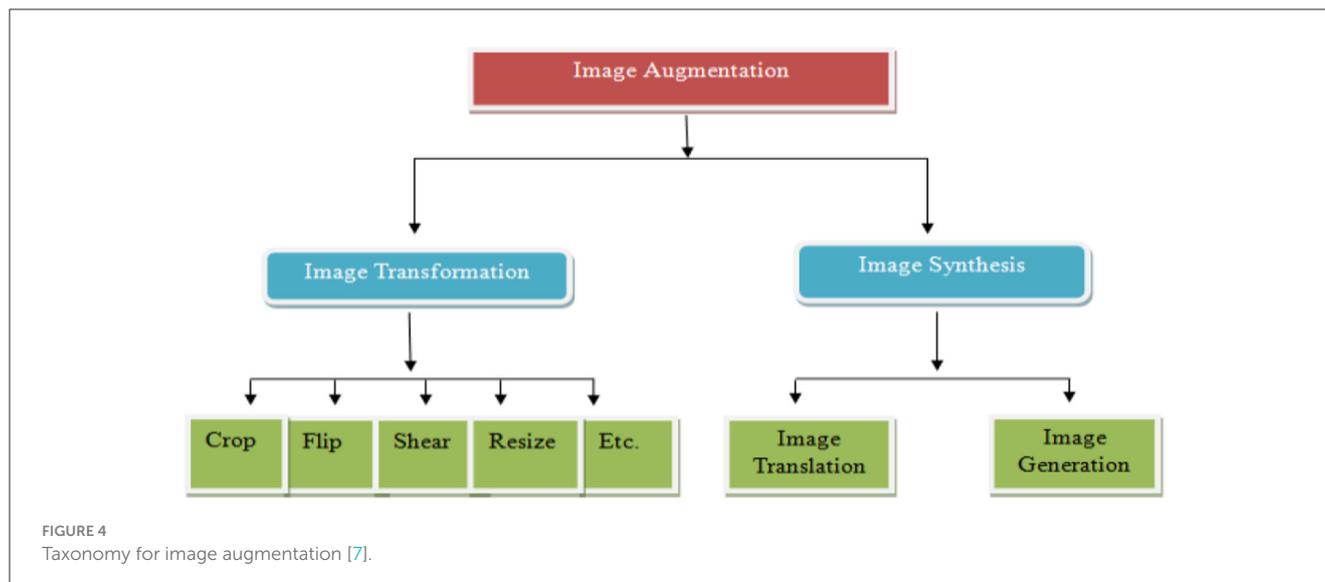
before performing classification. To know the effectiveness of models including Convolutional Neural Network (CNN), extraction was done from test images using several early trained models. Various pre-trained models, including VGG-19, VGG-16, Alexnet, and ResNet50, are currently employed by CNN to perform feature extraction from both training dataset and test dataset. For typical training data and testing data, these all models yield similar results. However, it is evident from Figure 5 that, among the CNN models, VGG19 exhibited superior accuracy and specificity, although ResNet50 showed the best sensitivity [10].

The figure clearly shows that in contrast to the CNN models, the scratch model's performance was not adequate. VGG16 and AlexNet models produced substantially worse outcomes overall than ResNet50 and VGG19 models.

In order to develop effective Deep Learning models, it's crucial for the validation error to decrease in tandem with the training error. Augmentation of data is a valuable for achieving this objective. By introducing a broader spectrum of potential

data instances, augmented data can effectively minimize the gap between the training data set and validation datasets and also any other future testing datasets [6]. Many imaging methods used in medicine use the GAN framework. A PGGAN (progressively growing generative adverserial network) was trained by Beers et al. [9] to synthesize medical images from photos of premature retinopathic vascular pathology (ROP) and multimodal MRI imaging of gliomas. Using GAN, Dai et al. [14] produced segmented pictures of the heart and lungs from a chest X-ray.

Zhao et al. [15] designed a multi-scale network i.e., VGG16 along with a model that is based on Forward and Backward Generative Adverserial Network (F&BGAN) to generate synthetic images for the classification of lung based nodules.

Using LBP feature extraction for texture features, significant features are extracted from the dataset in the second stage and used as new input data for the designated classification techniques. In the end, the results after experiment obtained without the use of feature extraction are computed to assess their compliance with

TABLE 1 Publicly available datasets based on chest images.

Name of dataset	Size of dataset	Type of chest images
Cohen Dataset	Updated Regularly	X-ray, CT scan images
Poul Mooney Dataset	5,856 images	X-ray images
Kaggle Dataset	97 images	X-ray, CT scan images
COVIDx Dataset	104,009 images	CT scan images
ChestXray—8 Dataset	108,948 images	X-ray images
CheXpert Dataset	224,316 images	Chest radiographs
Kaggle 2 Dataset	2,909 images	X-ray images

the performance criteria of Accuracy, Sensitivity, Specificity, and F-Score. Aslan et al. [16] used weighted classification methods like Ensemble, Kernel Naive Bayes, LD, QD, Cubic SVM, and KNN as input datasets The results are displayed in Table 1.

As it is clearly visible from Table 2, the Cubic SVM method's greatest estimation accuracy is 94.1%. Cubic SVM's performance metrics for sensitivity, specificity, and F_score are 91.25, 95.61, and 91.86%, respectively. These performance standards are also greater than those used by other categorization systems. While the Kernel Naive Bayes approach exhibits the lowest accuracy rate at 71.2%, both Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) methods demonstrate comparably high accuracy rates, standing at 87.0 and 86.7%, respectively. Table 2 displays the outcomes of the use of LBP feature extraction.

LBP feature extraction is used to enhance performance standards and raise the prediction accuracy of disease detection in chest X-ray images. It is clearly seen in Table 3 that the use of LBP has improved the performance of all models by increasing their accuracy.

LBP, HOG, and Haralick feature extraction algorithms were utilized in the work by Hasoon et al. [37] on 5,000 datasets. As classification techniques, KNN and SVM were employed. By averaging the 5-fold cross-validation prediction values, the performance criteria values were determined. In the KNN weighted classification by making use of Local Binary Pattern based feature extraction, the highest accuracy was recorded i.e., 98.66%. In Support Vector Machine classification with Histogram of Oriented Gradients (HOG) feature extraction, the obtained accuracy was low, reaching 89.20%. Additionally, the hybrid (HOG + SVM) technique exhibited lowest sensitivity (recall) and specificity—78.61 and 65.30%, respectively.

For the first time, a new sustainable COVID-19 medical waste supply chain network is established for allocation and location of the detoxification center of the COVID-19 medical wastes by Ghasemi et al. [38]. After that, the COVID-19 illness predictions are made using the fuzzy inference algorithm. In addition, individuals are categorized into four groups: the healthy, suspicious, mildly suspicious, and strongly suspicious.

A new set of production, allocation, placement, inventory management, and distribution issues for a new sustainable-resilient healthcare network (SRHCN) were developed in relation to the COVID-19 patients by Goodarzian et al. [3]. In order to manage

the distribution centers and warehouses, the management of the medicine distribution and inventory, and the control of the medication flows, a new multi-objective multi-period multi-level multi-commodity Mixed-Integer Linear Programming (MILP) mathematical model was developed [3]. To make location and allocation decisions that take waste management into account, a new fuzzy sustainable model for the COVID-19 medical waste supply chain network was designed by Goodarzian et al. [39]. In order to reduce supply chain costs, the environmental impact of medical waste, develop detoxification centers, and manage social responsibility centers during the COVID-19 outbreak are addressed. Sensitivity analysis is carried out on crucial parameters to demonstrate how well the suggested model performs. To verify the given paradigm, a real-world case study in Tehran, Iran, is advised.

By applying Deep Learning models to chest X-ray (CXR) images, researchers have achieved success in diagnosing pulmonary disorders, including COVID-19 infection or pneumonia. For instance, Rajpurkar et al. [40] introduced a novel architecture which is a Convolutional Neural Network (CNN) based on DenseNet121, to classify 14 different lung disorders. They trained it on approximately 100,000 X-ray images, naming it CheXNet. Their approach outperformed typical radiologists, particularly in terms of the F1 metric. Similarly, another study [41] proposed a method for using pre-trained convolutional neural networks to automatically diagnose COVID-19 pneumonia from CXR images with an accuracy of 99%. Furthermore, Vaid et al. [42] suggested a deep learning based model for COVID-19 detection from CXR images and images related to clustering based on the model's output. Lastly, another method for COVID-19 detection was presented in Togaçar et al.[30].

Chowdhury et al. [43] employed X-radiation (X-ray) images of chest for development of unique structure called PDCOVID Net, which relies on CNN which is dilated in parallel. They achieved a detection accuracy of 96.58% with this approach, using CNN in the parallel stack to abduct and extend the required information effectively. Abbas et al. [20] proposed DeTraC, a deep convolutional neural network, for identifying COVID-19 patients from their chest X-ray images. Their approach included a decomposition method to analyze class boundaries, resulting in high accuracy (93.1%) and sensitivity (100%) in detecting abnormalities from the dataset.

Che Azemin et al. [44] utilized DL technique based on the residual network named as ResNet-101 which is a CNN based model. This methodology involved pre-training with thousands of images to identify significant objects and abnormalities in chest X-ray images, achieving a precision rate of 71.9%.

The layers—patient layer, cloud layer and the hospital layer—make up framework described by authors El-Rashidy et al. [45]. A collection of data was gathered from the patient layer utilizing a mobile app and some wearable sensors. The patient X-ray pictures were used to train a deep learning model based on neural networks to recognize COVID-19. The proposed model attained 98.85% specificity and 97.9% accuracy.

Khan and Aslam [46] devised a novel architecture for carrying out the diagnoses of radiological (X-ray) images as either COVID-19 infected or normal, utilizing deep learning based models like residual network (ResNet50), Visual Geometry

TABLE 2   Results of classification without utilizing LBP (Local Binary Pattern) feature extraction [16].

Method	Accuracy	Sensitivity/recall	Specificity	F_score
Cubic Support Vector Machine (SVM)	94.1%	91.25%	95.61%	91.86%
Linear Discriminant Analysis (LDA)	87.0%	80.87%	89.91%	81.52%
Quadratic Discriminant Analysis (QDA)	86.7%	86.54%	91.64%	82.29%
Ensemble method	88.4%	79.93%	90.43%	83.12%
Kernel Naïve Bayes	71.2%	72.66%	83.89%	64.48%
KNN weighted classifier	90.0%	82.82%	91.37%	85.51%

TABLE 3   Results of classification utilizing LBP (Local Binary Pattern) feature extraction [16].

Method	Accuracy	Sensitivity/recall	Specificity	F_score
Cubic Support Vector Machine (SVM)	98.05%	90.99%	95.39%	91.71%
Linear Discriminant Analysis (LDA)	95.36%	80.56%	89.78%	81.24%
Quadratic Discriminant Analysis (QDA)	95.71%	86.41%	91.62%	82.10%
Ensemble method	95.80%	79.38%	90.06%	82.61%
Kernel Naïve Bayes	89.03%	72.46%	83.78%	64.25%
KNN Weighted Classifier	96.31%	82.27%	91.12%	84.94%

Group (VGG16), VGG19 and densely-connected-convolutional networks (DenseNet121) which were pre-trained. Among these models, VGG16 and VGG19 exhibited highest accuracy rate. Their suggested technique consisted of two stages: (i) Data preprocessing (ii) data augmentation. These stages were followed by transfer learning, resulting in an impressive accuracy of 99.3%.

In a different approach by Loey et al. [33], a dataset containing 307 images categorized into four classes—normal, pneumonia bacterial, pneumonia virus and COVID-19 was employed to train deep learning based transfer models: GoogleNet, AlexNet, and ResNet18. To optimize memory usage and execution time, they conducted study in three scenarios. Remarkably, GoogleNet achieved 100% accuracy for testing and its accuracy was 99.9% for validation. Barstugan et al. [47] structured their classification based on the quantity of patches, utilizing 150 data points and conducting a 10-fold cross-validation. The GLCM-SVM approach yielded the highest predicted accuracy at 98.91%, while the LDP-SVM approach exhibited lower accuracy, at 50.70%, respectively.

Rohmah and Bustamam's study [48], involving 2,200 data points, achieved an accuracy rate of 97.5% using the LBP-SVM approach. In a model designed by Wang et al. [22], two models ResNet-101, ResNet-151 were combined in fusion to dynamically improve their weight ratio. This model categorized chest X-ray images into three categories: normal, viral pneumonia and COVID-19. This method gained testing accuracy—96.1%.

Yoo et al. [49] utilized Chest X-ray radiography (CXR) images for COVID-19 classification using deep learning-based classifier named as decision-tree classifier. This classifier, implemented on the PyTorch framework, incorporated three binary decision trees, with the third tree achieving average accuracy—95% in classifying CXR images as– normal and abnormal. Sahlol et al. [50] proposed an enhanced ensembled classification strategy for classifying COVID-19 images. They utilized multiple CNN and the marine predators algorithm, with the marine predators method identifying the pertinent features from images after feature extraction using the CNNs' inception architecture. However, the study did not explore fusion approaches to further enhance COVID-19 image categorization and feature extraction.

Most of the research published so far has relied on chest X-ray images for COVID-19 diagnosis, highlighting the importance of analyzing these images as a dependable tool for physicians and radiographers. However, there are instances where achieving the required accuracy in classification results becomes challenging due to data imbalance and a deficiency in extracting essential features from the images. To address these limitations and enhance COVID-19 detection accuracy, Alam et al. proposed a solution involving the combination of features derived from HOG and CNN, followed by classification using CNN.

4 Evaluation metrics

We evaluate how well our suggested model performs using a number of benchmarked assessment measures for classification tasks. The correctness of any deep learning model is typically our main concern. The confusion matrix accomplishes this. It is a N*N matrix that aids in assessing how well a deep learning model performs when applied to a classification issue. In our study, we have considered the deep learning algorithms utilized to detect COVID-19 and diagnose using three measures of performance: accuracy, sensitivity, and specificity.

TABLE 4 Studies examining the performance of deep learning algorithms in the detection and diagnosis of COVID-19.

References	Population	Method	Model	Category of image data	Results of validation process (%)
Wang et al. [17]	5,372 (two datasets)	Deep learning (DL) method	DenseNet121-FPN	CT (Computed Tomography) images	Accuracy—87 and 88 Sensitivity/recall- 80.3 and 79.35 Specificity—76.3 and 81.1
Butt et al. [18]	618	Convolutional Neural Network (CNN)	Residual Network (ResNet-18)	CT (Computed Tomography) images	Sensitivity/recall—98.2 Specificity—92.2
Waheed et al. [19]	1,124	Generative Adverserial Network (CovidGAN)	ACGAN3, VGG-16	X-radiation (X-ray) images	Accuracy—95 Sensitivity/recall—90 Specificity—97
Abbas et al. [20]	6,523	CoroNet (Deep learning based method)	VGG-16	X-radiation (X-ray) images	Accuracy—97
Panwar et al. [21]	337	Deep learning (DL) method	VGG-16	X-radiation (X-ray) images	Sensitivity/recall—97.62 Specificity—78.57 Accuracy—88.10
Wang et al. [22]	181	Deep learning (DL) method	VGG-19	X-radiation (X-ray) images	Accuracy—96.3
Hasan et al. [23]	321	Deep transfer learning (TL) method	LSTM neural network classifier	CT (Computed Tomography) images	Accuracy—99.68
Ko et al. [24]	3,993	FCONet (Deep learning method)	Residual Network (ResNet-50)	CT (Computed Tomography) images	Sensitivity/recall—99.58 Specificity—100 Accuracy—99.87
Li et al. [25]	4,356	Deep learning (DL) method	Residual Network (ResNet-50)	CT (Computed Tomography) images	Sensitivity/recall—90 Specificity—96
Pereira et al. [26]	1,144	Convolutional Neural Network CNN (Deep learning method)	Inception-V3	X-radiation (X-ray) images	F1 score—89
Rahimzadeh and Attar [27]	11,302	Deep learning (DL) method	ResNet50V2 and Xception	X-radiation (X-ray) images	Accuracy—95.5 Overall average accuracy—91.4
Sethy et al. [28]	381	Convolutional Neural Network (CNN) and Support Vector Machine (SVM)	ResNet-50	X-radiation (X-ray) images	Sensitivity/recall—95.33
Ni et al. [29]	14,531	Deep learning (DL) method	Convolutional 3D U-Net and MVPNet	CT (Computed Tomography) images	Sensitivity/recall—100 for detecting patient Sensitivity/recall for per-lung Lobe lesion—0.96
Ardakani et al. [30]	194	Deep learning (DL) method	AlexNet, VGG-16, VGG-19, GoogLeNet, SqueezeNet	CT (Computed Tomography) images	Accuracy—99.51 Sensitivity/recall—100 Specificity—99.02
Apostolopoulos and Mpesiana [31]	455	CoroNet (Deep learning method)	MobileNetV2	X-radiation (X-ray) images	Sensitivity/recall—97.36 Specificity—99.42 Accuracy—99.18
Wu et al. [32]	495	CoroNet (Deep learning method)	VGG-19	CT (Computed Tomography) images	Accuracy—76.0 Sensitivity/recall—81.1 Specificity—61.15
Yang et al. [33]	295	Deep learning (DL) method	DenseNet	CT (Computed Tomography) images	Accuracy—92 Sensitivity/recall—97 Specificity—7
Saiz and Barandiaran, [34]	1,500	Convolutional Neural Network (CNN) using transfer learning	VGG-16 (SDD)	X-radiation (X-ray) images	Accuracy—94.92 Sensitivity/recall—94.92 Specificity—92 F1 score—97
Loey et al. [35]	306	Deep learning (DL) method	GoogLeNet	X-radiation (X-ray) images	Accuracy—100
Pathak et al. [36]	852	Transfer Learning (TL) technique	Residual Network (ResNet-50)	CT (Computed Tomography) images	Accuracy—93
Brunese et al. [2]	6,523	CoroNet (Deep learning method)	VGG-16	X-radiation (X-ray) images	Accuracy—96.3

4.1 Accuracy

The ability of a model to correctly assign samples to their appropriate labels is referred to as accuracy. It can be explained mathematically as follows:

$$Accuracy =$$
$$(TP + TN) / (TP + TN\ TP + TN + FP + FN) \qquad (1)$$

4.2 Recall/sensitivity

Recall, often referred to as sensitivity of a model, denotes the model's thoroughness. It can be stated mathematically as follows:

$$Recall = (TP) / (TP + FN) \qquad (2)$$

4.3 Specificity

Specificity can be defined as the algorithm or model's ability to correctly predict a true negative in each of the available categories.

$$Specificity = (TN) / (TN + FP) \qquad (3)$$

4.4 F1 score

The accuracy of any test is measured by F-measure called as F-score. It is evaluated from the recall and precision.

$$F1Score = 2^* (Precision \times Recall) / (Precision + Recall) \qquad (4)$$

where
$$Precision = (TP)/(TP + FP)$$

In this instance, the TP (True Positive) value indicates that the patient is COVID-19 positive and that the model properly recognized him as such. Similar to this, FP (False Positive) denotes a patient who does not actually have COVID-19 but whose COVID-19 positivity has been identified by the model. While the model had indicated that the patient would not have COVID-19, TN (True Negative) indicates that he does not. Similar to this, FN (False Negative) suggests that a patient has COVID-19 when the model has shown that he does not. Table 4 presents the details of different articles in terms of population, method, model, type of dataset and the accuracy of the model.

5 Research gaps in studied literature

(i) Compared to other prevalent lung disorders, COVID-19 is far less well-documented in chest X-ray and CT imaging. The systems under review revealed this issue with the data imbalance.

(ii) As the dataset, experimental environment and test cases for the systems discussed in this study are all somewhat distinct, it is difficult to single out one system in particular.

(iii) For COVID-19 patients, the available imaging data is mislabeled, noisy, fragmentary, and unclear. It is extremely difficult to train a deep learning architecture with such enormous and varied data sets. Numerous issues such as data redundancy, sparsity, and missing values must be fixed.

6 Limitations of the study

(i) A certain amount of domain-specific knowledge is assumed for this review work.

(ii) Certain details of the reviewed neural networks are not covered here. These details include the number of layers, layer specifications, learning rate, batch size, dropout layer, optimizer, and loss function. For such details readers are urged to consult relevant references.

(iii) This paper does not offer any qualitative findings of diagnosis in CT or X-ray pictures.

7 Conclusion

Deep learning is increasingly utilized in the realm of COVID-19 radiologic image analysis to reduce errors in detecting and diagnosing the illness, thereby providing patients with a unique opportunity for rapid, cost-effective, and secure diagnostic services. Since the COVID-19 pandemic began in the fourth quarter of 2019, there has been a scarcity of data to train deep learning models. To address this scarcity, researchers generated unique datasets by integrating multiple repositories. Our study involved the detailed enquiry of each model independently and then comparing their results. Among all the studied methods, it is clear that when GoogleNet (22 layers deep CNN model) is selected to be the main deep transfer model, it achieved 100% accuracy for X-ray images where as Residual Network (ResNet-50) has provided highest sensitivity—100% and specificity—99.2% for CT images. The ensemble model (AlexNet, VGG-16, VGG-19, GoogLeNet, SqueezeNet) reached to accuracy—99.51%, Sensitivity/Recall—100% and Specificity—99.02%.Among all the studied methods, deep transfer learning model (GoogleNet) and Residual Network (ResNet-50) performs the best. The ensemble model has also provided remarkable results and it's worth emphasizing that ensemble models can significantly enhance the performance measure of deep learning algorithms. More reliable deep learning models can be introduced by researchers to achieve 100% accuracy, 100% precision, 100% specificity, 100% sensitivity and 100% F1-score in future.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

KA: Conceptualization, Data curation, Investigation, Methodology, Software, Writing—original draft, Writing—review & editing. AK: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Writing—review & editing. YG: Conceptualization, Funding acquisition, Investigation, Resources, Supervision, Validation, Visualization, Writing—original draft, Writing—review & editing. YH: Conceptualization, Investigation, Methodology, Software, Supervision, Writing—review & editing. MM: Conceptualization, Investigation, Methodology, Software, Supervision, Writing—review & editing. AS: Conceptualization, Investigation, Supervision, Validation, Writing—review & editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Abhinav S. Deep Covid-coronavirus diagnosis using deep neural networks and transfer learning. *Medrxiv.* 2021–05.

2. Xu Y, Yang Z, Li X, Kang H, Yang X. Coronavirus disease 2019 (COVID-19): situation report, 75. Dynamic opposite learning enhanced teaching–learning based optimization. *Knowl Based Syst.* 188:104966. doi: 10.1016/j.knosys.2019.104966

3. Goodarzian F, Ghasemi P, Gunasekaren A, Taleizadeh AA, Abraham A. A sustainable-resilience healthcare network for handling COVID-19 pandemic. *Ann Operat Res.* (2022) 312:761–825. doi: 10.1007/s10479-021-04238-2

4. Islam M, Karray F, Alhajj R, Zeng J. A review on deep learning techniques for the diagnosis of novel coronavirus (COVID-19). In: *Special Section On AI And IOT Convergence For Smart Health.* Vol. 9. IEEE (2021).

5. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data.* (2019) 6:60. doi: 10.1186/s40537-019-0197-0

6. Trier D, Jain AK, Taxt T. Feature extraction methods for character recognition—a survey. *Pattern Recogn.* (1996) 29:641–62. doi: 10.1016/0031-3203(95)00118-2

7. Dumagpi JK, Jeong YJ. Evaluating GAN-based image augmentation for threat detection in large-scale X-ray security images. *Appl Sci.* (2020). doi: 10.3390/app11010036

8. Singh V, Kumar B, Patnaik T. Feature extraction techniques for handwritten text in various scripts: a survey. *Int J Soft Comp Eng.* (2013) 3.

9. Beers A, Brown J, Chang K, Campbell JP, Ostmo S, Chiang MF, et al. High-resolution medical image synthesis using progressively grown generative adversarial networks. *arXiv.* (2018).

10. Bhuyan HK, Chakraborty C, Shelke Y, Pani SK. COVID-19 diagnosis system by deep learning approaches. *Exp Syst.* (2022) 39:e12776. doi: 10.1111/exsy.12776

11. Hammoudi K, Benhabiles H, Melkemi M, Dornaika F, Arganda-Carreras I, Collard D, et al. Deep learning on chest X-ray images to detect and evaluate pneumonia cases at the era of COVID-19. *J Med Syst.* (2021) 45:75. doi: 10.1007/s10916-021-01745-4

12. Keidar D, Yaron D, Goldstein E, Shachar Y, Blass A, Charbinsky L, et al. COVID-19 classification of X-Ray images using deep neural networks. *Eur Radiol.* (2021) 31:9654−63. doi: 10.1007/s00330-021-08050-1

13. Alam N-A, Ahsan M, Based MA, Haider J, Kowalski M. COVID-19 Detection from chest X-ray images using feature fusion and deep learning. *Sensors.* (2021) 21:1480. doi: 10.3390/s21041480

14. Dai W, Doyle J, Liang X, Zhang H, Dong N, Li Y, et al. SCAN: Structure correcting adversarial network for chest X-rays organ segmentation. *arXiv.* (2017). doi: 10.1007/978-3-030-00889-5_30

15. Nie D, Trullo R, Lian J, Petitjean C, Ruan S, Wang Q, et al. Medical image synthesis with context-aware generative adversarial networks. *Proc Int Conf Med Image Comput Comput Assist Intervent.* (2017) 10435:417–25. doi: 10.1007/978-3-319-66179-7_48

16. Aslan N, Dogan S, Koca GO. Classification of Chest X-ray COVID-19 images using the local binary pattern feature extraction method. *Turk J Sci Technol.* (2022) 17:299–308. doi: 10.55525/tjst.1092676

17. Wang S, Zha Y, Lietal W. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *Eur Respir J.* (2020) 56:2000775. doi: 10.1183/13993003.00775-2020

18. Butt C, Gill J, Chun D, Babu BA. Deep learning system to screen coronavirus disease 2019 pneumonia. *Appl Intell.* (2020) 50:4874. doi: 10.1007/s10489-020-01714-3

19. Waheed A, Goyal M, Gupta D, Khanna A, Al-Turjman F, Pinheiro PR. CovidGAN: data augmentation using auxiliary classifier GAN for improved COVID-19 detection. *IEEE Access.* (2020) 8:91916–23. doi: 10.1109/ACCESS.2020.2994762

20. Abbas A, Abdelsamea MM, Gaber MM. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Appl Intell.* (2021) 51:854–64. doi: 10.1007/s10489-020-01829-7

21. Panwar H, Gupta PK, Siddiqui MK, MoralesMenendez R, Singh V. Application of deep learning for fast detection of COVID-19 in X-rays using nCOVnet. *Chaos Solit Fract.* (2002) 138. doi: 10.1016/j.chaos.2020.109944

22. Wang N, Liu H, Xu C. Deep learning for the detection of COVID-19 using transfer learning and model integration. In: *Proceedings of the 2020 IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China.* (2020). p. 281–4.

23. Hasan AM, Al-Jawad MM, Jalab HA, Shaiba H, Ibrahim RW, Al-Shamasneh AR. Classification of COVID-19 coronavirus, pneumonia and healthy lungs in CT scans using Q-deformed entropy and deep learning features. *Entropy.* (2020) 22:517. doi: 10.3390/e22050517

24. Ko H, Chung H, Kang WS, Kim KW, Shin Y, Kang SJ, et al. COVID-19 pneumonia diagnosis using a simple 2D deep learning framework with a single chest CT image: model development and validation. *J Med Int Res.* (2020) 22:e19569. doi: 10.2196/19569

25. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, et al. Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. *Radiology.* (2020) 296:200905. doi: 10.1148/radiol.2020200905

26. Pereira RM, Bertolini D, Teixeira LO, Silla CN, Costa YMG. COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. *Comp Methods Progr Biomed.* (2020) 194. doi: 10.1016/j.cmpb.2020.105532

27. Rahimzadehand M, Attar A. Amodifieddeepconvolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. *Informat Med Unlocked.* (2020) 19:100360. doi: 10.1016/j.imu.2020.100360

28. Sethy PK, Behera SK, Ratha PK, Biswas P. Detection of coronavirus disease (COVID-19) based on deep features and support vector machine. *Int J Math Eng Manag Sci.* (2020) 5:643–51. doi: 10.33889/IJMEMS.2020.5.4.052

29. Ni Q, Sun ZY, Qi L, Chen W, Yang Y, Wang L, et al. A deep learning approach to characterize 2019 coronavirus sisease (COVID-19) pneumonia in chest CT images. *Eur Radiol.* (2002) 30. doi: 10.1007/s00330-020-07044-9

30. Togaçar M, Ergen B, Cömert Z. COVID-19 Detection using deep learning models to exploit social mimic optimization and structured chest X-ray images using fuzzy color and stacking approaches. *Comput Biol Med.* (2020) 121:103805. doi: 10.1016/j.compbiomed.2020.103805

31. Apostolopoulos ID, Mpesiana TA. COVID-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Phys Eng Sci Med.* (2020) 43:635–40. doi: 10.1007/s13246-020-00865-4

32. Wu X, Hui H, Niu M, Li L, Wang L, He B, et al. Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: a multicentre study. *Eur J Radiol.* (2020) 128:109041. doi: 10.1016/j.ejrad.2020.109041

33. Loey M, Smarandache FM, Khalifa NE. Within the lack of chest COVID-19 X-ray dataset: a novel detection model based on GAN and deep transfer learning. *Symmetry.* (2020) 12:651. doi: 10.3390/sym12040651

34. Zhao D, Zhu D, Lu J, Luo Y, Zhang G. Synthetic medical images using F&BGAN for improved lung nodules classification by multi-scale VGG16. *Symmetry.* (2018) 10:51. doi: 10.3390/sym10100519

35. Brunese L, Mercaldo F, Reginelli A, Santone A. Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. *Comp Methods Progr Biomed.* (2002) 196:105608. doi: 10.1016/j.cmpb.2020.105608

36. Wang Z, Cui X, and Gao L. A hybrid model of sentimental entity recognition on mobile social media. *Eurasip J Wireless Commun Netw.* (2016) 253. doi: 10.1186/s13638-016-0745-7

37. Hasoon JN, Fadel AH, Hameed RS, Mostafa SA, Khalaf BA, Mohammed MA, et al. COVID-19 anomaly detection and classification method based on supervised machine learning of chest X-ray images. *Results Phys.* (2021) 31:105045. doi: 10.1016/j.rinp.2021.105045

38. Ghasemi P, Goodarzian F, Simic V, Tirkolaee EB. A DEA-based simulation-optimisation approach to design a resilience plasma supply chain network: a case study of the COVID-19 outbreak. *Int J Syst Sci.* 10:2224105. doi: 10.1080/23302674.2023.2224105

39. Goodarzian F, Ghasemi P, Gunasekaren A, Labibi A. A fuzzy sustainable model for COVID-19 medical waste supply chain network. *Fuzzy Optimiz Decis Making.* (2020).

40. Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, et al. CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv.* (2017).

41. Chowdhury MEH, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahbub ZB, et al. Can AI help in screening viral and COVID-19 pneumonia? *IEEE Access.* (2020) 8:132665–76. doi: 10.1109/ACCESS.2020.3010287

42. Vaid S, Kalantar R, Bhandari M. Deep learning COVID-19 detection bias: accuracy through artificial intelligence. *Int Orthopaed.* (2002) 44:1539–42. doi: 10.1007/s00264-020-04609-7

43. Chowdhury NK, Rahman MM, Kabir MA. PDCOVIDNet: a parallel-dilated convolutional neural network architecture for detecting COVID-19 from chest X-ray images. *Health Inf Sci Syst.* (2020) 8:1–14. doi: 10.1007/s13755-020-00119-3

44. Che Azemin MZ, Hassan R, Mohd Tamrin MI, Md Ali MA. COVID-19 deep learning prediction model using publicly available radiologist-adjudicated chest X-ray images as training data: preliminary findings. *Int J Biomed Imaging.* (2020) 22:512–15. doi: 10.1155/2020/8828855

45. El-Rashidy N, El-Sappagh S, Islam SMR, El-Bakry HM, Abdelrazek S. End-to-end deep learning framework for coronavirus (COVID-19). *Detect Monit Electron.* (2020) 9:1439. doi: 10.3390/electronics9091439

46. Khan IU, Aslam N. Deep-learning-based framework for automated diagnosis of Covid-19 using X-ray images. *Information.* (2020) 11:419. doi: 10.3390/info11090419

47. Barstugan M, Ozkaya U, Ozturk S. Coronavirus (covid-19) classification using CT images by machine learning methods. *arXiv.* (2020). doi: 10.48550/arXiv.2003.09424

48. Rohmah LN, Bustamam A. Improved classification of coronavirus disease (covid-19) based on combination of texture features using CT scan and x-ray images. In: *2020 3rd International Conference on Information and Communications Technology (ICOIACT).* IEEE (2020). p. 105−9.

49. Yoo SH, Geng H, Chiu TL Yu SK, Cho DC, Heo J, Choi MS, et al. Deep learning-based decision-tree Classifier for COVID-19 diagnosis from chest X-ray imaging. *Front Med.* (2020) 7:427. doi: 10.3389/fmed.2020.00427

50. Sahlol AT, Yousri D, Ewees AA, Al-Qaness MA, Damasevicius R, Abd Elaziz M. COVID-19 image classification using deep features and fractional-order marine predators algorithm. *Sci Rep.* (2020) 10:1–15. doi: 10.1038/s41598-020-71294-2