Deep learning framework for early detection of COVID-19 using X-ray images

Kainat Khero[1] · Muhammad Usman[1] · Alvis Fong[2]

Abstract

The whole world is imposing efforts to combat the deadly COVID-19 virus that continues to have a disastrous effect on health, economy, education, transport & communication, and many other sectors. The crucial action taken to control its rapid spread is first to detect the infected person. Deep learning-based algorithms utilize mathematical models to detect Covid-19 cases. Deep learning approach is applied to track and diagnose Covid-19 and help radiologists and medical doctors enhance prognosis performance. X-ray images are popularly used deep learning methods for Covid-19 detection. However, the existing techniques suffer from several limitations that need to be addressed to detect Covid-19 cases more accurately: Firstly, there is a small number of Covid-19 images. Secondly, an unbalanced dataset. Thirdly, model overfitting, and fourthly, correct detection of Covid-19 and pneumonia cases sometimes does not provide accurate results because COVID-19 and pneumonia symptoms are similar. Therefore, this paper aimed to develop an automated solution to classify the detected Covid-19 into two classes to overcome the small and unbalanced dataset, and model overfitting problems. This study compared nine state-of-the-art CNN architectures through a transfer learning approach. Our approach achieved better results in comparison to the work done on this benchmark dataset yielding 99.86% accuracy with 99.9% recall using VGG-16 using deep learning model. The proposed framework presents a transfer learning technique to increase the performance of the deep learning-based Covid-19 detection method. Moreover, the Comparative evaluation presents that the proposed framework outperforms existing methods. Close results show that VGG-16 on a large dataset correctly identified COVID-19.

Keywords Covid-19 · Deep learning · Prediction · Transfer learning · Chest X-ray

✉ Alvis Fong
   alvis.fong@wmich.edu

   Muhammad Usman
   dr.usman@szabist-isb.edu.pk

[1] Department of Computer Science, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology (SZABIST), Islamabad Campus, Street # 09, Plot # 67, Sector H-8/4, Islamabad, Pakistan

[2] Department of Computer Science, Western Michigan University, 1903 W. Michigan Ave, Kalamazoo, MI 49008, USA

1 Introduction

COVID-19 first arose in December 2019 in Hubei province China. However, it has extended briskly worldwide since the start of 2020. This new virus is concomitant with various signs; some common symptoms are fever, dry cough, tiredness, pains, and severe symptoms are breathing or shortness of inhalation, chest pain, and in extreme cases, organ failure and decease.

On March 11, 2020, the World Health Organization (WHO) professed the COVID-19 spate a universal pandemic. As a result, travel restrictions, curfews, and contact bans have been enforced in copious countries worldwide, and all non-essential public institutions have been padlocked to passive the blowout of the COVID-19. The widespread of this deadly virus has the worst effect on human life quality and cripples many industries. The new variant of (the SARS-CoV) family rapidly transmits humans to humans, causing an ongoing pandemic of respiratory illness.

It takes the lives of billions of people from January 23 to December 21, 2020. Over 196 million confirmed cases with over 4.16 Million deaths from COVID-19 worldwide. The most recorded deaths observed in countries are the USA with 627 K, Brazil with 551 K, India with 422 K, and Mexico with 239 K death cases. On the other hand, Pakistan and Sri Lanka have minor death case records. Pakistan has over 1 million confirmed cases with 23 Thousand deaths, and Sri Lanka has over 299 Thousand confirmed cases with 4 Thousand deaths records.

The most challenging task to rheostat the feast of COVID-19 is to identify asymptomatic cases resulting in the virus spreading. In addition, to quick and accurate determine who is currently infected and who has been infected with COVID-19. There are three basic types of COVID-19 tests: nucleic acid, antigens, and serological (antibodies). COVID-19 tests are divided into two categories: Molecular and Serological assays, both of which use different kinds of test samples. Molecular uses nasal or oral swabs as a sample, and Serological uses blood as a sample. Molecular testing is done to identify COVID-19 patients who are currently infected. Serological testing identifies individuals who have developed antibodies to the virus.

With the increase in the cases of corona patients, Artificial Intelligence is necessary to spot COVID-19 infected sufferer at the initial or immediate stage and segregate them so that they would not infect other healthy populations. Therefore, the remedial diligence is regarding the latest expertise to combat COVID-19. AI is one such technology. The recent years' advancement in Artificial Intelligence (AI) in the medical specialty can automatic detection of many diseases such as the discovery of cancer types, the capacity of air in lungs, advanced bone clampdown in x-rays, glaucoma, harm to the tiny vessels in retina classification, protuberance classification pores and skin gaze classification M.rahizadeh et al. [22].

AI plays an essential role in this pandemic, which can easily track the extent of COVID-19, categorize high menace patients, and detect the infected virus. AI can support us to combat COVID-19 by populace screening, therapeutic help, report, and offers about infection virus control. Narin et al. [18] proposed this technology can mend the formation, treatment, and testified consequences of the COVID-19 patient. Artificial Intelligence (AI) Algorithms can be applied to identifying diseases using medical imaging technologies. Initial research becomes evaluates that chest images of infested patients are evidence of COVID-19 Gunraj et al. [10]. Computer Tomography (CT), Magnetic resonance imaging (MRI), and X-ray images can be occupied quickly to diagnose COVID-19 [9].

Regarding chest CT and X-ray medical images techniques used for diagnosis of COVID-19, a few advantages differentiate between both techniques. Likewise, the CT provides

more details about the patient's status and is comparatively quick compared with the X-ray technique. While chest X-rays can get the results at a lower price and lower radiation.

With this inspiration, numerous researchers are working on deep learning and machine learning techniques applied to medical images, e.g., Chest CT and X-ray images, to create accurate, fast identification frameworks for COVID-19 detection. The authors in [9, 12, 30] utilized the chest CT technique for the diagnosis of COVID-19, while authors in papers [4, 10, 24, 26] exploited X-ray images. The authors of [19] used both CT and X-ray chest datasets in their paper. In [4, 9, 11] authors utilized Deep learning model, ResNet-50, to detect coronavirus and non-coronavirus patients from CT scan images based on two levels: 2D and 3D [9]. The authors of [4] recommended COVID-19 detection method built on VGG-16 CNN by exploiting transfer learning. This method achieved high accuracy on model-1 96.0% and model-2 98.0%. The authors of [11] Employed ResNet-50 and achieve an accuracy of 96.0%.

The authors of [3, 11, 15, 22] utilized hybrid methodology by combining the features of two different models. In [11], Resnet50 and VGG16 deep learning models were used and achieved an accuracy of 91.24%. The authors of [22] proposed the CNN extracted features using Xception and classification by Res-Net50V2. They achieved an accuracy of 91.4%. The authors of [15] applied SVM classifier with GLSZM feature extraction and achieved a high accuracy of 99.68%. In [3], the authors evaluated from the experiment that DenseNet121 plus Bagging tree produced 99% accuracy, and on the other hand, ResNet50 plus LightGBM achieved 98%. Authors of [1, 24] combined method of DL and classical ML to extract features and classification on SVM to achieve accuracy 98.97% and 98.66%.

Ghoshal et al. [8] proposed a deep learning-based method to evaluate the improbability and explicable in identifying COVID-19. The author applied the Bayesian convolutional neural network (BCNN) technique. Open source chest X-ray images Dataset is composed of GitHub and Kaggle repositories. This proposed model found prediction uncertainty is exceptionally correlated with prediction accuracy. This model has improved results in diagnosis accuracy from 85.2% to 92.9% utilizing the pre-trained VGG-16 architecture. Narin et al. [18] suggested an approach based on deep learning using transfer learning techniques for COVID-19 prediction. The model categorizes two modules (i) COVID-19 and (ii) Normal. The author implements three pre-trained models, namely, InceptionV3, ResNet50, and InceptionResNetV2. The experimented results show that ResNet50 for binary classification achieved a higher accuracy of 98% compared with other models.

The Data insufficiency is due to the diverse nature of source data of datasets, like electronics health care records, regional, even per hospital level. The challenge is to combine these sources and overcome practical differences across each source [5]. As we have many images about pneumonia and other bacteria, we have very few images of COVID-19, which creates the problem of data unbalance, while we select pneumonia and other bacteria images equal to COVID-19 images that affect on accuracy [22]. The challenge COVID-19 and other pneumonia cases sometimes do not provide correct results because symptoms of COVID-19 and bacteria are similar. Deep learning models are applied to large data to get correct results. A model initially seems to learn well, predicting with >99% accuracy on the original dataset, but later, when you evaluate your model on a new dataset accuracy drops to 50%.

As general problems, we observed from literature 1. Unbalance ratio of datasets, 2. Correct detection of COVID-19, 3.Small dataset, 4. Outfitting & overfitting datasets, 5.Efficient feature extractor, 6. Accurate COVID-19 Classifier. Collecting and interpreting more data is costly but a biased or incomplete dataset will only produce similarly flawed trained models.

The core objectives of this study make this framework novel. In earlier work, the focus is on the COVID-19 detection model from medical imagining. However, there is no concentration

on developing an accurate prediction model. Moreover, the practice of Deep Learning for effective prediction model building is never investigated in earlier research. There is no systematic approach in the literature to investigate the following issues:

- An improved model is developed to incorporate COVID-19 detection. The system will learn and classify COVID-19 patients and Non-Covid-19 patients.
- Previously techniques have never been critically evaluated the Deep Learning studies, and their impact is not studied with insights about accurate prediction models to increase COVID-19 detection accuracy.
- A generic methodology has been proposed and implemented that can accurately predict COVID-19 patients. A complete automated framework, a prediction model has been designed based upon the chest X-ray as input.
- The proposed framework has been validated through real-world datasets of chest X-ray images from 2020 to 2021 to study the efficacy of Deep Learning procedures.

The Exemplary is constructed on Deep Learning techniques. We compared our models on two different learning rates, 0.001 and 0.005. Once the best learning rate is identified, we compare the best learning rate on different epochs, numbers 15 and 25. We utilized nine models VGG19, ResNet-50, Inception-v3, Xception, DenseNet-201, MobileNet, VGG16, NASNet-Mobile and EfficientNetB0. This experiment is implemented on three datasets, (i) Dataset.

has 422 images, (ii) Dataset has a total of 4608 images, and (on two classes Covid-19 and NonCovid-19. Dataset is split into 70% training and 30% testing.

Our model is analyzed based on Different learning rates, epoch sizes, and accuracy. It is observed from the results that Dataset 1 provides better results on Inception-v3 and Xception; on the other hand, Dataset 2 produces better results on VGG-16.

The primary focus of the contribution is towards a framework design, an automated analysis tool to help in this pandemic for early COVID-19 detection. This work presented experimental results on the real-world benchmark datasets, which are open source and available on famous repositories. In addition, a complete framework has been designed to expand the prediction of COVID-19 patients expending chest X-ray images. Finally, the work represented the observation result on the real datasets to validate the proposed methodology and revealed the efficacy of the devised architecture.

The structure of the paper is as follows. First, Section 2 introduces the related work essential to understanding our research. Section 3 presents the dataset used for the experiments. After that, Section 4 describes the proposed framework to classify Covid-19 into its categories, the architectures used, and the training process carried out in the study. Finally, in Section 5 we present the experimental results, followed in Section 6 by the conclusion and future work.

2 Related work

Deep Learning (DL), a subset of ML, reaches efficiency and suppleness related to conventional Machine Learning prototypes by representing spur from biotic neural networks to decipher an eclectic multiplicity of complex chores. The Deep Learning model involves several deposits such as input, output, and hidden layers. These layers modify the data with convolutional filters. In this era, DL is considered an efficient technique to detect COVID-19 from medical imaging. The most common Deep learning approach for COVID-19 detection is Convolutional Neural Network. CNN architecture has many layers containing

convolutionary, nonlinear, pooling, and fully connected layers with parameters that can be modified as needed. It utilized various algorithms to advance an expert method that can identify COVID-19 patients. It provides an efficient result, mainly dealing with image data such as classification in large image datasets, computer vision, and natural language processing (NLP).

This research review focused on Deep learning approach of Convolutional Neural Network CNN to diagnose COVID-19 from medical images Chest CT and X-ray images.

The authors of [9, 12, 30] utilized Chest CT scan images datasets to diagnose COVID-19. In [30], RT-PCR method was compared with chest CT images to evaluate infected patients of the corona virus. RT-PCR has loss of responsiveness, inadequate stability, and comparatively, lengthy processing time changed into unfavorable to the management of the ailment epidemic. RT-PCR means that many COVID-19 sufferers won't be recognized and obtain appropriate remedies on time. Insufficient sources kits. While Chest CT is stress-free to perform and can produce fast detection of corona. As associated with RT-PCR, chest CT imaging can be a more consistent, concrete, high responsiveness, accuracy, and speedy technique to identify and evaluate COVID-19.CT scan changed into advance than RT-PCR. It is observed the enhancement by implementing the chest CT images datasets. The authors of [9] utilized the ResNet-50 CNN model to detect coronavirus and non-coronavirus patients from CT scan images and quantify the disease burden and tracking the disease progression. The scheme has two separate levels: Subsystem A: 3D study nodules and focal opacities, where Subsystem B: 2D detects and localizes ground glass infiltrates of huge-sized verbose of each slice. The previous methods can detect smaller-sized focal GG opacities, unlike Ophir et al. method that introduced detecting more diffuse global GG opacities. This paper used multiple international private datasets. As ResNet-50 is good at saving computing resources, high accuracy and training time, but the network is 50 layers deep, we can apply more deep layers e.g. 101, 152, and 1202 to get more enhanced performance.

The authors of [4, 10, 11, 14, 24, 26] utilized X-ray images COVID-19 datasets, with papers [10, 26] utilizing the same datasets. Hall et al. [11] proposed a method by adding CNN model features to COVID-19 classification. In this paper Resnet 50 and VGG 16 deep learning techniques are applied to a small dataset of X-ray images ($n = 135$) of COVID-19 cases and pneumonia ($n = 320$). To form a balanced training dataset, first COVID-19 ($n = 102$) personal belongings and ($n = 102$) pneumonia personal belongings to get two upon three being viral pneumonia and one upon three bacterial pneumonia cases. Five classifiers are utilized a snapshot ensemble because it needs lots of data to set parameters, and we have insufficient datasets of corona. Compared with the existing Resnet50 model, our proposed method improved accuracy and achieved an overall accuracy of 91.24%. For testing, the model 10-fold cross-validation is utilized. There are some limitations of our method 1. It reduced the resolution of the COVID-19 images, 2. its incompetent to extricate disease in patients who do not have human evident Chest X-ray findings on presentation. 3. Lack of information on outcomes 4. It is only feasible for breathing issues.

Unlike [9, 11], authors of [10, 22, 26] contribute the same methodology. Gunraj et al. [10] proposed a lightweight structure referred to as the COVID-Net network that merges varied layers of deep CNN on different sizes of kernel size model. This paper designed dataset called COVIDx by utilizing five open source datasets containing COVID-19, normal and non-COVID19 pneumonia. The process united "Human-Driven" strategy prototyping with "Machine-Driven" strategy exploration for the detection of COVID-19 cases. Human-driven strategy was used to make three estimates: 1) Normal, 2) Disease, 3) COVID-19. Design macro-architecture and micro-architecture optimal that classify and acquired by machine-driven approach to building the final deep neural network

architecture. This method has great description ability, operation training, and memory efficiency. PPV is higher in recommended method on COVID-19 patients, which causes incorrect results where mean counterfeit COVID-19 infected cases are enhanced. To improve training methodology to generalize better from scenarios. An open source network design COVID-Net and Dataset COVIDx are available to the general public.

Similarly, Ucara et al. [26] proposed a Deep-Bayes-SqueezeNet-based-COVIDiagnosis-Net model based on deep learning. SqueezNet has more additional benefits than AlexNet. SqueezNet is: 1) Fewer parameters, 2) Reduce Communication transactions, 3) lightweight model, 4) contrivance on Embedded systems. In this paper, the Convolution network consists of a chain of a Conv, ReLu, activation functions, pooling, and batch regularization procedure. The author selected the Bayesian method for optimization, an efficient algorithm handling an unknown feature that calls for an approximation with the assistance of a few acknowledged samples, i.e. earlier knowledge. The architecture was designed on three stages 1. data preprocessing using offline augmentation 2.Model building 3. Model testing. Obtained dataset called COVIDx by combined COVID chest Xray and Kaggle chest Xray pneumonia datasets. The method achieved overall accuracy is 98.26%.

In the similar context, Rahizadeh et al. [22] designed a succession neural network by combining the extracted features of X-ception and Res-Net50V2. With the usage of each of the inception- primarily layers and residual- primarily layers, the superiority of the generated semantic features would be improved. Succession data input to the classifier. 1×1 kernel size of the convolutional layer was added after concatenating features with 1024 filter. Xception plus Res-Net50V2 very helpful in learn from succession extracted features and achieving high accuracy of 91.4% compared to others. To make the dataset balanced, they selected an equal number of covid19 and non-covid19 cases. The model is trained on 3-fold; each fold has 8 stages and total images of 3783. In each stage total images 633 are divided as 250 of normal cases and 234 pneumonia cases and 149 of COVID-19. Where all the covid19 and non-covid19 metaphors are communal for the respective stage, but images for normal and pneumonia cases were change on each step. As writer prevails to gain true effects in detecting COVID-19 instances however it drops accuracy due to the small dataset.

In [4], the authors recommended a COVID-19 diagnosis method built on VGG-16 CNN by exploiting transfer learning. Desist the head of fully connected layers on pre-trained ImageNet and after that construct layers involving (Combining, Compress, Impenetrable, Idler and a last Dense with the softmax stimulation" as fully connected layers for fine-tuning. Afterward, to complete fine-tuning setup, training done on fully connected layers by interruptions the convolutional weights of VGG16. Their methodology has three fold methods: 1. To detect normal & disease 2. To detect COVID-19 & pneumonia 3.To concentrate the areas in the X-ray someplace COVID-19 exist. The paper is composed of three datasets which are: 1.COVID-19, 2. Automated-Detection-COVID-19, 3. Nih-Chest-Xray and then data preprocessing step performed where augmentation applied on COVID-19 images and got a new dataset with total 5100 images where health 2000 cases, disease 2000 cases and COVID-191100 cases. After that data is split for training 3000 images and testing 2100. To decrease training time for neural network model by implementation of transfer learning will help to decrease errors. Used visualizer "Grad-CAM" to track, debug and predict COVID- 19 area into chest X-ray images. With this technique 1. Healthy & disease has accuracy 0.96 and 2.COVID-19 & pneumonia has accuracy 0.98. While to overcome overfitting dropout layer is used, but place of Dropout layer is effect on performance where it increase or decrease performance. In this paper author place the dropout layer before last layer where network is not able to "correct" errors. This place is not efficient one, it drops accuracy.

By applying efficient feature extraction techniques on medical images models can achieve even more good accuracy. In [11], the authors designed a deep transfer learning algorithm for COVID-19 classification that helps in efficiently extract features by employed ResNet-50. The algorithm is divided into 6 steps: 1.Take input images, 2.Extract features using ResNet-50, 3.Classification on transfer learning, 4.CNN model is trained on optimized hyper-parameters, 5. Apply k-fold validation, 6. Return the results of classifier. Refrain the preliminary parameter using transfer learning on deep layers. The dataset is collected from 1.https://link.springer.com/article/10.1007%2Fs10096-020-03901-z, 2.https://arxiv.org/abs/2003.13145. There are COVID-19 ($n=413$) images and ($n=439$) non-covid19 infected patient's images. The data is spilt 60% training and 40% testing. Out of training 10% of that used for validation. "Cost sensitive smooth loss function" helps to improve noisy and imbalance data". Proposed model achieves accuracy for training 96.2264% and testing 93.0189%. As compared with other supervised learning our designed model achieved outstanding results. The ResNet-50 trained on "hyper parameters optimization" but hyper-parameters with finest collection is not acceptable. The spatial complexity of 3D volumes facing encounter for accurate COVID-19 screening such as the labeling difficulty of infection areas, and the slight discrepancy between COVID-19 and Pneumonia in chest CT. While a small number of inventive works have made massive progress, they may be both annoying guide annotations of contagion areas or interpretability loss.

In [12], the authors recommended Attention-base-deep-3D Multi-instance-learning (AD3D-MIL) by proposed an algorithm based on 3D CT. the decisions made on CNN directly. As this approach helps to remove mistakes which occurs in intermediate processes but then again still it has lacking interpretability of outcomes called as black box model. For interpretable of model author offers algorithms that would be supplementary useful and convincing. Purposed method can generate instances automatically and to identify virtually COVID-19 from CT chest by attention mechanism to discover key cases. Data collected from the designated COVID-19 "hospitals in Shandong Province". Total Chest CT images 460, where 230 COVID-19, 100 with pneumonia, and 130 without pneumonia. Studies used "Binary Classification" while proposed methodology work on "Multi-Class Classification". This method achieved accuracy of 97.9% and provides scalability and interpretability. Ordinary assumption of MIL used by paper that assume "where a bag is positive bag if it has a single instance is positive" but this procedure unsuccessful in this case, where sack essentials with several confident instances [19].

Likewise, the authors of [12, 25] took into consideration multi-objective differential evolution (MODE) algorithm implemented on deep learning to prognosis the COVID-19 and NonCOVID-19 sufferers commencing Chest CT images. Have fitness function called "Multi-objective" and MODE algorithm is recommended for hyper parameters optimization. Features are extracted on convolutional, max pooling and Rectified linear unit (ReLU) layer on three dimensions width, height, and depth. Intricacy features size of spatial is minimize with the help of this model. 20-fold cross-validation used to reduce overfitting issue. The specific task of fully connected layers is to classify and evaluate probability for object in the image. Training and testing data is built by using COVID chest-Xray Dataset. The distribution of dataset on the percentage of training on 20%,30%,40%,50%,60%,70%,80%,90% and testing 80%,30%,40%,50%,60%,70%,80%,90%. With the comparative of competitive CNN, ANFIS, ANN classification algorithms, this proposed model achieved higher accuracy 1.9789%. This algorithm can be help to use for real time COVID-19 diagnoses but limited their benefits on metaheuristic method so there is no guarantee of an optimal solution.

The authors of [1, 15, 24] utilized the same methods for diagnosis COVID-19. In [24], the authors proposed a methodology to achieve high accuracy, and rapid detection of

COVID-19 and NonCovid-19 classes. Their approach used deep learning CNN architecture and Machine learning algorithms. The CNN model trained on customized layers have five blocks, three fully connected layers and softmax layer. Where individual block has Convolution, ReLU, normalization, pooling layers that extract deep-features from COVID-19-Radiology dataset. The dataset consist of total 2905 images which spilt on training 2798 and test 872 numbers of Chest Xray metaphors and due to the short of covid19 images, we preprocess data and extend covid19 data using offline augmentation. Afterwards extracted features uses as input for ML algorithms which were used models k-Nearest-Neighbor (KNN), Support Vector Machine (SVM), and Decision Tree (DT) and optimized on "Bayesian". Experimental results shows that SVM perform well then all other Classifiers. Accuracy of this model is 98.97% on SVM which is greater than other classifiers. As the model decreased the wrong analysis but then use CNN bring some disadvantages while feature extraction. CNN face difficulty when image is rotated or gradient. If objects of an image are not in same position, the ConvNet drops information that might not be able to classify the image.

As with shortage of COVID-19 images, datasets are available in hundreds and the deep learning classifiers need large dataset. In [1], the authors recommended a method which combine deep learning Res-Net-50 and ML SVM models to perform classification on small dataset with accurate results. Author compared thirteen CNN models (Alex Net, VGG 16, VGG 19, Google Net, ResNet 18, ResNet 50, ResNet 101, Inception V3, InceptionResNet V2, DenseNet 201, Xception Net, MobileNet V2 and Shuffle Net and machine learning model SVM selected for model training with three different descriptors ("LBP, HOG, GLCM") to check accuracies. It observed from tries that SVM classifier with LBP perform better than other descriptors. Features are extracted from Deep learning model thru pre-trained network by means of fully connected layers with pooing layers on MAX operation and then applied machine learning classifier SVM to classify from three classes covid19, normal and pneumonia. The FitCeCoC function code is used to train SVM classifier on K (K-1)/2.The kaggle and GitHub repositories are used to collect 381 images from COVID-19127, normal 127 and pneumonia127 images. Dataset is divided into training 80% and 20% testing and selected data for train and test on randomize function. Experimental outcomes done on 20 independent executions. It observed from tries that SVM plus LBP perform better than other descriptors and SVM plus ResNet50 achieved 98.66% accuracy which is higher than other CNN models but working on small dataset cause drop accuracy. It should performance on large dataset.

The authors of [15] selected Support Vector Machines (SVM) classifier and compared five different feature extractor on four different patches size 16×16, 32×32, 48×48, 64×64. As images are taken from different CT tools provides different information, so to overcome such diversity author created images sets on different types of images. The Experiment is done on four subsets, subset-1 Normal ($n = 5912$) and Disease ($n = 6940$), subset-2 Normal ($n = 942$) and Disease ($n = 1122$), subset-3 Normal ($n = 255$) and Disease ($n = 306$), subset-4 Normal ($n = 76$) and Disease ($n = 107$). Classification on two altered methods is done. Where (i) Four subsets takes as input and convert into vector without future extraction and classify it on SVM, (ii) Four subsets takes as input and extracted features on Grey Level Co Occurrence Matrix (GLCM), Local Directional Patterns (LDP), Grey Level Run Length Matrix (GLRLM), Grey Level Size Zone Matrix (GLSZM), and Discrete Wavelet Transform (DWT) and subsequently classify COVID-19 and non-COVID-19 on SVM model. SVM is one of efficient binary classifier. Validation done on, 2 fold, 5 fold and 10 folds. Manually dataset is created from CT images. Total numbers of images are 150, where disease cases are 53 taken from Italian hospital. Highest accuracy achieved 99.68% on SVM plus GLSZM feature extraction on 10-fold validation. This method increased the accuracy but it limited on small dataset and Non-standard dataset.

In [3], the authors compared fifteen deep learning feature extractors such as Mobile Net, Dense Net, Xception, ResNet, InceptioV3, Inception, ResNetV2, VGGNet, NASNet and six machine learning classifiers are Decision Tree (DT), Random Forest (RF), XG Boost, Ada Boost, Bagging and Light GBM classifiers. Dataset collection of CT scan & Xray metaphors from Dr. Joseph Cohen, kaggle Chest X-Ray and RSNA Pneumonia Detection datasets, where COVID-19 ($n = 137$) and Normal ($n = 137$) images. CNN is pre-trained model can be used to extract deep features from different type's images where images are converted into vector sparse descriptors. Overfitting problem can be occurring due to shortage of training dataset but by applying transfer learning method overfitting and computational resources can be reduced. It has generalization capability for unknown data. To increase generalization capability of dataset there is no data augmentation technique while Resizing, Normalization data pre-processing techniques are applied. It is evaluated from experiment that DenseNet121 plus Bagging tree produced 99% accuracy and on other hand, ResNet50 plus LightGBM achieved 98% accuracy on COVID-19 and Normal cases patients. This approach has reduced generality errors and analysis time. To increase ResNet50 layers most efficient deep features can be extracted.

The authors of [24] suggested the method to correctly analysis the covid19 infected patients by implementing artificial intelligence sub fields' machine learning and deep learning. This paper proposed Deep Learning Methodology (DLM) which works on X-ray images. To extract features from Xray metaphors the Convolutional Neural Network (CNN) deep learning model is used called as Deep Learning based Methodology (DLM). Pre-trained CNN model is used. The features are extracted from CNN model will served as input to machine learning model to classification. SVM model is used to train the model and made discussion for COVID-19 or NonCOVID-19 patients' cases. SVM is simplest and achieved higher accuracy but it only work on small dataset and not work on noisy data. Shuai Wang [5] proposed Deep Learning Algorithm (DLA) that extracts features correctly to diagnosis the disease. This model detects COVID-19 and NonCOVID-19 from extracted features. Deep Learning Algorithm (DLA) aimed to save the clinical time and it distinguish COVID-19 and other disease accurately but this method still need to improve the optimal accuracy.

In [7], the authors recommended a process of identification for COVID-19 patients. This paper classified based on probabilistic method. CT images dataset is used and then features are extracted from CT images. The certain features used to serve the classification model. For model training and testing Stack Hybrid Classification (SHC) method is implemented. Stack Hybrid Classification (SHC) aimed to enhanced performance by combining some models. It is proved that performance of Stack Hybrid Classification (SHC) is outstanding. The authors of [23] recommended the EMCNet framework that achieved higher accuracy than other frameworks to detect the COVID-19 patients. This framework has Xray dataset that used to extract features from CNN deep model. Subsequently, extracted features served to classifier. ML binary classifiers are chose to construct and experiment the model. The nominated models are "random forest, Support Vector Machine (SVM), Decision Tree (DT), and AdaBoost". This framework achieved correct identification and quicker results for covid19 & noncovid19 patients. The accuracy of EMCNet model is 98.91%. Unlike [23], Apostolopoulos et al. [2] suggested pre-trained CNN model that can work on small datasets. This paper compared several CNN models with transfer learning which are not customized designed. Here are two X-ray images datasets are used (i)Dataset-1 contains total 1427 images where pneumonia 700, normal 504 and COVID-19224 patient cases (ii) Dataset-2 also contains equal quantity of images for respectively class. By way of consequential the MobileNetV2 pre = trained CNN model achieved higher accuracy 96.78% on small dataset as compared with other

pre-trained models. Oh et al. [20] proposed a CNN which utilized the less parameters patch-based framework. The device decided on the very last type selection through majority balloting from consequences at numerous course locations. Dataset is designed on total images 15,043 where health ($n = 8851$), pneumonia ($n = 6012$) and COVID-19 ($n = 180$) images. This CNN based framework achieved accuracy 88.9%.

Islam et al. [14] recommended a varied architecture framework by best features of CNN deep model and Long Short Term Memory (LSTM) to support simple and less time to analysis of COVID-19 patients. The deep network applied to mine deep features from x-ray dataset and to done classification LSTM method is utilized. Work on X-ray dataset, total number of images are 4575 where COVID-19 has 1525 images. This method reached to accuracy of 99.4%. We can utilize big dataset by increasing COIVD-19 images.

Ghoshal et al. [8] proposed a deep learning-based method to evaluate the improbability and explicable in identifying COVID-19. The author applied the Bayesian convolutional neural network (BCNN) technique. Open source chest X-ray images Dataset is composed from GitHub and Kaggle repositories. In this paper, 68 Posterior-Anterior (PA) X-ray images of lungs with COVID-19 virus from Dr. Joseph Cohen's Github [6]. The improved the dataset with Kaggle's Chest X-Ray images (Pneumonia) from well patients, a total of 5941 PA chest radiography images across four classes: Normal 1583, Bacterial Pneumonia 2786, non-COVID-19 Viral Pneumonia 1504, and COVID-19 68. This proposed model found prediction uncertainty is exceptionally correlated with prediction accuracy. This model has improved results in diagnosis accuracy from 85.2% to 92.9% utilizing the pre-trained VGG-16 architecture.

Narin et al. [18] suggested an approach based on deep learning utilizing transfer learning techniques for COVID-19 prediction. The model categorizes two modules (i) COVID-19 and (ii) Normal. Open source chest X-ray and CT scan metaphors dataset gathered where 341 COVID-19 metaphors were composed of Cohen et al. [6]. Normal 2800 chest X-ray images were designated from the "ChestX-ray8" database. In accumulation, 2772 bacterial and 1493 viral pneumonia chest X-ray images were applied from the Kaggle repository called "Chest X-Ray Images (Pneumonia)" [17]. The author implements three pre-trained models, namely, InceptionV3, ResNet50, and InceptionResNetV2. The experimented results show that ResNet50 for binary classification achieved a higher accuracy of 98% compared with other models.

Ozturk et al. [21] proposed object detection based on the technique, unlike the common deep learning-based identification models. The author has built DarkNet architecture for the classification of CXR images for COVID-19 identified. Dataset is composed of two different open-source depositories (i) COVID-19 X-ray images gathered from Cohen [6] (ii) normal and pneumonia images taken from ChestX-ray8 database [27]. To evade the unbalanced data problem, we applied 500 no-findings and 500 pneumonia class frontal chest X-ray images indiscriminately from this database. The investigational results displayed that binary classification accomplished high accuracy of 98.08%. Nevertheless, the model demonstrated rather low performance for multiclass classification and made accuracy of 87.02% only.

Hasan j et al. [13] utilized the Convolutional Neural Network CNN deep learning-based practice. The author of this paper matched the performance of the three CNN models, explicitly, Inception V3, Xception, and ResNeXt models, and examined their accuracy. Open-source repositories occupy the PA view of chest X-ray metaphors. The model classifies COVID-19 affected patients as well as fit patients. Experimental results indicate that the Xception model has the utmost accuracy of 97.97%.

- CNN applied on the small dataset that drops accuracy [1, 11, 26].

- The dropout layer is placed right before the last layer, which hurts performance [4].
- Not appropriate choice of hyperparameters [11].
- PPV is higher in recommended COVID-19 patients that cause incorrect results, which means false COVID-19 infected cases are enhanced [10].
- Used SVM and CNN for feature extractions that face difficulty when the image is rotated or gradient. If objects of an image are not in the same position, ConvNet drops information that might not be able to classify the images [24].
- Increase ResNet50 layers' most efficient deep features can be extracted without using a standard dataset [15].

In short, every approach used different classifiers and algorithms for COVID-19 detection in their corresponding work. Each approach has revealed one classifier that outperformed other classifiers in the accuracy, and each approach has some limitations, too.

Moreover, this study concluded from the literature that the combined Deep Learning and Machine Learning algorithms that have been excessively used in the past for various works and performed efficiently. Where DL is used to extract deep features from images, the machine learning model accurately classifies the COVID-19 and NonCOVID-19 patients.

3 Dataset

Availability of a high-quality large COVID-19 dataset is crucial for developing an efficient deep learning architecture to successfully classify the COVID-19 images through an automated solution. Generally, to develop a decent and high-performing deep learning model, large datasets such as ImageNet, and Microsoft's COCO, containing millions of hand-.

annotated images with object classification, highlight, and labeling, are extensively used. However, creating such a high-quality large dataset in this pandemic is a challenging and expensive task with regard to the finance and expertise needed. In last year, few public and private datasets for Covid-19 detection and classification are released, but the size of data available is not as large; therefore, several studies have collected their own private dataset for the purpose. Some benchmark repositories such as Kaggle and GitHub contain many COVID-19 datasets from 2020 to 2021. In this study, two datasets were used for COVID-19 detection. (i)Dataset 1 contribution of Hall et al. [11] that contains 102 COVID-19 and 320 NonCOVID-19 images, (ii)Dataset 2 contribution of Ucara et al. [26] that contains 1536 COVID-19 and 3072 NonCOVID-19 images. Datasets contains X-ray images and were collected from standard repositories GitHub and Kaggle. Two input images from each class are shown in Fig. 1. Details of three datasets are summarized in Table 1.
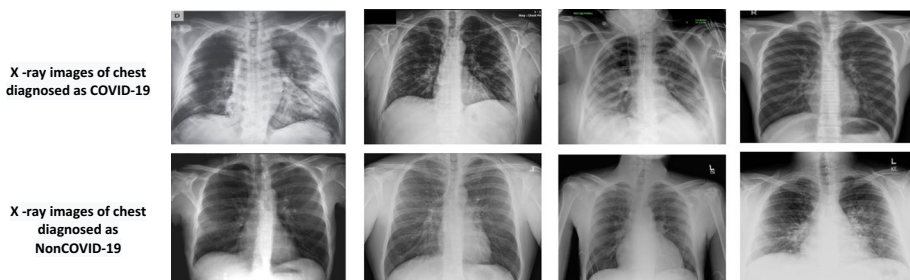

Fig. 1 Sample input Covid-19 from the different class in dataset

Table 1 Datasets details

Datasets		Class Label		Training & Testing split		
Dataset#	Dataset Name	COVID-19	NonCovid-19	No. of images	Train 70%	Test 30%
Dataset 1	Chest X-ray CXR	102	320	422	295	127
Dataset 2	COVIDx	1536	3072	4608	3226	1382

4 Deep learning framework for COVID-19 classification

Proposed framework for Covid-19 classification is shown in Fig. 2, and description of each part is given in this section.

4.1 Data preprocessing

As explained above, the real-world datasets are gathered in the first step. Fully automated image segmentation technqiues do not generally provide sufficiently high accuracy necessary for downstream processing. As such, interaction with human in the loop is required [31]. Initially, each dataset's observation was sorted into two classes, COVID-19 and Non-Covid-19. Data pre-processing has been performed once the sorting step is complete to remove noise and class imbalance issues. Next, data augmentation applies operations to increase COVID-19 images and obtain a suitable dataset. Below Table 2 describe data pre-processing parameters such as rescale, shear range, and horizontal flip. Next, extraction has been performed in this step, where several algorithms have been utilized to mine features from X-ray images and convert them into numerical parts. Finally, feature Selection has been performed through a random function after feature extraction.

4.2 Transfer learning and training process

Prediction model building has been performed using different Deep Learning algorithms. A total of nine different DL algorithms have been applied; each model has its depth. Similar to the previous chapter, one with higher accuracy on each source dataset would be used. They mentioned in Table 3 that images that give input to the models are size 224×224. Deep Learning algorithms with high accuracy have been carried forward
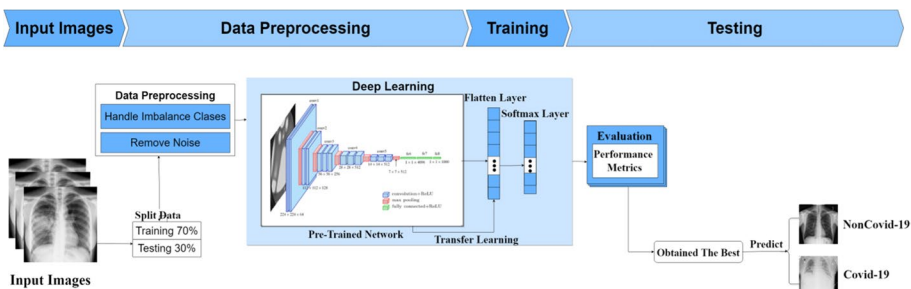

Fig. 2 Architecture for Covid-19 Classification

Table 2 Data Pre-processing Parameters

Pre-processing Parameter	Values/types
Rescale	1./255
Shear range	0.2
Zoom range	0.2
Horizontal flip	True
Clockwise	15 degree
Anti-clockwise	15 degree

for actual prediction on an unseen dataset for evaluation. Further details are given in below Table 3. The following sub-section explains the results of Deep Learning algorithms on two benchmark datasets.

Below, Fig. 3 describes the step-by-step process of how Deep Learning models are building. The first step model pre-trained the network, where each layer learned low-level features like edges, blocks, and colors. After that, in the Second step, replace the final layer where new layers learn features specific to the data set. In the third step, images are trained on new data with the same weights as pre-trained models. Finally, in the last step, the trained model sends for evaluation.

4.3 Model tuning

Hyperparameter tuning is vital before training any algorithm to control the behavior of the model. The experimental setup for nine Deep Learning algorithms was tested on 25 and 15 epochs. To reduce the losses by changing the attributes of neural networks such as learning rate and weights. To minimize the losses, we chose ADAM, RMSprop, and SGD optimizers. Apart from the optimizer, two different learning rates were utilized for the architecture. The learning rate can impact (i) how fast the algorithm learns and (ii) whether the cost functions is minimized or not.

Moreover, the learning rate influences how hastily our model can converge to local minima. As if we select too large a learning rate, it causes the model to connect too promptly to a suboptimal solution, while conversely, too small a learning rate drives the process to get stuck and take more training time. For the proposed model, two different learning rates were selected for the model (i) Lr-0.001 and (ii) Lr-0.005 that give a more optimal solution and provide efficient results. Categorical Cross-entropy loss function has been used to predict neural network error. Table 4 describes experimental settings.

Table 3 Pre-trained CNN architecture details

Architecture	Depth	Parameters (Millions)	Image Input size
ResNet-50	–	25,636,712	224×224
VGG19	26	143,667,240	224×224
Inception-v3	159	23,851,784	224×224
Xception	126	22,910,480	224×224
DenseNet-201	201	20,242,984	224×224
MobileNet	88	4,253,864	224×224
VGG16	23	138,357,544	224×224
NASNetMobile	–	5,326,716	224×224
EfficientNetB0	–	5,330,571	224×224

Fig. 3 Process of Model building
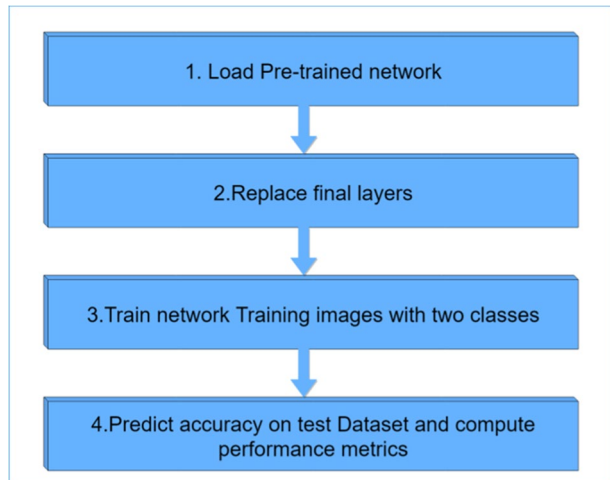
Datasets-1 and dataset-2 experimented on 50 epochs. As the dataset-1 was trained and tested successfully on 50 epochs, but when dataset-2 was trained, it froze after 40 epochs and took lots of time to train the model. The reason has been evaluated is that dataset-1 size is smaller than dataset-2. So, it has been observed that increasing the epoch's size on a large dataset can cause the problem in the training model.

4.4 Model evaluation

Testing has been performed separately for each dataset. The train and test have been performed on gathered datasets of 2020 and 2021 using a 30% dataset for testing. However, the evaluation of the model has been performed on an unseen dataset.

5 Experimental results

This section is composed of the investigational operation of the proposed method, "Deep Learning framework for early detection of COVID-19 using X-Ray images". All the experiments on Deep Learning were executed on a GPU-Based workstation with an i7 processor, 16GB RAM, 1 TB HDD, and 3GB GTX 1060 graphics.

Table 4 Experimental Settings

Network	Optimizer function	Loss Function	Learning Rates	Epochs	Batch Size
VGG19	ADAM		0.001	15	
ResNet-50	RMSprop			25	
Inception-v3	SGD				
Xception		Categorical Cross-entropy	0.005	15	32
DenseNet-201					
MobileNet				25	
VGG16					
NASNetMobile					
EfficientNetB0					

Table 5 ADAM optimizer, learning rate: 0.001 and 25 epochs on Dataset 1

Model(s)	ADAM			
	Accuracy	Precision	Recall	F1 score
ResNet-50	31.69	39	30	28
VGG19	11.97	6	39	11
Inception-v3	84.51	42	50	46
Xception	84.51	42	50	46
DenseNet-201	80.99	61	59	60
MobileNet	52.11	49	48	43
VGG16	11.97	21	29	12
NASNetMobile	17.61	58	51	16
EfficientNetB0	15.49	8	50	13

5.1 Dataset 1

In experiments on Dataset-1, selected the best model performance. Values of epochs chosen were 25 for a learning rate of 0.001 with the ADAM optimizer. Once epochs number was selected based on the results, best performing deep learning architecture was identified. The results obtained are shown in Table 5, and their respective graphs are shown in Fig. 4. Experimental results on 25 numbers of epochs show that InceptionV3 and Xception generate the highest results with 84.51% accuracy, 42% precision, 46% F1-score, and 50% recall on 25 epochs in all the settings except 0.001 learning rate and 25 epochs where ADAM performs better.

In Table 5, we observed that the ADAM optimizer effects the results of some of the state-of-art models. The optimizers affects the accuracy, such as VGG16 compatible with SDG optimizer and Inception-v3 compatible with ADAM optimizer.

Further experiments were performed to select the number of epochs that generated the best results. Values of epochs chosen were 15 and 25 for learning rates 0.001 and 0.005 with the ADAM optimizer. Comparison of results mentioned in Table 6 and Fig. 5 shows that 25 epochs produce better results, 84.51% accuracy and 61% recall as compared to 15 epochs that give 84.51% accuracy and 50.00% recall regardless of the learning rate value.
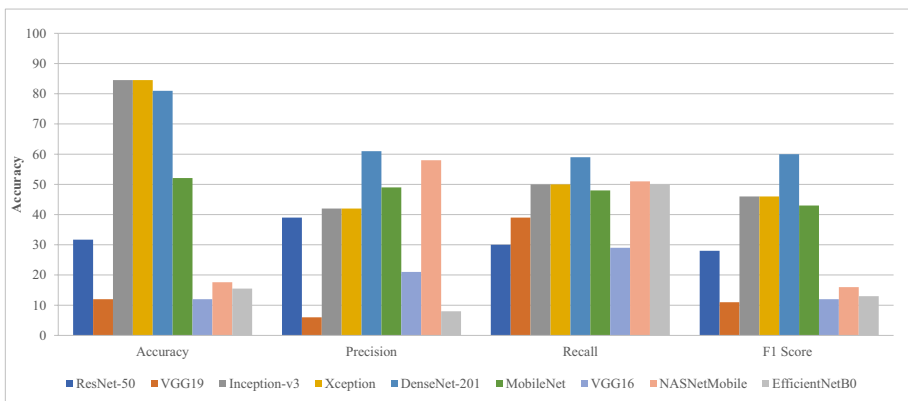

Fig. 4 Best performance Model on Dataset

Table 6 Results with Optimizer: ADAM, Learning Rate 0.001 and 0.005, Number of Epochs 15 and 25

Model(s)	Accuracy			
	LR: 0.001		LR: 0.005	
	15 Epochs	25 Epochs	15 Epochs	25 Epochs
ResNet-50	81.69	31.69	16.20	38.03
VGG19	35.21	11.97	34.51	12.68
Inception-v3	84.51	84.51	84.51	84.51
Xception	83.10	84.51	83.80	84.51
DenseNet-201	83.10	80.99	84.51	71.83
MobileNet	33.80	52.11	33.10	23.94
VGG16	14.08	11.97	12.68	48.59
NASNetMobile	15.49	17.61	16.90	19.01
EfficientNetB0	15.49	15.49	15.49	84.51
Model(s)	Precision			
	LR: 0.001		LR: 0.005	
	15 Epochs	25 Epochs	15 Epochs	25 Epochs
ResNet-50	15	39	35	41
VGG19	37	6	35	58
Inception-v3	42	42	42	42
Xception	42	42	42	42
DenseNet-201	42	61	42	42
MobileNet	46	49	55	38
VGG16	27	21	23	58
NASNetMobile	08	58	58	58
EfficientNetB0	08	8	08	42
Model(s)	Recall			
	LR: 0.001		LR: 0.005	
	15 Epochs	25 Epochs	15 Epochs	25 Epochs
ResNet-50	39	30	43	34
VGG19	25	39	22	52
Inception-v3	50	50	50	50
Xception	49	50	50	50
DenseNet-201	49	59	50	50
MobileNet	44	48	57	30
VGG16	36	29	32	52
NASNetMobile	50	51	51	52
EfficientNetB0	50	50	50	50
Model(s)	F1 score			
	LR: 0.001		LR: 0.005	
	15 Epochs	25 Epochs	15 Epochs	25 Epochs
ResNet-50	12	28	15	32
VGG19	28	11	26	18
Inception-v3	46	46	46	46
Xception	45	46	46	46

Table 6 (continued)

DenseNet-201	45	60		46
MobileNet	32	43	33	20
VGG16	14	12	12	18
NASNetMobile	13	16	15	18
EfficientNetB0	13	13	13	46

Next step, the last parameter to decide is the learning rate. According to the selected optimizer and number of epochs, results produced by both learning rates were compared. Results presented in Table 7 and Fig. 6 show that a 0.005 learning rate has better results than a 0.001 learning rate.

5.2 Dataset-2

In experiments on Dataset-2, selected the best model performance. Values of epochs chosen were 25for a learning rate of 0.001 with the ADAM Optimizer. Once the epochs number was selected based on the results, the best performing deep learning architecture was identified. The results obtained are shown in Table 8, and their respective graphs are shown in Fig. 7. Experimental results on 25 epochs show that VGG19 and DenseNet201 generate the highest results with 99.97% accuracy, 9.99% precision, 9.99% F1-score, and 9.99%
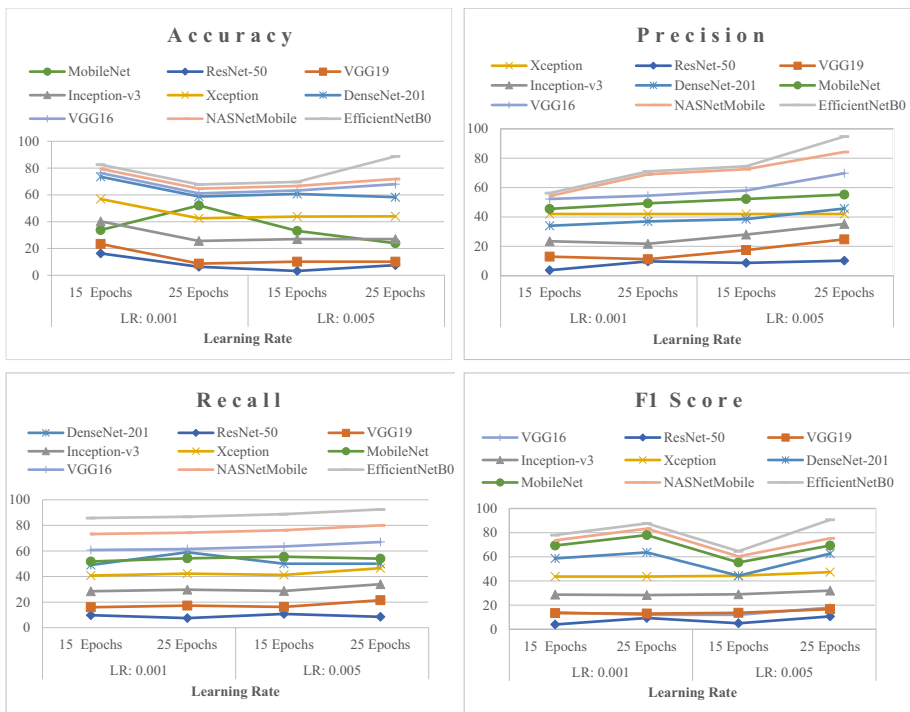

Fig. 5 Results with Optimizer: ADAM, Learning Rate 0.001 and 0.005, Number of Epochs 15 and 25

Table 7 Results with Optimizer: ADAM, Learning Rate 0.001 and 0.005, Number of Epochs 25

Model(s)	0.001				0.005			
	Accuracy	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score
ResNet-50	31.69	39	30	28	38.03	41	34	32
VGG19	11.97	6	39	11	12.68	58	52	18
Inception-v3	84.51	42	50	46	84.51	42	50	46
Xception	84.51	42	50	46	84.51	42	50	46
DenseNet-201	80.99	61	59	60	71.83	42	50	46
MobileNet	52.11	49	48	43	23.94	38	30	20
VGG16	11.97	21	29	12	48.59	58	52	18
NASNetMobile	17.61	58	51	16	19.01	58	52	18
EfficientNetB0	15.49	8	50	13	84.51	42	50	46

recall on 25 epochs in all the settings except 0.001 learning rate and 25 epochs where ADAM performs better.

Next step experiments were performed to select the number of epochs that generated the best results. Values of epochs chosen were 15 and 25 for learning rates 0.001 and 0.005 with the ADAM Optimizer. A comparison of results mentioned in Table 9 and Fig. 8 shows that 25 epochs produce better results with 99.01% accuracy and 99% recall as compared to 25 epochs that give 99.62% accuracy and 50.00% recall regardless of the learning rate value.

Now, the last parameter to decide is the learning rate. According to the selected optimizer, ADAM and number of epochs 25, results produced by both learning rates were compared. Results are shown in Table 10 and Fig. 9 tell that a 0.001 learning rate has better results.

Results of this experiment show that Dataset size affects the performance of the deep learning models. Results generated by both experiments conducted with experimental settings listed in Table 1 are shown in Table 11, the best settings for obtaining the accurate results.

Inceptin-v3 and Xception performed best on Dataset 1 with 84.51% accuracy, 42.00% precision, 50.00% recall, and 46.00% F1 score with ADAM optimizer, learning
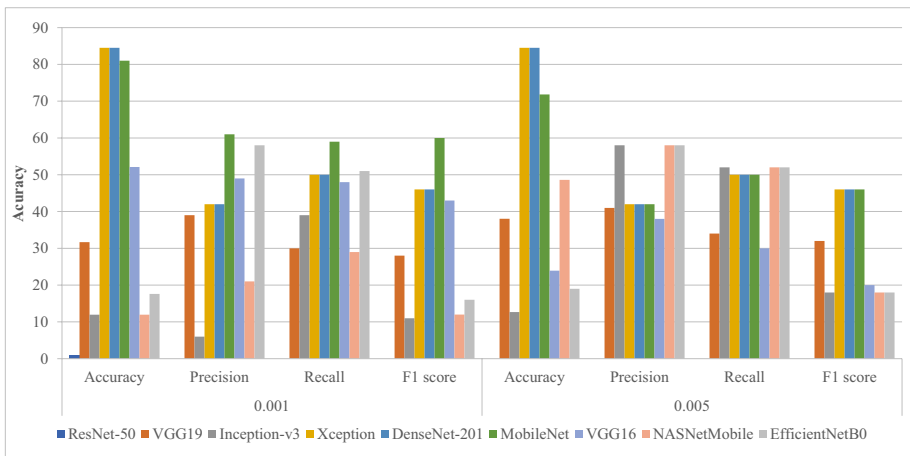

Fig. 6 Comparison of Learning Rates 0.001 and 0.005 at 25 epochs and ADAM Optimizer

Table 8 ADAM optimizer, learning rate: 0.001 and 25 epochs on Dataset 2

Model(s)	ADAM			
	Accuracy	Precision	Recall	F1 score
ResNet-50	95.92	97	97	97
VGG19	99.41	99	99	99
Inception-v3	98.91	19	50	28
Xception	97.88	19	50	28
DenseNet-201	98.85	19	50	28
MobileNet	96.95	32	49	28
VGG16	99.97	9.99	9.99	9.99
NASNetMobile	96.23	19	50	28
EfficientNetB0	57.33	19	50	28

rate 0.001 and 25 epochs. On Dataset 2, VGG-16 gives the most accurate results with the same learning rate and the number of epochs as of dataset 1 experiment; however, the optimizer that performs better is ADAM yielding 98.86%, 9.99% precision, 9.99% recall, and 9.99% F1 score. It has been observed that VGG16 achieved higher performance because VGG16 is a classification model which is suitable for classifying 1000 different categories and large image datasets. The aim of different results on two different datasets is that either they come from different distributions or they were obtained utilizing separate approaches, and at least one or two methods are biased (Table 12).

In terms of prediction model building, firstly, much work has been done in the past for COVID-19 prediction. However, most existing techniques were not accurately predicting results. Table 12 summarizes the results of the most promising existing methods by Hall et al. [11] and Ucara et al. [26], versus ours. Similarly, the AI models work on data, while in this endemic, there is a shortage of COVID-19 image datasets that causes class imbalance problems. The class imbalance ensues when there is a large amount of pneumonia, healthy class images, and few COVID-19 images. Secondly, there is a need for comparative analysis of ML and DL exiting techniques. For example, evaluating the
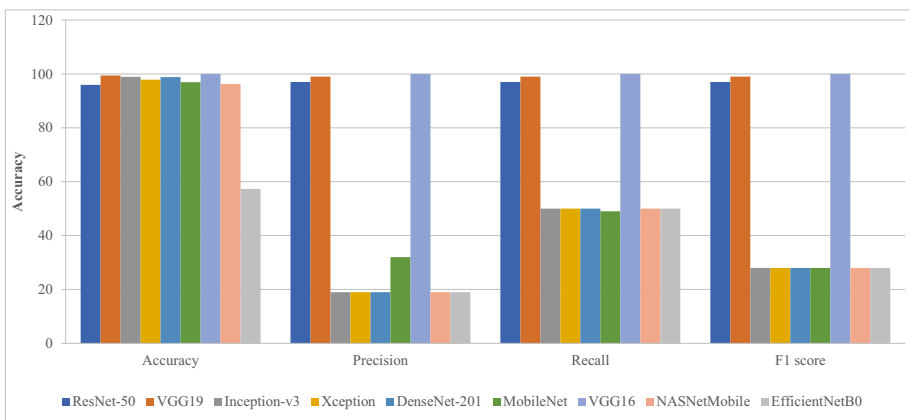

Fig. 7 Best performance Model on Dataset

Table 9 Results with Optimizer: ADAM, Learning Rate 0.001 and 0.005, Number of Epochs 15 and 25

Model(s)	Accuracy			
	LR: 0.001		LR: 0.005	
	15 Epochs	25 Epochs	15 Epochs	25 Epochs
ResNet-50	96.38	97.02	93.54	96.16
VGG19	99.50	99.29	97.59	99.01
Inception-v3	38.07	38.07	38.07	38.07
Xception	38.07	38.07	38.07	38.07
DenseNet-201	38.07	38.07	38.07	38.07
MobileNet	38.07	37.43	32.95	35.94
VGG16	99.64	99.86	98.08	98.86
NASNetMobile	61.93	38.07	38.07	38.07
EfficientNetB0	38.07	38.07	61.93	61.93
Model(s)	Precision			
	LR: 0.001		LR: 0.005	
	15 Epochs	25 Epochs	15 Epochs	25 Epochs
ResNet-50	97	97	95	97
VGG19	9.99	99	98	97
Inception-v3	19	19	19	19
Xception	19	19	19	19
DenseNet-201	19	19	19	19
MobileNet	19	32	28	21
VGG16	31	9.99	31	96
NASNetMobile	31	19	19	19
EfficientNetB0	19	19	31	31
Model(s)	Recall			
	LR: 0.001		LR: 0.005	
	15 Epochs	25 Epochs	15 Epochs	25 Epochs
ResNet-50	97	97	92	95
VGG19	99	99	98	95
Inception-v3	50	50	50	50
Xception	50	50	50	50
DenseNet-201	50	50	50	50
MobileNet	50	49	42	47
VGG16	50	9.99	50	94
NASNetMobile	50	50	50	50
EfficientNetB0	50	50	50	50
Model(s)	F1 score			
	LR: 0.001		LR: 0.005	
	15 Epochs	25 Epochs	15 Epochs	25 Epochs
ResNet-50	97	97	93	96
VGG19	99	99	90	96
Inception-v3	28	28	28	28
Xception	28	28	28	28

Table 9 (continued)				
DenseNet-201	28	28	28	28
MobileNet	28	28	27	27
VGG16	28	9.99	38	96
NASNetMobile	38	28	28	28
EfficientNetB0	28	28	38	38

ML and DL techniques based on their limitations, benefits, and approach is efficient for COVID-19 diagnosis. Thirdly, much work is done to create unique models for COVID-19 detection, but this work does not optimize their models' behavior on different parameter settings. Moreover, the model performance can be evaluated on various evaluation measures, but the existing approaches focus on accuracy and ignore other advanced evaluation measures, which are crucial for assessing model performance.

In this paper, an automated, accurate prediction methodology has been proposed to overcome the limitations above. The proposed method has been validated through Deep Learning algorithms. The proposed approach provides an automated solution to diagnose COVID-19 cases to stop the extent of COVID-19 in healthy populations. Thus, the proposed model helps for effective treatment. Experiments have been conducted on DL approaches on two live datasets taken from well-known and quality repositories. The supremacy of the proposed methodology has been observed by accurate prediction.

The existing models are validated mainly with one dataset, which is insufficient to provide an accurate model; there is a need to experiment with the model on more than one
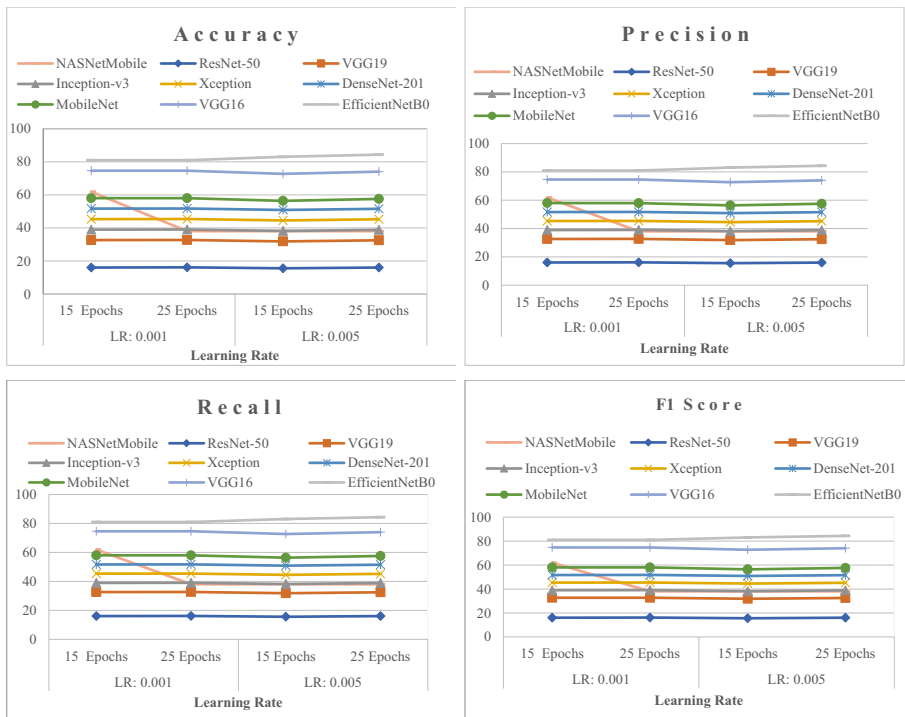

Fig. 8 Results with Optimizer: ADAM, Learning Rate 0.001 and 0.005, Number of Epochs 15 and 25

Table 10 Results with Optimizer: ADAM, Learning Rate 0.001 and 0.005, Number of Epochs 25

Model(s)	0.001				0.005			
	Accuracy	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score
ResNet-50	97.02	97	97	97	96.16	97	95	96
VGG19	99.29	99	99	99	99.01	97	95	96
Inception-v3	38.07	19	50	28	38.07	19	50	28
Xception	38.07	19	50	28	38.07	19	50	28
DenseNet-201	38.07	19	50	28	38.07	19	50	28
MobileNet	37.43	32	49	28	35.94	21	47	27
VGG16	99.86	9.99	9.99	9.99	98.86	96	94	96
NASNetMobile	38.07	19	50	28	38.07	19	50	28
EfficientNetB0	38.07	19	50	28	61.93	31	50	38

dataset. So, this study uses two benchmark datasets. Model building is done on the training set, and prediction has been made on unseen data, unlike the previous studies. The focus of this research is to estimate the performance of the model on advanced evaluation measures.

Finally, the evaluation of the model is done on advanced evaluation measures. The most crucial factor is the strategy based upon hyperparameters tuning with guidance provided by research in existing literature [16, 28, 29]. The results have been evaluated to see the effectiveness of the proposed methodology and, finally, an effective prediction model which predicts COVID-19 patients accurately on real-world chest X-ray datasets. Model tuning and using ADAM optimizer in the proposed model achieved higher accuracy than previous models.

6 Critical discussion

Experiments have been conducted to diagnosis the COVID-19 infected cases from chest X-ray images. A Deep Learning model-based method is proposed that identified between COVID-19 and Non-COVID-19 cases. They are Deep Learning neural networks that extracted automatic features, while Machine Learning cannot directly classify models.
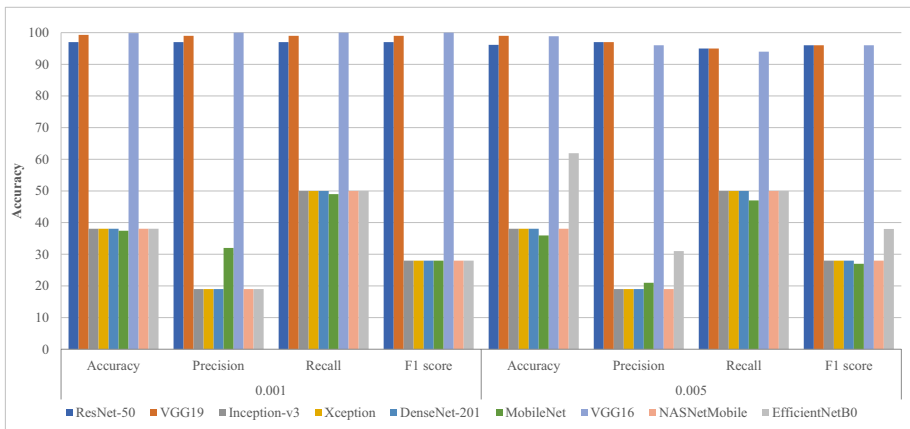

Fig. 9 Comparison of Learning Rates 0.001 and 0.005 at 15 epochs and ADAM Optimizer

Table 11 Performance on both Datasets

		Model	Optimizer	No of epochs	Learning Rate	Accuracy	Precision	Recall	F1 score
2 class	**Dataset 1**	Inception-V3 and Xception	ADAM	25	0.001	84.51	42	50	46
	Dataset 2	VGG 16	ADAM	25	0.001	99.86	99.99	99.99	99.99

When the model is appropriately trained, the learned model makes discussion about COVID-19 or Non-Covid-19. Our study selected VGG-16, which is more accurate than other classifiers on large datasets, and Inception-v3 and Xception classifiers have higher accuracy on a small dataset. Experiments are accompanied on real-world datasets occupied from famous and quality repositories Kaggle and GitHub to assess the proposed framework. Most of the existing methods work on Machine learning models or combined convolutional neural networks with Machine Learning procedures.

Our model is purely a Deep Learning-based technique, and we get the highest accuracy of these previous models. The proposed experimented results show that recommended framework outperformed than existing methods used in the COVID19 detection. Particularly in the task of classification of COVID-19, the mentioned method has accuracy on Dataset-1: 84.51% and Dataset-2: 99.86% that shows how correctly it identified COVID-19 patients.

As it has been discussed above that training on dataset-2 freezing after 40 epochs. The limitation of this proposed model is that it cannot run more than 40 epochs if dataset size is large.

7 Conclusion and future work

This research was motivated by the fact that there was minimal work in the Artificial Intelligence domain regarding the accurate diagnosis of COVID-19 patients concerning X-ray images. Therefore, it is imperative to design effective strategies timely to lessen the spread of the COVID-19. Numerous prediction models were proposed in the past; nevertheless, the previously proposed models suffered from many confines. Those limitations are explored in this research work, and a framework has been submitted and validated using two live COVID-19 chest X-ray images datasets. The outcomes have been critically evaluated to see the supremacy and effectiveness of our proposed methodology.

In terms of prediction model building, firstly, much work has been done in the past for COVID-19 prediction. However, the existing techniques were not accurately predicting results. Similarly, the AI models work on data, while in this endemic; there is a shortage

Table 12 Comparison table with existing Methods

Study	Methods	Class	Accuracy
Hall et al. [11]	VGG16, Resnet50	2	91.24%
Ucara et al. [26]	SqueezeNet	3	98.26%
Proposed method	Inception-V3 and Xception	2	84.51%
	VGG 16	2	99.86%

of COVID-19 image datasets that causes class imbalance problems. The class imbalance ensues when there is a large amount of pneumonia, healthy class images, and few COVID-19 images. Secondly, much work is done to create unique models for COVID-19 detection, but this work does not optimize their models' behavior on different parameter settings. Moreover, the model performance can be evaluated on various evaluation measures, but the existing approaches focus on accuracy and ignore other advanced evaluation measures, which are crucial for assessing model performance.

In this paper, an automated, accurate prediction methodology has been proposed to overcome the limitations above. Thus, the proposed model helps for effective treatment. The existing models are validated mainly with one dataset, which is insufficient to provide an accurate model; there is a need to experiment with the model on more than one dataset. So, this studies using three benchmark datasets. Model building is done on the training set, and prediction has been made on unseen data, unlike the previous studies. The focus of this research, to estimate the performance of the model on advanced evaluation measures.

Finally, the evaluation of the model is done on advanced evaluation measures. The most crucial factor is the strategy based upon hyperparameters tuning guided by research in existing literature [16, 28, 29]. The results have been evaluated to see the effectiveness of the proposed methodology and, finally, an effective prediction model which predicts COVID-19 patients accurately on real-world chest X-ray datasets.

In future work, the proposed model has been implemented on Real-world application. Artificial Intelligence is an approaching and helpful apparatus to detect timely COVID-19 infected cases and also reliefs in observing the condition of the COVID-19 infested persons. There is a need for a real-time application to prevent COVID-19 from spreading in the healthy population in this pandemic. The real-time application is supportive in the cure of COVID-19 infested persons and their appropriate wellbeing monitoring. It can track the predicament of COVID-19 at district measures such as therapeutic, molecular, and epidemiological uses. It is correspondingly supportive to enable the research on this disease utilize considering the accessible data. Furthermore, it can expressively mend treatment reliability and decision-making by building concrete procedures. This application can advantage in evolving appropriate treatment procedures, timely diagnosis, and vaccine enlargement.

Data availability Data sharing not applicable to this article as no publicly available datasets were generated or analyzed during the current study.

Declarations

Conflicts of interests The authors have no relevant financial or non-financial interests to disclose.

References

1. Alimadadi A, Aryal S, Manandhar I, Munroe PB, Joe B, Cheng X (2020) Artificial intelligence and machine learning to fight covid-19. Physiol Genomics 52(4):200–202. https://doi.org/10.1152/physiolgenomics.00029.2020
2. Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Australas Phys Eng Sci Med 43(2):635–640. https://doi.org/10.1007/s13246-020-00865-4
3. Barstugan M, Ozkaya U, Ozturk S (2020) Coronavirus (COVID-19) classification using CT images by machine learning methods. arXiv 5:1–10

4. Brunese L, Mercaldo F, Reginelli A, Santone A (2020) Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. Comput Methods Prog Biomed 196:105608. https://doi.org/10.1016/j.cmpb.2020.105608

5. Carbonneau M-A, Cheplygina V, Granger E, Gagnon G (2018) Multiple instance learning: a survey of problem characteristics and applications. Pattern Recognit 77:329–353. https://doi.org/10.1016/j.patcog.2017.10.009

6. Cohen, JP (2020) Open database of covid-19 cases, https://github.com/ieee8023/covid-chestxray-dataset. Accessed 15 Oct 2021

7. Farid AA, Selim GI, Khater HAA (2020) A novel approach of CT images feature analysis and prediction to screen for corona virus disease (COVID-19). Int J Sci Eng Res 11(03):1141–1149. https://doi.org/10.14299/ijser.2020.03.02

8. Ghoshal B, Tucker, A. (2020) Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection. 1–14. http://arxiv.org/abs/2003.10769

9. Gozes O et al. (2020) Rapid AI development cycle for the coronavirus (COVID-19) Pandemic: Initial results for automated detection & patient monitoring using deep learning CT image analysis, [Online]. Available: http://arxiv.org/abs/2003.05037. Accessed 15 Oct 2021

10. Gunraj H, Wang L, Wong A (2020) COVIDNet-CT: A tailored deep convolutional neural network design for detection of COVID-19 Cases from Chest CT Images, pp. 1–12, [Online]. Available: http://arxiv.org/abs/2009.05383. Accessed 15 Oct 2021

11. Hall LO, Paul R, Goldgof DB, Goldgof GM (2020) Finding Covid-19 from chest x-rays using deep learning on a small dataset, pp. 1–8, [Online]. Available: http://arxiv.org/abs/2004.02060. Accessed 15 Oct 2021

12. Han Z et al (2020) Accurate screening of COVID-19 using attention-based deep 3D multiple instance learning. IEEE Trans Med Imaging 39(8):2584–2594. https://doi.org/10.1109/TMI.2020.2996256

13. Hasan MJ, Alom MS, Ali MS (2021) Deep learning based detection and segmentation of COVID-19 & pneumonia on chest x-ray image. In: 2021 international conference on information and communication technology for sustainable development (ICICT4SD), Dhaka, Bangladesh, pp 210–214. https://doi.org/10.1109/ICICT4SD50815.2021.9396878

14. Islam MZ, Islam MM, Asraf A (2020) A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. Inf Med Unlocked 20:100412. https://doi.org/10.1016/j.imu.2020.100412

15. Kassania SH, Kassasni PH, Wesolowski MJ, Schneider KA, Deters R (2020) Automatic detection of coronavirus disease (COVID-19) in X-ray and CT images: A machine learning based approach, arXiv, vol. 2019, pp. 1–18

16. Liao L, Li H, Shang W, Ma L (2022) An empirical study of the impact of Hyperparameter tuning and model optimization on the performance properties of deep neural networks. ACM trans. Softw. Eng. Methodol. 31, 3, article 53 (July 2022), 40 pages. https://doi.org/10.1145/3506695

17. Mooney P (2018) Chest X-Ray Images (Pneumonia), https://www.kaggle.com/paultimothymooney/chest-xraypneumonia. Accessed 15 Oct 2021

18. Narin, A, Kaya, C, Pamuk, Z (2020) Department of Biomedical Engineering, Zonguldak Bulent Ecevit University, 67100, Zonguldak, Turkey. ArXiv Preprint ArXiv:2003.10849. https://arxiv.org/abs/2003.10849. Accessed 15 Oct 2021

19. Nour, M, Cömert, Z, Polat, K (2020) A novel medical diagnosis model for COVID-19 infection detection based on deep features and Bayesian optimization. Appl Soft Comput J xxxx, 106580. https://doi.org/10.1016/j.asoc.2020.106580

20. Oh Y, Park S, Ye JC (2020) Deep learning COVID-19 features on CXR using limited training data sets. arXiv 39(8):2688–2700

21. Ozturk, T, Talo, M, Azra, E, Baran, U, Yildirim, O (2020) 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. Computers in Biology and Medicine, January

22. Rahimzadeh M, Attar A (2020) A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. Inf Med Unlocked 19:100360. https://doi.org/10.1016/j.imu.2020.100360

23. Saha P, Sadi MS, Islam MM (2021) EMCNet: automated COVID-19 diagnosis from X-ray images using convolutional neural network and ensemble of machine learning classifiers. Inf Med Unlocked 22:100505. https://doi.org/10.1016/j.imu.2020.100505

24. Sethy PK, Behera SK, Ratha PK, Biswas P (2020) Detection of coronavirus disease (COVID-19) based on deep features and support vector machine. Int J Math Eng Manag Sci 5(4):643–651. https://doi.org/10.33889/IJMEMS.2020.5.4.052

25. Singh D, Kumar V, Vaishali, Kaur M (2020) Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based convolutional neural networks. Eur J Clin Microbiol Infect Dis 39(7):1379–1389. https://doi.org/10.1007/s10096-020-03901-z

26. Ucara F, Korkmaz D (2020) 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 , January

27. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM (2017) ChestX-Ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: 2017 IEEE Conference on computer vision and pattern recognition (CVPR), Honolulu, HI, USA, pp 3462–3471. https://doi.org/10.1109/CVPR.2017.369

28. Wu J, Chen X-Y, Zhang H, Xiong L-D, Lei H, Deng S-H (2019) Hyperparameter optimization for machine learning models based on Bayesian Optimization. J Electron Sci Technol 17(1):26–40. https://doi.org/10.11989/JEST.1674-862X.80904120

29. Yang L, Shami A (2020) On hyperparameter optimization of machine learning algorithms: theory and practice. Neurocomputing 415:295–316, ISSN 0925-2312. https://doi.org/10.1016/j.neucom.2020.07.061

30. Zhong L, Gong P, Biging GS (2012) Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases 78, no. May, pp. 1–15

31. Zhou T, Li L, Bredell G, Li J, Unkelbach J, Konukoglu E (2023) Volumetric memory network for interactive medical image segmentation, Medical Image Analysis, Volume 83, 102599, ISSN 1361-8415, https://doi.org/10.1016/j.media.2022.102599