


RESEARCH

Open Access



Machine-learning classification of texture features of portable chest X-ray accurately classifies COVID-19 lung infection

Lal Hussain^{1,2*} , Tony Nguyen³, Haifang Li³, Adeel A. Abbasi¹, Kashif J. Lone¹, Zirun Zhao³, Mahnoor Zaib², Anne Chen³ and Tim Q. Duong³

*Correspondence:

lall_hussain2008@live.com

¹ Department of Computer Science and IT, King Abdullah Campus, University of Azad Jammu and Kashmir, Muzaffarabad 13100, Azad Kashmir, Pakistan
Full list of author information is available at the end of the article

Abstract

Background: The large volume and suboptimal image quality of portable chest X-rays (CXRs) as a result of the COVID-19 pandemic could post significant challenges for radiologists and frontline physicians. Deep-learning artificial intelligent (AI) methods have the potential to help improve diagnostic efficiency and accuracy for reading portable CXRs.

Purpose: The study aimed at developing an AI imaging analysis tool to classify COVID-19 lung infection based on portable CXRs.

Materials and methods: Public datasets of COVID-19 ($N = 130$), bacterial pneumonia ($N = 145$), non-COVID-19 viral pneumonia ($N = 145$), and normal ($N = 138$) CXRs were analyzed. Texture and morphological features were extracted. Five supervised machine-learning AI algorithms were used to classify COVID-19 from other conditions. Two-class and multi-class classification were performed. Statistical analysis was done using unpaired two-tailed t tests with unequal variance between groups. Performance of classification models used the receiver-operating characteristic (ROC) curve analysis.

Results: For the two-class classification, the accuracy, sensitivity and specificity were, respectively, 100%, 100%, and 100% for COVID-19 vs normal; 96.34%, 95.35% and 97.44% for COVID-19 vs bacterial pneumonia; and 97.56%, 97.44% and 97.67% for COVID-19 vs non-COVID-19 viral pneumonia. For the multi-class classification, the combined accuracy and AUC were 79.52% and 0.87, respectively.

Conclusion: AI classification of texture and morphological features of portable CXRs accurately distinguishes COVID-19 lung infection in patients in multi-class datasets. Deep-learning methods have the potential to improve diagnostic efficiency and accuracy for portable CXRs.

Keywords: Texture, Morphological, Machine learning, Feature extraction, Classification, COVID-19

Background

In December 2019, in the Wuhan Hubei province of China, a cluster of cases of pneumonia with an unknown cause was reported [1]. Eventually, it was discovered as severe



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

acute respiratory syndrome coronavirus-2 (SARS-CoV-2, previously named as 2019 novel coronavirus or COVID-19) which has then caused major public health issues and became a large global outbreak. According to the recent statistics, there are millions of confirmed cases in United States and India, and the number is still increasing. The WHO also declared on January 13, 2020 that COVID-19 was the sixth public health emergency of international concern following H1N1 (2009), polio (2014), Ebola in West Africa (2014), Zika (2016) and Ebola in the Democratic Republic of Congo (2019) [2]. It was also found that the novel coronaviral pneumonia is similar to another severe acute respiratory syndrome caused by the Middle East respiratory syndrome (MERS) coronavirus and that it was also capable of causing a more severe form known as acute respiratory distress syndrome (ARDS) [3, 4]. Consensus, criteria, and guidelines were being established with the aim to prevent transmission and facilitate diagnosis and treatment [2, 5, 6]. The rapid incidences of infection are due in part by the relatively slow onset of symptoms, thus enabling widespread transmission by asymptomatic carriers [7]. Along with the global connectivity of today's travel society, this infection readily spread worldwide [7], giving rise to a pandemic [8, 9].

Radiological imaging of the COVID-19 pneumonia reveals the destruction of pulmonary parenchyma which includes extensive consolidation and interstitial inflammation as previously reported in other coronavirus infections [10, 11]. In total, interstitial lung disease (ILD) comprises of more than 200 different types of chronic lung disorders that is characterized by inflammation of lung tissue, usually referred to as pulmonary fibrosis. The fibrosis causes lung stiffness, and this reduces the ability of the air sacs (i.e., spaces within an organism where there is the constant presence of air) to carry out and deliver oxygen into the bloodstream. This eventually can lead to the permanent loss of the ability to breathe. The ILDs are also heterogeneous diseases histologically but mostly contain similar clinical manifestations to each other or with other different lung disorders. This makes determining the differential diagnosis difficult. In addition, the large quantity of radiological data that radiologists are required to scrutinize (with lack of strict clinical guidelines) leads to a low diagnostic accuracy and high inter- and intra-observer variability, which was reported as great as 50% [12].

The most commonly used diagnosis for COVID-19 infections is through reverse transcription-polymerase chain reaction (RT-PCR) assays of nasopharyngeal swabs [13]. However, the high false-negative rate [14], length of test, and shortage of RT-PCR assay kits for the early stages of the outbreak can restrict a prompt diagnosis of infected patients. Computed tomography (CT) and chest X-ray (CXR) are well suited to image the lung of COVID-19 infections. In contrast to the swab test, CT and CXR reveals a spatial location of the suspected pathology as well as the extent of damages. The hallmark pathology of CXR are bilateral distribution of peripheral hazy lung opacities include air space consolidation [15]. The advantage of imaging is that it has good sensitivity, a fast turnaround time, and it can visualize the extent of infection in the lung. The disadvantage of imaging is that it has low specificity, challenging to distinguish different types of lung infection especially when there is severity in the lung infection.

Computer-aided diagnostic (CAD) systems can assist radiologists to increase diagnostic accuracy. Currently, researchers are using the hand-crafted or learning features which are based on the texture, geometry, and morphological characteristics of the lung

for detection. However, it is often crucial and challenging to choose the appropriate classifier that can optimally handle the property of the feature spaces of the lung. The traditional image recognition methods are Bayesian networks (BNs), support vector machine (SVM), artificial neural networks (ANNs), k-nearest neighbors (kNN), and Adaboost, decision trees (DTs). These machine-learning methods [16, 17] require hand-crafted features to compute such as texture, SIFT, entropy, morphological, elliptic Fourier descriptors (EFDs), shape, geometry, density of pixels, and off-shelf classifiers as explained in [18]. In addition, the machine-learning (ML) feature-based methods are known as non-deep learning methods. There are many applications for these non-deep learning methods such as uses in neurodegenerative diseases, cancer detection, and psychiatric diseases. [17, 19–22]. However, the major limitations of non-deep learning methods are that they are dependent on the feature extraction step and this makes it difficult to find the most relevant feature which are needed to obtain the most effective result. To overcome these difficulties, the use of artificial intelligence (AI) can be employed. The AI technology in the field of medical imaging is becoming popular especially for the technology advancement and development of deep learning [23–32]. Recently, [33] used Inf-net for automatic detection of COVID-19 lung infection segmentation from CT images. Moreover, [18] employed momentum contrastive learning for few shot COVID-19 diagnosis from chest CT images. There are vast applications of deep convolutional neural network (DCNN) and machine-learning algorithms in medical imaging problems [32, 34–38]; however, this study is specifically aimed to apply machine-learning algorithms with feature extraction approach. The main advantage of this method is the ability to learn the adaptive image features and classification, which are able to be performed simultaneously. The general goals are to develop automated tools by employing and optimizing machine-learning models along with texture and morphological features to detect early, to distinguish coronavirus-infected patients from non-infected patients. This proposed method will help the healthcare clinicians and radiologists for further diagnosis and tracking the disease progression. The AI-based system, once verified, and tested can lead towards crucial detection and control of patients affected from COVID-19. Furthermore, the machine-learning image analysis tools can potentially support the radiologists by providing an initial read or second opinion.

In this study, we employed machine-learning methods to classify texture features of portable CXRs with the aim to identify COVID-19 lung infection. Comparison of texture and morphological features on COVID-19, bacterial pneumonia, non-COVID-19 viral pneumonia, and normal CXRs were made. AI-based classification methods were used for differential diagnosis of COVID-19 lung infection. We tested the hypothesis that AI classification of texture features of CXR can accurately detect the COVID-19 lung infection.

Results

We applied five supervised machine-learning classifiers (XGB-L, XGB-Tree, CART, KNN and Naïve Bayes) to classify COVID-19 from bacterial pneumonia, non-COVID-19 viral pneumonia, and normal lung CXRs.

Table 1 shows the results of AI classification of texture and morphological features for COVID-19 vs normal utilizing five different classifiers: XGB-L, XGB-Tree, CART

Table 1 Performance of AI classification of texture and morphological features utilizing five different classifiers of COVID-19 (N = 130) vs normal (N = 138)

Classifier	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	AUC (LB, UP)	p value
XGB-L	100	100	100	100	100	1.00	2.00e−16
XGB-Tree	100	100	100.3	100	100	1.00	2.00e−16
CART (DT)	100	100	100	100	100	1.00	2.00e−16
KNN	89.71	100	100	91.11	95	0.99	2.00e−16
Naïve Bayes	100	100	100	100	100	1.00	2.00e−16

(DT), KNN, and Naïve Bayes. All classifiers yielded essentially 100% accuracy by all performance measures along with top four ranked features (i.e., compactness, thin ratio, perimeter, standard deviation), indicating that there is significant difference between the two groups.

Table 2 shows the results of AI classification of texture and morphological features for COVID-19 vs bacterial pneumonia. All classifiers except KNN performed well by all performance measures. Specifically, the XGB-L and XGB-Tree classifier yielded the highest classification accuracy (96.34% and 91.46%, respectively), while KNN classifier performed the worst (accuracy of 71.95%). While with the top four ranking features, the XGB-L and XGB-tree classifiers yielded highest accuracy of 85.37% and 86.59%, respectively.

Table 3 shows the results of AI classification of texture and morphological features for COVID-19 vs non-COVID viral pneumonia. All classifiers except KNN performed well by all performance measures. Specifically, the XGB-L and XGB-Tree classifier yielded the highest classification accuracy (97.56% and 95.12%, respectively), while KNN classifier performed the worst (accuracy of 79.27%).

Table 2 Performance of AI classification of texture and morphological features utilizing five different classifiers of COVID-19 (N = 130) vs bacterial pneumonia (N = 145)

Classifier	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	AUC (LB, UP)	p value
XGB-L	95.35	97.44	97.62	95.00	96.34	0.98 (0.96, 1.00)	2.19e−14
XGB-Tree	88.37	94.87	95.00	88.10	91.46	0.97 (0.93, 1.00)	2.19e−14
CART (DT)	74.42	97.44	96.97	77.55	85.37	0.92 (0.87, 0.98)	3.22e−10
KNN	86.05	56.41	68.52	78.57	71.95	0.83 (0.74, 0.92)	2.40e−04
Naïve Bayes	90.70	79.49	82.98	88.51	85.37	0.92 (0.85, 0.98)	3.22e−10

Table 3 Performance of AI classification of texture and morphological features utilizing five different classifiers of COVID-19 (N = 130) vs viral pneumonia (N = 145)

Classifier	Sensitivity	Specificity	PPV	NPV	Accuracy	AUC (LB, UP)	p value
XGB-L	97.44	97.67	97.44	97.67	97.56	0.98 (0.96, 1.00)	2.00e−16
XGB-Tree	94.87	95.35	94.87	94.87	95.12	0.98 (0.95, 1.00)	2.00e−16
CART (DT)	94.87	93.02	92.50	95.24	93.90	0.94 (0.89, 0.99)	2.00e−16
KNN	69.63	88.37	84.38	76.00	79.27	0.85 (0.77, 0.94)	4.38e−07
Naïve Bayes	64.10	95.35	92.59	74.55	80.49	0.93 (0.88, 0.98)	1.20e−07

Table 4 Two-class classification using XGB-linear with texture + morphological features for COVID-19 (N = 130) vs bacterial pneumonia (N = 145), COVID-19 vs non-COVID-19 viral (N = 145) and COVID-19 vs normal (N = 138)

Classification	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	AUC
COVID-19 vs bacterial pneumonia	95.35	97.44	97.62	95.00	96.34	0.98
COVID-19 vs non-COVID-19 viral pneumonia	97.44	97.67	97.44	97.67	97.56	0.98
COVID-19 vs normal	100	100	100	100	100	1.00

Table 5 Multi-class classification using XGB-linear with texture + morphological features

Dataset	Sensitivity	Specificity	PPV	NPV	Combined accuracy	Combined AUC
XGB-linear						
Bacterial pneumonia	74.42%	86.18%	65.31%	90.60%	79.52%	0.87
COVID-19	74.49%	95.28%	83.78%	93.80%		
Normal	95.12	100	100	98.43		
Viral pneumonia	67.77%	91.06%	73.17%	89.60%		

Table 4 shows the two-class classification using the XGB-L classifier. The result showed that model classified COVID-19 from normal patients most accurately, followed by COVID-19 from bacterial pneumonia, and lastly by COVID-19 from viral pneumonia.

Table 5 shows the results of the multi-class classification using the XGB-L classifier. For multi-class classification problem, the average accuracy for classification of all four classes is used to measure the performance of the classifier (i.e., combined accuracy and AUC). Multi-class classification was able to classify COVID-19 amongst the four groups, with a combined AUC of 0.87 and accuracy of 79.52%. While with the top two ranked features, the combined AUC of 0.82 and accuracy of 66.27% was obtained. Sensitivity, specificity, positive predictive value, and negative predictive value were similarly high. As reflected in Tables 1, 2, 3 and 4, the two-class classification performance (i.e., COVID-19 vs normal, COVID-19 vs bacterial pneumonia, COVID-19 vs viral pneumonia) in terms of sensitivity and PPV was higher than 95%, while these measures using multi-class (COVID-19 vs normal vs bacterial vs viral pneumonia) could achieve performance greater than 74% and 83% to detect COVID-19, respectively.

Feature ranking algorithms are mostly used for ranking features independently without using any supervised or unsupervised learning algorithm. A specific method is used for feature ranking in which each feature is assigned a scoring value, then selection of features will be made purely on the basis of these scoring values [39]. The finally selected distinct and stable features can be ranked according to these scores and redundant features can be eliminated for further classification. We first extracted first extracted texture features based on GLCM and morphological features from COVID-19, normal, viral and bacterial pneumonia CXR images and then ranked them based on empirical receiver-operating characteristic curve (EROCC) and random classifier slope [40], which ranks features based on the class separability criteria of the area between EROCC and random classifier slope. The ranked features show the features importance based on their

ranking which can be helpful for distinguish these different classes for improving the detection performance and decision making by the radiologists.

Figure 1 shows the ranking features of COVID-19 vs bacterial infection, COVID-19 vs normal, and their multi-class features. The top four features from COVID-19 vs bacterial CXR based on AUC were: skewness, entropy, compactness, and thin ratio. The top four features from COVID-19 vs normal CXR based on AUC were: compactness, thin ratio, perimeter, and standard deviation. The top feature from the multi-class was by far perimeter.

Discussion

We employed an automated supervised learning AI classification of texture and morphological-based features on portable CXRs to distinguish COVID-19 lung infections from normal, and other lung infections. The major finding was that the multi-class classification was able to accurately identify COVID-19 from amongst the four groups with a combined AUC of 0.87 and accuracy of 79.52%.

The hallmarks of COVID-19 lung infection on CXR are bilateral and peripheral hazy lung opacities and air space consolidation [15]. These features of COVID-19 lung infection likely stood out compared to other pneumonia, giving rise to distinguishable texture features. Our AI algorithm was able to distinguish COVID-19 vs normal CXR with 100% accuracy, COVID-19 vs bacterial pneumonia with 96.34% accuracy, and COVID-19 vs non-COVID-19 viral infection with 92.68% accuracy. These findings suggest that it is trivial to distinguish COVID-19 from normal CXR and the two viral infections were more similar than bacterial infection.

With the multi-class classification, all performance measures dropped significantly (except normal CXR) as expected. Nonetheless, the combined AUC and accuracy remained high. These findings are encouraging and suggest that the multi-class classification is able to distinguish COVID-19 lung infection from other similar lung infections.

The top four features from COVID-19 vs bacterial infection were skewness, entropy, compactness, and thin ratio. The top four features from COVID-19 vs normal were: compactness, thin ratio, perimeter, and standard deviation. The top feature from the multi-class was perimeter. Perimeter is the total count of pixels at the boundary of an image. It showed that the perimeter of COVID-19 lung CXRs differed significantly from other bacterial and viral infections as well as normal lung X-rays. These results together suggest that perimeter is a key distinguishable feature, consistent with a key observation that COVID-19 lung infection tends to be more peripheral and lateral together the boundaries of the lung.

A few studies have reported CNN analysis of CXR and CT for classification of COVID-19 [41–45]. Li et al. performed a retrospective multi-center study using a deep-learning model to extract visual features from chest CT to distinguish COVID-19 from community acquired pneumonia (CAP) and non-pneumonia CT with a sensitivity of 90%, specificity 95%, and AUC 0.96 (p value < 0.001) [41]. Hurt et al. performed a retrospective study using a U-net (CNN), to predict pixel-wise probability maps for pneumonia only from a public dataset that comprised of 22,000 radiographs. For their classification of pneumonia, the area under the receiver-operator characteristic curve was 0.854 with a sensitivity of 82.8% and specificity of 72.6 [46]. Wang et al. developed a deep CNN to

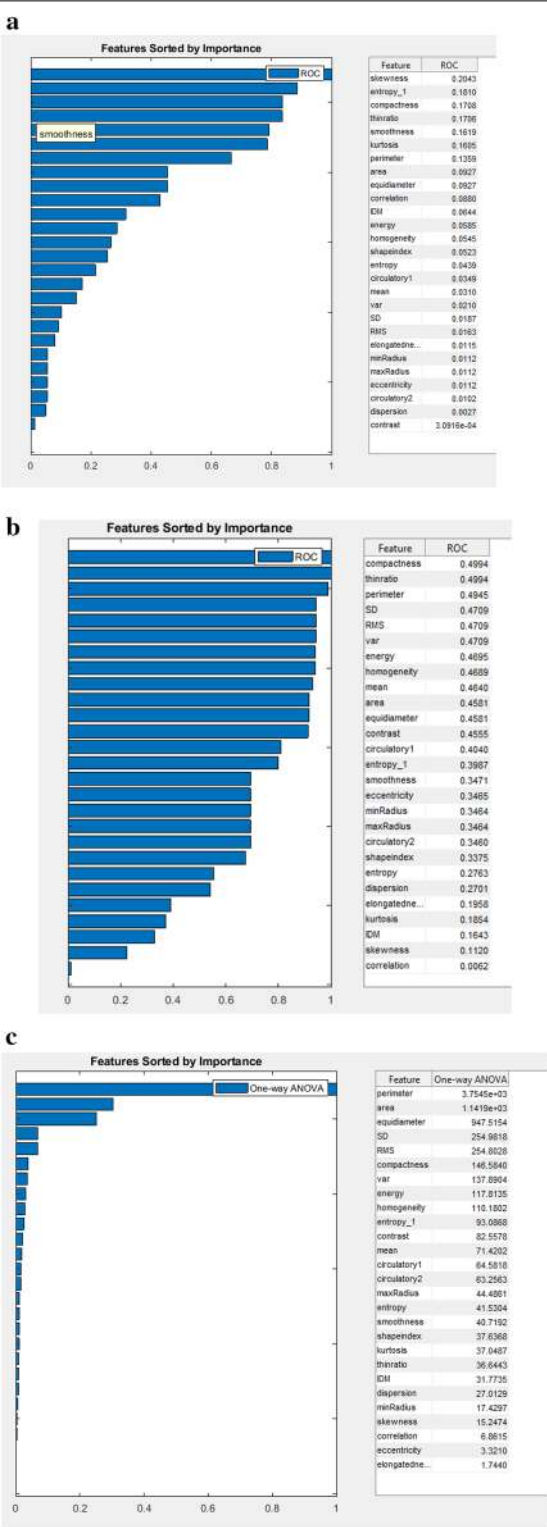


Fig. 1 Ranking parameters: **a** COVID vs bacterial infection, **b** COVID-19 vs normal, and **c** multi-class feature ranking

detect COVID-19 cases from non-COVID CXR. This study used interpretable AI to visualize the location of the abnormality and was able to distinguish COVID-19 from non-COVID-19 viral infection, bacterial infection, and normal with a sensitivity of 81.9%, 93.1%, and 73.9%, respectively, with an overall accuracy of 83.5% [38]. Gozes et al. developed a deep-learning algorithm to analyze CT images to detect COVID-19 patients from non-COVID-19 cases with 0.996 AUC (95% CI 0.989–1.00), 98.2% sensitivity and 92.2% specificity [43]. Apostolopoulos and Mpesiana [45] used deep learning with a transfer learning approach to extract features from X-rays to distinguish between COVID-19 and bacterial pneumonia, viral pneumonia, and normal with a sensitivity of 98.66%, specificity of 96.46%, and accuracy of 96.78%. Overall, most of these studies used two-class comparison (i.e., pneumonia vs COVID-19, or pneumonia vs normal) mostly on CT which is less suitable for contagious diseases. In these previous studies, two-class prediction performance was computed and yielded fine results but could not achieve the highest performance as compared to our approach. The aim of this research was to improve the prediction performance by extracting texture and morphological features from CXR images. As the machine-learning performance is still a challenging task to extract the most relevant and appropriate features by the researchers. The results reveal that features extracted using our approach contain the most pertinent and appropriate hidden information present in the COVID-19 lung infection which improved the two-class and multi-class classification. These features are then used as input to the robust machine-learning classifiers. The results obtained outperformed than these previously traditional methods.

There are several limitations of this study. This is a retrospective study with a small COVID-19 sample size. Portable CXR is sensitive but not specific as the phenotypes of different lung infections are similar on CXR. We used only four classes (disease types). Future studies should expand to include additional lung disorders.

Conclusion

In conclusion, deep learning of texture and morphological-based features accurately distinguish CXR of COVID-19 patients from normal subjects and patients with bacterial and non-COVID-19 viral pneumonia. This approach can be used to improve workflow, facilitate in early detection and diagnosis of COVID-19, effective triage of patients with or without the infectious disease, and provide efficient tracking of disease progression.

Limitation and future directions

This study is specifically aimed to extract the texture features and apply the machine-learning algorithms to predict the COVID-19 from multi-class. The texture features correctly predict the COVID-19 from multi-class; however, in future, we will employ and optimize the deep convolutional neural network models including ResNet101, GoogLeNet, AlexNet, Inception-V3 and use will use some other modalities, clinical profiles and bigger datasets.

Methods

Dataset

In this study, we used publicly available data of COVID-19 and non-COVID and normal chest CXR images. The COVID-19 images were downloaded from <https://github.com/>

iee8023/covid-chestxray-dataset [47] on Mar 31, 2020. The original download contained 250 scans of COVID-19 and SARS of CT and CXR taken in multiple directions. Two board-certified chest radiologists (one with 20+ years of experience) and one 2nd year radiology resident evaluated the images for quality and relevance. Only CXR from COVID-19 taken at anterior–posterior (AP) direction was included in this study, resulting in a final sample size of 130. The other dataset was taken from the Kaggle chest X-ray image (pneumonia) dataset (<https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>) [42]. Although the Kaggle database has a large sample size, we randomly selected a sample size comparable to that of COVID-19. The sample chosen for the bacterial pneumonia, non-COVID-19 viral pneumonia, and normal CXR were 145, 145, and 138, respectively. We first split the dataset into training and testing data with a 70% and 30% ratio using a stratified sampling method. Then for feature selection, we only used the training data instead of the whole dataset. Figure 2 below outlines the workflow and steps used in this study.

Figure 2 outlines the workflow with the initial input of lung CXRs going through feature extraction for texture + morphological analysis followed by the AI classifiers to determine the sensitivity, specificity, PPV, NPV, accuracy, and AUC of the four groups of interest (COVID-19, bacterial and viral pneumonia, and normal). These calculations are further outputted for data validation with fivefold cross-validation technique. Finally, data are statistically analyzed for significance using MATLAB 2018b and RStudio 1.2.5001.

Texture features

The texture features are estimated from the Grey-level Co-occurrence Matrix (GLCM) covering the pixel (image) spatial correlation. Each GLCM input image (u, v)th defines how often pixels with intensity value u co-occur in a defined connection with pixels with intensity value v . We extracted second-order features consisting of contrast, correlation, mean, entropy, energy, variance, inverse different moment, standard deviation, smoothness, root mean square, skewness, kurtosis, and homogeneity previously used in [48–54].

Morphological features

Morphological feature plays an important role in the detection of malignant tissues. Morphological features convert image morphology into a set of quantitative values that can be used for classification [55]. Morphological feature-extracting method (MFEM) is a nonlinear filtering process and its basic purpose is to search and find valuable information from an image and transform it morphologically according to the requirements for segmentation [56] and so on. The MFEM takes binary cluster as an input and finds the associated components in the clusters having an area greater than a certain threshold. There are several features that can be extracted from an image and area can be calculated from the number of pixels of an image. Area and perimeter combined helps to calculate the values of other different morphology features. The following formulas in [50] can be used to calculate the values of morphological features.

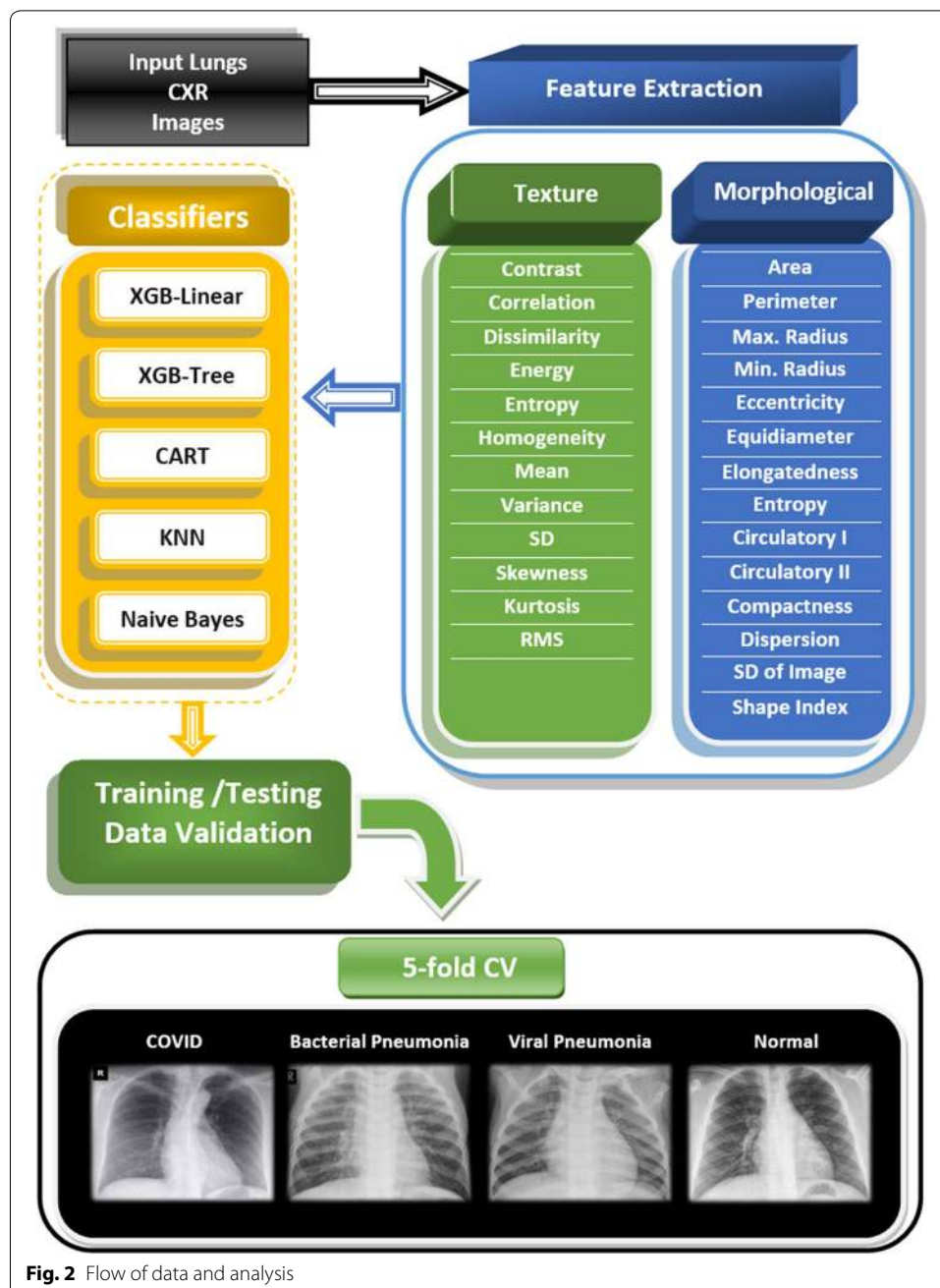


Fig. 2 Flow of data and analysis

Classification

We applied and compared five supervised machine-learning classification algorithms: XG boosting linear (XGB-L), XG boosting tree (XGB-tree), classification and regression tree (CART), k-nearest neighbor (KNN) and Naïve Bayes (NB). We used XGB ensemble methods in this study. In machine learning, ensemble is the collection of multiple models and is one of the self-efficient methods as compared to other basic models. Ensemble technique combines different hypothesis to hopefully provide best hypothesis. Basically, this method is used for obtaining a strong learner with the help of combination of weak learners. Experimentally, ensembles methods provide more

accurate results even there is considerable diversity between the models. Boosting is a most common types of ensemble method that works by discovering many weak classification rules using subset of the training examples simply by sampling again and again from the distribution.

XGBoost algorithms

Chen and Guestrin proposed XGBoost a gradable machine-learning system in 2016 [57]. This system was most popular and became the standard system when it was employed in the field of machine learning in 2015 and it provides us with better performance in supervised machine learning. The Gradient boosting model is the original model of XGBoost, which combine and relates a weak base with stronger learning models in an iterative manner [58]. In this study, we used XGBoost linear and tree with following optimization parameters.

We used the following parameter of each model in this study. For XGB-linear we initialized the parameters as $\lambda=0$, $\alpha=0$ and $\eta=0.3$, where λ and α are the regularization term on weights and η is the learning rate. For XGB-Tree, we initialized the parameters with maximum depth of tree i.e., $\text{max-depth}=30$, learning rate $\eta=0.3$, maximum loss reduction i.e., $\gamma=1$, minimum child weight = 1, subsample = 1. The nearest neighbor $k=5$ was used. For CART, we initialized parameters with $\text{minsplit}=20$, complexity parameter, i.e., $\text{cp}=0.01$, maximum depth = 30. For Naïve Bayes, we initialized the parameters with search method = grid, laplace = 0, and adjust = 1.

Classification and regression tree (CART)

A CART is a predictive algorithm used in the machine learning to explain how the target variable values can be predicted based on the other values. It is a decision tree where each fork is a split in a predictor variable and each node at the end has a prediction for the target variable. Decision tree (DT) algorithm was first proposed by Breiman in 1984 [59], is a learning algorithm or predictive model or decision support tool of Machine Learning and Data Mining for the large size of input data, which predict the target value or class label based on several input variables. In decision tree, the classifier compares and checks the similarities in the dataset and ranked it into distinct classes. Wang et al. [60] used DTs for classifying the data based on choice of an attribute which maximizes and fix the data division. Until the conclusion criteria and condition is met, the attributes of datasets are split into several classes. DT algorithm is constructed mathematically as:

$$\bar{X} = \{X_1, X_2, X_3, \dots, X_m\}^T,$$

$$X_i = \{x_1, x_2, x_3, \dots, x_{ij}, \dots, x_{in}\},$$

$$S = \{S_1, S_2, \dots, S_i, \dots, S_m\}.$$

Here the number of observations is denoted by m in the above equations, n represent number of independent variables, S is the m -dimension vector spaces of the variable forecasted from \bar{X} in the above equation. X_i is the i th module of n -dimension autonomous variables $x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}$ are autonomous variable of pattern vector X_i and T is the transpose symbol.

The purpose of DTs is to forecast the observations of \bar{X} . From \bar{X} , several DTs can be developed by different accuracy level; although, the best and optimum DT construction is a challenge due to the exploring space has enormous and large dimension. For DT, appropriate fitting algorithms can be developed which reflect the trade-off between complexity and accuracy. For partition of the dataset \bar{X} , there are several sequences of local optimum decision about the feature parameters are used using the Decision Tree strategies. Optimal DT, T_{k0} is developed according to a subsequent optimization problem:

$$\hat{R}(T_{k0}) = \min \left\{ \hat{R}(T_k) \right\}, k = 1, 2, 3, \dots, K,$$

$$\hat{R}(T) = \sum_{t \in T}^k \{r(t)p(t)\}.$$

In the above equation, $\hat{R}(T)$ represents an error level during the misclassification of tree T_k , T_{k0} represented the optimal DT that minimizes an error of misclassification in the binary tree, T represent a binary tree $\in \{T_1, T_2, \dots, T_k, t_1\}$, the index of tree is represented by k , tree node with t , root node by t_1 , resubstituting an error by $r(t)$ which misclassify node t , probability that any case drop into node t is represented with $p(t)$. The left and right sets of partition of sub trees are denoted by T^L and T^R . The result of feature plan portioning the tree T is formed.

Naïve Bayes (NB)

The NB [61] algorithm is based on Bayesian theorem [62] and it is suitable for higher dimensionality problems. This algorithm is also suitable for several independent variables whether they are categorical or continuous. Moreover, this algorithm can be the better choice for the average higher classification performance problem and have minimal computational time to construct the model. Naïve Bayes classification algorithm was introduced by Wallace and Masteller in 1963. Naïve Bayes relates with a family of probabilistic classifier and established on Bayes theorem containing compact hypothesis of independence among several features. Naïve Bayes is most ubiquitous classifier used for clustering in Machine Learning since 1960. Classification probabilities are able to compute using Naïve Bayes method in machine learning. Naïve Bayes is utmost general classification techniques due to highest performance than the other algorithm such as decision tree (DT), C-means (CM) and SVM. Bayes decision law is used to find the predictable misclassification ratio whereas assuming that true classification opportunity of an object belongs to every class is identified. NB techniques were greatly biased because its probability computation errors are large. To overcome this task, the solution is to reduce the probability valuation errors by Naïve Bayes method. Conversely, dropping probability computation errors did not provide the guarantee

for achieving better results in classification performance and usually make it poorest because of its different bias-variance decomposition among classification errors and probability computation error [63]. Naïve Bayes is widely used in present advance developments [64–67] due to its better performance [68]. Naïve Bayes techniques need a large number of parameters during learning system or process. The maximum possibility of Naïve Bayes function is used for parameter approximation. NB represents conditional probability classifier which can be calculated using Bayes theorem: problem instance which is to be classified, described by a vector $Y = \{Y_1, Y_2, Y_3, \dots, Y_n\}$ shows n features spaces, conditional probability can be written as:

$$S(N_k|Y_1, Y_2, Y_3, \dots, Y_n).$$

For each class N_k or each promising output, statistically Bayes theorem can be written as:

$$S(N_k|Y) = \frac{S(N_k)S(Y|N_k)}{S(Y)}.$$

Here, $S(N_k|Y)$ represents the posterior probability while $S(N_k)$ represents the preceding probability, $S(Y|N_k)$ represents the likelihood and $S(Y)$ represents the evidence. NB is represented mathematically as:

$$S(N_k|Y_1, Y_2, Y_3, \dots, Y_n) = \frac{1}{T} S(N_k) \prod_{i=1}^n S(Y_i|N_k).$$

Here $T = S(y)$ is scaling factor which is depends upon $(Y_1, Y_2, Y_3, \dots, Y_n)$, $S(N_k)$ is a parameter used for the calculation of marginal probability and conditional probability for each attribute or instances is represented by $S(Y_i|N_k)$. Naïve Bayes become most sensitive in the presence of correlated attributes. The existence of extremely redundant or correlated objects or features can bias the decision taken by Naïve Bayes classifier [67].

K-nearest neighbor (KNN)

KNN is most widely used algorithm in the field of machine learning, pattern recognition and many other areas. Zhang [69] used KNN for classification problems. This algorithm is also known as instance based (lazy learning) algorithm. A model or classifier is not immediately built but all training data samples are saved and waited until new observations need to be classified. This characteristic of lazy learning algorithm makes it better than eager learning, that construct classifier before new observation needs to be classified. Schwenker and Trentin [70] investigated that this algorithm is also more significant when dynamic data are required to be changed and updated more rapidly. KNN with different distance metrics were employed. KNN algorithm works according to the following steps using Euclidean distance formula.

Step I: To train the system, provide the feature space to KNN.

Step II: Measure distance using Euclidean distance formula:

$$d(x_i, y_i) = \sum_{i=1}^n \sqrt{(x_i - y_i)^2}.$$

Step III: Sort the values calculated using Euclidean distance using $d_i \leq d_{i+1}$, where $i = 1, 2, 3, \dots, k$.

Step IV: Apply means or voting according to the nature of data.

Step V: Value of K (i.e., number of nearest Neighbors) depends upon the volume and nature of data provided to KNN. For large data, the value of k is kept as large, whereas for small data the value of k is also kept small.

In this study, these classification algorithms were performed using RStudio with typical default parameters for each of the classifiers (XGB-L, GXB-tree, CART, KNN, NB) with a fivefold cross-validation. As we divided our dataset into train and test sets, so while training a classifier on train data we used the K-fold cross-validation technique, which shuffles the data and splits it into k number of folds (groups). In general, K-fold validation is performed by taking one group as the test data set, and the other $k-1$ groups as the training data, fitting and evaluating a model, and recording the chosen score on each fold. As we used fivefold cross-validation, so the train set is equally divided into five parts from which one is used as validation and the other four used for training of classifier on each fold.

Performance evaluation measures

The performance was evaluated with the following parameters.

Sensitivity The sensitivity measure also known as TPR or recall is used to test the proportion of people who test positive for the disease among those who have the disease. Mathematically, it is expressed as:

$$\text{Sensitivity} = \frac{\sum \text{True positive}}{\sum \text{Condition positive}},$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

i.e., the probability of positive test given that patient has disease.

Specificity The TNR measure also known as specificity is the proportion of negatives that are correctly identified. Mathematically, it is expressed as:

$$\text{Specificity} = \frac{\sum \text{True negative}}{\sum \text{Condition negative}},$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}},$$

i.e., probability of a negative test given that patient is well.

Positive predictive value (PPV) PPV is mathematically expressed as:

$$\text{PPV} = \frac{\sum \text{True positive}}{\sum \text{Predicted condition positive}},$$

$$PPV = \frac{TP}{TP + FP},$$

where TP denotes that the test makes a positive prediction and subject has a positive result under gold standard while FP is the event that test make a positive prediction and subject make a negative result.

Negative predictive value (NPV) NPV can be computed as:

$$NPV = \frac{\sum \text{True negative}}{\sum \text{Predicted condition negative}},$$

$$NPV = \frac{TN}{TN + FN},$$

where TN indicates that test make negative prediction and subject has also negative result, while FN indicate that test make negative prediction and subject has positive result.

Accuracy The total accuracy is computed as:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}.$$

Receiver-operating characteristic (ROC) curve

Based on sensitivity, i.e., true-positive rate (TPR) and specificity, i.e., false-positive rate (FPR) values of COVID-19 and non-COVID subjects. The mean values for COVID-19 subjects are classified as 0 and for non-COVID subjects are classified as 1. Then obtained vector is passed through ROC function, which plots each value against sensitivity and specificity values. ROC is considered as one of the standard methods for computation and graphical representation of the performance of a classifier. ROC plots FPR against x -axis and TPR against y -axis, while part of a square unit is represented by area under the curve (AUC). The value of AUC lies between 0 and 1 where $AUC > 0.5$ indicates the separation. Higher area under the curve represents the better and improved diagnostic system [71]. The number of correct positive cases divided by the total number of positive cases represents TPR. While the number of negative cases predicted as positive cases divided by the total number of negative cases represent FPR [72].

Training/testing data formulation

The Jack-knife fivefold cross-validation (CV) technique was applied for the training and testing of data formulation and parameter optimization. It is one of the most well known, commonly practiced, and successfully used methods for validating the accuracy of a classifier using fivefold CV. The data are divided into fivefold in training, the four-fold participate, and classes of the samples for remaining folds are classified based on the training performed on fourfold. For the trained models, the test samples in the test fold are purely unseen. The entire process is repeated five times and each class sample is

classified accordingly. Finally, the unseen samples classified labels that are to be used for determining the classification accuracy. This process is repeated for each combination of each systems' parameters and the classification performance have been reported for the samples as depicted in the Tables 1, 2, 3 and 4.

Statistical analysis and performance measures

Analyses examining differences in outcomes used unpaired two-tailed *t* tests with unequal variance. Receiver-operating characteristic (ROC) curve analysis was performed with COVID-19, normal, bacterial, and non-COVID-19 viral pneumonia as ground truth. The performance was evaluated by standard ROC analysis, including sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, area under the receiver-operating curve (AUC) with 95% confidence interval, and significance with the *P* value. AUC with lower and upper bounds and accuracy were tabulated. MATLAB (R2018b, MathWorks, Natick, MA) and RStudio 1.2.5001 were used for statistical analysis.

Abbreviations

MERS: Middle East respiratory syndrome; WHO: World Health Organization; ARDS: Acute respiratory distress syndrome; RT-PCR: Polymerase chain reaction; CT: Computed tomography; CXR: Chest X-ray; CAD: Computer-aided diagnostic; BNs: Bayesian networks; SVM: Support vector machine; ANNs: Artificial neural networks; kNN: K-nearest neighbors; DTs: Adaboost, decision trees; SIFT: Scale-invariant Fourier transform; EFDs: Elliptic Fourier descriptors; ML: Machine learning; DCNN: Deep convolutional neural network; CV: Cross-validation; ROC: Receiver-operating characteristic; PPV: Positive predictive value; NPV: Negative predictive value.

Acknowledgements

None.

Authors' contributions

LH conceptualized the study, analyzed data and wrote the paper. TN conceptualized the study, and edited the paper. AAA edited the paper. KJL edited the paper. ZZ edited the paper. MZ edited and reviewed the paper. AC edited the paper. TQ conceptualized the study and edited the paper. All the authors read and approved the final manuscript.

Funding

None.

Availability of supporting data and materials

These data are already available via <https://github.com/ieee8023/covid-chestxray-dataset>.

Ethical approval and consent to participate

Not applicable. Data were obtained from a publicly available, deidentified dataset. <https://github.com/ieee8023/covid-chestxray-dataset>

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹ Department of Computer Science and IT, King Abdullah Campus, University of Azad Jammu and Kashmir, Muzaffarabad 13100, Azad Kashmir, Pakistan. ² Department of Computer Science and IT, Neelum Campus, University of Azad Jammu and Kashmir, Athmuqam 13230, Azad Kashmir, Pakistan. ³ Department of Radiology, Renaissance School of Medicine at Stony Brook University, 101 Nicolls Rd, Stony Brook, NY 11794, USA.

Received: 19 August 2020 Accepted: 17 November 2020

Published online: 25 November 2020

References

1. Lu H, Stratton CW, Tang Y. Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. *J Med Virol*. 2020;92:401–2.
2. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med*. 2020;382:2001316.

3. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395:497–506.
4. Graham RL, Donaldson EF, Baric RS. A decade after SARS: strategies for controlling emerging coronaviruses. *Nat Rev Microbiol*. 2013;11:836–48.
5. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395:507–13.
6. Chen Z-M, Fu J-F, Shu Q, Chen Y-H, Hua C-Z, Li F-B, et al. Diagnosis and treatment recommendations for pediatric respiratory infection caused by the 2019 novel coronavirus. *World J Pediatr*. 2020;16:240–6.
7. Biscayart C, Angeleri P, Lloveras S, Chaves TSS, Schlagenhauf P, Rodríguez-Morales AJ. The next big threat to global health? 2019 novel coronavirus (2019-nCoV): what advice can we give to travellers?—Interim recommendations January 2020, from the Latin-American society for Travel Medicine (SLAMVI). *Travel Med Infect Dis*. 2020;33:101567.
8. Carlos WG, Dela Cruz CS, Cao B, Pasnick S, Jamil S. Novel Wuhan (2019-nCoV) coronavirus. *Am J Respir Crit Care Med*. 2020;201:P7–8.
9. Munster VJ, Koopmans M, van Doremalen N, van Riel D, de Wit E. A novel coronavirus emerging in China—key questions for impact assessment. *N Engl J Med*. 2020;382:692–4.
10. Chung M, Bernheim A, Mei X, Zhang N, Huang M, Zeng X, et al. CT imaging features of 2019 novel coronavirus (2019-nCoV). *Radiology*. 2020;295:202–7.
11. Fang Y, Zhang H, Xu Y, Xie J, Pang P, Ji W. CT manifestations of two cases of 2019 novel coronavirus (2019-nCoV) pneumonia. *Radiology*. 2020;295:208–9.
12. Sluimer I, Schilham A, Prokop M, van Ginneken B. Computer analysis of computed tomography scans of the lung: a survey. *IEEE Trans Med Imaging*. 2006;25:385–405.
13. Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing. *Radiology*. 2020;296:200343.
14. Chan JF-W, Yuan S, Kok K-H, To KK-W, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet*. 2020;395:514–23.
15. Wong HYF, Lam HYS, Fong AH-T, Leung ST, Chin TW-Y, Lo CSY, et al. Frequency and distribution of chest radiographic findings in COVID-19 positive patients. *Radiology*. 2019;296:201160.
16. Fehr D, Veeraraghavan H, Wibmer A, Gondo T, Matsumoto K, Vargas HA, et al. Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images. *Proc Natl Acad Sci*. 2015;112:E6265–73.
17. Orrù G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neurosci Biobehav Rev*. 2012;36:1140–52.
18. Chen X, Yao L, Zhou T, Dong J, Zhang Y. Momentum contrastive learning for few-shot COVID-19 diagnosis from chest CT images. *arXiv preprint arXiv:2006.13276*.
19. Parmar C, Bakers FCH, Peters NHGM, Beets RGH. Deep learning for fully-automated localization and segmentation of rectal cancer on multiparametric. *Sci Rep*. 2017;7:1–9.
20. Oakden-Rayner L, Carneiro G, Bessen T, Nascimento JC, Bradley AP, Palmer LJ. Precision radiology: predicting longevity using feature engineering and deep learning methods in a radiomics framework. *Sci Rep*. 2017;7:1648.
21. Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. *Cancer Inform*. 2006;2:117693510600200.
22. Doyle S, Hwang M, Shah K, Madabhushi A, Feldman M, Tomaszewski J. Automated grading of prostate cancer using architectural and textural image features. In: 2007 4th IEEE International Symposium on Biomedical Imaging From Nano to Macro. IEEE; 2007. pp 1284–7.
23. Wang J, Ding H, Bidgoli FA, Zhou B, Iribarren C, Molloy S, et al. Detecting cardiovascular disease from mammograms with deep learning. *IEEE Trans Med Imaging*. 2017;36:1172–81.
24. Shen D, Wu G, Suk H-I. Deep learning in medical image analysis. *Annu Rev Biomed Eng*. 2017;19:221–48.
25. Lee H, Tajmir S, Lee J, Zissen M, Yeshiwas BA, Alkasab TK, et al. Fully automated deep learning system for bone age assessment. *J Digit Imaging*. 2017;30:427–41.
26. Gao XW, Hui R, Tian Z. Classification of CT brain images based on deep learning networks. *Comput Methods Programs Biomed*. 2017;138:49–56.
27. Forsberg D, Sjöblom E, Sunshine JL. Detection and labeling of vertebrae in MR images using deep learning with clinical annotations as training data. *J Digit Imaging*. 2017;30:406–12.
28. Zhang Q, Xiao Y, Dai W, Suo J, Wang C, Shi J, et al. Deep learning based classification of breast tumors with shear-wave elastography. *Ultrasonics*. 2016;72:150–7.
29. Ortiz A, Munilla J, Górriz JM, Ramírez J. Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease. *Int J Neural Syst*. 2016;26:1650025.
30. Nie D, Zhang H, Adeli E, Liu L, Shen D. 3D deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients. *Cham: Springer*; 2016. p. 212–20.
31. Ithapu VK, Singh V, Okonkwo OC, Chappell RJ, Dowling NM, Johnson SC. Imaging-based enrichment criteria using deep learning algorithms for efficient clinical trials in mild cognitive impairment. *Alzheimer's Dement*. 2015;11:1489–99.
32. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE Trans Med Imaging*. 2016;35:1207–16.
33. Fan D-P, Zhou T, Ji G-P, Zhou Y, Chen G, Fu H, et al. Inf-Net: automatic COVID-19 lung infection segmentation from CT images. *IEEE Trans Med Imaging*. 2020;39:2626–37.
34. Cha KH, Hadjiiski L, Samala RK, Chan H-P, Caoili EM, Cohan RH. Urinary bladder segmentation in CT urography using deep-learning convolutional neural network and level sets. *Med Phys*. 2016;43:1882–96.
35. Ghafoorian M, Karssemeijer N, Heskes T, Bergkamp M, Wissink J, Obels J, et al. Deep multi-scale location-aware 3D convolutional neural networks for automated detection of lacunes of presumed vascular origin. *Neurolmage Clin*. 2017;14:391–9.
36. Lekadir K, Galimzianova A, Betriu A, del Mar VM, Igual L, Rubin DL, et al. A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound. *IEEE J Biomed Heal Inform*. 2017;21:48–55.

37. Rajkomar A, Lingam S, Taylor AG, Blum M, Mongan J. High-throughput classification of radiographs using deep convolutional neural networks. *J Digit Imaging*. 2017;30:95–101.
38. Samala RK, Chan H-P, Hadjiiski L, Helvie MA, Wei J, Cha K. Mass detection in digital breast tomosynthesis: deep convolutional neural network with transfer learning from mammography. *Med Phys*. 2016;43:6654–66.
39. Wang H, Raton B. A comparative study of filter-based feature ranking techniques. *IEEE IRI*. 2010;1:43–8.
40. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit*. 1997;30:1145–59.
41. 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.
42. Wang L, Wong A. COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-Ray images. *arXiv Prepr arXiv200309871*. 2020.
43. Gozes O, Frid-Adar M, Greenspan H, Browning PD, Zhang H, Ji W, et al. Rapid AI Development cycle for the coronavirus (COVID-19) pandemic: initial results for automated detection & patient monitoring using deep learning CT image analysis. *arXiv preprint arXiv:2003.05037*. 2020.
44. Narin A, Kaya C, Pamuk Z. Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks. *arXiv Prepr arXiv200310849*. 2020.
45. 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.
46. Hurt B, Yen A, Kligerman S, Hsiao A. augmenting interpretation of chest radiographs with deep learning probability maps. *J Thorac Imaging*. 2020;1.
47. Cohen JP, Morrison P, Dao L. COVID-19 Image Data Collection. *arXiv Prepr arXiv200311597*. 2020.
48. Khalvati F, Wong A, Haider MA. Automated prostate cancer detection via comprehensive multi-parametric magnetic resonance imaging texture feature models. *BMC Med Imaging*. 2015;15:27.
49. Haider MA, Vosough A, Khalvati F, Kiss A, Ganeshan B, Bjarnason GA. CT texture analysis: a potential tool for prediction of survival in patients with metastatic clear cell carcinoma treated with sunitinib. *Cancer Imaging*. 2017;17:4.
50. Guru DS, Sharath YH, Manjunath S. Texture features and KNN in classification of flower images. *Int J Comput Appl. Special Issue on RTIPPR (1)* 2010;21–9.
51. Yu H, Scalera J, Khalid M, Touret A-S, Bloch N, Li B, et al. Texture analysis as a radiomic marker for differentiating renal tumors. *Abdom Radiol*. 2017;42:2470–8.
52. Castellano G, Bonilha L, Li LM, Cendes F. Texture analysis of medical images. *Clin Radiol*. 2004;59:1061–9.
53. Khuzi AM, Besar R, Zaki WMDW. Texture features selection for masses detection in digital mammogram. *IFMBE Proc*. 2008;21:629–32.
54. Esgiar AN, Naguib RNG, Sharif BS, Bennett MK, Murray A. Fractal analysis in the detection of colonic cancer images. *IEEE Trans Inf Technol Biomed*. 2002;6:54–8.
55. Masseroli M, Bollea A, Forloni G. Quantitative morphology and shape classification of neurons by computerized image analysis. *Comput Methods Programs Biomed*. 1993;41:89–99.
56. Li YM, Zeng XP. A new strategy for urinary sediment segmentation based on wavelet, morphology and combination method. *Comput Methods Programs Biomed*. 2006;84:162–73.
57. Chen T, Guestrin C. XGBoost. *Proc 22nd ACM SIGKDD Int Conf Knowl Discov Data Min—KDD '16*. New York: ACM Press; 2016. pp. 785–94.
58. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat*. 2001;29:1189–232.
59. Ariza-López FJ, Rodríguez-Avi J, Alba-Fernández MV. Complete control of an observed confusion matrix. In: *International Geoscience Remote Sensors Symposium*. IEEE; 2018. pp. 1222–5.
60. Wang R, Kwong S, Wang X, Jiang Q. Continuous valued attributes. 45.
61. Nahar J, Chen Y-PP, Ali S. Kernel-based Naive Bayes classifier for breast cancer prediction. *J Biol Syst*. 2007;15:17–25.
62. Yamauchi Y, Mukaidono M. Probabilistic inference and Bayesian theorem based on logical implication. *Lecture notes on computer science*. Berlin: Springer; 1999. p. 334–42.
63. Fang X. Naïve Bayes: inference-based Naïve Bayes cost-sensitive turning. *Nai*. 2013;25:2302–14.
64. Zaidi NA, Du Y, Webb GL. On the effectiveness of discretizing quantitative attributes in linear classifiers. *J Mach Learn Res*. 2017;01.
65. Zhang J, Chen C, Xiang Y, Zhou W, Xiang Y. Internet traffic classification by aggregating correlated naive bayes predictions. *IEEE Trans Inf Forensics Secur*. 2013;8:5–15.
66. Chen C, Zhang G, Yang J, Milton JC, Alcántara AD. An explanatory analysis of driver injury severity in rear-end crashes using a decision table/Naïve Bayes (DTNB) hybrid classifier. *Accid Anal Prev*. 2016;90:95–107.
67. Bermejo P, Gámez JA, Puerta JM. Knowledge-based systems speeding up incremental wrapper feature subset selection with Naive Bayes classifier. *Knowl-Based Syst*. 2014;55:140–7.
68. Huang T, Weng RC, Lin C. Generalized Bradley-Terry models and multi-class probability estimates. *J Mach Learn Res*. 2006;7:85–115.
69. Zhang P, Gao BJ, Zhu X, Guo L. Enabling fast lazy learning for data streams. In: *Proceedings of IEEE international conference on data mining, ICDM*. 2011; pp. 932–41.
70. Schwenker F, Trentin E. Pattern classification and clustering: a review of partially supervised learning approaches. *Pattern Recognit Lett*. 2014;37:4–14.
71. Hussain L, Ahmed A, Saeed S, Rathore S, Awan IA, Shah SA, et al. Prostate cancer detection using machine learning techniques by employing combination of features extracting strategies. *Cancer Biomark*. 2018;21:393–413.
72. Rathore S, Hussain M, Khan A. Automated colon cancer detection using hybrid of novel geometric features and some traditional features. *Comput Biol Med*. 2015;65:279–96.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.