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

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

Knowledge-Based Systems 241 (2022) 108207

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

SAM: Self-augmentation mechanism for COVID-19 detection using chest X-ray images

Usman Muhammad [a,*], Md. Ziaul Hoque [a], Mourad Oussalah [a,b], Anja Keskinarkaus [a], Tapio Seppänen [a], Pinaki Sarder [c]

[a] Center for Machine Vision and Signal Analysis, Faculty of Information Technology and Electrical Engineering, University of Oulu, Finland
[b] Medical Imaging, Physics, and Technology (MIPT), Faculty of Medicine, University of Oulu, Finland
[c] Department of Pathology and Anatomical Sciences, University at Buffalo, USA

ARTICLE INFO

Article history:
Received 12 August 2021
Received in revised form 7 January 2022
Accepted 8 January 2022
Available online 17 January 2022

Keywords:
COVID-19 detection
Feature augmentation
Transfer learning
BiLSTM
RICA

ABSTRACT

COVID-19 is a rapidly spreading viral disease and has affected over 100 countries worldwide. The numbers of casualties and cases of infection have escalated particularly in countries with weakened healthcare systems. Recently, reverse transcription-polymerase chain reaction (RT-PCR) is the test of choice for diagnosing COVID-19. However, current evidence suggests that COVID-19 infected patients are mostly stimulated from a lung infection after coming in contact with this virus. Therefore, chest X-ray (i.e., radiography) and chest CT can be a surrogate in some countries where PCR is not readily available. This has forced the scientific community to detect COVID-19 infection from X-ray images and recently proposed machine learning methods offer great promise for fast and accurate detection. Deep learning with convolutional neural networks (CNNs) has been successfully applied to radiological imaging for improving the accuracy of diagnosis. However, the performance remains limited due to the lack of representative X-ray images available in public benchmark datasets. To alleviate this issue, we propose a self-augmentation mechanism for data augmentation in the feature space rather than in the data space using reconstruction independent component analysis (RICA). Specifically, a unified architecture is proposed which contains a deep convolutional neural network (CNN), a feature augmentation mechanism, and a bidirectional LSTM (BiLSTM). The CNN provides the high-level features extracted at the pooling layer where the augmentation mechanism chooses the most relevant features and generates low-dimensional augmented features. Finally, BiLSTM is used to classify the processed sequential information. We conducted experiments on three publicly available databases to show that the proposed approach achieves the state-of-the-art results with accuracy of 97%, 84% and 98%. Explainability analysis has been carried out using feature visualization through PCA projection and t-SNE plots.

© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Coronavirus disease (COVID-19) is a viral respiratory disease that initially emerged in China when a cluster of patients with unknown pneumonia was reported in the capital of Hubei Province (Wuhan). The virus that caused the disease was identified to be severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the International Committee on Taxonomy of Viruses based on phylogeny, taxonomy, and established practice [1]. At the time of writing, the World Health Organization (WHO) reported that approximately 185 million people are affected and 92,798 deaths worldwide [2]. Moreover, animals such as cats and dogs have also been reported to be infected with SARS-CoV-2 in many countries, including the United States. Thus, WHO declared this virus a "Public health emergency of international concerns" and classified it as a pandemic on March 2020 [3]. Most infected people develop mild to moderate illness and common symptoms are runny nose, body aches, cough, fever, sore throat, and shortness of breath [4].

Since the beginning of the pandemic, several diagnostic methods have been approved by several international and country-specific agencies. However, there is no clear consensus on the correct tests to be used related to any acute complaints to yield a correct diagnosis in a timely constraint. In EU member countries, there are 365 different commercialized devices that have been used for conducting a such research. Among them, 168 are Immunoassays, three are sequencing-based methods, 192 are PCR-based methods, and two commercialized tools are based on

* Corresponding author.
E-mail addresses: muhammad.usman@oulu.fi (U. Muhammad), mdziaul.hoque@oulu.fi (Md.Z. Hoque).

https://doi.org/10.1016/j.knosys.2022.108207
0950-7051/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

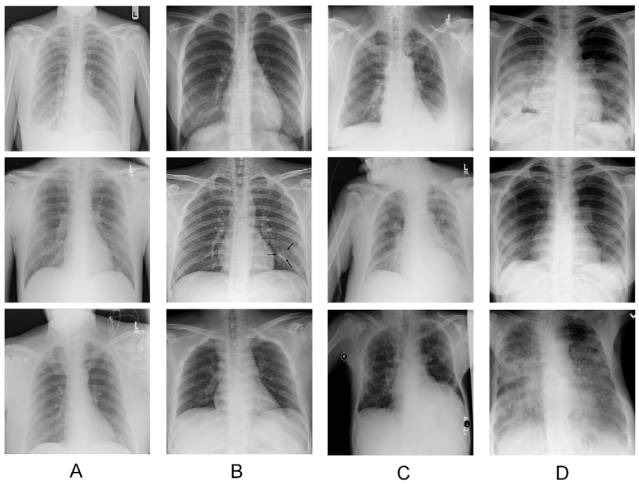


Fig. 1. Samples images from chest X-ray database: (A) Normal, (B) COVID-19, (C) Pneumonia, (D) Bacterial pneumonia.

different medical devices [5]. WHO recommended RT-PCR test (developed by Corman), which is nowadays considered as the current standard for detecting a coronavirus infection. However, the false-negative rate was found to be approximately 20% to 40% in the infected cases in China due to inappropriate sample collection, faulty operation, storage, and low sensitivity test kits [6].

Along with laboratory testing, chest CT scans with the help of a radiologist can be considered as a complementary tool with RT-PCR [7]. COVID-19 infected patients show ground-glass opacities (GGO) in the periphery of both lungs, and appear more gray or hazy as opposed to the normally dark-appearing lungs. It is also stated that those patients who recovered from COVID-19 pneumonia, lung disease was observed ten days after the onset of symptoms [8]. In the early days of the pandemic, clinical centers in Wuhan were working with an insufficient number of often malfunctioning test kits, resulting in a concerning amount of false negatives. To counteract these challenges, doctors were persuaded to make diagnoses based only on laboratory and chest CT results [9]. In developing countries, such as India, where the number of test kits remains low, CT is also used for COVID-19 detection.

In addition to CT scans, chest X-ray scanning machines are easily accessible in almost all hospitals, and have a potential role in the diagnosis because X-ray images represent visual indexes linked with COVID-19 [7]. In Fig. 1, we visualize example images from normal, COVID-19, pneumonia, and bacterial pneumonia classes, taken from two X-ray image databases [8,10,11]. Thus, radiologic images obtained from COVID-19 cases with laboratory results may help in the early detection of infection. The study conducted on CT images by Kong et al. [12] demonstrates acute bilateral airspace opacities in infected patients. Zhao et al. [13] reported that most patients had a fever as the onset symptom. Based on the result of X-ray scans, most patients with COVID-19 pneumonia are reported with ground-glass opacities (GGO) (87%) and traction bronchiectasis (53%). Moreover, authors show that lesions present on CT images are more likely to have a peripheral distribution. Li et al. observed that chest CT had a small rate of missed diagnosis of COVID-19. GGOs and consolidation with or without vascular enlargement are common CT features of COVID-19 and may be useful as a standard method for the rapid diagnosis of COVID-19 [7]. Similarly, Zu et al. [14] concluded from 1649 cases that chest CTs depict rounded lung opacities.

Machine learning (ML) techniques are attracting substantial interest in the medical field, where deep learning-based models have been successfully utilized in many healthcare applications such as depression detection [15], pain estimation [16], breast cancer detection [17], Alzheimer's disease classification [18], and pneumonia detection from chest X-ray images [19]. Due to the increase in COVID-19 cases, healthcare systems have been overwhelmed and require alternative solutions for the automated diagnosis of COVID-19. In this regard, many attempts have been put forward to address such problems using radiology images [7–9,12,20,21]. However, it is not feasible to build a large labeled database for every disease, i.e., viral pneumonia, COVID-19, bacterial pneumonia, aspiration pneumonia, etc. Thus, the bias in small datasets and the lack of representative training and tuning data impair the performance of such deep learning models.

A simple way to deal with these challenges consists in applying data augmentation techniques, which enable researchers to significantly increase the diversity of data, without collecting new data. However, augmented data that could be borrowed from unlabeled data [22], random erasing [23] or randomly masking regions [24] are heavily dependent on training parameters. For instance, a slight rotation between 1 to 30 or a random cropping ($(288, 288) \rightarrow (224, 224)$), could be useful on digit recognition tasks such as MNIST, but as the rotation degree increases, the label of the data is no longer preserved under post-transformation [25]. The dominant approaches such as Generative adversarial networks (GANs) [26], Bidirectional GANs [27], the DCGAN [28], Progressively Growing GANs [29], the CycleGAN [30], generate synthetic images but require careful domain adaptation to transfer the knowledge and features to the real image domain. A study based on combined CNN-BiLSTM reveals that training samples of COVID-19 needs to be enlarged to test the generalizability of the developed systems [31]. The authors claim that COVID-19 images might be associated with multiple disease symptoms, and demand computer-aided diagnostic systems (CAD) to detect them accurately and rapidly. Most of the existing methods either consider 2-class (normal vs. COVID-19) i.e., binary classification) or 3-class classification (normal vs. pneumonia vs. COVID-19). To overcome this issue, Asif et al. [32] proposed a deep learning model based on Xception architecture for 4-class cases (COVID vs. Pneumonia bacterial vs. pneumonia viral vs. normal). Wang et al. [33] introduces COVID-Net and achieved 83.5% accuracy in classifying pneumonia-bacterial, COVID-19, normal, and pneumonia-viral classes. However, the performance for multi-class disease predictions remains unclear. Except ensemble-CNNs [7], none of the methods discussed to treat 5-class cases.

On the other hand, existing CNN-LSTM based COVID-19 detection methods [34] treat the convolutional features as equally important and ignore the interference information (e.g., mutual exclusion and redundancy), which can prevent learning of long data sequences. Moreover, the high dimensional vector generated by CNN can increase the network parameters of LSTM and make the network difficult to optimize. Therefore, motivated by the urgent need to develop an artificial intelligence (AI) solution to aid in rapid evaluation of different lung diseases with COVID-19 detection, inspired by open source available databases, we propose a unified architecture that consists of a deep convolutional neural network (CNN), a feature augmentation mechanism, and a bidirectional LSTM (BiLSTM) for the detection of 5-cases including COVID-19 from X-ray images. Specifically, the feature augmentation mechanism based on reconstruction independent component analysis (RICA) [35] is designed in such a way that it improves the performance of CNN-based BiLSTM architecture by approximating the real distribution in feature space rather than in data space, where the generated features are mutually

independent and promise diversity. To the best of our knowledge, this is the first work in COVID-19 literature that implements feature augmentation without performing any training data augmentation strategy. An illustration of a such concept is provided in Fig. 2. By employing this strategy, the interference information or redundancy is significantly eliminated by selecting the low-dimensional augmented features.

In addition, it is worth mentioning that chest radiography analysis is known to have inherent limitation in early stages of COVID-19 detection, due to low sensitivity in ground-glass opacity detection [14]. Moreover, recovered patients are likely to be protected against reinfection for several weeks but may still transmit the virus. However, well-trained deep learning methods can focus on anomalies that are not visible to human eyes, and may encourage their applications in a health care system. Overall, our main contributions in this paper are summarized as follows:

1. We introduce a deep feature augmentation framework to improve COVID-19 detection mitigating the current lack of sufficient annotated data.
2. We employ a combined CNN-BiLSTM network to show that the proposed low-dimensional augmented features are more compact and more powerful than raw CNN features for the diagnosis of COVID-19 in a robust manner.
3. The effectiveness and validation of our proposed method have been extensively explored on three publicly available datasets and compared with state-of-the-art results.
4. PCA and t-SNE feature visualization has been utilized to demonstrate the explainability of the proposed learning model. Moreover, a detailed experimental analysis is conducted in terms of specificity, sensitivity, F1-score, accuracy, confusion matrix, and receiver operating characteristic (ROC) to determine the performance of the proposed method.

The rest of this paper is organized as follows. Section 2 deals with literature review in the field. Section 3 details the proposed method highlighting the different phases of our proposal, including deep-feature extraction, augmentation-based learning module and the associated recurrent neural network. Section 4 emphasizes the experiment results, including dataset description, evaluation metrics, implementation details, results and explainability analysis. Finally conclusive statements and perspective works are provided in Section 5.

2. Literature review

In recent months, researchers have evaluated SARS-CoV-2 infected chest X-ray images using convolutional neural networks (CNNs), recurrent neural networks (RNNs), generative adversarial networks (GANs) and encoder–decoder models. A brief overview about recent developments is provided in this section.

Narin et al. [36] utilized five pre-trained deep learning models (InceptionV3, Inception-ResNetV2, ResNet-50, ResNet-101 and ResNet-152) to detect COVID-19, bacterial pneumonia and viral pneumonia. Their best performing model achieved an accuracy of 98% using a ResNet-50 CNN pre-trained on COVID-19 images. Rajaraman et al. [37] proposed a method that increases training data using weakly labeled data augmentation. A stage-wise approach was used to train the CNN. The authors concluded weakly labeled data augmentation is superior in comparison to baseline non-augmented training. A correlation learning mechanism is proposed, and the images are augmented by flipping, image nosing, and rotation in [38].

Ozyurt et al. [39] used a traditional machine learning descriptor (LBP), and a feature selector method that selects most informative features together to achieve a better performance, achieving a 95.84% classification accuracy on CT images. Rahimzadeh and Attar [40] introduced a combined deep CNN to identify 11 302 chest X-ray images. In their study, Xception and ResNet-50V2 were used and claimed as a new strategy to address the unbalanced dataset problem. They reported an accuracy rate of 99.56%. Castiglioni et al. [41] utilized an independent dataset of 110 patients suspected for COVID-19 infection, and developed a ten convolutional neural networks (CNNs) to evaluate the performance.

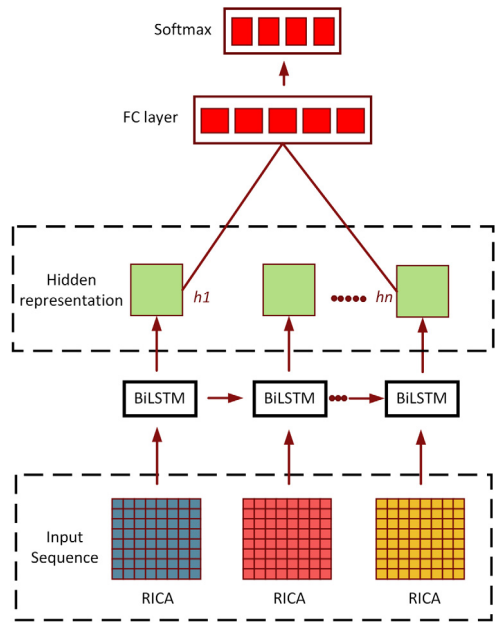


Fig. 2. The main idea of the proposed work. RICA is designed to augment the feature space. The augmented features are used as a sequence input to train the BiLSTM network for final detection. Three different colors in RICA represent three augmented features.

Multiple state-of-the-art deep learning models including DenseNet201, Resnet-50V2 and Inceptionv3, were fine-tuned individually to make independent predictions in [42]. Then, a weighted average ensembling technique was used to combine them to achieve a classification accuracy of 91.62%. Similarly, Wang et al. [33] proposed a tailored deep convolutional neural network for classifying chest X-ray images. Hemdan et al. [43] examined seven different CNN architectures in their experiment, including DenseNet-121, VGG-19, ResNetV2, Inception-V3, Xception, MobileNetV2 and InceptionResNetV2. Their work revealed that the DenseNet and VGG-19 models achieved the best performance with 91% accuracy for detecting COVID-19 and non-COVID-19 infections.

Berrimi et al. [44] fine-tuned two pre-trained models, DenseNet and InceptionV3 to classify both X-ray and CT chest scans. To increase the diversity of the training data, the images were rotated, zoomed, horizontally flipped, and shifted. Nour et al. [45] designed a five convolution layers (CNN) from the scratch. The extracted CNN features are then evaluated with traditional machine learning classifiers such as k-nearest neighbor, support vector machine (SVM), and decision tree. The authors concluded that the SVM classifier with an accuracy of 98.97% performs the best among all of them. Giacomo et al. [46] detect lung disorder by using X-ray images. Specifically, a fuzzy logic segmentation method combined with a neural network is proposed, and accuracy of 92.56% is reported in their work. Yoo et al. [47] used the pre-trained ResNet-18 and different decision trees are utilized to detect CXR images as normal, tuberculosis, and COVID-19. Aslan et al. [34] proposed a hybrid architecture based on CNN–BiLSTM

for COVID-19 detection. Moreover, the authors employed two deep learning architectures including Artificial Neural Networks (ANN) and a hybrid structure containing a BiLSTM layer to utilize the temporal properties. The accuracy of 98.14% and 98.70% were achieved using first and second architecture, respectively. Nayaar et al. [48] showed thoracic (chest) imaging are found to be effective in the diagnosis of coronavirus disease (COVID-19). Mukherjee et al. [49] utilized a lightweight (9 layered) CNN-tailored deep neural network to detect COVID-19 positive cases, and achieved an overall accuracy of 96.28%. A federated learning is proposed to detect COVID-19, and 98.72% accuracy was reported in [50]. Challenges, innovations and opportunities to detect COVID-19 are discussed in [51]. Mukherjee et al. [52] proposed a light-weight CNN-tailored shallow architecture to detect COVID-19. The proposed model was designed with fewer parameters as compared to other deep learning models and validated using 321 COVID-19 positive Chest X-ray images with an accuracy 99.69%. Marcin et al. [53] designed a method for diseased tissues detection over input X-ray images.

M. Turkoglu [54] employed the transfer learning approach by using the AlexNet architecture. To choose the most effective features, the Relief feature selection algorithm is used. Finally, the Support Vector Machine (SVM) is applied to detect COVID-19 and Pneumonia disease. An accuracy of 99.18% was reported. Sahlol et al. [55] proposed a combined approach where Inception model is utilized to extract the features and a swarm-based feature selection algorithm is applied to choose the most relevant features. Two public COVID-19 X-ray datasets are used and 99.18% accuracy was reported. Mesut et al. [56] utilized MobileNetV2 and SqueezeNet models to extract the deep features. Then, the Social Mimic optimization method is proposed and the features were combined and classified using SVM classifier. To overcome the limitation of chest X-ray samples, Karbhari et al. [57] proposed an Auxiliary Classifier Generative Adversarial Network (ACGAN) to generate synthetic images. Based on obtained images, Convolutional Neural Networks (CNNs) is utilized to detect COVID-19 in the CXRs.

Loey et al. [58] utilized a GAN architecture to synthesize auxiliary images as a motivation to overcome the issue of lack of datasets especially in chest X-rays images. Three deep transfer models are selected to detect four classes, i.e., the COVID-19, normal, pneumonia bacterial, and pneumonia virus. Googlenet performed the best in their work. The network consists of encoder and decoder is proposed in [59], to show that CORONA-Net performs the best for COVID-19 detection. MASC-Net consists of a multi-input encoder–decoder, and introduced to automatically detect infected lung regions from COVID-19 chest CT scans [60]. 3D U-Net is proposed as encoder–decoder method in [61], where the multi-task learning is applied and compared with four transfer learning strategies. The authors concluded that using multiple lung lesion datasets can extract more general features.

Therefore, previous research showed that chest X-ray images have been commonly used in most of the current works and have an important role in the diagnosis of COVID-19 detection. However, learning from imbalanced data or lack of necessary extracted features obtained from limited X-ray training samples cannot provide the expected performance in the COVID-19 detection. Thus, the proposed work focuses on a data augmentation strategy where the label preserved features are generated to improve the performance of deep learning model.

3. The proposed method

The general framework of the proposed approach is divided into three components: (1) extraction of deep features (2) an augmentation-based learning module and (3) a BiLSTM based sub-network. We first describe the procedure of feature extraction for guiding the process of feature generation. Next, we explain the procedure of augmenting the training data in feature space. Finally, the structure of BiLSTM network is discussed. The overall procedure of the proposed approach is illustrated in Fig. 3.

3.1. Deep features extraction

Inspired by the performance of deep learning models, we adapt ResNet-50, a CNN architecture known for its stability and performance, to extract high-quality features for our task [62]. The model is fine-tuned by replacing the last fully-connected layer with a new fully connected layer and setting the number of outputs equal to the number of classes in the dataset. We freeze the weights of the first ten layers so that the gradients of these do not need to be computed. This is motivated by the fact that earlier features of ResNet contain more generic features (e.g. color blob detectors or edge detectors) and make the remaining layers more specific to the details of the classes contained in the original dataset. The weights of the new fully-connected layer are increased by a learning factor 10 and a bias factor 20. By biasing the weight updates in the new fully-connected layer, the influence of each training sample in the new dataset is magnified and training time is reduced. We utilize cross-entropy loss to adjust model weights during training. The purpose is to decrease the loss and motivate the network towards accurate predictions. It is defined as

$$L = -\sum_{i=1}^{k} t_i \, log(s_i), \quad For\ k\ classes, \tag{1}$$

where t_i denotes the true label, s_i the softmax probability for the ith class, and k the total number of classes.

3.2. An augmentation-based learning module

Data augmentation in the image space is a well established technique that enhances the size and quality of training datasets such that deep learning models can robustly model the training data. However, feature augmentation has not yet acquired the same level of attention. This is crucial for applications like COVID-19 detection, where the number of training samples remain limited. To accomplish this, feature augmentation is conducted based on reconstruction independent component analysis (RICA) [35]. The latter was designed to overcome the drawbacks of independent component analysis (ICA) by replacing ICA's orthonormality constraint with a soft reconstruction penalty, which turns out to be very useful in learning sparse features. Therefore, the idea behind our proposed mechanism is to extract more meaningful information from generated ones to correctly classify target samples.

In our case, RICA receives data as input from the last pooling layer of RestNet-50, then it converts it into a new lower-dimension representation. In order to apply transformation, RICA is calculated by using the following equation:

$$v = Zk \tag{2}$$

where v is the vector representing the CNN features, Z denotes the matrix, and k are the independent components for dimensionality reduction. The goal of RICA is to define the observed data v by mixing the components k. We need to determine both Z and k from the data v because we cannot directly extract the sources k, nor know the mixing matrix Z. Let W be the inverse of Z, then the model can be expressed as:

$$k = Wv \tag{3}$$

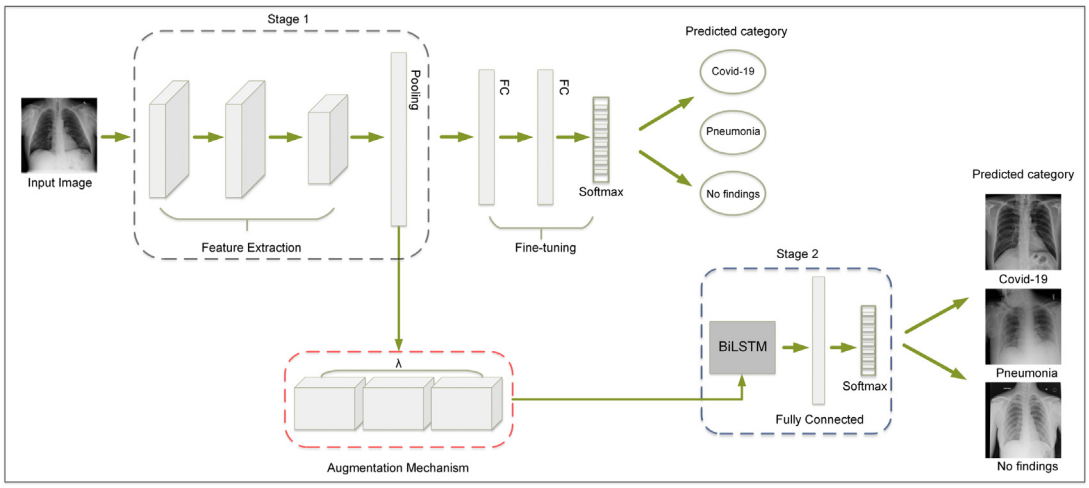


Fig. 3. Flowchart of the proposed method. First, fine-tuning is performed to train a convolutional neural network (CNN) for COVID-19 detection. An augmentation mechanism based on reconstruction independent component analysis is used to leverage the most relevant CNN features and form a sequence of augmented features, which are then fed into a bidirectional Long Short-Term Memory (LSTM) network for the final representation.

Hence, using the original data v, the goal is to determine a set of vectors (corresponding to the column vectors of matrix W) that will form the features k sparse; while being an orthonormal basis. In this regard, our matrix W will assign the data v to features k. The optimization problem defined by RICA becomes [35]:

$$\min_{W} \ \lambda\|(Wv)\|_1 + \frac{1}{2}\|W^T Wx - x\|_2^2 \tag{4}$$

where λ evaluates the objective and non-linear convex function and $W \in \mathbb{R}^{q\times n}$ (where q denotes the features and n is the number of data vectors in v). $L1$ denotes the sparsity penalty and has a tied reconstruction matrix W. To decrease the computational cost of the optimization, limited-memory BFGS (LBFGS) algorithm method [63] is used as a constrained optimizer that results in fast convergence. Moreover, RICA can manage data with approximate whitening or even without whitening [35].

As illustrated in the formulation of RICA in equation (4), in the first part, λ represents the weight assigned to the sparsity constraint in relation to the recreation condition. The second part emphasizes accurate recreation of the original features by minimizing the recreation error $\left\|W^T Wx - x\right\|$. In this regard, we fix the feature dimension of pooling layer features equal to 400, and empirically set the λ weights to 80, 100, and 120 because the higher the weight we give to the sparsity constraint the less precise will the recreation be, and vice-versa. Hence, we obtained three augmented feature vectors with different weights by keeping the same dimension 400 as mentioned above. Similarly, we repeat the same procedure to obtain three augmented features sets for the second dataset by setting the feature dimension to 500. The representations learned by the augmentation mechanism contain discriminative information related to the classes, which allows the network to accurately predict them.

3.3. Recurrent Neural Network (RNN)

Mainstream CNN frameworks are related to conventional statistical models, thus lacking the capacity to map sequences to sequences. BiLSTM [64] is one kind of RNN, which has the ability to process sequences of arbitrary length, and has yielded surprising performance in natural language processing [65]. However, the high dimensionality and sparsity of the data are one of the major challenges that limit its performance. Taking advantage of the low-dimensional RICA features, BiLSTM performs better than using the raw CNN features (we further discuss this argument in Section 4.4). The BiLSTM is implemented similarly to the standard bi-directional LSTM except that the input is based on three augmented features. We found that the proposed strategy calculated on each time step resulted in improved reconstructions, which we found to be vital to accomplish our feature augmentation process.

BiLSTM Networks capture each sequence vector based on the memory cell (C_e), and compete for retaining dependencies between the elements in the input sequence. It is comprised of an input gate (i_e), an output gate (o_e) and a forget gate (g_e). The input gate governs the information flow into the cell by multiplying the cell's non-linear transformation of inputs m_e. The output gate decides how much information from the cell is used to compute the output activation of the LSTM unit. The forget gate regulates the extent to which a value remains in the cell. The LSTM unit updates for time step e are:

$$\begin{bmatrix} g_e \\ i_e \\ m_e \\ o_e \end{bmatrix} = \begin{bmatrix} \sigma \\ \sigma \\ \tanh \\ \sigma \end{bmatrix} H \cdot [p_{e-1}, x_e] \tag{5}$$

$$C_e = g_e \odot C_{e-1} + m_e \odot i_e \tag{6}$$

$$p_e = \tanh(C_e) \odot o_e \tag{7}$$

where x_e is the input at the current time-step, i_e is the current cell state, g, i and m is the input gate activation, forget gate activation and output gate activation respectively, σ illustrates the logistic sigmoid function and $\odot$ represents element-wise multiplication.

4. Experiments and results

In this section, we first provide a brief description of three databases that are used to evaluate our method. Then, we present evaluation metrics, implementation details, and experimental results, which are discussed later in comparison to state-of-the-art methods.

4.1. Datasets

X-ray image dataset: The first dataset used in the experiments is obtained from open access sources provided by Cohen JP [10] and Wang et al. [11]. It contains three classes: pneumonia, "no-findings", and COVID-19, where pneumonia and "no-findings" have 500 images each and COVID-19 has 125 X-ray scan images. A

Table 1
Number of images per class for each dataset.

Datasets	Classes	Train set	Validation set	Test set
X-ray image dataset	3	900	–	225
COVID-19 X-ray scan dataset	3	1212	300	621
COVID-19 X-ray scan (II)	5	2020	500	1035
SARS-CoV-2 CT-scan dataset	2	2232	–	249

detailed explanation (https://github.com/muhammedtalo/COVID-19, (accessed on June 05, 2021)) can be found in [8].

COVID-19 X-ray scan database: The second dataset is collected from the open access source provided by Vantaggiato et al. [7], where two scenarios are examined. In the first scenario, three classes (Normal, COVID-19, and Pneumonia) are provided in the dataset. For training, each class has 404 images. Validation and testing sets contain 100 and 207 images, respectively. In the second scenario, two more classes were added by the authors yielding a five class model: Normal, COVID-19, Viral-Pneumonia, Bacterial-Pneumonia and Lung-Opacity. This dataset (https://github.com/Edo2610/Covid-19_X-ray_Two-proposed-Databases, (accessed on June 11, 2021)) is acquired from different open access sources [10,66–68].

The SARS-CoV-2 CT-scan dataset: The third dataset used in our work is acquired from [69]. It consists of 2481 CT scan images and collected from hospitals of Sao Paulo, Brazil. In total 1252 patients were infected with SARS-CoV-2 and (1230) were reported as normal. This dataset (https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset, (accessed on Nov 15, 2021)) is also publicly available, and we summarize number of classes and images of each dataset in Table 1.

4.2. Evaluation metrics

The performance of the proposed method is evaluated with respect to Sensitivity, Specificity, Precision, F-Score and Accuracy, defined using the equations below:

$$Sensitivity = \frac{TP}{TP + FN} * 100 \tag{8}$$

$$Specificity = \frac{TN}{TN + FP} * 100 \tag{9}$$

$$Precision = \frac{TP}{TP + FP} * 100 \tag{10}$$

$$F1 - Score = 2 * \frac{(Precision * Sensitivity)}{Precision + Sensitivity} * 100 \tag{11}$$

$$Accuracy = \frac{Number\ of\ correct\ prediction}{Total\ number\ of\ prediction} * 100 \tag{12}$$

where TP, FP and FN represent the True Positive, False Positive and False Negative, respectively.

In addition to the above, performance analysis using ROC (Receiver Operating Characteristic) curves was conducted. These consists of graphs showing the variation of True Positive Rate (TPR) or, equivalently, *Sensitivity* as defined in Eq. (8), with respect to False Positive Rate (FPR), where FPR is defined as:

$$FPR = \frac{FP}{FP + TN} * 100 \tag{13}$$

The ROC exhibits the performance of the underlined classification model at all classification thresholds.

4.3. Implementation details

All the X-ray images are resized to 224 × 224 based on the size requirement of the model. No image data augmentation was applied, the features are augmented only in the feature space. The CNN is fine-tuned using stochastic gradient descent (SGD) with a learning rate of $3e^{-4}$, mini-batch size of 32 and epochs of 5, and with shuffling of samples between every epoch. To build the feature augmentation mechanism, the feature vector is extracted from the output of the last pooling layer of ResNet-50 with a size of 2048. RICA is utilized to augment the pooling layer features into three augmented feature sets with dimension of 400 for X-ray image dataset, and 500 for COVID-19 X-ray scan and the SARS-CoV-2 CT-scan databases, respectively. The performance also varies by varying the number of iterations done by RICA before stopping, and 80 to 120 iterations were used to extract the augmented features.

We treat these augmented features as three sequences, and each sequence is an A-by-Z array, where A is the number of features (the output size of RICA) and Z is the number of samples. Therefore, BiLSTM takes the input of three sequence features sets. For training the BiLSTM, the Adam optimizer is used by setting the learning rate up to 0.0001. A BiLSTM layer with 60 hidden units, a fully connected layer, a softmax layer, and the number of epochs are fixed to 150. Initializing the BiLSTM sub-network with random initialization can be challenging because large random-valued weights may lead to the problem of exploding gradients. Therefore, we set the recurrent weights with *He* initializer [75] which performs the best in all scenarios of our experiments.

4.4. Experimental results

In order to detect COVID-19 on the first X-ray image database [10,11], we split the original image dataset to eighty percent for training and twenty percent for testing as in the same spirit as [8]. Rather than proposing the CNN model from the scratch, we leverage transfer learning using pre-trained CNN models that have shown outstanding results in classification tasks of a wide variety of classes/types/applications. Specifically, four state-of-the-art pre-trained models, such as ResNet-50 [62], SqueezeNet [76], GoogleNet [77], and DenseNet-201 [78] were fine-tuned for the COVID-19 detection task and their results are reported in terms of specificity, sensitivity, precision, F1-score, and accuracy in Table 2. Then, we evaluate the performance of deep models by combining them with BiLSTM (CNN-BiLSTM) network. Finally, we report the efficacy of the proposed augmentation mechanism with a combined CNN-BiLSTM. It can be noted that all these deep learning models exhibit a limited performance to detect three classes: COVID-19, No-Findings, and Pneumonia. One of the reasons is that a number of training samples in the COVID-19 class resembles Pneumonia class and this is not enough to compare to the other two classes (No-Findings and Pneumonia), which might cause overfitting of the model. However, when these models are connected with BiLSTM, the performance improves but remains limited due to ignoring the interference information.

Table 3 summarizes the comparison with other state-of-the-art works using three-class X-ray image database. Since COVID-19 is an emerging disease, the first dataset (X-ray image) used in our work is being updated regularly with the new images. Thus, making a fair comparison with other works would not be possible except the previous work [8] in comparison to us. However, we compare our method with other three class methods which are specifically designed for COVID-19 detection. At the time of writing this paper, the database contained a total of 125 COVID-19 chest X-ray images. The best results were obtained from the ResNet-50 and validated using a 5-fold cross-validation

Table 2
The performance comparison results for the three-class X-ray image database.

Method	Performance metrics (%)				
	Sensitivity	Specificity	Precision	F1-score	Accuracy
SqueezeNet	70.00	80.90	82.87	69.59	75.66
SqueezeNet+BiLSTM	89.00	93.60	93.30	90.40	91.89
SqueezeNet+SAM+BiLSTM	**94.66**	**95.73**	**95.40**	**94.63**	**95.25**
Googlenet	84.66	88.93	83.41	83.06	87.25
Googlenet+BiLSTM	92.00	94.10	93.58	93.49	93.55
Googlenet+SAM+BiLSTM	**93.66**	**96.73**	**96.67**	**94.66**	**96.10**
DenseNet201	89.66	92.03	87.04	88.18	90.81
DenseNet201+BiLSTM	95.33	94.86	91.27	92.42	94.16
DenseNet201+SAM+BiLSTM	**98.00**	**95.20**	**96.02**	**94.99**	**96.40**
ResNet-50	83.00	91.20	88.29	85.56	87.55
ResNet-50+BiLSTM	99.80	88.80	87.71	93.45	93.77
ResNet-50+SAM+BiLSTM	**99.93**	**92.80**	**91.74**	**95.69**	**97.26**

Table 3
The general comparison of the proposed method with other 3-class state-of-the-art methods.

Study	Methods	Class	Overall accuracy (%)
Ucar et al. [70]	COVIDiagnosis-Net	3	98.26
Apostolopoulos et al. [71]	The pretrained CNNs	3	93.48
Li and Zhu [72]	DenseNet	3	88.9
Wang and Wong [33]	Tailored CNN	3	92.3
Chowdhury et al. [73]	Sg-SqueezeNet	3	98.3
Ucar et al. [70]	Bayes-SqueezeNet	3	98.3
Ozturk et al. [74]	DarkCovidNet	3	87.02
Proposed method	ResNet-50+SAM+BiLSTM	3	97.26

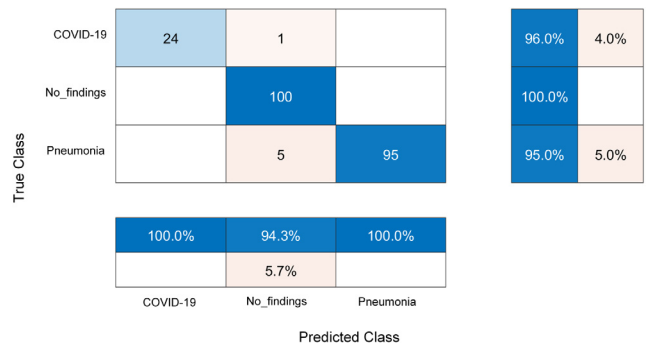


Fig. 4. Confusion matrix of the three-class using ResNet-50+SAM+BiLSTM for X-ray image dataset [8]. The horizontal and vertical axis is for predicted and true classes, respectively.

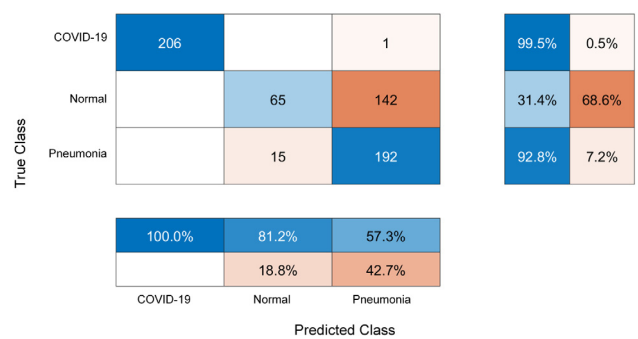


Fig. 5. Confusion matrix of the three-class using ResNet-50+SAM+BiLSTM for COVID-19 X-ray scan database [7]. The horizontal and vertical axis is for predicted and true class, respectively.

procedure. The proposed method provides a 97% accuracy which is 10% higher than the previously proposed method [8] on the same dataset (three classes), and 99% accuracy for the two-class scenario. For further analysis, a confusion matrix is shown in Fig. 4. It can be observed that the proposed approach classified COVID-19 better than the other two classes.

Table 4 illustrates the results of the three-class scenario on the second COVID-19 database [7]. The training set, validation set, and testing set are provided separately. Using the self-augmentation mechanism, the proposed method achieves 79% accuracy and improves 4% performance from the previous study [7]. In addition to the three-class scenario, the results of the five-class scenario are also reported in Table 5. It can be observed that the proposed method provides 84% accuracy which is 3% better than the Ensemble-CNNs [7]. For further analysis, Figs. 5 and 6 represent the confusion matrices of the three-class and the five-class, respectively. The main observation is that the proposed method attained 99.5% accuracy for the detection of COVID-19 samples. A confusion matrix in Fig. 7 is illustrated only for COVID-19 and Normal class with two rows and two columns showing the number of true positives, false negatives, false positives and true negatives. It shows that the model predicted all 207 COVID-19 X-ray images correctly, and no false negative are detected. In the case of normal class, 98 images are misclassified while 109 images were correctly classified. All correct predictions are located in the diagonal of the table (highlighted in light blue and dark blue), so it is easy to visually inspect the table for prediction errors. Therefore, we can observe that the proposed framework is proficient in distinguishing the COVID-19 samples in both datasets.

As our main focus is the classification of COVID-19 samples, we present ROC curves for a two-class detection problem (COVID-19 vs. Normal), in which only the true positive rate (TPR) and false positive rate (FPR) are needed. The best possible detection method would allow a learning curve in the upper left corner or coordinate (0,1) of the ROC space, depicting 100% sensitivity (no false negatives) and 100% specificity (no false positives). In Fig. 8, the curves are visualized with raw CNN features, CNN with BiLSTM-based network, and finally with an augmentation mechanism. It can be observed that the proposed augmentation mechanism clearly improves the performance of a CNN-based BiLSTM architecture as exhibited by higher sensitivity rate.

In Table 6, we evaluate our proposed method for the COVID-19 CT-scan dataset with state of-the-art methods. The results are obtained by dividing the dataset into 90% as training and 10% as testing dataset [80]. In contrast to previous deep learning methods [69,79,80], our proposed method explicitly takes advantage of augmented features and efficiently detect COVID-19 cases by achieving the accuracy of 98.38%. Thus, based on experimental analysis on all three datasets, we concluded that neither a single CNN model nor CNN-based BiLSTM achieves the best results for all the evaluation metrics. Therefore, the proposed augmentation is essential to produce a robust feature representation for COVID-19 detection.

Table 4
The performance comparison results for the three-class COVID-19 database.

Method	Performance metrics (%)				
	Accuracy	Precision	Sensitivity	Specificity	F1-score (%)
SqueezeNet	66.60 ± 0.44	61.25	49.75	74.87	48.57
SqueezeNet+BiLSTM	73.26 ± 0.82	59.90	70.30	83.91	56.68
SqueezeNet+SAM+BiLSTM	**76.57 ± 1.47**	**64.57**	**67.28**	**84.01**	**63.29**
Googlenet	72.38 ± 1.47	70.82	67.08	80.54	65.57
Googlenet+BiLSTM	76.49 ± 1.47	67.24	73.97	85.62	62.19
Googlenet+SAM+BiLSTM	**77.61 ± 2.47**	**74.42**	**73.79**	**86.15**	**73.97**
DenseNet201	75.19 ± 2.47	71.17	67.69	81.34	64.65
DenseNet201+BiLSTM	76.19 ± 2.47	61.03	67.19	82.55	62.30
DenseNet201+SAM+BiLSTM	**79.90 ± 3.05**	**75.48**	**78.39**	**81.36**	**69.75**
ResNet-50	74.53 ± 1.25	73.88	73.10	86.55	68.97
ResNet-50+BiLSTM	77.53 ± 3.53	71.49	71.93	87.71	69.64
Ensemble-CNNs [7]	75.23 ± 3.40	78.28	73.20	87.60	73.43
ResNet-50+SAM+BiLSTM	**79.53 ± 1.41**	**74.55**	**77.52**	**88.42**	**71.96**

Table 5
The performance comparison results for the five-class COVID-19 database.

Method	Performance metrics (%)				
	Accuracy	Precision	Sensitivity	Specificity	F1-score (%)
SqueezeNet	72.88 ± 1.17	63.42	58.45	89.61	55.62
SqueezeNet+BiLSTM	79.55 ± 2.17	61.73	67.98	91.22	55.41
SqueezeNet+SAM+BiLSTM	**81.78 ± 2.41**	**67.43**	**75.10**	**92.75**	**61.27**
Googlenet	73.98 ± 1.88	68.01	63.28	90.82	62.55
Googlenet+BiLSTM	77.68 ± 1.47	70.53	74.87	81.87	62.43
Googlenet+SAM+BiLSTM	**80.77 ± 1.13**	**73.62**	**77.40**	**94.03**	**71.36**
DenseNet201	82.51 ± 2.40	73.02	72.36	93.09	70.14
DenseNet201+BiLSTM	81.10 ± 1.54	84.54	77.41	82.29	83.95
DenseNet201+SAM+BiLSTM	**84.10 ± 1.90**	**83.76**	**80.11**	**91.22**	**83.25**
ResNet-50	79.64 ± 1.84	75.94	71.49	92.87	68.29
ResNet-50+BiLSTM	83.64 ± 1.84	67.14	77.93	92.41	60.71
Ensemble-CNNs [7]	81.00 ± 2.39	82.99	82.96	85.24	81.49
ResNet-50+SAM+BiLSTM	**84.64 ± 2.55**	**76.32**	**82.19**	**90.82**	**73.88**

Table 6
The performance comparison results for the SARS-CoV-2 CT-scan dataset.

Method	Performance metrics (%)				
	Accuracy	Precision	Sensitivity	Specificity	F1-score
xDNN [69]	97.38	99.16	95.53	–	97.31
GoogleNet [69]	91.73	90.20	93.50	–	91.82
Inception ResNet [79]	90.90	90.15	92.06	89.72	91.09
DenseNet [79]	96.25	96.29	96.29	96.21	96.29
ResNet-152V2 [79]	94.91	92.92	97.35	92.43	95.09
VGG16 [80]	96.39	96.80	96.03	96.21	96.41
COV-CAF [80]	97.59	96.88	98.41	97.82	97.63
ResNet-50+SAM+BiLSTM (Ours)	**98.38**	**96.85**	**99.97**	**96.80**	**98.40**

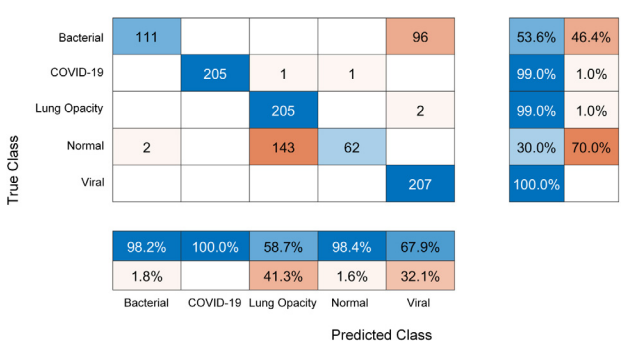


Fig. 6. Confusion matrix of the five-class using ResNet-50+SAM+BiLSTM for COVID-19 X-ray scan database [7]. The horizontal and vertical axis is for predicted and true class, respectively.

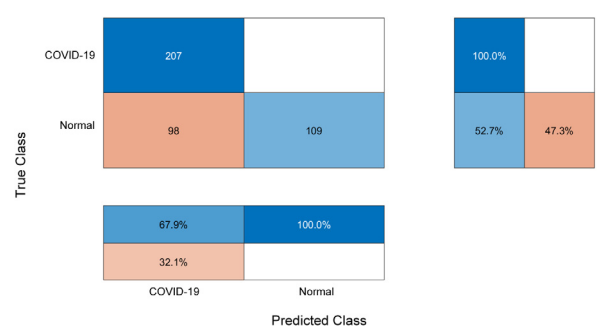
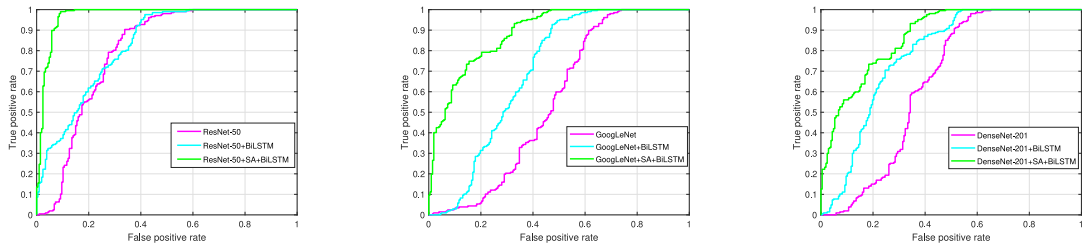


Fig. 7. Two class confusion matrix for COVID-19 X-ray scan database. The horizontal and vertical axis is for predicted and true class, respectively.

4.5. Ablation study

We conduct the ablation study to present how the weight (λ) assigned to the sparsity constraint can have influences on performance. We note the effectiveness of each augmented feature set on all three datasets. The computational time is also calculated which was required to train and test the model. Moreover, we utilize other dimensional reduction techniques such as principal component analysis (PCA) [81] and factor analysis [82] to compare the performance with our propose method. All the

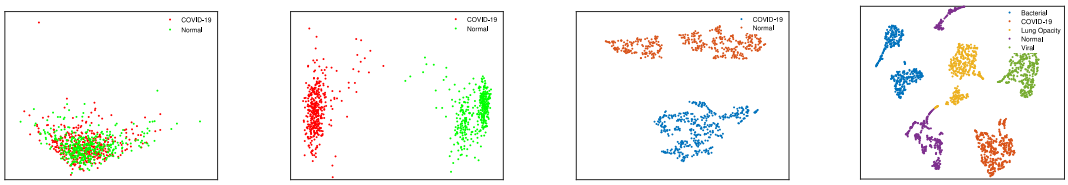


(a) The performance evaluation with the ResNet-50 model.

(b) The performance evaluation with the GoogLeNet model.

(c) The performance evaluation with the DenseNet-201 model.

Fig. 8. The Receiver Operating Characteristics (ROC) curves of the COVID-19 X-ray scan database, where the true positive rate represents the *y*-axis and false positive rate represents the *x*-axis.



(a) ResNet-50 features with PCA visualization.

(b) ResNet+BiLSTM with PCA visualization.

(c) ResNet+BiLSTM with t-SNE visualization.

(d) ResNet+BiLSTM with t-SNE visualization.

Fig. 9. Two-dimensional scatter plots of ResNet features with PCA and t-SNE over COVID-19 X-ray database. The clusters correspond to two and five different classes available in the data.

experiments are performed on a workstation with 3.5 GHz Intel Core i7-5930k and 64 GB RAM memory.

From Table 7 findings, it is evident that the ResNet-50+λ_1+ BiLSTM obtains good accuracy by achieving 95.60% accuracy. When we increase the weights(λ_2, λ_3), the model further improves the performance. By combining all the weighted features, we achieved the best performance on X-ray image dataset (3-class case). Similarly, for COVID-19 X-ray scan database (5-class case) and SARS-CoV-2 CT-scan dataset (2-class case), we obtained the highest performance with the combination of λ_1, λ_2, and λ_3. In our experiments, the proposed method takes 352.4 s for training and 1.1 s for testing X-ray image dataset. In addition, COVID-19 X-ray scan database and SARS-CoV-2 CT-scan dataset take 528.4 s and 478.1 s for training, and 1.9 s and 1.1 s for testing, respectively.

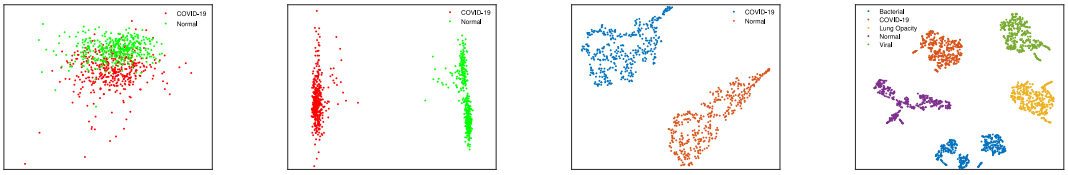
In Table 8, we can see the detailed classification results after using dimensional reduction techniques. The obtained results show that PCA significantly drops the performance on all the datasets. It might be possible that PCA fails to sustain feature transformation when reduce into a small number of components, i.e. linear combinations of the original features. Surprisingly, factor analysis provides better performance on COVID-19 X-ray scan database and achieves state-of-the-art performance when ResNet and BiLSTM networks are combined. However, it decreases the performance when tested on SARS-CoV-2 CT-scan and X-ray image databases. In contrast to PCA and factor analysis, the proposed mechanism maintains the best performance on all three datasets.

4.6. Explainability analysis

One of the advantages of the proposed approach is that we can interpret the detection process of the model. For each stage, we can see how the features are structured into the high-dimensional (ResNet) and the impact of the augmented feature space along the different classification stages. Taking this into consideration, we employed the PCA projection [83] and the (t-SNE) algorithm [84]. PCA offers a nice explanatory framework since its axes are made of a linear combination of the original dimensions, allowing comprehension of high dimensional patterns. Similarly, t-distributed Stochastic Neighbor Embedding (t-SNE) creates a low-dimensional representation of complex high dimensional data through a series of transformation and fine-tuned optimization procedures. In this respect, the projection results of both PCA and t-SNE provide a rough indication of the quality of the separation and supporting explainability through visual exploration.

The ResNet features in Fig. 9(a) indicate that both classes (COVID-19 vs. Normal) are strongly correlated, which makes it hard for the BiLSTM to separate them as shown in Fig. 9(b) and (c). We also visualize the five class features in Fig. 9(d) and observed that the Normal class is still correlated with the Lung Opacity class that causes overfitting of CNN-BiLSTM architecture. The derived clusters indicate that the prior information obtained from raw CNN features causes to decrease the performance. On the other side, augmented features generated by RICA reduce the correlation between similar classes as shown in Fig. 10(a) and are able to capture more variability in the feature space. Moreover, it can be noticed from Fig. 10(b) and (c) that data points corresponding to Normal and COVID-19 are linearly separable which could potentially lead to better performance when training BiLSTM on low-dimensional data. Thus, the proposed mechanism helps to overcome the overfitting issue and also separate the five classes efficiently in comparison to raw CNN features as shown in Fig. 10(d).

It should be noted that our exploration through visualization as a way to achieve explainability can be further expanded in different directions. First, projection quality metric can be used to assess the quality of each projection by using PCA or t-SNE. This includes global measures such that Normalized Stress, Distance Consistency, ClustMe [85], or local measures such as projection precision score [86] can contribute to shedding light on the quality of such projections. Nevertheless, it should be noted that such assessment may also be misleading and cannot contribute towards comprehending *why such results occurred*. In this context, one shall mention the interesting work of Fujiwara et al. [87]



(a) ResNet-50+RICA with PCA visualization. (b) ResNet+RICA+BiLSTM with PCA visualization. (c) ResNet+RICA+BiLSTM with t-SNE visualization. (d) ResNet+RICA+BiLSTM with t-SNE visualization.

Fig. 10. Two-dimensional scatter plots of augmented features with PCA and t-SNE over COVID-19 X-ray database. The clusters correspond to two and five different classes available in the data.

Table 7
Quantitative analysis on each dataset.

	Quantitative analysis on each dataset	Accuracy (%)	Training (s)	Testing (s)
X-ray image dataset				
λ_3	ResNet-50+λ_3+BiLSTM	96.88	313.4	1.2
λ_2	ResNet-50+λ_2+BiLSTM	96.41	312.6	1.2
λ_1	ResNet-50+λ_1+BiLSTM	95.60	313.3	1.2
All	ResNet-50+SAM+BiLSTM	97.26	352.4	1.2
COVID-19 X-ray scan database				
λ_3	ResNet-50+λ_3+BiLSTM	84.59	494.7	1.9
λ_2	ResNet-50+λ_2+BiLSTM	82.10	492.8	1.9
λ_1	ResNet-50+λ_1+BiLSTM	83.71	493.2	1.9
All	ResNet-50+SAM+BiLSTM	84.64	528.4	1.9
SARS-CoV-2 CT-scan dataset				
λ_3	ResNet-50+λ_3+BiLSTM	98.10	510.7	1.3
λ_2	ResNet-50+λ_2+BiLSTM	97.56	514.8	1.3
λ_1	ResNet-50+λ_1+BiLSTM	97.75	511.4	1.3
All	ResNet-50+SAM+BiLSTM	98.38	539.1	1.3

Table 8
Performance evaluation of dimensionality reduction methods for COVID-19 datasets.

Dataset	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
X-ray image dataset	ResNet+PCA+BiLSTM	83.10	76.22	89.17
	ResNet+FAC+BiLSTM	91.20	89.64	93.66
	ResNet+SAM+BiLSTM	97.26	99.93	92.80
COVID-19 X-ray scan database	ResNet+PCA+BiLSTM	59.27	52.88	63.07
	ResNet+FAC+BiLSTM	81.88	79.22	84.48
	ResNet+SAM+BiLSTM	79.53	77.52	88.42
The SARS-CoV-2 CT-scan dataset	ResNet+PCA+BiLSTM	81.27	77.10	87.15
	ResNet+FAC+BiLSTM	89.36	84.24	92.14
	ResNet+SAM+BiLSTM	98.38	99.97	96.80

who proposed a *contrasting clusters in PCA* (ccPCA) as a way to find out which dimensions contributed more to the formation of a selected cluster and why it differed from the rest of the dataset, based on information on separation and internal versus external variability.

5. Conclusion

In this study, we address the problem of COVID-19 detection from chest CT and X-ray images. For this purpose, a unified architecture consisting of a deep convolutional neural network, an augmentation mechanism, and a bidirectional-LSTM is proposed. The CNN provides the high-level features extracted at the pooling layer where the augmentation mechanism selects the most relevant features and generates low-dimensional augmented features. Finally, BiLSTM is used to classify the processed sequential information. The proposed method provides an end-to-end structure without the need for manual feature extraction. We showed that the detection of COVID-19 was improved by using the low-dimensional augmented features through a reconstruction independent component analysis method. Extensive experiments on three publicly available COVID-19 X-ray image datasets using state-of-the-art network architectures including Squeez-eNet, GoogleNet, and DenseNet-201 and recently published works showed that our newly designed CNN-based BiLSTM architecture outperformed several state-of-the-art models.

Our model achieved a 97% accuracy which is 10% higher than the best performing model published so far in the literature [74] on the three classes, and 99% accuracy for the two-classes dataset. In the five class case, our model achieved 84% accuracy which is 3% better than the previously proposed method in [7]. In some other scenarios, the developed model has demonstrated the ability to achieve 100% accuracy for the detection of COVID-19 samples. On the other hand, we also showed the possibility to utilize the componentwise property of the overall architecture where each stage (component) can be used to generate explanations that can be employed to comprehend the actions of the model. Explainability through PCA and t-SNE have also been explored and duly commented as well as highlighting the potential deficiencies that may restrict the ability of PCS or t-SNE projection to provide an answer to the "why" question in the explainability, while the prospect of a newly introduced ccPCA has been recognized. In the future, we plan to further robustify the feature selection method and RICA analysis in the convolutional layer of a CNN in a way to enhance the explainability

of the results and develop joint visualization approach that can comprehend both PCA, t-SNE projection outcomes with attention weights. The source code is available at the project webpage: https://github.com/ziaul55/COVID-19-Detection

CRediT authorship contribution statement

Usman Muhammad: Conceptualization, Methodology, Software, Writing – original draft. **Md. Ziaul Hoque:** Validation, Methodology, Visualization, Writing – original draft. **Mourad Oussalah:** Supervision, Investigation, Writing – review & editing. **Anja Keskinarkaus:** Investigation, Review & editing. **Tapio Seppänen:** Investigation, Review & editing. **Pinaki Sarder:** Investigation, Writing – review & editing.

Declaration of competing interest

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

Acknowledgments

The research work of this paper was supported by the Center for Machine Vision and Signal Analysis (CMVS) in the Faculty of Information Technology and Electrical Engineering (ITEE) at University of Oulu, Finland. The authors are grateful to the Academy of Finland Profi5 DigiHealth project (#326291). We are also thankful to Samuel P Border (University at Buffalo, USA) and Jiehua Zhang (University of Oulu, Finland) for their valuable suggestions.

References

[1] Coronaviridae Study Group of the International et al., The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2, Nat. Microbiol. 5 (4) (2020) 536.
[2] CDC Covid, Response Team, CDC COVID, Response Team, CDC COVID, Response Team, S.L. Burrer, M.A. de Perio, M.M. Hughes, D.T. Kuhar, et al., Characteristics of health care personnel with COVID-19—United States, February 12–April 9, 2020, Morb. Mortal. Wkly. Rep. 69 (15) (2020) 477.
[3] S. Albahli, Efficient GAN-based Chest Radiographs (CXR) augmentation to diagnose coronavirus disease pneumonia, Int. J. Med. Sci. 17 (10) (2020) 1439.
[4] T. Singhal, A review of coronavirus disease-2019 (COVID-19), Indian J. Pediatr. 87 (4) (2020) 281–286.
[5] L. Falzone, G. Gattuso, A. Tsatsakis, D.A. Spandidos, M. Libra, Current and innovative methods for the diagnosis of COVID-19 infection, Int. J. Mol. Med. 47 (6) (2021) 1–23.
[6] X. Dong, Y.-y. Cao, X.-x. Lu, J.-j. Zhang, H. Du, Y.-q. Yan, C.A. Akdis, Y.-d. Gao, Eleven faces of coronavirus disease 2019, Allergy 75 (7) (2020) 1699–1709.
[7] E. Vantaggiato, E. Paladini, F. Bougourzi, C. Distante, A. Hadid, A. Taleb-Ahmed, Covid-19 recognition using ensemble-cnns in two new chest x-ray databases, Sensors 21 (5) (2021) 1742.
[8] T. Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, U.R. Acharya, Automated detection of COVID-19 cases using deep neural networks with X-ray images, Comput. Biol. Med. 121 (2020) 103792.
[9] A. Bernheim, X. Mei, M. Huang, Y. Yang, Z.A. Fayad, N. Zhang, K. Diao, B. Lin, X. Zhu, K. Li, et al., Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection, Radiology (2020) 200463.
[10] J.P. Cohen, P. Morrison, L. Dao, K. Roth, T.Q. Duong, M. Ghassemi, Covid-19 image data collection: Prospective predictions are the future, 2020, arXiv preprint arXiv:2006.11988.
[11] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M. Summers, Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2097–2106.
[12] W. Kong, P.P. Agarwal, Chest imaging appearance of COVID-19 infection, Radiol.: Cardiothorac. Imaging 2 (1) (2020) e200028.
[13] W. Zhao, Z. Zhong, X. Xie, Q. Yu, J. Liu, Relation between chest CT findings and clinical conditions of coronavirus disease (COVID-19) pneumonia: a multicenter study, Am. J. Roentgenol. 214 (5) (2020) 1072–1077.
[14] Z.Y. Zu, M.D. Jiang, P.P. Xu, W. Chen, Q.Q. Ni, G.M. Lu, L.J. Zhang, Coronavirus disease 2019 (COVID-19): a perspective from China, Radiology 296 (2) (2020) E15–E25.
[15] W.C. de Melo, E. Granger, A. Hadid, A deep multiscale spatiotemporal network for assessing depression from facial dynamics, IEEE Trans. Affect. Comput. (2020).
[16] M. Tavakolian, A. Hadid, A spatiotemporal convolutional neural network for automatic pain intensity estimation from facial dynamics, Int. J. Comput. Vis. 127 (10) (2019) 1413–1425.
[17] N. Chouhan, A. Khan, J.Z. Shah, M. Hussnain, M.W. Khan, Deep convolutional neural network and emotional learning based breast cancer detection using digital mammography, Comput. Biol. Med. 132 (2021) 104318.
[18] J. Feng, S.-W. Zhang, L. Chen, J. Xia, A.D.N. Initiative, et al., Alzheimer's disease classification using features extracted from nonsubsampled contourlet subband-based individual networks, Neurocomputing 421 (2021) 260–272.
[19] A. Al-Waisy, M.A. Mohammed, S. Al-Fahdawi, M. Maashi, B. Garcia-Zapirain, K.H. Abdulkareem, S. Mostafa, N.M. Kumar, D.N. Le, Covid-deepnet: hybrid multimodal deep learning system for improving covid-19 pneumonia detection in chest x-ray images, Comput. Mater. Contin. 67 (2) (2021).
[20] J.F.-W. Chan, S. Yuan, K.-H. Kok, K.K.-W. To, H. Chu, J. Yang, F. Xing, J. Liu, C.C.-Y. Yip, R.W.-S. Poon, 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 395 (10223) (2020) 514–523.
[21] M.M. Islam, F. Karray, R. Alhajj, J. Zeng, A review on deep learning techniques for the diagnosis of novel coronavirus (covid-19), IEEE Access 9 (2021) 30551–30572.
[22] Y. Fu, T.M. Hospedales, T. Xiang, S. Gong, Transductive multi-view zero-shot learning, IEEE Trans. Pattern Anal. Mach. Intell. 37 (11) (2015) 2332–2345.
[23] Z. Zhong, L. Zheng, G. Kang, S. Li, Y. Yang, Random erasing data augmentation, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, 2020, pp. 13001–13008.
[24] T. DeVries, G.W. Taylor, Improved regularization of convolutional neural networks with cutout, 2017, arXiv:1708.04552.
[25] C. Shorten, T.M. Khoshgoftaar, A survey on image data augmentation for deep learning, J. Big Data 6 (1) (2019) 1–48.
[26] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial networks, 2014, arXiv preprint arXiv:1406.2661.
[27] J. Donahue, P. Krähenbühl, T. Darrell, Adversarial feature learning, 2016, arXiv preprint arXiv:1605.09782.
[28] A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, 2015, arXiv preprint arXiv:1511.06434.
[29] T. Karras, T. Aila, S. Laine, J. Lehtinen, Progressive growing of gans for improved quality, stability, and variation, 2017, arXiv preprint arXiv:1710.10196.
[30] J.-Y. Zhu, T. Park, P. Isola, A.A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2223–2232.
[31] M.Z. Islam, M.M. Islam, A. Asraf, A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images, Inform. Med. Unlocked 20 (2020) 100412.
[32] A.I. Khan, J.L. Shah, M.M. Bhat, CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images, Comput. Methods Programs Biomed. 196 (2020) 105581.
[33] L. Wang, Z.Q. Lin, A. Wong, Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images, Sci. Rep. 10 (1) (2020) 1–12.
[34] M.F. Aslan, M.F. Unlersen, K. Sabanci, A. Durdu, CNN-based transfer learning–BiLSTM network: A novel approach for COVID-19 infection detection, Appl. Soft Comput. 98 (2021) 106912.
[35] Q. Le, A. Karpenko, J. Ngiam, A. Ng, ICA with reconstruction cost for efficient overcomplete feature learning, Adv. Neural Inf. Process. Syst. 24 (2011) 1017–1025.
[36] A. Narin, C. Kaya, Z. Pamuk, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, Pattern Anal. Appl. (2021) 1–14.
[37] S. Rajaraman, S. Antani, Weakly labeled data augmentation for deep learning: A study on covid-19 detection in chest x-rays, Diagnostics 10 (6) (2020) 358.
[38] M. Woźniak, J. Siłka, M. Wieczorek, Deep neural network correlation learning mechanism for CT brain tumor detection, Neural Comput. Appl. (2021) 1–16.
[39] F. Ozyurt, T. Tuncer, A. Subasi, An automated COVID-19 detection based on fused dynamic exemplar pyramid feature extraction and hybrid feature selection using deep learning, Comput. Biol. Med. 132 (2021) 104356.

[40] M. Rahimzadeh, A. Attar, A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of xception and ResNet-50V2, Inform. Med. Unlocked 19 (2020) 100360.
[41] I. Castiglioni, D. Ippolito, M. Interlenghi, C.B. Monti, C. Salvatore, S. Schiaffino, A. Polidori, D. Gandola, C. Messa, F. Sardanelli, Artificial intelligence applied on chest X-ray can aid in the diagnosis of COVID-19 infection: a first experience from lombardy, Italy, MedRxiv (2020).
[42] A.K. Das, S. Ghosh, S. Thunder, R. Dutta, S. Agarwal, A. Chakrabarti, Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network, Pattern Anal. Appl. (2021) 1–14.
[43] E.E.-D. Hemdan, M.A. Shouman, M.E. Karar, Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images, 2020, arXiv preprint arXiv:2003.11055.
[44] M. Berrimi, S. Hamdi, R.Y. Cherif, A. Moussaoui, M. Oussalah, M. Chabane, COVID-19 detection from Xray and CT scans using transfer learning, in: 2021 International Conference of Women in Data Science at Taif University (WiDSTaif), IEEE, 2021, pp. 1–6.
[45] M. Nour, Z. Cömert, K. Polat, A novel medical diagnosis model for COVID-19 infection detection based on deep features and Bayesian optimization, Appl. Soft Comput. 97 (2020) 106580.
[46] G. Capizzi, G.L. Sciuto, C. Napoli, D. Połap, M. Woźniak, Small lung nodules detection based on fuzzy-logic and probabilistic neural network with bioinspired reinforcement learning, IEEE Trans. Fuzzy Syst. 28 (6) (2019) 1178–1189.
[47] S.H. Yoo, H. Geng, T.L. Chiu, S.K. Yu, D.C. Cho, J. Heo, M.S. Choi, I.H. Choi, C. Cung Van, N.V. Nhung, et al., Deep learning-based decision-tree classifier for COVID-19 diagnosis from chest X-ray imaging, Front. Med. 7 (2020) 427.
[48] N. Islam, S. Ebrahimzadeh, J.-P. Salameh, S. Kazi, N. Fabiano, L. Treanor, M. Absi, Z. Hallgrimson, M.M. Leeflang, L. Hooft, et al., Thoracic imaging tests for the diagnosis of COVID-19, Cochrane Database Syst. Rev. (3) (2021).
[49] H. Mukherjee, S. Ghosh, A. Dhar, S.M. Obaidullah, K. Santosh, K. Roy, Deep neural network to detect COVID-19: one architecture for both CT Scans and Chest X-rays, Appl. Intell. 51 (5) (2021) 2777–2789.
[50] M. Abdul Salam, S. Taha, M. Ramadan, COVID-19 detection using federated machine learning, Plos One 16 (6) (2021) e0252573.
[51] H. Mukherjee, A. Dhar, S.M. Obaidullah, K. Santosh, K. Roy, COVID-19: A necessity for changes and innovations, in: COVID-19: Prediction, Decision-Making, and its Impacts, Springer, 2021, pp. 99–105.
[52] H. Mukherjee, S. Ghosh, A. Dhar, S.M. Obaidullah, K. Santosh, K. Roy, Shallow convolutional neural network for COVID-19 outbreak screening using chest X-rays, Cogn. Comput. (2021) 1–14.
[53] M. Woźniak, D. Połap, Bio-inspired methods modeled for respiratory disease detection from medical images, Swarm Evol. Comput. 41 (2018) 69–96.
[54] M. Turkoglu, COVIDetectioNet: COVID-19 diagnosis system based on X-ray images using features selected from pre-learned deep features ensemble, Appl. Intell. 51 (3) (2021) 1213–1226.
[55] A.T. Sahlol, D. Yousri, A.A. Ewees, M.A. Al-Qaness, R. Damasevicius, M. Abd Elaziz, COVID-19 image classification using deep features and fractional-order marine predators algorithm, Sci. Rep. 10 (1) (2020) 1–15.
[56] M. Toğaçar, B. Ergen, Z. Cömert, COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches, Comput. Biol. Med. 121 (2020) 103805.
[57] Y. Karbhari, A. Basu, Z.-W. Geem, G.-T. Han, R. Sarkar, Generation of synthetic chest X-ray images and detection of COVID-19: A deep learning based approach, Diagnostics 11 (5) (2021) 895.
[58] M. Loey, F. Smarandache, N.E. M Khalifa, Within the lack of chest COVID-19 X-ray dataset: a novel detection model based on GAN and deep transfer learning, Symmetry 12 (4) (2020) 651.
[59] S. Elbishlawi, M.H. Abdelpakey, M.S. Shehata, M.M. Mohamed, CORONA-Net: Diagnosing COVID-19 from X-ray images using re-initialization and classification networks, J. Imaging 7 (5) (2021) 81.
[60] V. Vasilescu, A. Neacşu, E. Chouzenoux, J.-C. Pesquet, C. Burileanu, A deep learning approach for improved segmentation of lesions related to COVID-19 chest CT scans, in: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), IEEE, 2021, pp. 635–639.
[61] Y. Wang, Y. Zhang, Y. Liu, J. Tian, C. Zhong, Z. Shi, Y. Zhang, Z. He, Does non-COVID-19 lung lesion help? investigating transferability in COVID-19 CT image segmentation, Comput. Methods Programs Biomed. 202 (2021) 106004.
[62] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[63] M. Schmidt, Minfunc (A Matlab function for unconstrained optimization of differentiable real-valued multivariate functions using line-search methods), 2005, URL http://www.cs.ubc.ca/~schmidtm/Software/minFunc.html.
[64] M. Schuster, K.K. Paliwal, Bidirectional recurrent neural networks, IEEE Trans. Signal Process. 45 (11) (1997) 2673–2681.
[65] M. Sundermeyer, R. Schlüter, H. Ney, LSTM neural networks for language modeling, in: Thirteenth Annual Conference of the International Speech Communication Association, 2012.
[66] G. Shih, C.C. Wu, S.S. Halabi, M.D. Kohli, L.M. Prevedello, T.S. Cook, A. Sharma, J.K. Amorosa, V. Arteaga, M. Galperin-Aizenberg, et al., Augmenting the national institutes of health chest radiograph dataset with expert annotations of possible pneumonia, Radiol.: Artif. Intell. 1 (1) (2019) e180041.
[67] J. Irvin, P. Rajpurkar, M. Ko, Y. Yu, S. Ciurea-Ilcus, C. Chute, H. Marklund, B. Haghgoo, R. Ball, K. Shpanskaya, et al., Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, 2019, pp. 590–597.
[68] D.S. Kermany, M. Goldbaum, W. Cai, C.C. Valentim, H. Liang, S.L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, et al., Identifying medical diagnoses and treatable diseases by image-based deep learning, Cell 172 (5) (2018) 1122–1131.
[69] P. Angelov, E. Almeida Soares, SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification, MedRxiv (2020).
[70] F. Ucar, D. Korkmaz, COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images, Med. Hypotheses 140 (2020) 109761.
[71] I.D. Apostolopoulos, T.A. Mpesiana, Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks, Phys. Eng. Sci. Med. 43 (2) (2020) 635–640.
[72] X. Li, D. Zhu, Covid-xpert: An ai powered population screening of covid-19 cases using chest radiography images, 2020.
[73] M.E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al Emadi, et al., Can AI help in screening viral and COVID-19 pneumonia? IEEE Access 8 (2020) 132665–132676.
[74] T. Ozturka, M. Talob, E.A. Yildirimc, U.B. Baloglud, O. Yildirime, U. Acharyafgh, Automated detection of COVID-19 cases using deep neural networks with X-ray images, Comput. Biol. Med. 121 (103792) (2020) 10–1016.
[75] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1026–1034.
[76] F.N. Iandola, S. Han, M.W. Moskewicz, K. Ashraf, W.J. Dally, K. Keutzer, SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size, 2016, arXiv preprint arXiv:1602.07360.
[77] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.
[78] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.
[79] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, M. Kaur, Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning, J. Biomol. Struct. Dyn. (2020) 1–8.
[80] M.R. Ibrahim, S.M. Youssef, K.M. Fathalla, Abnormality detection and intelligent severity assessment of human chest computed tomography scans using deep learning: a case study on SARS-COV-2 assessment, J. Ambient Intell. Humaniz. Comput. (2021) 1–24.
[81] I. Jolliffe, Principal component analysis, in: Encyclopedia of Statistics in Behavioral Science, Wiley Online Library, 2005.
[82] D.J. Bartholomew, Spearman and the origin and development of factor analysis, Br. J. Math. Stat. Psychol. 48 (2) (1995) 211–220.
[83] I.T. Jolliffe, Principal components in regression analysis, in: Principal Component Analysis, Springer, 2002, pp. 167–198.
[84] L. Van der Maaten, G. Hinton, Visualizing data using t-SNE, J. Mach. Learn. Res. 9 (11) (2008).
[85] M.M. Abbas, M. Aupetit, M. Sedlmair, H. Bensmail, Clustme: A visual quality measure for ranking monochrome scatterplots based on cluster patterns, Comput. Graph. Forum 38 (3) (2019) 225–236.
[86] T. Schreck, T. Von Landesberger, S. Bremm, Techniques for precision-based visual analysis of projected data, Inf. Vis. 9 (3) (2010) 181–193.
[87] T. Fujiwara, O.-H. Kwon, K.-L. Ma, Supporting analysis of dimensionality reduction results with contrastive learning, IEEE Trans. Vis. Comput. Graphics 26 (1) (2019) 45–55.