Health and Technology
https://doi.org/10.1007/s12553-024-00879-y

ORIGINAL PAPER

Unsupervised generative learning-based decision-making system for COVID-19 detection

Neeraj Menon[1] · Pooja Yadav[1] · Vinayakumar Ravi[2] · Vasundhara Acharya[3] · V Sowmya[4]

Received: 13 September 2023 / Accepted: 30 April 2024

Abstract

Purpose The study aims to develop an unsupervised framework using COVGANs to learn better visual representations of COVID-19 from unlabeled X-ray and CT scans.

Methods We trained multiple-layer GANs to develop the COV-GAN framework on unlabeled X-ray and CT scans. We evaluated the quality of the learned representations using t-SNE visualization, K-means, and GMM clustering. The proposed unsupervised method's performance was compared with leading unsupervised methods for COVID-19 classification on X-ray and CT scans.

Results Our method achieved an accuracy of 75.1% on X-ray scans and 75.7% on CT scans, which is at least 13.9% and 12.3% higher than the leading unsupervised methods for COVID-19 classification on X-ray and CT scans, respectively. The t-SNE visualization, K-means, and GMM clustering showed that our method learned better visual representations of COVID-19 from unlabeled data.

Conclusions Our unsupervised framework using COV-GANs can learn better visual representations of COVID-19 from unlabeled X-ray and CT scans. The learned representations can improve the performance of COVID-19 classification. The outcomes show the potential of unsupervised learning methods to overcome the dearth of labelled data in the medical profession, particularly in times of public health crises like the COVID-19 epidemic.

Keywords COVID-19 · Unsupervised representation learning · Generative adversarial networks · Clustering

1 Introduction

In 2019, the world witnessed the emergence of a novel coronavirus disease, commonly referred to as COVID-19. It originated in the Wuhan city of China in 2019 and then turned into a pandemic by February 2020, reaching 229 million confirmed cases and 4.7 million deaths in more than 200 countries as of September 19th, 2021. On March 11, 2020, the World Health Organization (WHO) declared the outbreak of the novel coronavirus (COVID-19) a global pandemic [1]. The United States, India, Brazil, and the United Kingdom were the nations most affected by COVID-19 as of October 31, 2021. India had the second-highest number of confirmed cases, at 34,246,157, behind the United States [2]. The World Health Organization announced that as of April 2024, there were 775,335,916 confirmed COVID-19 cases worldwide, an increase of 40,529 cases in the preceding week [3]. The pandemic has had a catastrophic effect on people's health, life, and the global economy. As the number of cases is rising rapidly, early detection of positive cases is essential to help faster recovery and release unprecedented pressure on the healthcare systems. Real-time polymerase chain reaction (RT-PCR) is the technique for diagnosing COVID-19. This diagnostic procedure uses swabs to obtain samples from the throat and nasopharynx in order to detect viral nucleic acid. However, the unavailability, instability,

Neeraj Menon and Pooja Yadav contributed equally to this work.

✉ Vinayakumar Ravi
vinayakumarr77@gmail.com

V Sowmya
v_sowmya@cb.amrita.edu

1 VMware, Bangalore, India

2 Center for Artificial Intelligence, Prince Mohammad Bin Fahd University, Khobar, Saudi Arabia

3 Manipal Institute of Technology (MIT), Manipal Academy of Higher Education (MAHE), Manipal, India

4 Center for Computational Engineering and Networking (CEN), Amrita Vishwa Vidyapeetham, Coimbatore, India

Published online: 14 May 2024

and low reproducibility of RT-PCR testing may delay identifying potential patients [4]. Alternatively, medical imaging such as X-ray and CT scan, even though it cannot fully substitute the RT-PCR, is essential to complement it as one of the early diagnostics for COVID-19 cases and to determine the priority of patients' treatments. However, the manual screening of COVID-19 using CT scans and X-rays is laborious, time-consuming, and prone to variation among different observers and even within a single observer [5]. Moreover, considering the increased workload for radiologist departments, the inability of radiologists in remote villages to identify COVID-19 infection from medical images, and the necessity for high-output screening from CT scan and X-ray, automatic screening of COVID-19 is critically required in clinical practice.

Accordingly, several deep learning (DL) models for COVID-19 classification from X-ray and CT images have been actively explored [6–12]. Some approaches that use Convolutional Neural Networks (CNNs) have performed remarkably in medical imaging [6–10], but these methods usually require a vast amount of labeled data or fine-tuning of millions of parameters.

However, as noted in prior research like [13, 14], the amount and variety of data that may be used to train supervised learning models in the field of medical imaging is constrained. During public health catastrophes like the one we're currently experiencing, this constraint becomes even more obvious. Unsupervised representational learning algorithms, however, provide a benefit since they only use unlabeled data, obviating the necessity for big labelled datasets.

Generative Adversarial Networks (GANs) [15] is one such promising unsupervised method that generates synthetic images using a two-player minimax game and, as a result, extends the training dataset artificially.

An unsupervised generative adversarial networks (COV-GANs) model is proposed in this work to learn CT and X-ray image representations using only unlabeled data. The model would, though not replace, complement the existing COVID-19 methodologies for an automated, faster, and more accurate diagnosis.

Our proposed network architecture is an improved method of MARTA GANs [16] and LungGAN [17]. We demonstrate that COV-GANs produce higher-quality samples of X-ray and CT images and provide better classification accuracies, compared to the existing MARTA GAN on the same dataset.

With the aid of novel methods and advancements in the field of unsupervised representational learning shown in this paper, it is possible to quickly and accurately diagnose COVID-19 from X-ray and CT scan images. The following are the main innovations and contributions of this paper:

- Introduction of a novel unsupervised representational learning method created specifically for leveraging CT and X-ray images to quickly and accurately diagnose COVID-19.
- Two clustering methods namely K-means and Gaussian Mixture Models (GMM) are used to assess the caliber of the representations that the suggested model learns.
  • Extensive analyses and tests were performed on a large-scale publicly available Xray and CT scan dataset in order to fully assess and evaluate the efficacy of the suggested strategy.
- Comparative analysis against three other existing unsupervised learning approaches, demonstrating superior performance and notable improvements over the alternative methods.
- Evaluation of the proposed model's effectiveness at discriminating between"COVID" and"NON-COVID" examples and visualization of the features that were learned.

Dataset and code are accessible at https://github.com/yadavpa1/COV-GAN.

2 Literature survey

Recently, numerous studies have used DL-based methods to detect COVID-19 from chest X-ray images. For instance, Khan et al. [18] proposed CoroNet, a CNN-based model for COVID-19 screening with X-ray. Oh et al. [19] developed a patch-based CNN model to diagnose COVID-19 from X-rays. Final decisions are made by majority vote among many patches located throughout the lungs. They asserted that compared to other methods that need a bigger dataset for consistent training, their method produces superior sensitivity for COVID-19 situations.

Several automatic lesion segmentation methods for COVID-19 CT have also been studied. Wang et al. [20] developed a weakly supervised DL framework for classifying COVID-19 and localizing lesions using 3D CT volumes. To obtain better lesion localization results, they combined the activation regions produced by a 3D deep CNN (DeCoVNet) and an existing unsupervised lung segmentation method. Another weakly supervised lesion segmentation framework was proposed by Xu et al. [21], called GASNet, by embedding the process of GAN training into a segmentation network.

For COVID-19 classification from X-rays, Zhou et al. [22] applied the domain adaptation approach with semi-supervised learning. Their method overcomes the downsides of fine-tuning based transfer learning methods. Calderon-Ramirez et al. [23] analyzed the impact of data imbalance, proposed a pseudo-label-based balance correction approach into MixMatch [24], and improved the classification accuracy w.r.t the non-balanced MixMatch.

Alshazly et al. [25] employed advanced deep CNN models to analyze CT-scan images for COVID-19 detection, achieving an accuracy of 93.96%, a precision of 99.13%, and a sensitivity of 94%. Narin et al. [26] developed a method to detect COVID-19 from X-ray images using pre-trained CNNs, specifically ResNet50. They experimented on a dataset of 100 images (50 COVID-19, 50 healthy), achieving 98% accuracy through cross-validation. Born et al. [27] introduced a NasNetMobile model that identified COVID-19 from lung ultrasound images, attaining sensitivity and specificity rates of 81% and 96%. Wang et al. [28] analyzed 1065 CT images from confirmed COVID-19 cases and those with typical viral pneumonia using a modified inception transfer-learning model. This model underwent internal and external validation, achieving accuracies of 89.5% and 79.3%, respectively.

Kim et al. [29] developed a diagnostic model integrated into an automated triage system for assessing COVID-19 pneumonia severity and progression using chest radiographs. They utilized over 6,500 non-public images from the DICOM Image Analysis and Archive. The model was validated against internally held-out and external test sets, achieving a 95% diagnostic accuracy on 80 prospective radiographs.

Gour and Jain et al. [30] designed a new stacked convolutional neural network model that automatically diagnosed COVID-19 from chest X-ray and CT images by utilizing sub-models derived from VGG19 and Xception models. The model demonstrated a sensitivity of 97.62% for classifying X-ray images into COVID-19, Normal, and Pneumonia categories, and 98.31% sensitivity for binary classification of CT images into COVID-19 and No-Finding classes. Bermejo-Peláez et al. [31] developed a deep learning approach to automatically segment lesion subtypes in nonenhanced CT scans and used this quantification to predict clinical outcomes in a multicentric cohort of 103 patients. The technique was assessed using segmentation accuracy metrics and predictive performance metrics showing strong agreement with radiologists' severity scoring. The model effectively predicted mortality, ICU admission, and the need for mechanical ventilation with AUCs of 0.87, 0.73, and 0.68, respectively.

Hardy-Werbin et al. [32] developed a multimodal prediction algorithm using opensource DL algorithms to differentiate between COVID-19, heart failure, non-COVID pneumonia, and healthy control patients. Analyzing 8578 samples from 6123 patients, their algorithm achieved an overall accuracy of 84% with a mean AUC of 0.92 on the entire test set and outperformed radiologists' accuracy on 300 random test images, scoring 69.6%. Di Napoli et al. [33] developed a deep-learning model using 3D chest CT images and clinical data from 1051 COVID-19 patients to predict mortality, ICU admission, and intubation outcomes. They utilized a 3D residual neural network and achieved high accuracy and ROC-AUC scores in internal validation, with lower performance in external validation. The integration of laboratory data improved sensitivity and specificity for predicting severe outcomes. Oi et al. [34] developed a CT scoring method from axial images that was divided into three sections and rated from 0 to 5 to predict severe outcomes in COVID pneumonia patients. Of the 71 patients analyzed, the CT score at admission accurately predicted mortality or the need for ECMO with an ROC of 0.718. To overcome the challenges of acquiring large annotated CT datasets, the authors in [35] developed a three-level optimization method. They enhanced the utility of CT data from a source domain for use in a target domain by automatically downweighting low-quality source examples with significant domain discrepancies. Their approach yielded F1 scores of 91.8% and 92.4% for detecting pneumonia and other types on a dataset of 2218 target CT scans. Miyazaki et al. [36] evaluated an EfficientNet-based DL model's diagnostic performance on COVID-19 using 26,393 chest X-rays (CXRs) from public and private datasets and investigated its impact on radiologists' performance. In two reading sessions, radiologists' diagnostic accuracy improved from 0.696 to significantly higher AUC values when aided by the DL model. The DL model demonstrated superior diagnostic performance and improved the radiologists' ability to distinguish COVID-19 from normal and non-COVID-19 pneumonia cases. Abad et al. [37] conducted a detailed performance analysis of ResNet50, DenseNet121, and Inception-ResNet-v2 on a large dataset of chest X-ray images containing both positive and negative COVID-19 cases. The results showed significant reductions in accuracy across the models upon external validation, with DenseNet121 achieving the highest internal accuracy of 96.71% and Inception-ResNet-v2 reaching 76.70% externally. The ensemble method using uncertainty-based weighting significantly enhanced diagnostic accuracy by achieving 97.38% in internal validation and 81.18% in external validation. Sadeghi et al. [38] proposed multi-head attention squeeze and excitation residual (MASERes) neural network for patient-level analysis of entire CT scans, achieving a 100% accuracy in diagnosing COVID-19, while slide-level transfer learning models detected COVID-19 with over 99% accuracy.

Having achieved success in COVID-19 detection, the major drawback of the aforementioned approaches is that they depend on the availability of annotated data. On the other hand, unsupervised methods have the advantage of not requiring labeled data to train the networks. King et al. [39] presented an unsupervised method called SOFM to cluster COVID-19 X-ray images and extract the features that cause clustering. The Variational Autoencoder (VAE) based approach proposed by Mansour et al. [40] uses adaptive Wiener filtering (AWF) in the preprocessing step to

enhance the image quality. Further, the feature vectors extracted using Inception v4 with Adagrad are used by the unsupervised VAE model for COVID-19 classification. Rashid et al. [41] also use Autoencoders in a 2-stage training approach. Firstly, a deep convolutional autoencoder network is used for feature extraction from the X-ray images. The spatial characteristics are captured by the encoder through this unsupervised learning technique. In the next stage, a merging technique uses the features extracted from the different levels of the encoder for X-ray classification using a small dataset.

GANs have been one of the widely used unsupervised methods. Alizadehsani et al. [42] used GANs to initialize the parameters of a CNN which is then fine-tuned with limited labeled data. Another similar study proposed by Waheed et al. [43] employed an auxiliary classifier GAN-based model (CovidGAN) to generate synthetic X-ray images to enhance a CNN's performance on COVID-19 detection. Similarly, Loey et al. [44] used GANs to augment the dataset to prevent overfitting while training Convolutional Neural Networks.

The literature survey clearly shows the need for unsupervised learning methods during the outbreak of diseases like COVID -19 as supervised learning might not perform well with limited data. In the literature, it is observed that unsupervised learning techniques are used for learning spatial representations. Further, these features are used for classification using supervised learning. Other researchers have used unsupervised learning to generate synthetic data to improve performance of CNN. There has not been much research on employing clustering algorithms on these feature representations. Our research evaluates the quality of the feature representations using clustering algorithms like K-means and GMM.

The remainder of this paper is organized as follows: Section 3 provides a comprehensive explanation of the proposed network architecture, detailing its key components and design choices. Section 4 outlines the dataset utilized in this study, describing its characteristics, size, and any preprocessing steps employed. In Section 5, the results of this research are analyzed and discussed. This section presents a thorough examination of the experimental findings, including quantitative metrics, visualizations, and comparisons with existing methods. Finally, the conclusion of the paper is presented at the end, summarizing the main contributions and discussing the implications of the results.

3 Proposed methodology

When both models are multi-layer perceptrons, GANs are the most simple to implement. However, because the dataset for this study includes intricate X-ray and CT images, we will need a deeper network to identify patterns accurately in these images and generate new ones. Therefore, we have used CNNs for both discriminator and generator. The generator network is trained to produce samples $x = G_r(a;\theta^{(Gr)})$ from the data distribution, $p_{data}(x)$, with parameters $\theta^{(Gr)}$, by transforming vectors a which obey a prior noise distribution $p_a(a)$. At the same time, the discriminator network is trained to differentiate samples created by the generator from real data. $D_r(x;\theta^{(Dr)})$ denotes the probability of x being a real rather than a fake training sample. The discriminator model D_r is then used as a feature extractor for training a classifier. The generative model provides the discriminative model with additional training data and thus helps it learn better image representations.

Algorithm 1 COVID-19 Classification

Input: A set of X-ray or CT scan images img_1, img_2,., img_n. **Output:** Labels y_1, y_2,..., y_n.

```
while number_of_epochs ≥ 1 do
    imgs ← get images(batch size);
    imgs ← preprocess images(imgs);
    train discriminator(imgs); train generator(imgs);
    end while
for number of batches do
    imgs ← get images(batch size);
    imgs ← preprocess images(imgs);
    f ← extract features from discriminator(imgs);
    features ← append(f);
end for
features_train ← load features train();
labels train ← load labels train ();
features test ← load features test();
labels test ← load test labels();
train GMM classifier(features train, labels train);
pred labels ← test GMM classifier(features test);
compute model score(pred labels, labels test);
```

The objective of the discriminator (D_r) in GANs is to differentiate between real training examples and fake images generated by the generator (G_r). The discriminator is trained to achieve two goals:

1. Minimize $D_r(x)$ for fake data from the generator:
2. Maximize $D_r(x)$ for real training data

The generator G_r is simultaneously trained to produce images that cannot be differentiated from the real training images by D_r. In other words, G_r tries to maximize $D_r(G_r(a))$ or equivalently minimize $\log(1 - D_r(G_r(a)))$. We call it perceptual loss which is defined as follows:

$$lperceptual = \mathrm{E}a \sim pa(a)\left[\log(1 - Dr(Gr(a)))\right] \quad (1)$$

In a nutshell, the objective of a GAN is:

$$\min\max V\left(D_r, G_r\right) = \mathrm{E}_{x\sim p_{data}}(x)\log D_r(x) + G_r \rightarrow D_r$$

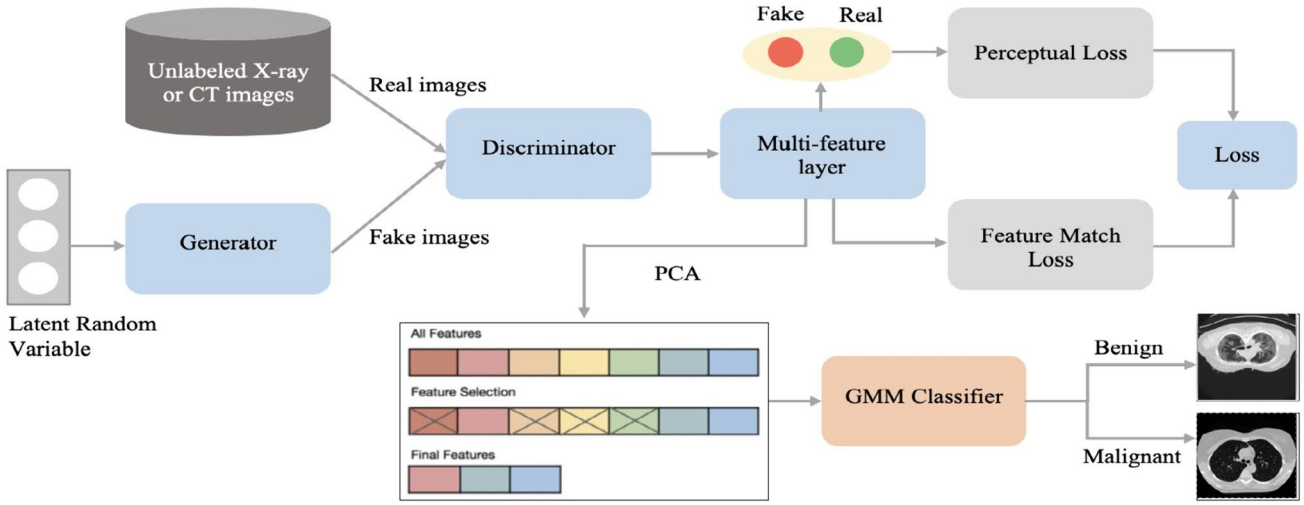


Fig. 1 Proposed end-to-end system using COV-GAN

$$\mathrm{E}_{a\sim pa_{(a)}}\left[\log\left(1-D_r\left(G_r(a)\right)\right)\right] \tag{2}$$

We train the generator to resemble the expected features on the multi-feature layer depicted in textbf Fig. 2b in order for it to output images that are identical to the training data. The generator's feature-matching loss can be described as follows:

$$l_{feature\_matching} = \| \mathbb{E}_{x\sim p_{data}(x)} f(x) - \mathbb{E}_{a\sim p_a(a)} f\left(G_r(a)\right) \|_2^2 \tag{3}$$

Here, $f(x)$ denotes activations on the multi-feature layer of D_r.

Thus, the overall objective of G_r is to minimize the sum of perceptual and featurematching loss i.e.

$$l_{final} = l_{perceptual} + l_{feature\_matching} \tag{4}$$

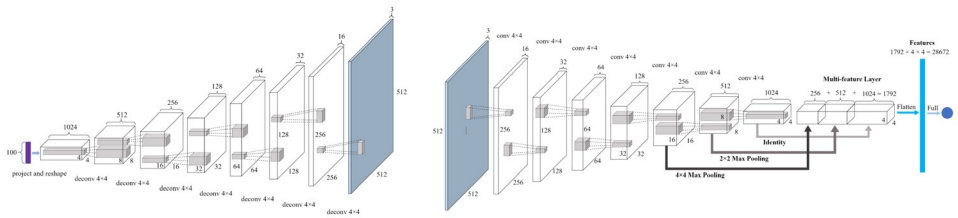
3.1 Network architecture

3.1.1 Feature extraction

We suggest an improved version of Lin et al. [16] generator and discriminator architectures, which have demonstrated promising results in the acquisition of multi-scale spatial information for high-resolution images. The generator receives a 100-dimensional, uniformly distributed noise vector a and transforms it into a 4D tensor. We then used a series of seven fractionally-strided convolutions (also called deconvolutions) in our generator to transform this high-level representation into an image of 512×512 resolution. Pooling layers were not used. Figure 2a shows a visualization of the generator.

The inputs to the discriminator are the real and synthesized images having 512×512 pixel resolution. We only used strided convolutions to allow the discriminator to learn its own spatial down sampling.

In the network architecture, the last third convolutional layer undergoes a max pooling operation with a size of 4*times*4, and the last second convolutional layer undergoes a second max pooling operation with a size of 2*times*2. The last layer is then subjected to the identity function, resulting in the creation of feature maps with a spatial size of 4*times*4.A multi-feature layer is created by concatenating these feature maps. A sigmoid output node is then fed with the flattened multi-feature layer. In addition to



(a) generator. (b) discriminator.

Fig. 2 COV-GAN's network architecture. **a** A generator (G_r) creates a 515×512 RGB image from a 100-dimensional uniformly distributed noise vector (a). **b** A discriminator (D_r) learns to distinguish between real and fake images by extracting hidden features from the input image (512×512)

being used for classification, the features recovered from the flattened multi-feature layer are also used to calculate the feature-matching loss (also known as the 3), which measures how similar the features retrieved from real and fake images are.

We set the stride and kernel size to 2 and 4 respectively in all deconvolutional layers to avoid the overlaps that cause checkerboard artifacts [45]. In the generator, we apply batch normalization and the ReLU activation to all the deconvolutional layers (except the output layer). he network's output layer employs the hyperbolic tangent activation function. Additionally, all convolutional layers other than the first layer employ batch normalisation. The LeakyReLU activation function with a slope of 0.2 is the activation function employed for the convolutional layers. Additionally, a decay factor of 0.9 is used.

3.1.2 Classification

The assumption made by GMMs, which are probabilistic models, is that all of the data points came from a finite number of Gaussian or normal distributions. They can model any data set that can be clustered into multiple normal distributions. In addition, they can estimate the probability of a new data point belonging to each cluster. GMMs consist of two parts - mean (μ) and covariance (Σ). For instance, we have three normal distributions - D1, D2, and D3.

Each of these distributions (represented by μ_1, μ_2, μ_3 for means and σ_1, σ_2, σ_3 for variances) corresponds to a specific set of characteristics in the given dataset. By utilizing Gaussian Mixture Models (GMM), we can determine the probability of each data point belonging to each of these distributions.

GMMs, in general, do not require knowing which cluster a data point belongs to, thereby facilitating the model to learn the groups automatically. As the cluster assignment is unknown, it can be considered a form of unsupervised learning. Moreover, they are also comparatively robust to outliers, making them powerful for clustering data.

The end-to-end proposed system is depicted in Fig. 1. Algorithm 1 details the proposed approach in a step-by-step manner. The algorithm generates labels for each image in a set that corresponds to the input of a set of chest X-ray or CT scan images (Fig. 2).

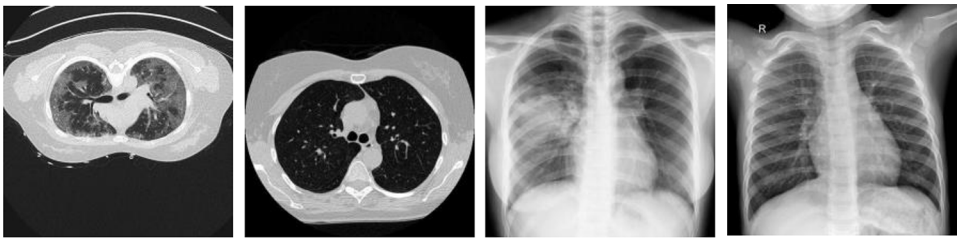
4 Description of dataset

The dataset used in this paper consists of 17,099 augmented images, including both CT and X-ray [46]. More precisely, 7,555 of them are CT scans and the rest are Xray images. Amongst X-ray images, there are 4,044 COVID and 5,500 Non-COVID images. Among CT-scans, 5,427 are images of COVID and 2,628 of Non-COVID. Figure 3 illustrates some of them. The images utilised in the study were of different sizes, colours, and pixel values, ranging from 0 to 255. The pixel intensities were normalised to the range [-1, 1] to ensure consistency. The images were also downscaled to 512*times*512 pixels in resolution and converted to the RGB format. The ratio used for the train-test split was 70:30, with 70% of the images being used for training and 30% being used for testing.

5 Results & discussion

We evaluate the efficacy of COV-GANs on CT scan and X-ray datasets against two popular clustering algorithms, K-means and GMM.We implemented COV-GANs using the TensorLayer library, which is a deep learning and reinforcement learning (RL) library built on top of Google TensorFlow (https://tensorlayer.readthedocs.io/en/latest/). Both the generator and discriminator models were trained by SGD algorithm with a *batch size* = 64, *learning rate* = 0.0002 and *momentum* = 0.5. Adam optimizer was used. An Nvidia Tesla K80 graphics processing unit (GPU) with 12 GB of RAM was used for the training procedure. A total of 2,418 CT scans and 2,518 X-ray images were held back for testing in all tests.

On X-ray and CT-scan images, classification can be accomplished with the pretrained models. The proposed architecture can also be trained on datasets not used in our study, which would first require training the generator and



(a) COVID CT scan (b) Non-COVID CT scan (c) COVID X-ray (d) Non-COVID X - ray

Fig. 3 Class-wise images from X-ray and CT dataset [46]

Table 1 Parameters for k-means and GMM clustering using COV-GAN features

Clustering Methods	CT-scan dataset		X-ray dataset	
	tolerance	max iterations	tolerance	max iterations
K-means	$9e-02$	10,000	$3e-02$	5,000
GMM	$7e-03$	5,000	$2e-02$	5,000

the discriminator. The cost of implementation implicitly depends on the computing infrastructure. In our case, the Nvidia Tesla K80 graphical processing unit (GPU) having 12 GB of memory was used, and for every 2500 images in each dataset, training the generator and discriminator took around 2 h. Performing Principal Component analysis on the extracted features from the X-ray and CT-scan datasets took 650.87 and 505.66 s, respectively. Fitting and predicting the features after the dimensionality reduction using GMM took less than 3 s for both datasets.

5.1 Performance analysis

We verified the relevance of the features learned by the discriminator in the multifeature layer shown in Fig. 2b. Since a 28,672 dimension vector is too large for clustering, we reduced it by applying Principal Component Analysis (PCA) before passing it as an input to K-means and GMM. To perform GMM on the COV-GANs model on both datasets, we passed 0.94 as a parameter to the PCA, which means that PCA will hold 94% of the variance and the number of components required to capture 94% variance will be used. Similarly, for K-means, we used a variance of 0.89 for the CT scan dataset and 0.84 for the X-ray dataset. We set *n clusters* and *n components* as 2 for K-means and GMM respectively. Table 1 details tolerance (*tol*) and maximum iteration (*max iter*) parameter values used for performing K-means and GMM algorithms on the features obtained from COV-GANs for CT and X-ray dataset.

Table 2 Performance comparison of proposed COV-GAN and MARTA-GAN

Networks	CT-scan dataset		X-ray dataset	
	k-means	GMM	k-means	GMM
MARTA GAN [16]	66.4%	75.0%	55.4%	66.1%
COV-GAN	73.4%	**75.7%**	66.6%	**75.1%**
DEC [47]	59.5%		58.7%	
IDEC [48]	63.4%		61.2%	

The findings from two GAN architectures, MARTA GAN and COV-GAN, on the CT scan and X-ray datasets are summarized in Table 2. COV-GANs achieved a clustering accuracy of 66.6% using K-means and 75.1% using GMM on the X-ray dataset, which is approximately 11% and 9% better than MARTA GANs respectively. On the CT dataset, MARTA GANs reached a clustering accuracy of 66.4% on K-means and 75.0% on GMM that increased to 73.4% and 75.7% respectively in the case of COV-GANs. The suggested network can produce chest X-ray or CT scan images with a resolution of 512*times*512 by adding an additional transposed convolutional layer, but MARTA GAN is only capable of producing images with a maximum resolution of 256*times*256. The outcomes show that the suggested model's additional layer makes it easier to generate significant features, improving its overall performance.

The confusion matrices of classification results from GMM for CT and X-ray dataset are plotted in Fig. 4. The confusion matrix in Fig. 4a shows that GMM classifier trained on COV-GAN features can predict 1,042 out of 1,629 covid CT images accurately. The number reduces to 1,025 when it is trained on the features obtained from MARTA GAN as illustrated in Fig. 4b. Similarly, Fig. 4c depicts that GMM classifier correctly predicts 755 out of 1,128 covid X-ray images on COV-GAN features compared to 435 as shown in Fig. 4d on MARTA GAN features. Moreover, the

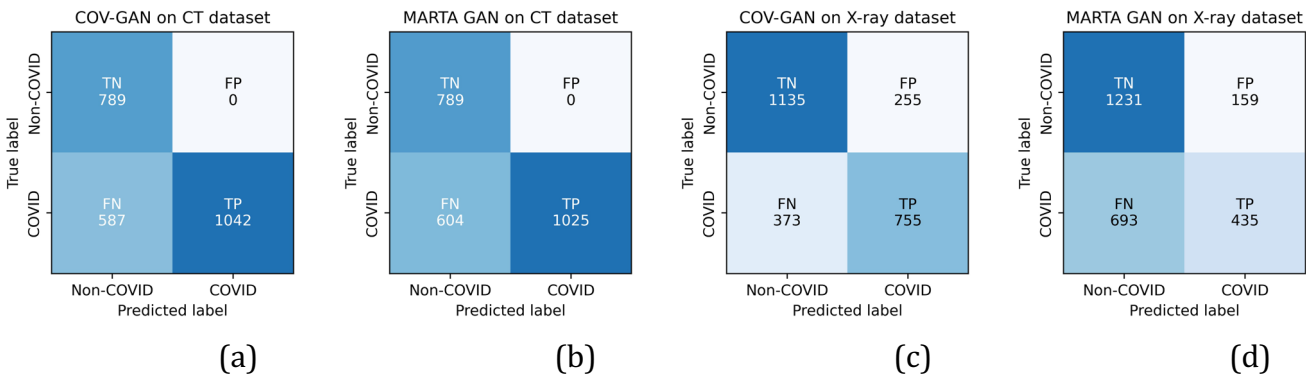
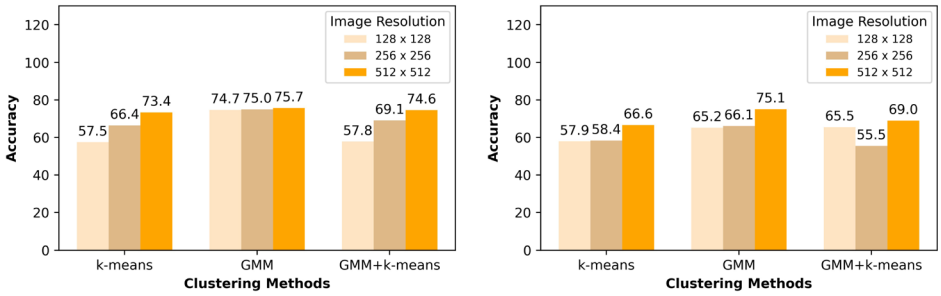


Fig. 4 Confusion matrices for COVID-19 detection with GMM



(a) CT scan. (b) X-ray.

Fig. 5 Clustering results with different image resolutions of CT scan and X-ray images

number of false negatives (i.e. the number of times a covid patient is falsely predicted as non-covid) reduces by approx. 3% on CT and 86% on X-ray dataset from COV-GAN to MARTA GAN.

We performed several clustering experiments with different input image resolutions of CT and X-ray images. Figure 5a, b illustrates the results obtained from three distinct clustering methods mainly, K-means, GMM and GMM + K-means (i.e. GMM initialized with K-means) for CT scan dataset and X-ray dataset, respectively. We observed that the clustering accuracies are proportional to the resolution of input images for all three methods.

Figure 6a, b respectively show the ROC curves for the CT scan and X-ray datasets.

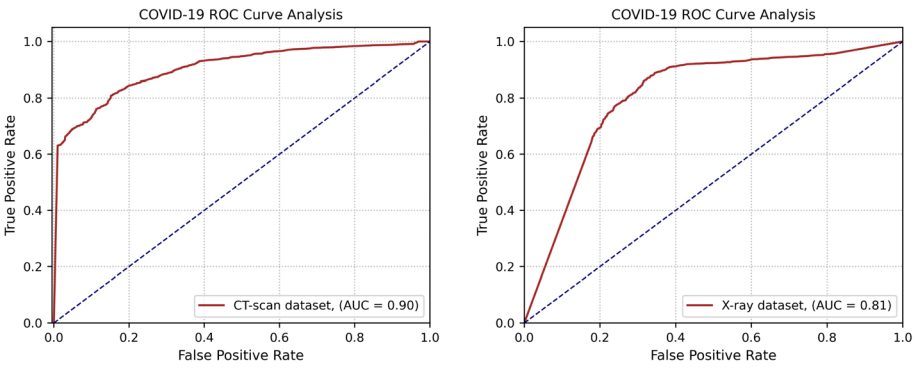
The performance of the classifier is visually represented by these curves. Performance is better indicated by a curve that is closer to the top-left corner.

Based on that finding, Fig. 6a shows that the characteristics learnt by COV-GANs on CT scan images are more effective than those learned on X-rays. The CT scan dataset's Area under the ROC curve (AUC) is 0.90, while the X-ray dataset's AUC is 0.81. Better model performance is indicated by a higher AUC value.

5.2 Comparison with different networks

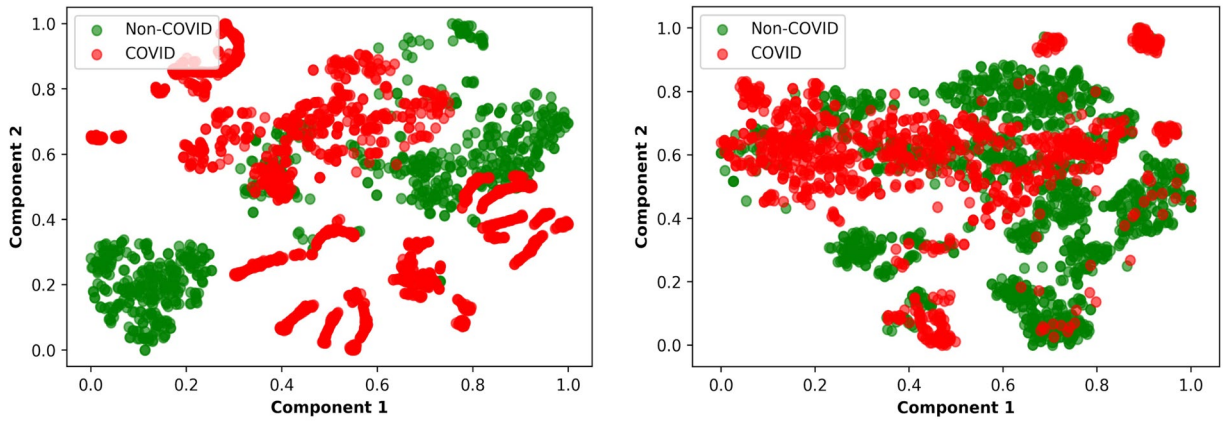
We compared our proposed COV-GANs with two alternative unsupervised techniques already in use. Our method yields the highest clustering accuracy amongst all of them. According to a study done on CT scan and X-ray datasets, our suggested approach performs better than IDEC (Improved Deep Embedded Clustering) algorithm [48]. In terms of clustering accuracy, our technique outperforms IDEC by 12.3% on the CT scan dataset and 13.9% on the X-ray dataset.

In addition, it produces better results compared to DCEC [47] (a more effective convolutional neural network-based unsupervised deep clustering algorithm). Table 2 shows a quantitative comparison of MARTA-GANs, COV-GANs, DEC and IDEC on the X-ray and CT dataset. It can be observed that our model achieved consistently better performances on both datasets. GANs have shown excellent results on computer vision datasets. We have integrated the



(a) CT-scan. (b) X-ray.

Fig. 6 COVID-19 ROC curve analysis for COV-GANs on CT and X-ray dataset



(a) CT-scan. (b) X-ray.

Fig. 7 t-SNE visualization for COV-GANs on CT and X-ray dataset

feature matching methodology to increase the efficacy of our method. This method involves training the generator to match the anticipated feature values on an intermediate layer of the discriminator, as opposed to only concentrating on increasing the output of the discriminator. This strategy has led to higher stability during GAN training and, ultimately, superior performance. The feature matching strategy has demonstrated to be a beneficial improvement, enhancing the efficacy and dependability of our system.

5.3 Visualization using t-SNE

In order to enhance the visualization of CT and X-ray image representations stored in the features learned by COV-GANs, we used the statistical method known as t-SNE [49]. After computing the high-dimensional characteristics, we mapped them into a 2D plane using t-SNE. We used a perplexity of 30, a learning rate of 200 and an early exaggeration of 12 for 100 gradient descent iterations for the visualization presented in Fig. 7. The visualization in Fig. 7a, b reveals that the features or identified patterns in the X-ray dataset are more complex than those in the CT scan dataset. This justifies the better accuracies of COV-GANs on the CT scan dataset when compared to the X-ray dataset.

6 Conclusion

The automation of COVID-19 patient triage utilising X-ray and CT scans can expedite testing and aid in the reduction of the ongoing coronavirus (COVID-19) pandemic burden. As a result, several DL-based approaches for COVID-19 diagnosis have been proposed. However, DL models generally require vast amounts of labeled data to provide better generalization. The collection of such amounts of labeled data is burdensome, time-consuming, and requires experts. We tackle this problem by presenting an unsupervised COV-GAN architecture that learns meaningful CT and X-ray image representations using only unlabelled data. Our model generates high-quality synthetic X-ray and CT images of COVID-19 patients and extends the available dataset to improve the classification performance of the model. Experimental analysis on a 17,099 image dataset demonstrates that COV-GANs outputs significantly improved accuracies compared with the current leading unsupervised methods for COVID-19 classification. Despite that, COV-GAN does not aim to replace the existing COVID-19 testing methodologies but complements them in faster and more accurate diagnosis. In our future work, our primary focus will be on enhancing the performance and classification accuracy of COV-GAN by incorporating semi-supervised approaches.

Author contributions Neeraj Menon: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing, Validation. Pooja Yadav: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing, Validation. Vinayakumar Ravi: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing, Validation, Supervision. Vasundhara Acharya: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing, Validation. Sowmya V: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing, Validation.

Funding None.

Availability of data and material The data that support the findings of this study are available from the first author upon reasonable request.

Code availability The code is available from the first author upon reasonable request.

Declarations

Informed consent None.

Conflict of interest The authors declare no conflict of interest.

References

1. World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19. January 30 2022.
2. WHO Coronavirus (COVID-19) Dashboard. http://covid19.who.int/?gclid=CjwKCAjw8KmLBhB8EiwAQbqNoAQF669E29xtQxZeOrBjBOEg0WZ3X_2OgH4h32GnFPEmy8bqhY3nPBoCkfMQAvD_BwE. Accessed 27 Apr 2024.
3. WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int/. Accessed 27 Apr 2024.
4. Wang S, Kang B, Ma J, Zeng X, Xiao M, Guo J, Cai M, Yang J, Li Y, Meng X, et al. A deep learning algorithm using ct images to screen for corona virus disease (covid-19). Eur Radiol. 2021;1–9.
5. Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, He K, Shi Y, Shen D. Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19. IEEE Rev Biomed Eng. 2020.
6. Pham TD. A comprehensive study on classification of covid-19 on computed tomography with pretrained convolutional neural networks. Sci Rep. 2020;10(1):1–8.
7. Wang L, Lin ZQ, Wong A. Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. Sci Rep. 2020;10(1):1–12.
8. Duran-Lopez L, Dominguez-Morales JP, Corral-Jaime J, Vicente-Diaz S, Linares-Barranco A. Covid-xnet: a custom deep learning system to diagnose and locate covid-19 in chest x-ray images. Appl Sci. 2020;10(16):5683.
9. Apostolopoulos ID, Aznaouridis SI, Tzani MA. Extracting possibly representative covid-19 biomarkers from x-ray images with deep learning approach and image data related to pulmonary diseases. J Med Biol Eng. 2020;40:462–9. https://doi.org/10.1007/s40846-020-00529-4.
10. Farooq M, Hafeez A. Covid-resnet: a deep learning framework for screening of covid19 from radiographs. arXiv preprint arXiv:2003.14395. 2020.
11. Afshar P, Heidarian S, Naderkhani F, Oikonomou A, Plataniotis KN, Mohammadi A. Covid-caps: a capsule network-based framework for identification of covid-19 cases from x-ray images. Pattern Recogn Lett. 2020;138:638–43.
12. Wu Y-H, Gao S-H, Mei J, Xu J, Fan D-P, Zhang R-G, Cheng MM. Jcs: an explainable covid-19 diagnosis system by joint classification and segmentation. IEEE Trans Image Process. 2021.
13. Roth HR, Lu L, Liu J, Yao J, Seff A, Cherry K, Kim L, Summers RM. Improving computer-aided detection using convolutional neural networks and random view aggregation. IEEE Trans Med Imaging. 2015;35(5):1170–81.
14. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J. Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Trans Med Imaging. 2016;35(5):1299–312.
15. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial networks. arXiv preprint arXiv:1406.2661. 2014.
16. Lin D, Fu K, Wang Y, Xu G, Sun X. Marta gans: unsupervised representation learning for remote sensing image classification. IEEE Geosci Remote Sens Lett. 2017;14(11):2092–6.
17. Yadav P, Menon N, Ravi V, Vishvanathan S. Lung-gans: unsupervised representation learning for lung disease classification using chest ct and x-ray images. IEEE Trans Eng Manag. 2021. https://doi.org/10.1109/TEM.2021.3103334.
18. Khan AI, Shah JL, Bhat MM. Coronet: a deep neural network for detection and diagnosis of covid-19 from chest x-ray images. Comput Methods Programs Biomed. 2020;196:105581.
19. Oh Y, Park S, Ye JC. Deep learning covid-19 features on cxr using limited training data sets. IEEE Trans Med Imaging. 2020;39(8):2688–700.
20. Wang X, Deng X, Fu Q, Zhou Q, Feng J, Ma H, Liu W, Zheng C. A weakly-supervised framework for covid-19 classification and lesion localization from chest ct. IEEE Trans Med Imaging. 2020;39(8):2615–25.
21. Xu Z, Cao Y, Jin C, Shao G, Liu X, Zhou J, Shi H, Feng J. Gasnet: weakly-supervised framework for covid-19 lesion segmentation. arXiv preprint arXiv:2010.09456. 2020.
22. Zhou J, Jing B, Wang Z. Soda: detecting covid-19 in chest x-rays with semisupervised open set domain adaptation. arXiv preprint arXiv:2005.11003. 2020.
23. Calderon-Ramirez S, Moemeni A, Elizondo D, Colreavy-Donnelly S, Chavarria-Estrada LF, Molina-Cabello MA, et al. Correcting data imbalance for semi-supervised covid-19 detection using x-ray chest images. arXiv preprint arXiv:2008.08496. 2020.
24. Berthelot D, Carlini N, Goodfellow I, Papernot N, Oliver A, Raffel C. Mixmatch: A holistic approach to semi-supervised learning. arXiv preprint arXiv:1905.02249. 2019.
25. Alshazly H, Linse C, Abdalla M, Barth E, Martinetz T. Covidnets: deep CNN architectures for detecting COVID-19 using chest CT scans. medRxiv. 2021.
26. Narin A, Kaya C, Pamuk Z. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. Pattern Anal Appl. 2021. https://doi.org/10.1007/s10044-021-00984-y.
27. Born J, Wiedemann N, Cossio M, Buhre C, Brändle G, Leidermann K, Goulet J, Aujayeb A, Moor M, Rieck B, et al. Accelerating detection of lung pathologies with explainable ultrasound image analysis. Appl Sci. 2021;11(2):672. https://doi.org/10.3390/app11020672.
28. Wang S, Kang B, Ma J, Zeng X, Xiao M, Guo J, Cai M, Yang J, Li Y, Meng X, Xu B. A deep learning algorithm using CT images to screen for corona virus disease (COVID-19). Eur Rad. 2021;6096–104.
29. Kim CK, Choi JW, Jiao Z, Wang D, Wu J, Yi TY, Halsey KC, Eweje F, Tran TM, Liu C, Wang R. An automated COVID-19 triage pipeline using artificial intelligence based on chest radiographs and clinical data. NPJ Digit Med. 2022;5(1):5.
30. Gour M, Jain S. Automated COVID-19 detection from X-ray and CT images with stacked ensemble convolutional neural network. Biocybern Biomed Eng. 2022;42(1):27–41. https://doi.org/10.1016/j.bbe.2021.12.001.
31. Bermejo-Peláez D, San José Estépar R, Fernández-Velilla M, Palacios Miras C, Gallardo Madueño G, Benegas M, Gotera Rivera C, Cuerpo S, Luengo-Oroz M, Sellarés J, Sánchez M. Deep learning-based lesion subtyping and prediction of clinical outcomes in COVID-19 pneumonia using chest CT. Sci Rep. 2022;12(1):9387. https://doi.org/10.1038/s41598-022-13298-8.

32. Hardy-Werbin M, Maiques JM, Busto M, Cirera I, Aguirre A, Garcia-Gisbert N, Zuccarino F, Carbullanca S, Del Carpio LA, Ramal D, Gayete Á. MultiCOVID: a multi modal deep learning approach for COVID-19 diagnosis. Sci Rep. 2023;13(1):18761. https://doi.org/10.1038/s41598-023-46126-8.
33. Di Napoli A, Tagliente E, Pasquini L, Cipriano E, Pietrantonio F, Ortis P, Curti S, Boellis A, Stefanini T, Bernardini A, Angeletti C. 3D CT-inclusive deep-learning model to predict mortality, ICU admittance, and intubation in COVID-19 patients. J Digit Imaging. 2023;36(2):603–16. https://doi.org/10.1007/s10278-022-00734-4.
34. Oi Y, Ogawa F, Yamashiro T, Matsushita S, Oguri A, Utada S, Misawa N, Honzawa H, Abe T, Takeuchi I. Prediction of prognosis in patients with severe COVID-19 pneumonia using CT score by emergency physicians: a single-center retrospective study. Sci Rep. 2023;13(1):4045.
35. Xie P, Zhao X, He X. Improve the performance of CT-based pneumonia classification via source data reweighting. Sci Rep. 2023;13(1):9401.
36. Miyazaki A, Ikejima K, Nishio M, Yabuta M, Matsuo H, Onoue K, Matsunaga T, Nishioka E, Kono A, Yamada D, Oba K. Computer-aided diagnosis of chest X-ray for COVID-19 diagnosis in external validation study by radiologists with and without deep learning system. Sci Rep. 2023;13(1):17533.
37. Abad M, Casas-Roma J, Prados F. Generalizable disease detection using model ensemble on chest X-ray images. Sci Rep. 2024;14(1):5890.
38. Sadeghi A, Sadeghi M, Sharifpour A, Fakhar M, Zakariaei Z, Sadeghi M, Rokni M, Zakariaei A, Banimostafavi ES, Hajati F. Potential diagnostic application of a novel deep learning-based approach for COVID-19. Sci Rep. 2024;14(1):280.
39. King, B., Barve, S., Ford, A., Jha, R.: Unsupervised clustering of covid-19 chest x-ray images with a self-organizing feature map. In: 2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS). IEEE. 2020. p. 395–98.
40. Mansour RF, Escorcia-Gutierrez J, Gamarra M, Gupta D, Castillo O, Kumar S. Unsupervised deep learning based variational autoencoder model for covid-19 diagnosis and classification. Pattern Recogn Lett. 2021;151:267–74. https://doi.org/10.1016/j.patrec.2021.08.018.
41. Rashid N, Hossain MAF, Ali M, Islam Sukanya M, Mahmud T, Fattah SA. Autocovnet: unsupervised feature learning using autoencoder and feature merging for detection of covid-19 from chest x-ray images. Biocybern Biomed Eng. 2021;41(4):1685–701. https://doi.org/10.1016/j.bbe.2021.09.004.
42. Alizadehsani R, Sharifrazi D, Izadi NH, Joloudari JH, Shoeibi A, Gorriz JM, Hussain S, Arco JE, Sani ZA, Khozeimeh F, et al. Uncertainty-aware semi-supervised method using large unlabelled and limited labeled covid-19 data. arXiv preprint arXiv:2102.06388. 2021.
43. Waheed A, Goyal M, Gupta D, Khanna A, Al-Turjman F, Pinheiro PR. Covidgan: data augmentation using auxiliary classifier gan for improved covid-19 detection. Ieee Access. 2020;8:91916–23.
44. Loey M, Smarandache FM, Khalifa NE. Within the lack of chest covid-19 x-ray dataset: a novel detection model based on gan and deep transfer learning. Symmetry. 2020;12(4).
45. Odena A, Dumoulin V, Olah C. Deconvolution and checkerboard artifacts. Distill. 2016;1(10):3.
46. El-Shafai W. Extensive COVID-19 X-Ray and CT chest images dataset. Mendeley. 2020. https://doi.org/10.17632/8H65YWD2JR.3. https://data.mendeley.com/datasets/8h65ywd2jr/3.
47. Guo X, Liu X, Zhu E, Yin J. Deep clustering with convolutional autoencoders. In: Liu D, Xie S, Li Y, Zhao D, El-Alfy ES, editors. Neural information processing. ICONIP 2017. Lecture Notes in Computer Science, vol. 10635. Cham: Springer; 2017. https://doi.org/10.1007/978-3-319-70096-0_39.
48. Guo X, Gao L, Liu X, Yin J. Improved deep embedded clustering with local structure preservation. In: IJCAI. 2017. p. 1753–59.
49. Maaten L, Hinton G. Visualizing data using t-sne. J Mach Learn Res. 2008;9(11).