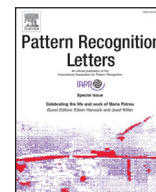




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

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.



Unsupervised Deep Learning based Variational Autoencoder Model for COVID-19 Diagnosis and Classification



Romany F. Mansour^a, José Escorcía-Gutiérrez^b, Margarita Gamarra^c, Deepak Gupta^d, Oscar Castillo^e, Sachin Kumar^{f,*}

^a Department of Mathematics, Faculty of Science, New Valley University, El-Kharga 72511, Egypt

^b Electronic and telecommunications program, Universidad Autónoma del Caribe, Barranquilla, Colombia

^c Department of Computational Science and Electronic, Universidad de la Costa, CUC, Barranquilla, Colombia

^d Department of Computer Science & Engineering, Maharaja Agrasen Institute of Technology, Delhi, India

^e Tijuana Institute Technology, Tijuana, Mexico

^f Department of Computer Science, South Ural State University, Chelyabinsk, Russian Federation

ARTICLE INFO

Article history:

Received 10 April 2021

Revised 24 July 2021

Accepted 26 August 2021

Available online 22 September 2021

Edited by: Maria De Marsico

Keywords:

COVID-19

Deep learning

Unsupervised learning

Variational autoencoder

Image classification

ABSTRACT

At present times, COVID-19 has become a global illness and infected people has increased exponentially and it is difficult to control due to the non-availability of large quantity of testing kits. Artificial intelligence (AI) techniques including machine learning (ML), deep learning (DL), and computer vision (CV) approaches find useful for the recognition, analysis, and prediction of COVID-19. Several ML and DL techniques are trained to resolve the supervised learning issue. At the same time, the potential measure of the unsupervised learning technique is quite high. Therefore, unsupervised learning techniques can be designed in the existing DL models for proficient COVID-19 prediction. In this view, this paper introduces a novel unsupervised DL based variational autoencoder (UDL-VAE) model for COVID-19 detection and classification. The UDL-VAE model involved adaptive Wiener filtering (AWF) based preprocessing technique to enhance the image quality. Besides, Inception v4 with Adagrad technique is employed as a feature extractor and unsupervised VAE model is applied for the classification process. In order to verify the superior diagnostic performance of the UDL-VAE model, a set of experimentation was carried out to highlight the effective outcome of the UDL-VAE model. The obtained experimental values showcased the effectual results of the UDL-VAE model with the higher accuracy of 0.987 and 0.992 on the binary and multiple classes respectively.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

Recently, 2019 novel coronavirus is referred as COVID-19 originated from Wuhan city, China in December 2019, and it has been progressed as a dreadful disease and considered a communal health problem around the globe. The COVID-19 is a global disorder and evolved from Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) [1]. Coronaviruses (CoV) belong to the virus family and the disease is named COVID-19 which mimics the virus SARS-CoV as well as Middle East Respiratory Syndrome (MERS-CoV). COVID-19 has defined a novel viral infection developed in 2019 which has not been predicted in human beings. Also,

it is considered to be a zoonotic disorder as the germs have affected both animals and users. Developers have found that SARS-CoV virus is formed out of decomposed cats and it has been spread among human beings while MERS-CoV virus is desecration from dromedaries and distributed from Arabian camel to humans. COVID-19 virus is actually dispersed from bats to human beings. This type of virus is transferred through respiratory organs from human to human and results in rapid virus transmission. It provokes slight symptoms for most of the patients and some people have severe infection.

The physicians have applied X-ray images for pneumonia analysis, lung disorder, abscesses, and swollen lymph nodes. Every hospital has X-ray imaging device and it is possible to take X-rays test for COVID-19 without any special testing tools. The demerits of X-ray testing are that it requires radiology expert, time-consuming, expensive. Therefore, developing an automatic analysis model is extremely significant to limit the overhead of medical professionals. In addition, positive findings are attained from X-ray as well

* Corresponding Author

E-mail addresses: romanyf@scinv.au.edu.eg (R.F. Mansour), jose.escorcias23@gmail.com (J. Escorcía-Gutiérrez), mrgamarra22@gmail.com (M. Gamarra), deepakgupta@mait.ac.in (D. Gupta), ocastillo@tectijuana.mx (O. Castillo), sachinagnihotri16@gmail.com (S. Kumar).

as computed tomography (CT) images. However, it can be unfit for further diagnosis due to the prolonged diagnosis period of time. The accurate diagnosis rate of this method varies to greater extent. Hence, maximum count of repeated tests has to be conducted for gaining appropriate outcomes. But, it signifies a spherical allocation of images, it depicts the similar features with alternate viral lung disorder.

Previously, in case of fast progression of coronavirus, numerous works have been performed by many developers. In recent times, deep learning (DL) is considered to be the well-known model related to medical sector for diagnosing process [2]. DL is considered to be the combination of Machine Learning (ML) models which is extremely concentrated on computing automated feature extraction as well as image classification and capable of performing object prediction as well as medicinal image classification. The ML and DL techniques are assumed to be well-organized method used in mining, analyzing, and pattern identification from images. Regaining the betterment of clinical decision making as well as computer-aided detection (CAD) is considered to be non-trivial due to the generation of new data. Moreover, DL is defined as the process in which Deep Convolutional Neural Networks (DCNN) was employed to perform automatic feature extraction, which exploits the convolution task and layers perform on non-linear data. A layer is operated by data transformation from high to abstract level. Basically, DL represents novel deep networks when compared with classical ML methods with the help of big data.

The contribution of the study can be defined as follows. This paper devises an efficient unsupervised DL based variational autoencoder (UDL-VAE) model for COVID-19 detection and classification. The UDL-VAE model performs adaptive Wiener filtering (AWF) based preprocessing, Inception v4 with Adagrad based feature extraction, and unsupervised VAE based classification. The application of Adagrad technique helps to adjust the hyperparameters of Inception v4 model, and thereby the classification performance can be improved. For facilitating the effective detection performance of the UDL-VAE method, a comprehensive experimental validation takes place to make sure the proficient performance of the UDL-VAE model.

2. Literature Review

Generally, ML methods are considered as a sub-section of Artificial Intelligence (AI) and used prominently for medical domains in feature extraction and image analysis. Sorensen et al. [3] processed the heterogeneity from collective Regions of Interest (ROI). Then, the features are classified by using a standard vector space-relied classifier. Zhang and Wang [4] offered a CT classification with 3 classical features called grayscale values, shape, as well as texture, and symmetric features. It can be achieved using Radial Basis Function Neural Network (RBFNN) and perform feature classification. Homem et al. [5] proposed a relative study utilizing Jeffries–Matusita ($J\tilde{m}$) distance and Karhunen–Loève transformation feature extraction processes. Albrecht et al. [6] developed a classifier with the average grayscale measure of images in several image classifiers. Yang et al. [7] suggested an automatic classifier for classifying the breast CT photographs utilizing morphological properties. Followed by, the performance is limited when same process is determined with diverse datasets. In addition, hand-engineered models are simulated to develop CNN as well as automated feature extraction methods. Thus, CNN structure is referred as DL structure used in extraction and classification of images.

Ozyurt et al. [8] employed a hybrid mechanism known as fused perceptual hash depends on CNN to reduce the classification time of liver CT photographs and maintain the performance. Xu et al. [9] applied a Transfer Learning (TL) procedure to overcome the clinical image imbalance problem. Followed by, the performance

of GoogleNet, ResNet101, Xception, as well as MobileNetv2 are compared to gain better results. Lakshmanprabu et al. [10] analyzed the CT scan of lung images utilizing optimal deep neural network (DNN) and Linear Discriminate Analysis (LDA). Gao et al. [11] changed original CT images to maximum and minimum attenuation pattern rescale. Consequently, images are subjected to re-sampling and classified by applying CNN.

Shan et al. [12] developed DL based model for automatic segmentation of lung as well as defected regions using chest CT images. Xu et al. [13] concentrated on making basic screening models for differentiating COVID-19 pneumonia and Influenza-A viral pneumonia from healthy cases under the application of pulmonary CT images as well as DL module. Wang et al. [14] processed on the basis of COVID-19 radiographic modification from CT images and developed a DL method for extracting graphical features of COVID-19 to provide clinical analysis prior to reach pathologic condition and eliminate the fatal state of a patient. From Hamimi [15], MER-SCoV has depicted that features in chest X-ray as well as CT are depictions of pneumonia. Xie et al. [16] utilized Data Mining (DM) models for classifying SARS and pneumonia use of X-ray images.

In [17], a new AI-powered pipeline using DL model is developed to automatically detect COVID-19 and classification process takes place using CT scans. In addition, the segmentation network is integrated to the classifier model for COVID-19 identification and lesion categorization. In [18], a novel Multiple Kernels-ELM-based Deep Neural Network (MKs-ELM-DNN) model is projected to detect COVID-19 using CT images with CNN model. In [19], a set of 9 classifier models are employed for COVID-19 diagnosis such as VGG16, DenseNet121, DenseNet169, DenseNet201 and MobileNet, artificial neural network (ANN), decision tree (DT), and random forest (RF) using CT images. Some other COVID-19 diagnosis models are existed in the literature [20–37]. Even though there are numerous methods developed in diagnosing COVID-19, still there is a requirement of identifying COVID-19 from Chest X-ray images. Followed by, X-ray devices are used in scanning defected regions of a body such as fractures, bone misplacement, lung disorder, and lesions. The CT scanning is defined as the expanded version of X-ray used for examining soft tissues in the internal organs. The time delay in detecting COVID-19 pneumonia results in mortality. Hence, an efficient COVID-19 diagnosing method is applied for prediction and classification of disease.

3. Unsupervised Learning based COVID-19 Diagnosis Model

The proposed unsupervised learning based COVID-19 detection and classification model is demonstrated in Fig. 1. The figure indicated that the medical image quality can be raised by the use of AWF technique. Then, Inception v4 with Adagrad model extracts out the suitable set of feature vectors from the preprocessed image. At last, unsupervised VAE method is implemented to define the appropriate class labels of the input medical images.

3.1. Adaptive Wiener Filtering based Preprocessing

The remarkable development of Wiener Filter (WF) has performed the image processing due to the low pass features and intends to develop inevitable blurring of lines as well as edges. When the signal is defined as a non-Gaussian process like natural images, WF has surpassed the nonlinear estimators. The count of iterations for resolving the problem has applied a nonstationary scheme in which the features of signal as well as noise are enabled to modify. The WF expressed is describes a shift-invariant filter, and a similar filter is applied for the entire image. The filter can be made as spatially variant by applying local spatially differing method of noise

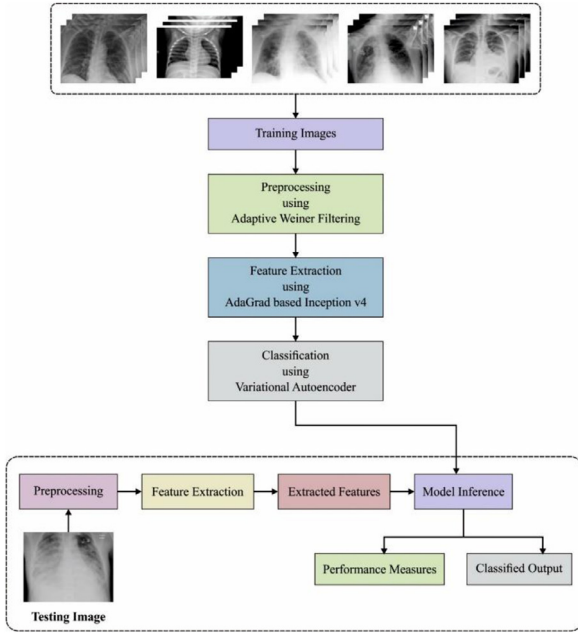


Fig. 1. The working process of proposed UDL-VAE model.

attribute σ_n and represented as,

$$H = \frac{S_{ff}}{S_{ff} + \sigma_n^2(x)} \quad (1)$$

The filter development is costlier due to the filter is changed from pixel to pixel in an image.

Lee suggested an effective execution of noise-adaptive WF by changing the signal as stationary procedure and leads in the filter formation,

$$\hat{f}(x) = m_f(x) + \frac{\sigma_f^2(x)}{\sigma_f^2(x) + \sigma_n^2} (g(x) - m_f(x)) \quad (2)$$

Where m_f denotes the local mean of signal g , and σ_f^2 implies local signal variance. Actually, a local mean operator is adopted by normalized local production function (LPF) in which support region of filter describes the position of mean operator. It can be pointed that the filter development is depicted as same as unsharp masking, even though the latter is not adaptive locally.

The WF result from a minimization relied on mean square error (MSE) criteria is related to 2nd order statistics of input data. The application of WF is to filter signals of interest and offers suboptimal results. The stationary white noise Wiener solution is developing visible function, $0 \leq \alpha \leq 1$, based on the magnitude of image gradient vector, in which 0 for “huge” gradients and 1 for absence of gradient. As a result, the generalized Backus-Gilbert condition has attained better result:

$$H_\alpha = \frac{S_{ff}}{S_{ff} + \alpha\sigma^2} \quad (3)$$

The Backus-Gilbert model is defined as a regularization model that varies from alternate models and prefers to enhance the stability and smoothness of the solution. Even though the Backus-Gilbert statement is varied from standard linear regularization models, the variations among the models are minimum.

Eq. (3) indicates the explicit trade-off among resolution as well as stability. $\alpha = 0$ provides a filter which is identity mapping while $\alpha = 1$ offers a smoother Wiener solution. Under the selection of spatially variant solution, a function of image location is $\alpha(x)$. It is comprised of a method with undesired properties where the

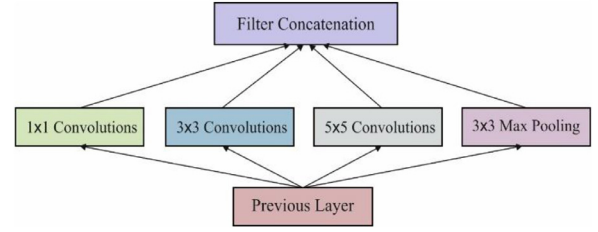


Fig. 2. The inception module.

filter is changed from point to point and it results in computation overhead. The “signal equivalent” model with a filter in linear integration of stationary WF as well as identity map:

$$H_\alpha = \alpha \frac{S_{ff}}{S_{ff} + \sigma^2} + (1 - \alpha) \quad (4)$$

It can be pointed that H_α is similar to Wiener solution for $\alpha = 1$, and in case of $\alpha = 0$ a filter is considered as identity map. It is evident that Eq. (4) is represented as

$$H_\alpha = \frac{S_{ff}}{S_{ff} + \sigma^2} + (1 - \alpha) \frac{\sigma^2}{S_{ff} + \sigma^2} \quad (5)$$

Eq. (5) implies that method is considered as a linear integration of stationary LPF elements as well as nonstationary high pass units [38]. The addition of H in stationary Wiener solution is depicted in Eq. (5) as

$$H_\alpha = H + (1 - \alpha)(1 - H) \quad (6)$$

3.2. Inception v4 with Adagrad based Feature Extraction

Here, preprocessed medical image is subjected to feature extraction in which the required features are extracted from given image using Inception v4 method Szegedy et al. [39]. Basically, CNN is comprised of numerous layers namely, convolutional layers, down-sampling layers, as well as activation layers. The inputs of CNN are determined by means of 1D, 2D as well as 3D. CNNs having 1D inputs classify the images directly in spectral domain whereas, in 2D, the inputs extract features from adjacent pixels and using neighbors of pixel to be divided as input; and finally, CNNs with 3D inputs filter complicated features from spectral as well as spatial domains. Then, CNNs with spatial data is capable of accomplishing optimal performance by means of classification accuracy. Fig. 2 shows the structure of Inception module. Here, 2 CNNs were developed on the basis of Directed Acyclic Graphs (DAG) structure in which fundamental layers are defined in the following.

Assume x as the vector of pixels in input image X , and single neuron computes the process x and generates the simulation outcome a . Hence, neuron function is described in the following:

$$a = \sigma (fx + b) \quad (7)$$

Where f denotes the weight filter, b implies the bias, and (\cdot) indicates the activation function called nonlinear function. A neuron is commonly related to particular spatial position (i, j) as well as a dimension d . It refers that convolutional block is executed n entire position of spectral dimensionality. For DAG, the input feature $x \in \mathbb{R}^{H \times W \times D}$ is defined as the position (i, j, d) with multi-dimensional filter $f \in \mathbb{R}^{H' \times W' \times D \times D'}$ as well as bias b , the output $a \in \mathbb{R}^{H'' \times W'' \times D''}$ is depicted as given below:

$$a_{i'',j'',d''} = \sigma \left(b_{d''} + \sum_{i'=1}^{H'} \sum_{j'=1}^{W'} \sum_{d=1}^D f_{i'j'd} \times x_{i'+i'-1,j'+j'-1,d',d''} \right) \quad (8)$$

It is pointed that activation function is performed with square images under different image-processing issues, however, it is operated in random inputs and filters. When the top-bottom and left-right paddings ($P_h^-, P_h^+, P_w^-, P_w^+$) as well as down-sampling strides (S_h, S_w) are depicted in Eq. (8):

$$a_{i''j''d''} = \sigma \left(b_{d''} + \sum_{i'=1}^{H'} \sum_{j'=1}^{W'} \sum_{d=1}^D f_{i'j'd} \times X_{S_h(i''-1)+i'-P_h^-, S_h(j''-1)+j'-P_w^-, d', d''} \right) \quad (9)$$

The size of output for convolutional layer in DAG structure is represented as $H'' = 1 + [(H - H' + P_h^- + P_h^+)/S_h]$.

For every layer, a single activation function has been employed. Also, a sigmoid function as well as Rectified Linear Unit (ReLU) are the prominently applied activation functions. And ReLU is expressed in Eq. (10).

$$\sigma(x) = \max(0, x) \quad (10)$$

Additionally, the convolutional layers, down-sampling process are included in the layers for enhancing the receptive field of neurons. Initially, down sampling is achieved by using pooling layers and stride for skipping few convolutions [40]. The major responsibility of CNNs in image classification is to detect the class labels of sample by reducing the loss function \mathcal{L} . A typically employed log-loss function is used and depicted as:

$$\mathcal{L}(x, c) = -\log x \quad (11)$$

Where x_c represents the positive label values. Here, a softmax function has been employed for top layer and generate the result with probability distribution which means that $x_k = p(k)$, $k = 1, \dots, C$. After applying \mathcal{L} , weights and biases are computed by loss reduction. Finally, optimization is estimated by using a Gradient Descent (GD) method. The primary derivative of loss function is attained by means of upgrading weight with a learning rate λ in all iteration ad shown in the following:

$$f_i = f_i + \lambda(\partial\mathcal{L}/\partial f_i) \quad (12)$$

The primary derivative $\partial\mathcal{L}/\partial f_i$ is attained by using Back Propagation (BP) chain rule.

The former models of Inception schemes are applied for training diverse blocks where the repeated blocks are divided as massive subnetworks and place the entire storage region. Therefore, Inception method is simply tuned and depicted that massive credible modifications are performed on the basis of filter count in different layers which do not affect the supremacy of fully equipped system. The optimization of training rate is performed by proper layer sizes tuned in reaching optimal trade-off in processing under diverse subnetworks. In contrast, by Tensor Flow, present Inception models are described without the repeated partitions. It is due to the application of present storage area to perform backpropagation (BP), accomplished by activating the required tensors for gradient processing and demonstrate estimation for reducing the tensors.

Then, Inception-v4 is projected to drop out the unwanted estimation and develop estimation for Inception blocks of different grid sizes. Fig. 3 demonstrates the process of Inception module with dimensionality reduction. In the residual inception blocks of residual versions from Inception method, minimum Inception blocks are projected across common Inception. Between the distinct versions of Inception techniques, the step time of Inception-v4 is found to be low. Batch normalization is found to be the actual variance among the residual as well as non-residual Inception blocks.

Under residual scaling, the residual version begins to provide instability and the network becomes inactive in the presence of filter count >1000, representing that the last layer earlier to the average pooling generates only 0's under various number of iterations. It can be discarded by diminishing the rate of learning or

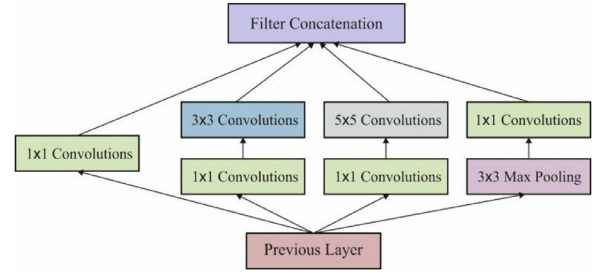


Fig. 3. The Inception Module with Dimensionality Reduction.

adjusting the extra batch normalization layer. On the other hand, hyperparameters undergo tuning for controlling the effective performance of the Inception method. The procedure of selecting the hyperparameters is the main feature of the DL approaches. In this case, the hyperparameters of the Inception-v4 model are tuned by the use of Adagrad optimization model.

It refers the parameter-based learning rates and corresponding learning rates of variables that are improvised sequentially as small and large parameters which are upgraded irregularly. Hence, update rule for Adagrad is provided as,

$$v_0 = 0 \text{ (initialize squared gradient accumulator)}$$

$$v_t = v_{t-1} + (\nabla J(\theta_{t-1}))^2 \text{ (accumulate squared gradient)} \quad (13)$$

$$\theta_t = \theta_{t-1} - \frac{\eta}{\sqrt{v_t} + \epsilon} \circ \nabla J(\theta_{t-1}) \text{ (apply update)} \quad (14)$$

Where 0 defines Hadamard (element-wise) product as well as $(\nabla J(\theta_{t-1}))^2$ implies the element-wise square of applied gradient. Also, division and square root $\frac{\eta}{\sqrt{v_t} + \epsilon}$ are evaluated from element-wise model. Then, i th component $\frac{\eta}{\sqrt{\sum_{k=1}^t (\partial J(\theta_{k-1}))^2 + \epsilon}}$ of secondary vector refers the learning rate applied for upgrading the parameter i in iteration t . A major inefficiency of Adagrad optimizer is static development of accumulator v in complete training process and every iteration of i th coordinate the respective squared partial derivative of cost function J has been included and concludes in infinite learning rates like ≈ 0 and terminates the training operations.

3.3. Unsupervised VAE based Classification

At the last stage, VAE model is applied to allocate the proper class label of the applied medical images. Basically, VAE is relied on the traditional Auto-Encoder (AE). In prior to developing VAE, it is crucial to learn the AE process. In general, AE is a type of unsupervised learning method which desires to extract secondary features concealed in the actual data. Hinton had applied unsupervised automated encoder to pre-train and resolve the issues of gradient diminishing.

VAE is a type of deep Bayesian system which is the combination of NN in conjunction with statistics. When compared with former AE, it enforces the latent codes and follows certain distribution like Gaussian distribution. Actually, encoder portion of entire NN is developed to imply the conditional probability $q_\phi(z|x)$, in which x denotes the actual data, ϕ refers the weights of encoder, and z shows the latent codes. Inversely, VAE promotes the distribution of latent codes z to the standard normal distribution. Finally, latent codes z are comprised of stable statistical features which result in better convenience for decoder and efficacy of this model [41]. Next, decoder manages to reform the actual data by applying latent codes z . Consequently, NN having the weight 0 indicates a

conditional probability $p_{\theta}(x|z)$. In order to retain the uncertainty within the system, inputs of decoder undergo sampling from scattering of latent codes z . Therefore, encoder is modeled in 2 resultant parameters namely, mean as well as variance vector of latent codes z . Followed by, 2 parameters describe the normal distribution for sample.

3.4. Kullback-Leibler (KL) divergence

To make distribution of latent code model for specific distribution, a major problem is to estimate the space among 2 distribution q and p . *KL* divergence is established for estimating the variations among 2 probability distributions. Hence, the closer distributions are smaller the measure of *KL*. When there are 2 unseen distribution $p(x)$ and $q(x)$, then *KL* divergence is attained by applied expression:

$$D_{KL}((p(x)||q(x))) = - \int p(x) \log q(x) dx - (- \int p(x) \log p(x) dx) = \int p(x) \log \left\{ \frac{p(x)}{q(x)} \right\} dx. \tag{15}$$

This is referred as Kullback Leibler or *KL* divergence from $p(x)$ and $q(x)$.

3.5. Establishment of loss function

In this model, there are 2 objectives in VAE namely, reform the actual data and create latent codes in certain distribution. Hence, the loss function has been classified into 2 portions. The initial part is to estimate the distance from produced and actual data. Afterward, MSE function has been applied and referred as desired value of squared difference among 2 parameters. MSE is applicable to measure the difference between actual and reformed data. The latter portion is the loss to relate the distance among distribution of latent code z as well as remarkable Gaussian distribution. The *KL* divergence is used and resulted in final loss process of VAE has been devised as,

$$\mathcal{L}_{VAE} = \lambda \text{MSE}(x, y) + E_{p_{data}(x)} [-\text{KL}(q_{\phi}(z|x)||p(z))], \tag{16}$$

Where x indicates the actual data, y defines the reformed data, λ represents a parameter, $q_{\phi}(z|x)$ defines the distribution of latent codes produced from actual data x by encoding device, and $p(z)$ depicts the distribution of latent codes. Therefore, $E_{p_{data}(x)}[\cdot]$ demonstrates the numerical expression.

4. Experimental Validation

The performance validation of the presented UDLVAE model has been tested against COVID Chest X-ray dataset [42]. It contains the images under distinct classes namely Normal, COVID-19, SARS, ARDS, and Streptococcus. The parameter setting of the proposed model is given as follows. Batch size: 500, max. epochs:15, learning rate: 0.05, dropout rate: 0.2, and momentum: 0.9. A few sample test images are illustrated in Fig. 4. Besides, the dataset is split is five different folds.

Table 1 and Fig. 5 tabulates the binary classification results of the UDLVAE model interms of distinct performance measures with varying folds. The obtained values denoted that the UDLVAE model has reached to effective classification outcome under diverse folds. On the applied fold_1, the UDLVAE model has accomplished a sens. of 0.981, spec. of 0.988, acc. of 0.984, and F-measure of 0.983. Besides, on the applied fold_3, the UDLVAE method has demonstrated a sens. of 0.987, spec. of 0.992, acc. of 0.991, and F-measure of 0.991. Also, on the applied fold_5, the UDLVAE technique has exhibited a sens. of 0.983, spec. of 0.988, acc. of 0.986, and F-measure of 0.988.

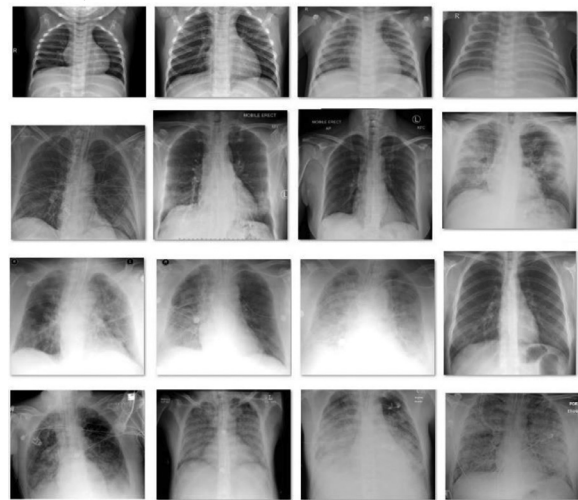


Fig. 4. The Sample Images.

Table 1 The Performance analysis of UDLVAE Model for Binary Class under different folds

Cross Folds	Sens.	Spec.	Acc.	F-measure
Fold_1	0.981	0.988	0.984	0.983
Fold_2	0.985	0.983	0.984	0.990
Fold_3	0.987	0.992	0.991	0.991
Fold_4	0.987	0.989	0.989	0.990
Fold_5	0.983	0.988	0.986	0.988
Average	0.985	0.988	0.987	0.988

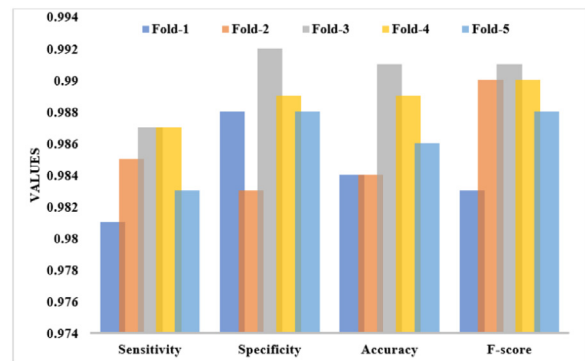


Fig. 5. The Binary-class results analysis of the UDLVAE model.

Table 2 The Performance analysis of UDLVAE Model for Multi-Class under different folds.

Cross Folds	Sens.	Spec.	Acc.	F-measure
Fold_1	0.997	0.993	0.995	0.994
Fold_2	0.996	0.995	0.995	0.995
Fold_3	0.990	0.990	0.990	0.991
Fold_4	0.989	0.992	0.987	0.987
Fold_5	0.997	0.997	0.995	0.994
Average	0.994	0.993	0.992	0.992

Table 2 and Fig. 6 examines the Multi-class classification results of the UDLVAE model with respect to distinct performance measures with varying folds. The attained values referred that the UDLVAE model has achieved effective classification results under diverse folds. On the applied fold_1, the UDLVAE method has showcased a sens. of 0.997, spec. of 0.993, acc. of 0.995, and F-measure of 0.994. Additionally, on the applied fold_3, the UDLVAE technique

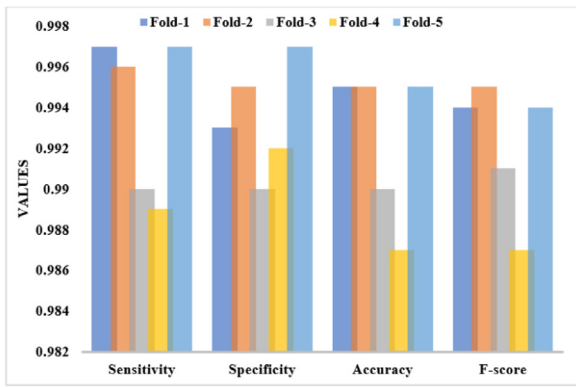


Fig. 6. The Multi-class results analysis of the UDLVAE model.

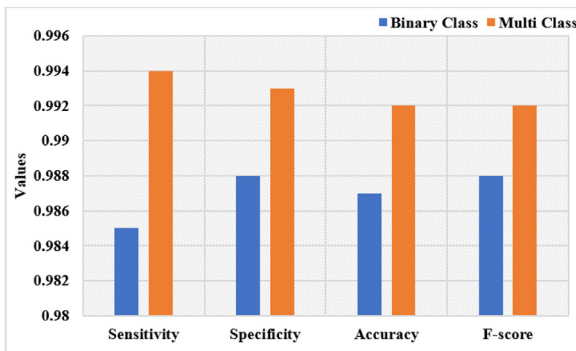


Fig. 7. The Average Analysis of Proposed UDLVAE Model.

Table 3
The Comparison study of UDLVAE with other existing methods.

Methods	Sensitivity	Specificity	Accuracy	F-measure
UDLVAE (Binary Class)	0.985	0.988	0.987	0.988
UDLVAE (Multi Class)	0.994	0.993	0.992	0.992
DWS-CNN (Binary Class)	0.984	0.986	0.985	0.986
DWS-CNN (Multi Class)	0.991	0.992	0.991	0.990
FR-CNN	0.977	0.955	0.974	0.985
ResNet-50	0.930	0.677	0.896	0.939
Inception V3	0.910	0.742	0.887	0.933
AlexNet	0.925	0.714	0.905	0.946
CovxNet	0.905	0.958	0.917	0.911
CapsNet	0.842	0.918	0.892	0.842
VGG19	0.971	0.960	0.963	0.942
Deep Transfer Learning	0.896	0.920	0.908	0.904
Multi-Layer Perceptron	0.930	0.872	0.931	0.930
Logistic Regression	0.930	0.903	0.921	0.920
K-Nearest Neighbour	0.890	0.907	0.889	0.890
Decision Tree	0.870	0.889	0.867	0.870

has accomplished a sens. of 0.990, spec. of 0.990, acc. of 0.990, and F-measure of 0.991. Eventually, on the applied fold_5, the UDLVAE approach has outperformed a sens. of 0.997, spec. of 0.997, acc. of 0.995, and F-measure of 0.994. An average classification results analysis of the UDLVAE model with other existing methods takes place, as given in Fig. 7. The resultant values depicted that the UDLVAE model has accomplished a higher average sens. of 0.985, spec. of 0.988, acc. of 0.987, and F-measure of 0.988 on the classification of binary classes. On the other hand, the UDLVAE model has resulted in a maximum sens. of 0.994, spec. of 0.993, acc. of 0.992, and F-measure of 0.992.

Table 3 and Fig. 8 Fig. 9 provide an elaborate comparative results analysis of the UDLVAE model with other existing techniques. On examining the predictive outcome interms of sensitivity, the CapsNet and DT models have achieved minimal sensitivity values

of 0.842 and 0.87 respectively. Afterward, the KNN, DTL, and CovxNet models have reached slightly increased and nearer sensitivity values of 0.89, 0.896, and 0.905 respectively. Afterward, the Inception v3 and AlexNet methods have reached a moderately closer sensitivity value of 0.91 and 0.925 respectively. At the same time, the ResNet-50, MLP, and LR models have accomplished an identical sensitivity value of 0.93. Along with that, the VGG19 and FR-CNN models have resulted in considerably improved sensitivity values of 0.971 and 0.977 respectively. Eventually, the DWS-CNN model has exhibited the higher result sensitivity values of 0.984 and 0.991 on the prediction of binary and multiple class labels respectively. Finally, the UDLVAE model has showcased superior outcomes by offering a maximum sensitivity of 0.985 and 0.994 on the prediction of binary and multiple class labels respectively.

On investigative the predictive outcome with respect to specificity, the ResNet-50 and AlexNet methods have achieved the worst specificity values of 0.677 and 0.714 correspondingly. Similarly, the Inception, MLP, and DT models have obtained slightly enhanced and nearer specificity values of 0.742, 0.872, and 0.889 correspondingly. Next, the LR and KNN techniques have reached a moderately closer specificity value of 0.903 and 0.907 correspondingly. Simultaneously, the CapsNet, DTL, and CovxNet models have accomplished even superior specificity values of 0.918, 0.920, and 0.958 respectively. Likewise, the FR-CNN and VGG19 models have resulted in considerably enhanced specificity values of 0.955 and 0.960 respectively. Also, the DWS-CNN technique has exhibited higher result specificity values of 0.986 and 0.992 on the prediction of binary and multiple class labels respectively. At last, the UDLVAE model has showcased superior outcomes by offering a maximum specificity of 0.988 and 0.993 on the prediction of binary and multiple class labels respectively.

On determining the predictive outcome interms of accuracy, the DT and Inception V3 models have reached lesser accuracy values of 0.867 and 0.887 respectively. Afterward, the KNN, CapsNet, and ResNet-50 techniques have reached somewhat increased and nearer accuracy values of 0.889, 0.892, and 0.896 correspondingly. Besides, the AlexNet and DTL approaches have reached a moderately closer accuracy value of 0.905 and 0.908 respectively. At the same time, the CovxNet, LR, and MLP models have accomplished even higher accuracy values of 0.917, 0.921, and 0.931 correspondingly. Similarly, the VGG19 and FR-CNN manners have resulted in considerably improved accuracy values of 0.963 and 0.974 respectively. But, the DWS-CNN approach has demonstrated the superior result accuracy values of 0.985 and 0.991 on the prediction of binary and multiple class labels respectively. Finally, the UDLVAE model has showcased superior outcomes by offering a maximum accuracy of 0.987 and 0.992 on the prediction of binary and multiple class labels correspondingly.

On examining the predictive result with respect to F-measure, the CapsNet and DT techniques have obtained minimum F-measure values of 0.842 and 0.87. Similarly, the KNN, DTL, and CovxNet techniques have attained slightly increased and nearer F-measure values of 0.89, 0.904, and 0.911 respectively. Afterward, the LR and MLP models have achieved a moderately closer F-measure value of 0.92 and 0.93 correspondingly. Simultaneously, the Inception V3, ResNet-50, and VGG19 methods have accomplished even higher F-measure values of 0.933, 0.939, and 0.942 correspondingly. At the same time, the AlexNet and FR-CNN techniques have resulted in considerably improved F-measure values of 0.946 and 0.985 respectively. Also, the DWS-CNN algorithm has portrayed the higher outcome F-measure values of 0.986 and 0.990 on the prediction of binary and multiple class labels correspondingly. Finally, the UDLVAE method has showcased superior outcomes by offering a superior F-measure of 0.988 and 0.992 on the prediction of binary and multiple class labels correspondingly.

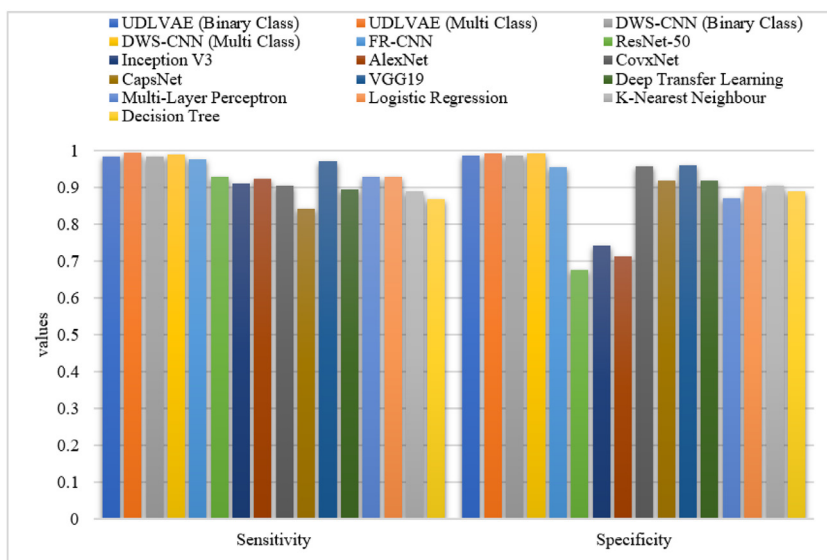


Fig. 8. The Sensitivity and Specificity analysis of UDLVAE with other models.

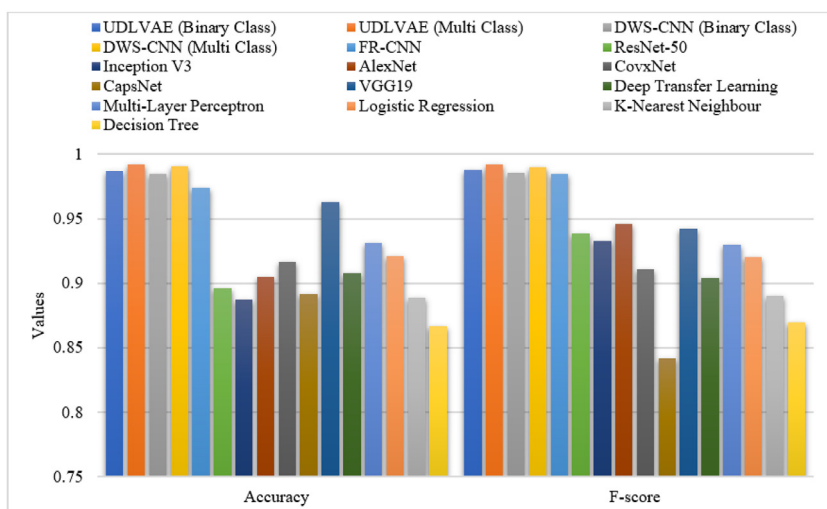


Fig. 9. The Accuracy and F-measure analysis of UDLVAE with other models.

5. Conclusion

This paper has developed a novel UDL-VAE method for COVID-19 detection and classification. The UDL-VAE model performs AWF based preprocessing, Inception v4 with Adagrad based feature extraction, and unsupervised VAE based classification. Primarily, the medical image quality can be raised by the use of AWF technique. Secondly, Inception v4 with Adagrad model extracts out the useful set of feature vectors from the preprocessed image. The application of Adagrad technique helps to adjust the hyperparameters of Inception v4 model, and thereby the classification performance can be improved. Lastly, unsupervised VAE model is applied to define the appropriate class labels of the input medical images. For facilitating the effective detection performance of the UDL-VAE method, a comprehensive experimental validation takes place to make sure the proficient performance of the UDL-VAE method. The obtained experimental values showcased the effectual results of the UDL-VAE model with the higher accuracy of 0.987 and 0.992 on the binary and multiple classes respectively. In future, metaheuristic optimization based learning rate schedulers can be designed for hyperparameter settings. In addition, the presented model can be

employed to diagnose COVID-19 using other imaging modalities like computed tomography (CT). As a part of future extension, it can be incorporated to the internet of things (IoT) and cloud based environment to enable e-healthcare applications.

Declaration of Competing Interest

All the authors of the manuscript declared that there are no potential conflicts of interest.

Acknowledgments

This work is financially supported by the Ministry of Science and Higher Education of the Russian Federation (Government Order FENU-2020-0022).

References

- [1] A Krizhevsky, I Sutskever, GE Hinton, ImageNet classification with deep convolutional neural networks, *Commun ACM* 60 (6) (2017) 84–90 May.
- [2] D.A. Pustokhin, I.V. Pustokhina, P.N. Dinh, S.V. Phan, G.N. Nguyen, G.P. Joshi, An effective deep residual network based class attention layer with bidirectional LSTM for diagnosis and classification of COVID-19, *Journal of Applied Statistics* (2020) 1–18.

- [3] D.N. Le, V.S. Parvathy, D. Gupta, A. Khanna, J.J. Rodrigues, K. Shankar, IoT enabled depthwise separable convolution neural network with deep support vector machine for COVID-19 diagnosis and classification, *International Journal of Machine Learning and Cybernetics* (2021) 1–14.
- [4] L. Sorensen, M. Loog, P. Lo, H. Ashraf, A. Dirksen, R.P. Duin, M. De Bruijne, Image dissimilarity-based quantification of lung disease from CT, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Berlin, Heidelberg, 2010, pp. 37–44.
- [5] W.L. Zhang, X.Z. Wang, Feature extraction and classification for human brain CT images, in: *In 2007 International Conference on Machine Learning and Cybernetics*, 2, IEEE, 2007, pp. 1155–1159.
- [6] M.R.P. Homem, N.D.A. Mascarenhas, P.E. Cruvinel, The linear attenuation coefficients as features of multiple energy CT image classification, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 452 (1–2) (2000) 351–360.
- [7] A. Albrecht, E. Hein, K. Steinh'fel, M. Taupitz, C.K. Wong, Bounded-depth threshold circuits for computer-assisted CT image classification, *Artificial Intelligence in Medicine* 24 (2) (2002) 179–192.
- [8] X. Yang, I. Sechopoulos, B. Fei, Automatic tissue classification for high-resolution breast CT images based on bilateral filtering, *In Medical Imaging 2011: Image Processing* 7962 (2011) 79623H International Society for Optics and Photonics.
- [9] F. Ozyurt, T. Tuncer, E. Avci, M. Koc, I. Serhatlioglu, A novel liver image classification method using perceptual hash-based convolutional neural network, *Arabian Journal for Science and Engineering* 44 (4) (2019) 3173–3182.
- [10] G. Xu, H. Cao, J.K. Udupa, C. Yue, Y. Dong, L. Cao, D.A. Torigian, A novel exponential loss function for pathological lymph node image classification, *In MIPPR 2019: Parallel Processing of Images and Optimization Techniques; and Medical Imaging* 11431 (2020) 114310A International Society for Optics and Photonics.
- [11] S.K. Lakshmanarabu, S.N. Mohanty, K. Shankar, N. Arunkumar, G. Ramirez, Optimal deep learning model for classification of lung cancer on CT images, *Future Generation Computer Systems* 92 (2019) 374–382.
- [12] ... & M. Gao, U. Bagci, L. Lu, A. Wu, M. Buty, H.C. Shin, Z. Xu, Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks, *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 6 (1) (2018) 1–6.
- [13] Shan, F., Gao, Y., Wang, J., Shi, W., Shi, N., Han, M., Xue, Z., and Shi, Y. Lung Infection Quantification of COVID-19 in CT Images with Deep Learning. arXiv preprint arXiv:2003.04655, 1-19, 2020.
- [14] Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., Yu, L., Chen, Y., Su, J., Lang, G., Li, Y., Zhao, H., Xu, K., Ruan, L., and Wu, W. Deep Learning System to Screen Coronavirus Disease 2019 Pneumonia. arXiv preprint arXiv:2002.09334, 1-29, 2020.
- [15] Wang, S., Kang, B., Ma, J., Zeng, X., Xiao, M., Guo, J., Cai, M., Yang, J., Li, Y., Meng, X., and Xu, B. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). medRxiv preprint doi: <https://doi.org/10.1101/2020.02.14.20023028>, 1-26, 2020.
- [16] A.MERS-CoV Hamimi, Middle East respiratory syndrome corona virus: Can radiology be of help? Initial single center experience, *The Egyptian Journal of Radiology and Nuclear Medicine* 47 (1) (2016) 95–106.
- [17] X. Xie, X. Li, S. Wan, Y. Gong, Mining X-ray images of SARS patients, in: *Graham J. Williams, Simeon J. Simoff (Eds.), Data Mining: Theory, Methodology, Techniques, and Applications*, Springer-Verlag, Berlin, Heidelberg, 2006, pp. 282–294. ISBN: 3540325476.
- [18] Pennisi, M., Kavasidis, I., Spampinato, C., Schininà, V., Palazzo, S., Rundo, F., Cristofaro, M., Campioni, P., Pianura, E., Di Stefano, F. and Petrone, A., 2021. An Explainable AI System for Automated COVID-19 Assessment and Lesion Categorization from CT-scans. arXiv preprint arXiv:2101.11943.
- [19] Turkoglu, M., 2021. COVID-19 Detection System Using Chest CT Images and Multiple Kernels-Extreme Learning Machine Based on Deep Neural Network. IRBM.
- [20] M. Agarwal, L. Saba, S.K. Gupta, A. Carriero, Z. Falaschi, A. Paschè, P. Danna, A. El-Baz, S. Naidu, J.S. Suri, A Novel Block Imaging Technique Using Nine Artificial Intelligence Models for COVID-19 Disease Classification, Characterization and Severity Measurement in Lung Computed Tomography Scans on an Italian Cohort, *Journal of Medical Systems* 45 (3) (2021) 1–30.
- [21] M. Jamshidi, A. Lalbakhsh, J. Talla, Z. Peroutka, F. Hadjilooei, P. Lalbakhsh, M. Jamshidi, L. La Spada, M. Mirmozafari, M. Dehghani, A. Sabet, Artificial intelligence and COVID-19: deep learning approaches for diagnosis and treatment, *IEEE Access* 8 (2020) 109581–109595.
- [22] Hemdan, E.E.D., Shouman, M.A. and Karar, M.E., 2020. Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images. arXiv preprint arXiv:2003.11055.
- [23] S.J. Fong, G. Li, N. Dey, R.G. Crespo, E. Herrera-Viedma, Composite Monte Carlo decision making under high uncertainty of novel coronavirus epidemic using hybridized deep learning and fuzzy rule induction, *Applied soft computing* 93 (2020) 106282.
- [24] H. Louati, S. Bechikh, A. Louati, C.C. Hung, L.B. Said, Deep Convolutional Neural Network Architecture Design as a Bi-level Optimization Problem, *Neurocomputing*, 2021.
- [25] V. Ganesan, P. Rajarajeswari, V. Govindaraj, K.B. Prakash, J. Naren, Post-COVID-19 Emerging Challenges and Predictions on People, Process, and Product by Metaheuristic Deep Learning Algorithm, in: *Machine Intelligence and Soft Computing*, Springer, Singapore, 2021, pp. 275–287.
- [26] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, M. Kaur, Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning, *Journal of Biomolecular Structure and Dynamics* (2020) 1–8.
- [27] M. Alazab, A. Awajan, A. Mesleh, A. Abraham, V. Jatana, S. Alhyari, COVID-19 prediction and detection using deep learning, *International Journal of Computer Information Systems and Industrial Management Applications* 12 (2020) 168–181.
- [28] D. Ezzat, A.E. Hassanien, H.A. Ella, An optimized deep learning architecture for the diagnosis of COVID-19 disease based on gravitational search optimization, *Applied Soft Computing* (2020) 106742.
- [29] P. Kasinathan, O.D. Montoya, W. Gil-González, R. Arul, M. Moovendan, S. Dhivya, R. Kanimozhi, S. Angalaeswari, APPLICATION OF SOFT COMPUTING TECHNIQUES IN THE ANALYSIS OF COVID-19: A REVIEW, *European Journal of Molecular & Clinical Medicine* 7 (6) (2020) 2480–2503.
- [30] A. Altan, S. Karasu, Recognition of COVID-19 disease from X-ray images by hybrid model consisting of 2D curvelet transform, chaotic salp swarm algorithm and deep learning technique, *Chaos, Solitons & Fractals* 140 (2020) 110071..
- [31] R. Kumar, R. Arora, V. Bansal, V.J. Sahayashela, H. Buckchash, J. Imran, N. Narayanan, G.N. Pandian, B. Raman, Accurate prediction of COVID-19 using chest x-ray images through deep feature learning model with smote and machine learning classifiers, *MedRxiv* (2020).
- [32] S. Roy, W. Menapace, S. Oei, B. Luijten, E. Fini, C. Saltori, I. Huijben, N. Chenakshava, F. Mento, A. Sentelli, E. Peschiera, Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound, *IEEE Transactions on Medical Imaging* 39 (8) (2020) 2676–2687.
- [33] T. Zhou, H. Lu, Z. Yang, S. Qiu, B. Huo, Y. Dong, The ensemble deep learning model for novel COVID-19 on CT images, *Applied Soft Computing* 98 (2021) 106885.
- [34] M. Nour, Z. Cömert, K. Polat, A novel medical diagnosis model for COVID-19 infection detection based on deep features and Bayesian optimization, *Applied Soft Computing* 97 (2020) 106580.
- [35] Ezzat, D. and Ella, H.A., 2020. GSA-DenseNet121-COVID-19: a hybrid deep learning architecture for the diagnosis of COVID-19 disease based on gravitational search optimization algorithm. arXiv preprint arXiv:2004.05084.
- [36] 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, *Informatics in medicine unlocked* 20 (2020) 100412.
- [37] MM Rahaman, C Li, Y Yao, F Kulwa, MA Rahman, Q Wang, S Qi, F Kong, X Zhu, X. Zhao, Identification of COVID-19 samples from chest X-Ray images using deep learning: A comparison of transfer learning approaches, *Journal of X-ray Science and Technology (Preprint)* (2020 Jan 1) 1–9.
- [38] M.K. Nath, A. Kanhe, M. Mishra, A Novel Deep Learning Approach for Classification of COVID-19 Images, in: *In 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, 2020, pp. 752–757. IEEE.
- [39] C.F. Westin, H. Knutsson, R. Kikinis, Adaptive image filtering, in: *Handbook of Medical Imaging Processing and Analysis*, Academic press, 2000, pp. 3208–3212.
- [40] C Szegedy, S Iofe, V Vanhoucke, AA Alemi, Inception-v4, inception-resnet and the impact of residual connections on learning, in: *In Thirty-First AAAI Conference on Artificial Intelligence*, Association for the Advancement of Artificial Intelligence, USA, 2017, pp. 1–3.
- [41] M.Y. Sikkandar, B.A. Alrasheedi, N.B. Prakash, G.R. Hemalakshmi, A. Mohanarathinam, K. Shankar, Deep learning based an automated skin lesion segmentation and intelligent classification model, *Journal of ambient intelligence and humanized computing* (2020) 1–11.
- [42] COVID-19 Image Data Collection: Prospective Predictions Are the Future Joseph Paul Cohen and Paul Morrison and Lan Dao and Karsten Roth and Tim Q Duong and Marzyeh Ghassemi arXiv:2006.11988, <https://github.com/ieee8023/covid-chestxray-dataset>, 2020