2022 19th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). Mexico City, Mexico. November 9-11, 2022

Autoencoder-based Unsupervised Anomaly Detection for Covid-19 Screening on Chest X-Ray Images

Nusrat Jahan
Department of Electrical & Computer Engineering
Rajshahi University of Engineering and Technology
Rajshahi-6204, Bangladesh
nusrat.ruet.ece15@gmail.com

Md. Al Mehedi Hasan
Department of Computer Science & Engineering
Rajshahi University of Engineering and Technology
Rajshahi-6204, Bangladesh
mehedi_ru@yahoo.com

***Abstract*—The COVID-19 outbreak is a major global catastrophe of our time and the largest hurdle since World War II. According to WHO, as of July 2022, there are more than 571 million confirmed cases of COVID-19 and over six million deaths. The issue of identifying unexpected inputs based on trained examples of normal data is known as anomaly detection. In the case of diagnosing covid-19, Chest X-ray disorders that are hardly apparent are extremely challenging to identify. Although various well-known supervised classification methods are being applied for that purpose, however in the real scenario, healthy patients' data is tremendously available but contaminated samples are scarce. The process of gathering samples from ill patients is troublesome and takes a lengthy time. To address the issue of data imbalance in anomaly detection, this research demonstrates an unsupervised learning technique using a convolutional autoencoder in which the training phase does not include any infected sample. Being trained only with the healthy data, The patterns of the healthy samples are preserved in latent vector space and can differentiate ill samples by observing substantial divergence from the distribution of healthy data. Higher reconstruction error and lower KDE (Kernel Density Estimation) indicate affected data. By contrasting the reconstruction error and KDE of healthy data with anomalous data, the suggested technique is feasible for identifying anomalous samples.**

***Index Terms*—Chest radiograph, Autoencoder, Deep learning, Unsupervised learning, Anomaly detection**

I. Introduction

Coronavirus disease (COVID-19) is a contagious disorder caused by the SARS-CoV-2 virus. More airways may fill with fluid when COVID-19 worsens because of leaks from the lungs' small blood veins. Breathing difficulty eventually develops and can result in lung failure. In the field of computer-aided diagnosis (CAD) for chest radiographs, deep neural network (DNN)-based techniques have made impressive strides in recent years. The majority of these studies have been conducted in supervised learning [1], a sort of instruction based on labels that correlate to the inputs. Yet, these supervised learning-based CAD systems have two issues. The trouble in providing training datasets is the first challenge. Even for specialists, accurately labeling numerous images with the details of disorders or lesions demands a significant amount of time and attention. So anomaly segmentation and classification techniques have to handle very imbalanced and messy input, which usually produces erroneous decisions. Secondly, to recognize different categories of abnormality, it is important to generate a variety of forms of anomalous data and annotate the anomalies, both of which are tough. Chest X-ray observations are often so delicate that there may be conflicting interpretations because healthy instances predominate significantly over anomalous cases. An unsupervised anomaly detection framework can be used to deal with these issues. such that, determining the properties of healthy data and identifying deviations from the features of healthy samples in the inspected samples. This approach allows for detecting any kind of anomaly and only requires normal samples for training. Another crucial use for anomaly detection is the assessment of uncommon samples. Regardless of the benefits mentioned above, unsupervised anomaly detection is a demanding technique that has not been extensively utilized for medical imaging. Fig. 1. shows a healthy sample. COVID-created pus and fluids trigger radiology sections (white regions) in the anomalous sample which is shown in Fig. 2.

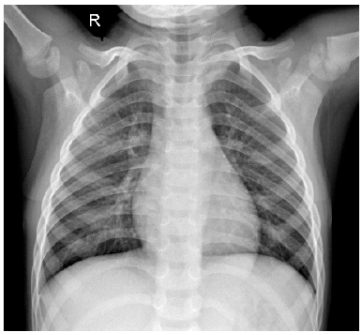

Fig. 1. A normal case of chest X-ray

This study suggests a Convolutional Autoencoder (CAE) as a framework of One Class Classification (OCC) [2]. The following are this research's contributions: The CAE is an encoder-decoder-based network in which the encoder creates a bottleneck that generates a representation of the original

978-1-6654-5508-4/22/$31.00 ©2022 IEEE

input's features in compressed form and the decoder tries to reproduce the input from that compressed portion. The CAE is assessed on both healthy and abnormal instances after training on only the healthy class.

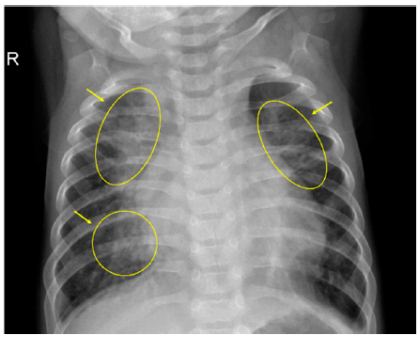

Fig. 2. Sample of Covid-19 case in which radiopaque segments are created by fluid accumulation

Based on Reconstruction error and kernel density estimator, the system can identify anomalous instances. The autoencoder produces a larger reconstruction error when the sample used as input is abnormal. According to KDE, samples that lie in lower probability density areas of normal data distribution are anomalous ones.

II. Related works

Even though several anomaly detection techniques have emerged in the past few years, only a few researchers have included medical data in their investigations [3] [4]. Latest deep learning innovations have been substantially used in the radiology imaging assessment of COVID-19 patients most of which are supervised techniques. A DarkCovidNet model was suggested by Hemdan et al. [5] using 17 convolutional layers with various filters on every layer. Arman et al. [6] used the transfer learning scheme and utilized the well-known CheXNet model to develop COVID-CXNet. The confidence-aware anomaly detection (CAAD) model, proposed by Zhang et al. [7], operates via a convolutional feature detector model entering into an anomaly monitoring system and a confidence projection section that combines. But positive viral-pneumonia CXR training examples are required for CAAD, which focuses on its confidence detection system. A deep autoencoder with gradually expanding blocks and residual connections is suggested by Tuluptceva et al. [8] where both typical and anomalous instances are used for training. To deal with the significant cost of misdiagnosis in COVID-19 situations, Li et al. [9] presented COVID-19 identification from CXR pictures as a cost-sensitive learning issue using a conditional center loss. Scarpiniti et al. [10] suggested an unsupervised Deep Denoising Convolutional Autoencoder (DDCAE) training on CT images which is very costly. By employing generative models, Motamed et al. [11] suggest a novel GAN architecture to improve chest X-rays for the semi-supervised identification of pneumonia and COVID-19. Mansour et al. [12] used Inception v4 as a feature extractor and an unsupervised VAE model is employed for classification. Gayathri et al. [13] employed a technique where features are taken from pre-trained networks, a sparse autoencoder for image compression, and a feed-forward neural network (FFNN) for COVID-19 identification.

In some of the approaches mentioned above, different models are used for feature extraction and classification respectively whereas, in the proposed approach, only a CAE is employed for both tasks. Moreover, all the mentioned strategies still rely on enough cases for COVID-19, which may not be accessible in the beginning phases of an outbreak. The literature study makes it obvious that there is a lack of investigations employing pure OCC or anomaly detection methods. The suggested model architecture is demonstrated in part after, along with data processing, experiments, and outcomes. A discussion and suggestions for new areas of investigation are included in the paper's conclusion.

III. Materials and Methods

A. Dataset

In this work, the COVIDX dataset introduced by Wang and Wong (2020) and the "Chest X-Ray Images(Pneumonia)" [14] dataset are used from Kaggle which are publicly available. From the "Chest X-Ray Images(Pneumonia)" dataset, only the normal images are retrieved. The NORMAL images from these two datasets are merged for the training and validation of the autoencoder with a larger number of normal data. COVID-positive images are taken from the COVIDX dataset for the assessment of anomaly detection. The test dataset is simply provided from the COVIDX dataset which includes 200 healthy samples and 200 anomalous samples of covid positive patients. Details appear in Table I.

TABLE I: Dataset Information

Dataset	Number of Normal Images	Number of Covid positive images
Training data	11041	0
Validation data	4732	0
Covid-19 data	0	4732
Test data	200	200

B. Convolutional Autoencoder

In the suggested system framework, the convolutional autoencoder (CAE) [15] is used to identify COVID-19 as an abnormality in CXR pictures. Fully connected autoencoders neglect the 2D imaging format and retrieve global features, which is why CAE architecture is chosen, CAE, on the other hand, maintains spatial locality and uses shared weights across all input regions. A pair of encoder-decoder makes up a standard CAE. Several convolutional layers are included in each encoder layer, which is followed by pooling layers. With layers of up-sampling and deconvolution, the decoder portion replicates the encoder design in reverse order. The decoder phase that recreates the input picture using the latent encoded form follows the encoder stage where the network attempts to extract spatial characteristics throughout training [16]. This

objective is attained by reducing the Mean Squared Error (MSE) between the input picture and the reconstructed image. The illustration of the encoder functions as the feature vector for quantifying the substance of an image. Using the chain rule to backpropagate the erroneous components initially via the decoder system and then via the encoder system makes it simple to acquire the necessary gradient. In an autoencoder, the dimensionality of the input data is minimized to a code space of a smaller one at the hidden layer because the volume of the hidden layer is smaller than the volume of the input data. The CAE learns to extract characteristics that are essential for reconstructing the normal instances since it only sees the normal examples during the training process. As a result, the CAE is anticipated to yield low reconstruction error for the test set's normal samples while delivering greater reconstruction error for samples that deviate from the norm or are abnormal. As an outcome, the score based on CAE's reconstruction error may be utilized to identify abnormalities.

C. Model Architecture and Training details

The encoder section uses batch normalization, LeakyReLU activation, and consecutive layers of 2D convolutions to compress normalized pixel-wise data from input pictures (128× 128 x 3 volume) into reduced dimensional feature maps.- 256 one-dimensional feature spaces are represented by a fully connected layer that receives the output from the convolutional blocks. In the decoder section, transposed convolutions that up-sample the features to the original input size are used to enlarge that 256 fully-connected output of the encoder. At each stage of the transposed convolution sequence, batch normalization and LeakyReLU activation are also included, and the dimensions of the encoder and decoder filters are mirrored. The feature vectors for each picture in the dataset are generated after the autoencoder has been trained by simply passing the image forward through the network, where the encoder's output (i.e., the latent-vector representation) acts as our feature space. After receiving input data, the autoencoder will condense it to the latent-space representation into a vector with a lot fewer dimensions (256), or essentially a compression of almost 99 percent of the source input data (128 x 128 x 3).

There are three convolution blocks and one fully connected block included in the encoder. Each convolution block consists of four layers: one convolution layer, one leaky ReLU layer, one batch normalization layer, and one Maxpool (size = 2*2) layer. The filter size is 3*3 and 'same' padding is used. The outputs from each convolution layer have 8, 16, and 32 channels, accordingly. A flattened layer of size (16384*1) and to limit the dimensionality of the data to 128 dimensions, there are two fully connected layers with leaky ReLU layers are added. The encoder section is pictured in Fig. 3.

In the decoder section, there are three deconvolution blocks and one fully connected block. The fully connected block (FC2) comprises two fully connected layers including leaky ReLU activation for both and it expands the dimensions of data from (128;) to (8192;). It is then reshaped into (16; 16; 32) and then sent to the three deconvolution blocks. A convolution transposed layer, a batch normalizing layer, a leaky ReLU layer, and an upsampling layer (factor=2) make up each of these blocks. Filter sizes of 3*3 and 'same' padding are used for every transposed convolution layer. The outputs from each deconvolution block have 32, 16, and 8 channels, accordingly. The decoder section is pictured in Fig. 4.

The suggested scheme is trained on normal chest x-ray samples for 70 epochs in the Kaggle platform with GPU. The Mean Squared Error (MSE) is used as the loss function and also as the evaluation metrics. Adam optimizer is utilized and the learning rate is 0.001.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 128, 128, 8)	224
batch_normalization (BatchNo	(None, 128, 128, 8)	32
leaky_re_lu (LeakyReLU)	(None, 128, 128, 8)	0
max_pooling2d (MaxPooling2D)	(None, 64, 64, 8)	0
conv2d_1 (Conv2D)	(None, 64, 64, 16)	1168
batch_normalization_1 (Batch	(None, 64, 64, 16)	64
leaky_re_lu_1 (LeakyReLU)	(None, 64, 64, 16)	0
max_pooling2d_1 (MaxPooling2	(None, 32, 32, 16)	0
conv2d_2 (Conv2D)	(None, 32, 32, 32)	4640
batch_normalization_2 (Batch	(None, 32, 32, 32)	128
leaky_re_lu_2 (LeakyReLU)	(None, 32, 32, 32)	0
max_pooling2d_2 (MaxPooling2	(None, 16, 16, 32)	0
flatten (Flatten)	(None, 8192)	0
dense (Dense)	(None, 128)	1048704

Fig. 3. Architechture of encoder

dense_1 (Dense)	(None, 8192)	1056768
leaky_re_lu_3 (LeakyReLU)	(None, 8192)	0
reshape (Reshape)	(None, 16, 16, 32)	0
conv2d_transpose (Conv2DTran	(None, 16, 16, 32)	9248
batch_normalization_3 (Batch	(None, 16, 16, 32)	128
leaky_re_lu_4 (LeakyReLU)	(None, 16, 16, 32)	0
up_sampling2d (UpSampling2D)	(None, 32, 32, 32)	0
conv2d_transpose_1 (Conv2DTr	(None, 32, 32, 16)	4624
batch_normalization_4 (Batch	(None, 32, 32, 16)	64
leaky_re_lu_5 (LeakyReLU)	(None, 32, 32, 16)	0
up_sampling2d_1 (UpSampling2	(None, 64, 64, 16)	0
conv2d_transpose_2 (Conv2DTr	(None, 64, 64, 8)	1160
batch_normalization_5 (Batch	(None, 64, 64, 8)	32
leaky_re_lu_6 (LeakyReLU)	(None, 64, 64, 8)	0
up_sampling2d_2 (UpSampling2	(None, 128, 128, 8)	0
conv2d_3 (Conv2D)	(None, 128, 128, 3)	219

Total params: 2,127,203
Trainable params: 2,126,979
Non-trainable params: 224

Fig. 4. Architechture of decoder

IV. Experiments and Results

The assessment of this work is based on Reconstruction Error and Kernel density estimation. The probability density function of the data points randomly distributed in a sample space is estimated using the kernel density estimation approach. If it is assumed that a dataset's norm should suit a particular type of probability distribution, an anomaly is something that we should only see extremely infrequently or with very low probability. A threshold for reconstruction error and another for the Kernel density estimation are set. If the reconstruction error of a sample is greater than the reconstruction error threshold or the kernel density is lower than the kernel density threshold, then that sample is detected to be anomalous.

The graphs of validation loss and training loss have presented in Fig. 5.

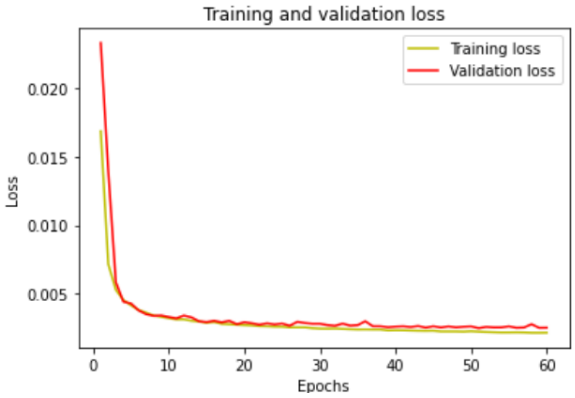

Fig. 5. Loss curves

The error scores of healthy and anomalous data are mentioned in Table II. The test dataset including 200 healthy

TABLE II: Result

Dataset	Reconstruction Error	KDE
Normal	0.0024	6845.03
COVID-positive	0.0059	5555.91

samples and 200 COVID samples is used to calculate the Recall and ROC-AUC score which is mentioned in Table III. I demonstrate the visualization of several instances of

TABLE III: Test Result

Dataset	Recall	ROC-AUC
Test Data	0.66	0.672

healthy and anomalous samples using the reconstruction error approach in Fig. 6 and Fig. 7.

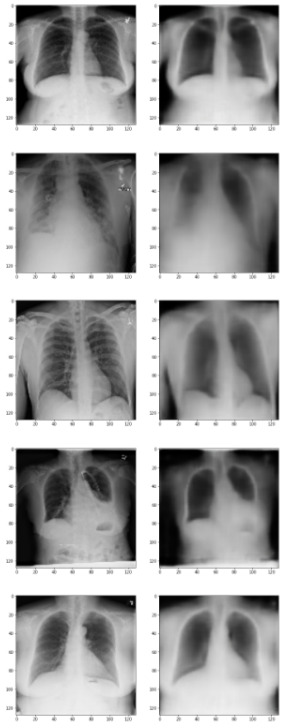

Fig. 6. Reconstruction of Normal data

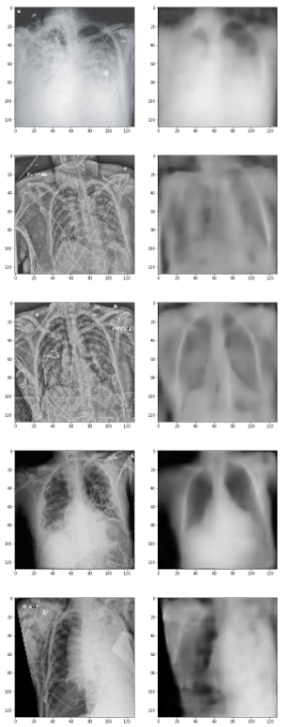

Fig. 7. Reconstruction of anomalous data

V. Comparison with similar approaches

Studies on the suggested method are quite scarce overall. Compared with the others, firstly, healthy samples of two datasets are combined for normal training data so that the

reconstruction process can correctly differentiate the anomalous samples. Secondly, none of the others investigated the KDE which is a crucial evaluator for unsupervised one-class classification. Furthermore, the mentioned CAE architecture is very simple, faster, and lightweight which is convenient for implementation.

TABLE IV: Comparison

Author	Number of healthy data in training	Recons--truction error of covid data	KDE of covid data	AUC	Recall
Tang et al. [17]	4,479	-	-	0.662	-
Khan et al. [18]	8851	-	-	0.69	-
Nakao et al. [19]	6853	-	-	0.630	-
Proposed work	11041	0.0059	5555.91	0.672	0.66

Although the AUC score is not significantly better than those of the works listed in Table IV, the amount of normal data is higher which is extremely significant for training the model and obtaining an accurate score of the reconstruction error.

VI. Conclusion and Future Work

Unlike supervised CAD systems, which demand target disease or lesion pictures and annotations for training, the proposed system just needs typical chest radiographs and no annotations, making it simple to build a training dataset. Additionally, unlike supervised systems, which often can only recognize particular lesions, our technique may detect a variety of lesions or abnormalities. The fact that this approach does not rely on any special target processing also makes it simple to apply to any objective, including any organ or modality in addition to chest radiographs. To aid with better monitoring, recognition, and prognosis of COVID-19, it is crucial to integrate imaging data with pathological changes and laboratory test findings for a system of this kind. It should be noted that imaging only offers a limited feature for COVID patients. Therefore, I will try to apply the fusion technique with these several sources of data. By utilizing other picture transformation networks, such as the U-Net, or applying attention-based networks, this design may be made even more advantageous. The exploration of several loss functions in combination is another scope for future improvement.

References

[1] N. Jahan, M. S. Anower, and R. Hassan, "Automated diagnosis of pneumonia from classification of chest x-ray im ages using efficientnet," in *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*. IEEE, 2021, pp. 235–239.

[2] S. S. Khan and M. G. Madden, "One-class classification: taxonomy of study and review of techniques," *The Knowledge Engineering Review*, vol. 29, no. 3, pp. 345–374, 2014.

[3] C. Baur, B. Wiestler, S. Albarqouni, and N. Navab, "Deep autoencoding models for unsupervised anomaly segmentation in brain mr images," in *International MICCAI brainlesion workshop*. Springer, 2018, pp. 161–169.

[4] K. Ouardini, H. Yang, B. Unnikrishnan, M. Romain, C. Garcin, H. Zenati, J. P. Campbell, M. F. Chiang, J. Kalpathy-Cramer, V. Chandrasekhar *et al.*, "Towards practical unsupervised anomaly detection on retinal images," in *Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data*. Springer, 2019, pp. 225–234.

[5] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images," *arXiv preprint arXiv:2003.11055*, 2020.

[6] A. Haghanifar, M. M. Majdabadi, Y. Choi, S. Deivalakshmi, and S. Ko, "Covid-cxnet: Detecting covid-19 in frontal chest x-ray images using deep learning," *Multimedia Tools and Applications*, pp. 1–31, 2022.

[7] J. Zhang, Y. Xie, G. Pang, Z. Liao, J. Verjans, W. Li, Z. Sun, J. He, Y. Li, C. Shen *et al.*, "Viral pneumonia screening on chest x-rays using confidence-aware anomaly detection," *IEEE transactions on medical imaging*, vol. 40, no. 3, pp. 879–890, 2020.

[8] N. Shvetsova, B. Bakker, I. Fedulova, H. Schulz, and D. V. Dylov, "Anomaly detection in medical imaging with deep perceptual autoencoders," *IEEE Access*, vol. 9, pp. 118 571–118 583, 2021.

[9] L. Wang, Z. Q. Lin, and A. Wong, "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images," *Scientific Reports*, vol. 10, no. 1, pp. 1–12, 2020.

[10] M. Scarpiniti, S. S. Ahrabi, E. Baccarelli, L. Piazzo, and A. Momenzadeh, "A novel unsupervised approach based on the hidden features of deep denoising autoencoders for covid-19 disease detection," *Expert Systems with Applications*, vol. 192, p. 116366, 2022.

[11] S. Motamed, P. Rogalla, and F. Khalvati, "Data augmentation using generative adversarial networks (gans) for gan-based detection of pneumonia and covid-19 in chest x-ray images," *Informatics in Medicine Unlocked*, vol. 27, p. 100779, 2021.

[12] R. F. Mansour, J. Escorcia-Gutierrez, M. Gamarra, D. Gupta, O. Castillo, and S. Kumar, "Unsupervised deep learning based variational autoencoder model for covid-19 diagnosis and classification," *Pattern Recognition Letters*, vol. 151, pp. 267–274, 2021.

[13] J. Gayathri, B. Abraham, M. Sujarani, and M. S. Nair, "A computer-aided diagnosis system for the classification of covid-19 and non-covid-19 pneumonia on chest x-ray images by integrating cnn with sparse autoencoder and feed forward neural network," *Computers in Biology and Medicine*, vol. 141, p. 105134, 2022.

[14] Z. K. Daniel, G. M. Kang, and G. Michael, "Labeled optical coherence tomography (oct) and chest x-ray images for classification," *Mendeley Data, v2*, 2018.

[15] J. Masci, U. Meier, D. Cireşan, and J. Schmidhuber, "Stacked convolutional auto-encoders for hierarchical feature extraction," in *International conference on artificial neural networks*. Springer, 2011, pp. 52–59.

[16] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *science*, vol. 313, no. 5786, pp. 504–507, 2006.

[17] Y.-X. Tang, Y.-B. Tang, M. Han, J. Xiao, and R. M. Summers, "Deep adversarial one-class learning for normal and abnormal chest radiograph classification," in *Medical Imaging 2019: Computer-Aided Diagnosis*, vol. 10950. SPIE, 2019, pp. 305–311.

[18] S. S. Khan, F. Khoshbakhtian, and A. B. Ashraf, "Anomaly detection approach to identify early cases in a pandemic using chest x-rays," *arXiv preprint arXiv:2010.02814*, 2020.

[19] T. Nakao, S. Hanaoka, Y. Nomura, M. Murata, T. Takenaga, S. Miki, T. Watadani, T. Yoshikawa, N. Hayashi, and O. Abe, "Unsupervised deep anomaly detection in chest radiographs," *Journal of Digital Imaging*, vol. 34, no. 2, pp. 418–427, 2021.