Computers, Materials & Continua
DOI:10.32604/cmc.2021.018040
Article

Tech Science Press

Pseudo Zernike Moment and Deep Stacked Sparse Autoencoder for COVID-19 Diagnosis

Yu-Dong Zhang[1], Muhammad Attique Khan[2], Ziquan Zhu[3] and Shui-Hua Wang[4,*]

[1]School of Informatics, University of Leicester, Leicester, LE1 7RH, UK
[2]Department of Computer Science, HITEC University Taxila, Taxila, Pakistan
[3]Science in Civil Engineering, University of Florida, Gainesville, Florida, FL 32608, Gainesville, USA
[4]School of Mathematics and Actuarial Science, University of Leicester, LE1 7RH, UK
[*]Corresponding Author: Shui-Hua Wang. Email: shuihuawang@ieee.org
Received: 22 February 2021; Accepted: 07 April 2021

Abstract: (*Aim*) COVID-19 is an ongoing infectious disease. It has caused more than 107.45 m confirmed cases and 2.35 m deaths till 11/Feb/2021. Traditional computer vision methods have achieved promising results on the automatic smart diagnosis. (*Method*) This study aims to propose a novel deep learning method that can obtain better performance. We use the pseudo-Zernike moment (PZM), derived from Zernike moment, as the extracted features. Two settings are introducing: (i) image plane over unit circle; and (ii) image plane inside the unit circle. Afterward, we use a deep-stacked sparse autoencoder (DSSAE) as the classifier. Besides, multiple-way data augmentation is chosen to overcome overfitting. The multiple-way data augmentation is based on Gaussian noise, salt-and-pepper noise, speckle noise, horizontal and vertical shear, rotation, Gamma correction, random translation and scaling. (*Results*) 10 runs of 10-fold cross validation shows that our PZM-DSSAE method achieves a sensitivity of 92.06% ± 1.54%, a specificity of 92.56% ± 1.06%, a precision of 92.53% ± 1.03%, and an accuracy of 92.31% ± 1.08%. Its F1 score, MCC, and FMI arrive at 92.29% ±1.10%, 84.64% ± 2.15%, and 92.29% ± 1.10%, respectively. The AUC of our model is 0.9576. (*Conclusion*) We demonstrate "image plane over unit circle" can get better results than "image plane inside a unit circle." Besides, this proposed PZM-DSSAE model is better than eight state-of-the-art approaches.

Keywords: Pseudo Zernike moment; stacked sparse autoencoder; deep learning; COVID-19; multiple-way data augmentation; medical image analysis

1 Introduction

COVID-19 has caused more than 107.45 m confirmed cases and 2.35 m deaths till 11/Feb/2021 in about 192 countries/regions and 26 cruise/naval ships [1]. Fig. 1 shows the top 10 countries of cumulative confirmed cases and deaths, respectively. The main symptoms of COVID-19 are low fever, a new and ongoing cough, a loss or change to taste and smell [2]. In the UK,

three vaccines are formally approved as Pfizer/BioNTech, Oxford/AstraZeneca, and Moderna. Two COVID-19 diagnosis methods are available. The former is viral testing to test the existence of viral RNA fragments [3]. The swab test shortcomings are two folds: (i) the swab samples may be contaminated, and (ii) it needs to wait from several hours to several days to get the test results. The latter is chest imaging. There are two main chest imaging available: chest computed tomography (CCT) [4] and chest X-ray (CXR) [5].

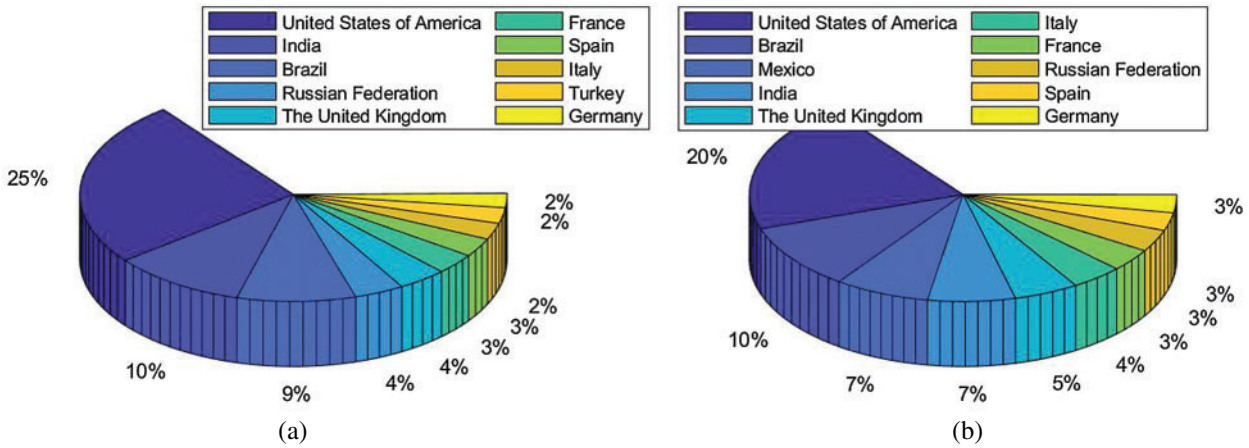


Figure 1: Data till 11/Feb/2021 (a) Cumulative confirmed cases (b) Cumulative deaths

CCT is one of the best chest imaging [6] techniques since it provides the finest resolution and can recognize extremely small nodules in the chest region. CCT employs computer-processed combinations of multiple X-ray observations taken from different angles [7] to produce high-quality 3D tomographic images (virtual slices). In contrast, CXR only provides one 2D image, which performs poorly on soft tissue contrast. This study focuses on the CCT images [8].

Currently, numerous studies are working on using machine learning (ML) and deep learning (DL) technologies [9,10]. For example, Guo et al. [11] employed ResNet-18 for classifying thyroid images. Lu [12] utilized an extreme learning machine (ELM) trained by bat algorithm (BA). Those two approaches were not developing for COVID-19, but they can be transferred to the COVID-19 dataset easily and used as comparison basis approaches in our experiments. For COVID-19 researches, Yao [13] proposed a wavelet entropy biogeography-based optimization (WEBBO) method for COVID-19 diagnosis. Wu [14] presented three-segment biogeography-based optimization (3SBBO) for recognizing COVID-19 patients. Wang et al. [15] presented a DeCovNet. Their accuracy achieved 90.1%. El-kenawy et al. [16] presented a novel feature selection voting classifier (FSVC) method for COVID-19 classification. Yu et al. [17] presented a GoogleNet-COD method to detect COVID-19. Chen [18] designed a gray-level co-occurrence matrix and support vector machine (GLCMSVM) method to classify COVID-19 images [19].

To further improve the performance of automatic COVID-19 diagnosis, this paper proposes a novel method that combines the traditional ML approach with the recent DL approach. We use the pseudo-Zernike moment (PZM) as the extracted features, and we use a deep-stacked sparse autoencoder (i.e., one of the deep neural networks) as the classifier. The combination achieves excellent results that overperform eight state-of-the-art approaches. The novelties of our paper lie in the following aspects

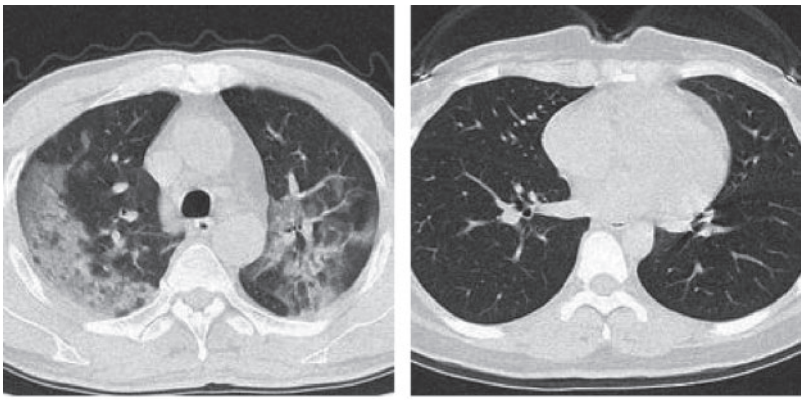
- We are the first to apply a pseudo-Zernike moment to COVID-19 image analysis.
- Deep stacked sparse autoencoder (DSSAE) works better than traditional classifiers.
- Our proposed "PZM-DSSAE" model is better than eight state-of-the-art approaches.

2 Dataset

We use the dataset in reference [20], which contains 148 COVID-19 patients and 148 healthy control (HC) subjects. Slice level selection [20] was employed to generate $C_1 = 320$ COVID-19 images and $C_2 = 320$ HC images. The raw images are with sizes of $1024 \times 1024 \times 3$. A four-step preprocessing was used on this dataset. First, the images are converted to grayscale to save storage amount. Second, histogram stretch is used to enhance the contrast. Third, border pixels are removed, which contains the text and ruler in the right side, and the check-up bed in the bottom. Finally, downsampling to width W and height H is carried out to further reduce the storage of the dataset. Fig. 2 display one example of COVID-19 patient and one sample of HC subject. Algorithm 1 itemizes the pseudocode of preprocessing.

Algorithm 1: Pseudocode of preprocessing

Input Import C_1 raw COVID-19 and C_2 raw HC images.
Step 1 Grayscale. All the images were converted to grayscale.
Step 2 Histogram Stretch: The minimum and maximum grayscale values are mapped to (g_{min}, g_{max}).
Step 3 Crop: $\left(m_1^c, m_2^c, m_3^c, m_4^c\right)$ Pixels are removed from the top, bottom, left, and right sides.
Step 4 Downsampling: All images are downscaled to the size of (W, H).
Output C_1 preprocessed COVID-19 image and C_2 preprocessed HC images.



(a) (b)

Figure 2: Example of preprocessed images (a) COVID-19 (b) HC

3 Methodology

3.1 Pseudo Zernike Moment

Tab. 1 displays the abbreviation list Image moment was firstly introduced by Hu [21], who used geometric moments to generate a set of invariants. Hu's moments have been widely used

in knee osteoarthritis classification [22], brain tumor classification [23], etc. However, geometric moments are sensitive to noise. Thus, Teague [24] introduced Zernike moments (ZMs) based on orthogonal Zernike polynomials. The orthogonal moments have been proven to be more robust in noisy conditions, and they can achieve a near-zero value of redundancy measure [25].

Table 1: Abbreviation list

Abbreviation	Meaning	Abbreviation	Meaning
AE	Autoencoder	HC	Healthy control
AUC	Area under the curve	IP	Image plane
CCT	Chest computed tomography	MCC	Matthews correlation coefficient
CM	Confusion matrix	ML	Machine learning
CXR	Chest X-ray	PZM	Pseudo Zernike moment
DL	Deep learning	ROC	Receiver operating characteristic
DSSAE	Deep stacked sparse autoencoder	UC	Unit circle
FMI	Fowlkes–Mallows index	ZM	Zernike moment

Later, pseudo Zernike moment (PZM) is derived from Zernike moment. PZMs have been proven to give better performances than other moment functions such as Hu moments, Zernike moments, etc. For example, for an order p, there are $(p+1)^2$ linearly independent pseudo-Zernike polynomials of orders $\leq p$, while there are only $\frac{1}{2}(p+1)(p+2)$ Zernike polynomials. Hence, PZM is more expressive and offers more feature vectors than ZM.

The kernel of PZMs is a set of orthogonal pseudo-Zernike polynomials defined over the polar coordinate inside a unit circle (UC). The 2D PZM of order p with repetition q of an image $g(r, \theta)$ is defined as [26]

$$Z_{pq} = \frac{p+1}{\pi} \int_{-\pi}^{\pi} \int_{0}^{1} W_{pq}^{*}(r, \theta)\, g(r, \theta)\, r\, \mathrm{d}r\, \mathrm{d}\theta, \tag{1}$$

where the pseudo-Zernike polynomials $W_{pq}(r, \theta)$ of order p are defined as

$$W_{pq}(r, \theta) = R_{pq}(r)\, e^{jq\theta}, \quad j = \sqrt{-1} \tag{2}$$

$$R_{pq}(r) = \sum_{k=0}^{p-|q|} (-1)^k \frac{(2p+1-k)!}{k!\,(p+|q|+1-k)!\,(p-|q|-k)!} r^{p-k} \tag{3}$$

where $0 \leq |q| \leq p$. In practice, pseudo Zernike functions (https://www.mathworks.com/matlabcentral/fileexchange/33644-pseudo-zernike-functions) are used for simplicity and fast calculation. Fig. 3 displays pseudo Zernike functions of orders $p \leq 5$.

Note that PZM are defined in terms of polar coordinates (r, θ) with $|r| \leq 1$. Therefore, the computation of PZM requires a linear transformation of the image plane (IP) coordinates (w, h), $1 \leq w \leq W$, $1 \leq h \leq H$ to the UC domain $(x, y) \in \mathbb{R}^2$. There are two commonly used transformations as shown in Fig. 4: (i) IP over UC; and (ii) IP inside UC. In this study, we use

the former (IP over UC), because the lesions will not occur within the four corners of the CCT image.

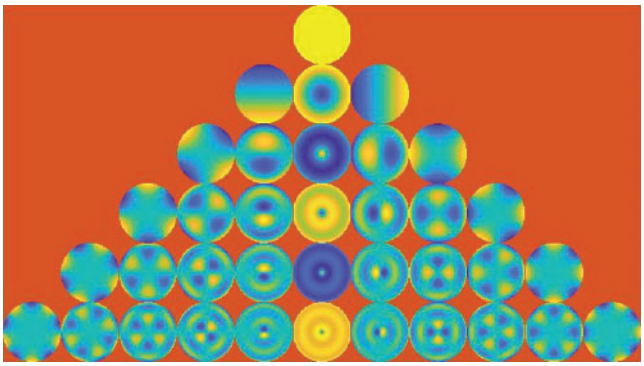


Figure 3: Pseudo Zernike functions of orders $p \leq 5$

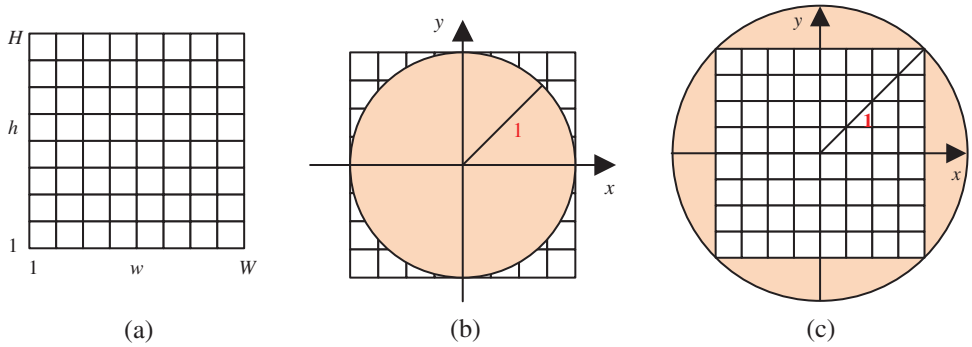


Figure 4: Two transformation (IP: image plane; UC: unit circle) (a) Raw image plane $W \times H$ (b) IP over UC (c) IP inside UC

3.2 Autoencoder

Traditionally, p-order PZMs are sent into shallow classifiers, such as multi-layer perceptron [27], adaptive differential evolution wavelet neural network (Ada-DEWNN) [28], linear regression classifier (LRC) [29], kernel support vector machine (KSVM) [30]. In this study, we introduced a customized deep-stacked sparse autoencoder (DSSAE). DSSAE is a type of deep neural network technologies, and we expect DSSAE to achieve better performances than shallow models.

The fundamental element of DSSAE in the autoencoder (AE), which is a typical shallow neural network that learns to map its input X to output Y. There is an internal code output I_N That represents the input X. The whole AE can be divided into two parts: An encoder part (A_X, B_X) that maps the input X to the code I_N, and a decoder part (A_Y, B_Y) that maps the code to a reconstructed data Y.

The structure of AE is displayed in Fig. 5, where the encoder part is with weight A_X and bias B_X, and the decoder part is with weights A_Y and bias B_Y. We have

$$I_N = z_{LS}(A_X X + B_X), \tag{4}$$

$$Y = z_{LS}(A_Y I_N + B_Y), \tag{5}$$

where the output Y is an estimate of input X, and z_{LS} is the log sigmoid function

$$z_{LS}(x) = \frac{1}{1+\exp(-x)}. \tag{6}$$

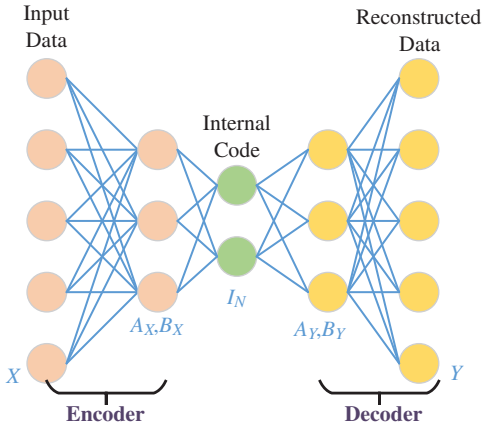


Figure 5: Structure of an AE

3.3 Sparse Autoencoder

The sparse autoencoder (SAE) is a variant of AE. SAE encourages sparsity into AE. SAE only allows a small fraction of the hidden neurons to be active at the same time. To minimize the error between the input vector X and the output Y, the raw loss function J_b of AE is deduced as:

$$J_b(A_X, A_Y, B_X, B_Y) = \frac{1}{N_S}\|Y - X\|^2, \tag{7}$$

where N_S means the number of training samples. From Eqs. (4) and (5), we find the output Y can be expressed in the way of

$$Y = z_{AE}(X \mid A_X, A_Y, B_X, B_Y), \tag{8}$$

where z_{AE} is the abstract of AE function [31]. Hence, Eq. (7) can be revised as

$$J_b(A_X, A_Y, B_X, B_Y) = \frac{1}{N_S}\|z_{AE}(X \mid A_X, A_Y, B_X, B_Y) - X\|^2. \tag{9}$$

To avoid over-complete mapping or learn a trivial mapping, we define one L_2 regularization term Γ_A of the weights (A_X, A_Y) and one regularization term Γ_s of the sparsity constraint. Therefore, the loss function J_l of SAE is derived as:

$$J_l(A_X, A_Y, B_X, B_Y) = \frac{1}{N_S} \|z_{AE}(X \mid A_X, A_Y, B_X, B_Y) - X\|^2 + a_s \times \Gamma_s + a_A \times \Gamma_A, \tag{10}$$

where a_s stands for the sparsity regulation factor, and a_A the weight regulation factor. The sparsity regularization term Γ_s is defined as:

$$\Gamma_s = \sum_{m=1}^{|I_N|} z_{KL}(\rho, \hat{\rho}_m) = \sum_{m=1}^{|I_N|} \rho \log \frac{\rho}{\hat{\rho}_m} + (1-\rho) \log \frac{1-\rho}{1-\hat{\rho}_m}, \tag{11}$$

where z_{KL} stands for the Kullback–Leibler divergence [32] function, $|I_N|$ is the number of elements of internal code output I_N, $\hat{\rho}_m$ is the m-th neuron's average activation value over all N_S training samples, and ρ is its desired value, viz., sparsity proportion factor. The weight regularization term Γ_A is defined as

$$\Gamma_A = \frac{1}{2} \times \left\| A_X \quad A_Y \right\|_2^2. \tag{12}$$

The training procedure is set to scaled conjugate gradient descent (SCGD) method.

3.4 Deep Stacked Sparse Autoencoder

We use SAE as the building block and establish the final deep-stacked sparse autoencoder (DSSAE) classifier by following three operations: (i) We include input layer, preprocessing layer, PZM layer; (ii) We stack four SAEs; (iii) We append softmax layer at the bottom of our AI model. The details of this proposed PZM-DSSAE model are listed in Tab. 2 and illustrated in Fig. 6. After processing, all the CCT images are normalized to fixed grayscaled images with the size of $W \times H$. Then, PZM is applied to obtain feature vector with size of $(p+1)^2 \times 1$. In the classification stage, four SAE blocks with number of neurons of (M_1, M_2, M_3, M_4) are employed. Finally, a softmax layer with neurons of M_c is appended, where M_c means the number of categories to be identified.

Table 2: Layer details of proposed PZM-DSSAE model

Layer	Trainable weights	Size
Input	None	$1024 \times 1024 \times 3$
Preprocessing	None	$W \times H$
PZM	None	$(p+1)^2 \times 1$
1st SAE	$(p+1)^2 \times M_1$	M_1
2nd SAE	$M_1 \times M_2$	M_2
3rd SAE	$M_2 \times M_3$	M_3
4th SAE	$M_3 \times M_4$	M_4
Softmax	$M_4 \times M_c$	$M_c \times 1$

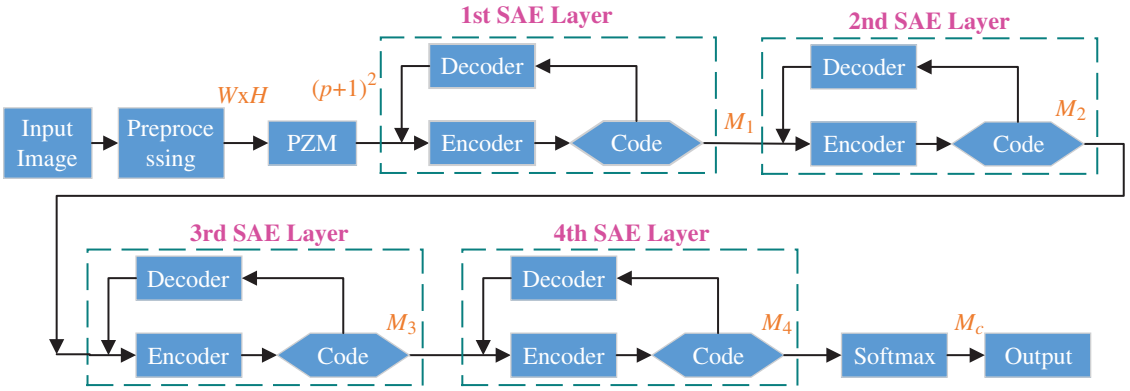


Figure 6: Structure of proposed PZM-DSSAE model

3.5 18-Way Data Augmentation

The small size of training images causes overfitting, one solution to data augmentation (DA) that creates fake training images. Multiple-way DA (MDA) is an enhanced method of DA. Wang [33] proposed a 14-way data augmentation, in which they employed seven different DA techniques on k-th training image $g(k)$ and its mirrored image $g^{(m)}(k)$.

In this study, we add two new DA techniques, speckle noise (SN) [34] and salt-and-pepper noise (SAPN). SN altered image is defined as

$$g^{SN}(k) = g(k) + N_{SN} * g(k), \tag{13}$$

where N_{SN} is uniformly distributed random noise. The mean and variance of N_{SN} is set to m_{SN} and v_{SN}, respectively.

For the k-th training image $g(k)$, the SAPN altered image [35] is defined as $g^{SAPN}(k)$ with its values are set as

$$\begin{cases} \mathrm{P}\left(g^{SAPN} = g\right) = 1 - \gamma_d^{sapn}, \\ \mathrm{P}\left(g^{SAPN} = g_{\min}\right) = \dfrac{\gamma_d^{sapn}}{2}, \\ \mathrm{P}\left(g^{SAPN} = g_{\max}\right) = \dfrac{\gamma_d^{sapn}}{2}, \end{cases} \tag{14}$$

where γ_d^{sapn} stands for noise density, and P the probability function. $g_{\min}$ and $g_{\max}$ correspond to black and white colors, respectively. The definitions of g_{min} and g_{max} can be found in Algorithm 1.

First, Q^D different DA methods as shown in Fig. 7 are applied to $g(k)$. Let $\mathbf{H}_m, m = 1, \ldots, Q^D$ denote each DA operation, we have the augmented dataset on raw image $g(k)$ as:

$$\mathbf{H}_m[g(k)], \quad m = 1, \ldots, Q^D. \tag{15}$$

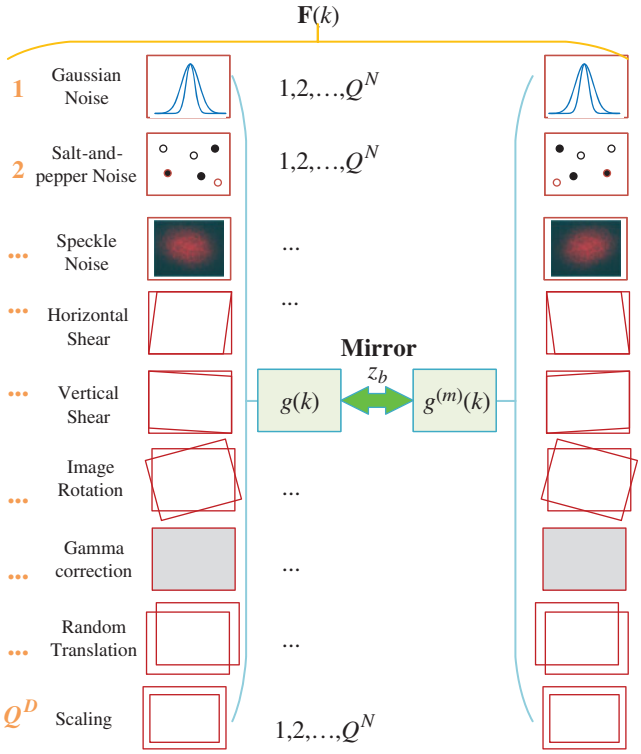


Figure 7: Diagram of proposed 16-way DA

Suppose Q^N stands for the size of generated new images for each DA method, we have

$$|\mathbf{H}_m[g(k)]| = Q^N. \tag{16}$$

Second, horizontal mirrored image is generated as:

$$g^{(m)}(k) = z_b[g(k)] \tag{17}$$

where z_b stands for horizontal mirror function.

Third, all the Q^D different DA methods are performed on the mirror image $p_c(k)$, and generate Q^D different dataset.

$$\begin{cases} \mathbf{H}_m\left[g^{(m)}(k)\right], & m = 1, \ldots, Q^D \\ \left|\mathbf{H}_m\left[g^{(m)}(k)\right]\right| = Q^N, & m = 1, \ldots, Q^D \end{cases} \tag{18}$$

Fourth, the raw image $g(k)$, the mirrored image $g^{(m)}(k)$, all the above Q^D-way results of raw image $\mathbf{H}_m[g(k)]$, and Q^D-way DA results of horizontal mirrored image $\mathbf{H}_m\left[g^{(m)}(k)\right]$, are

combined together. The final generated dataset from $g(k)$ is defined as $\mathbf{F}(k)$:

$$g(k) \mapsto \mathbf{F}(k) = z_a \begin{Bmatrix} g(k) & g^{(m)}(k) \\ \underbrace{\mathbf{H}_1[g(k)]}_{Q^N} & \underbrace{\mathbf{H}_1\left[g^{(m)}(k)\right]}_{Q^N} \\ \cdots & \cdots \\ \underbrace{\mathbf{H}_{Q^D}[g(k)]}_{Q^N} & \underbrace{\mathbf{H}_{Q^D}\left[g^{(m)}(k)\right]}_{Q^N} \end{Bmatrix} \tag{19}$$

where z_a stands for the concatenation function. Suppose augmentation factor is Q^A, which stands for the number of images in $\mathbf{F}(k)$, we have

$$Q^A = \frac{|\mathbf{F}(k)|}{|g(k)|} = \frac{\left(1 + Q^D \times Q^N\right) \times 2}{1} = 2 \times Q^D \times Q^N + 2 \tag{20}$$

Algorithm 2 summarizes the pseudocode of proposed 18-way DA method.

Algorithm 2: Pseudocode of proposed 18-way data augmentation on k-th training image

Step 1 Import raw preprocessed k-th training image $g(k)$.
Step 2 Q^D geometric or photometric or noise-injection DA transforms $\mathbf{H}_m$ are utilized on $g(k)$. We obtain $\mathbf{H}_m[g(k)], m = 1, \ldots, Q^D$. See Eq. (15).
Each enhanced dataset contains Q^N new images. See Eq. (16).
Step 3 A horizontal mirror image is generated as $g^{(m)}(k) = z_b[g(k)]$. See Eq. (17).
Step 4 Q^D-way data augmentation methods are carried out on $g^{(m)}(k)$, we obtain $\mathbf{H}_m\left[g^{(m)}(k)\right], m = 1, \ldots, Q^D$. See Eq. (18).
Step 5 The raw image, the mirrored image, all the above Q^D-way DA results of raw image, and Q^D-way DA results of a horizontal mirrored image are combined via z_a. See Eq. (19).
Step 6 A new dataset $\mathbf{F}(k)$ is generated with number of images as $Q^A = 2 \times Q^D \times Q^N + 2$. See Eq. (20).

3.6 Cross-Validation

F-fold cross-validation was used in this study. The whole dataset is divided into F folds. At f-th trial, $1 \le f \le F$, the f-th fold is selected as the test, and the rest $F - 1$ folds [36]: $[1, \ldots, f-1, f+1, \ldots, F]$ are selected as training set (Fig. 8). In this study, suppose $F = 10$, then each fold will contain 32 COVID-19 images and 32 HC images.

3.7 Evaluation

To avoid randomness, we run the whole above procedure N^R times with different initial random seeds and different cross-validation partitions. The ideal confusion matrix (CM) R^{ideal} is defined as

$$R^{ideal} = \left\{r^{ideal}(m, n)\right\} = N^R \times \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix}, \tag{21}$$

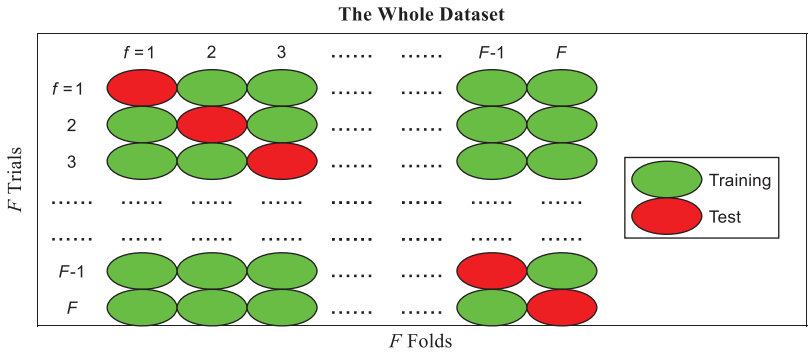


Figure 8: F-fold cross validation

Note here the off-diagonal entries of R^{ideal} are all zero, viz., $r^{ideal}(m, n) = 0, \forall m \neq n$. C_1 and C_2 are the number of samples of each category, which can be found in Algorithm 1. Seven measures are defined based on realistic CM [37] defined as:

$$R^{real} = \left\{ r^{real}(m, n) \right\} = \begin{bmatrix} TP & FN \\ FP & TN \end{bmatrix}, \tag{22}$$

The first four measures are sensitivity, specificity, precision and accuracy, common in most pattern recognition papers. The last three measures are F1 score, Matthews correlation coefficient (MCC) [38], and Fowlkes–Mallows index (FMI) [39]. They are defined as:

$$\begin{cases} F1 = \dfrac{2TP}{2TP + FP + FN} \\ MCC = \dfrac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}}, \\ FMI = \sqrt{\dfrac{TP}{TP + FP} \times \dfrac{TP}{TP + FN}} \end{cases} \tag{23}$$

Besides, the receiver operating characteristic (ROC) curve [40] is used to provide a graphical plot of our model. ROC curve is created by plotting the true positive rate against the false-positive rate at various threshold settings. The area under the curve (AUC) is also calculated.

4 Experimental Results

4.1 Parameter Setting

Tab. 3 displays the parameter setting of this study. The number of samples of each class is 320. The minimum and maximum grayscale values are set to (0, 255). For the crop operation, 200 pixels are removed from all four sides. The preprocessed image is with size of 256×256. The max order of PZM is set to 19, so we have $(19 + 1)^2 = 400$ PZM features. The weight regularization factor $a_A = 0.001$, the sparsity regulation factor $a_s = 1.1$, and the sparsity proportion factor is $\rho = 0.05$. The neurons of four SAEs are 300, 200, 100, and 50, respectively. The number of classes to be classified is set to 2. The number of folds in cross-validation is set to 10. The mean and variance of uniformly distributed random noise in SN are set to 0 and 0.05, respectively. The noise

density of SAPN is set to 0.05. The number of different DA methods is set to 9, and the number of the newly generated image is set to 30. The augmentation factor is obtained as $Q^A = 542$ (See Algorithm 2). The number of runs is set to 10.

Table 3: Parameter setting

Parameter	Value	Parameter	Value	Parameter	Value
(C_1, C_2)	(320, 320)	ρ	0.05	m_{SN}	0
(g_{min}, g_{max})	(0, 255)	M_1	300	v_{SN}	0.05
$(m_1^c, m_2^c, m_3^c, m_4^c)$	(200, 200, 200, 200)	M_2	200	γ_d^{sapn}	0.05
(W, H)	(256, 256)	M_3	100	Q^D	9
p	19	M_4	50	Q^N	30
a_A	0.001	M_c	2	Q^A	542
a_s	1.1	F	10	N^R	10

4.2 Illustration of 18-Way Data Augmentation

Fig. 9 shows the Q^D-way DA to the raw image. Due to the page limit, the mirrored image and its corresponding DA results are not displayed. As can be observed in Fig. 9, the multiple-way DA can increase our training images' diversity.

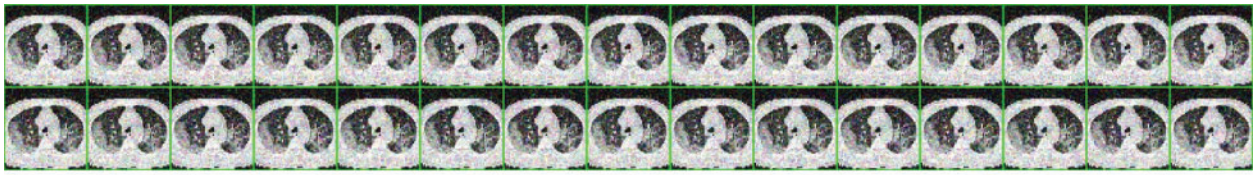
4.3 Statistical Analysis and Transformation Comparison

Tab. 4 gives the 10 runs of 10-fold cross-validation, where we can see our method achieves a sensitivity of 92.06% ± 1.54%, a specificity of 92.56% ± 1.06%, a precision of 92.53% ± 1.03%, and an accuracy of 92.31% ± 1.08%. Its F1 score, MCC, and FMI arrive at 92.29% ± 1.10%, 84.64% ± 2.15%, and 92.29% ± 1.10%, respectively. The AUC is 0.9576.

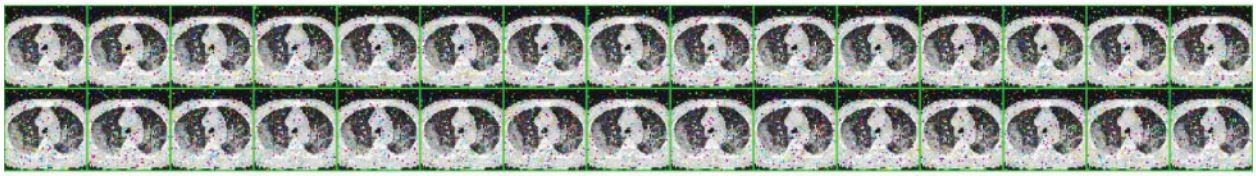
In addition, we compared the two transformation settings: IP over UC against IP inside UC (See Fig. 4). The IP inside the UC setting achieves a sensitivity of 91.84% ± 2.18%, a specificity of 92.44% ± 1.31%, and an accuracy of 92.14% ± 1.12%, which are worse than IP over UC setting. This comparison result demonstrates the reason why we choose IP over UC in this study. Particularly, the receiver operating characteristics (ROC) curves of both settings are displayed in Fig. 10.

4.4 Comparison to State-of-the-Art Methods

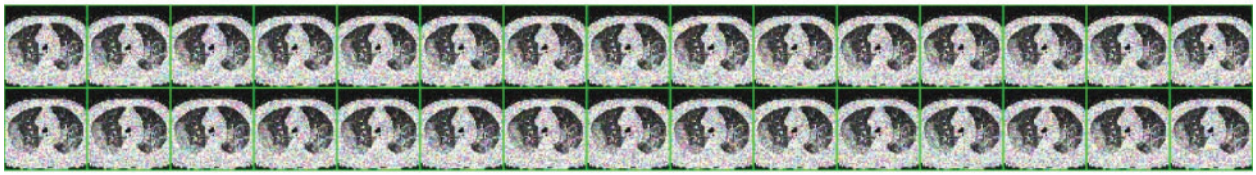
This proposed PZM-DSSAE method is compared with 8 state-of-the-art methods. The comparison results are carried out on the same dataset via 10 runs of 10-fold cross-validation, and the results are displayed in Tab. 5. Fig. 11 displays the error bar of the proposed method against 8 state-of-the-art methods. We can see that the proposed PZM-DSSAE gives the best performance among all the methods. The reason is three folds: (i) We try to use PZM as the feature descriptors, (ii) DSSAE is used as the classifier, (iii) 18-way DA is employed to solve the overfitting problem.



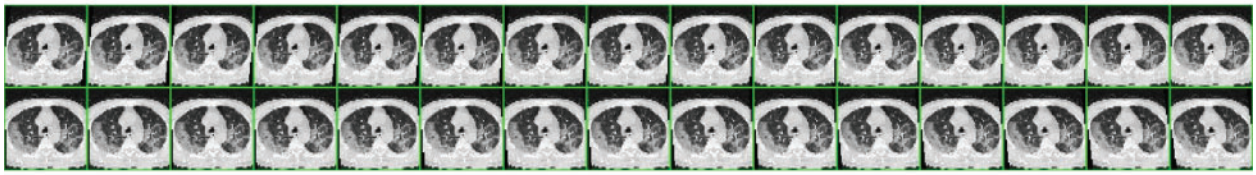
(a)



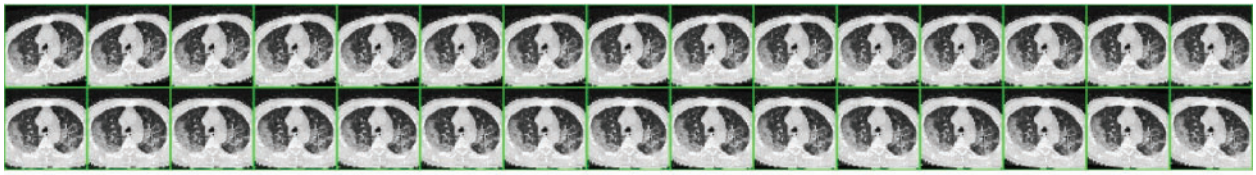
(b)



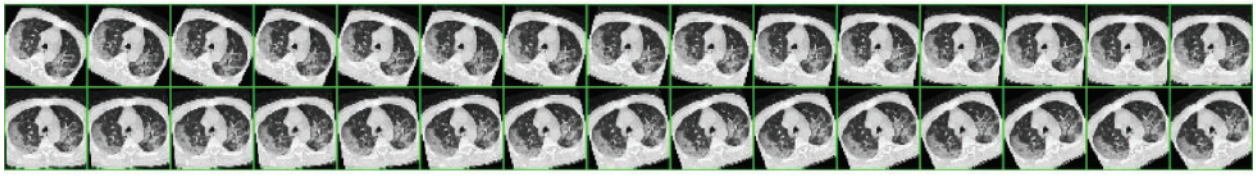
(c)



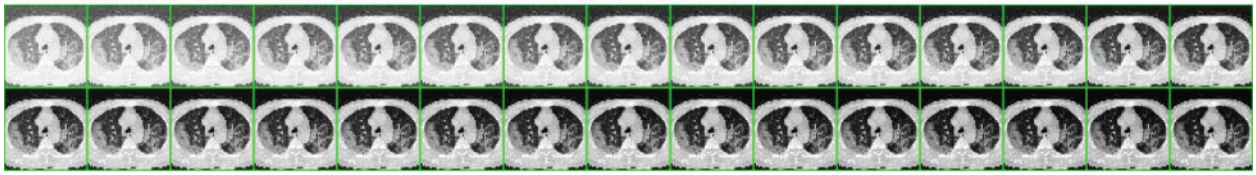
(d)



(e)



(f)



(g)

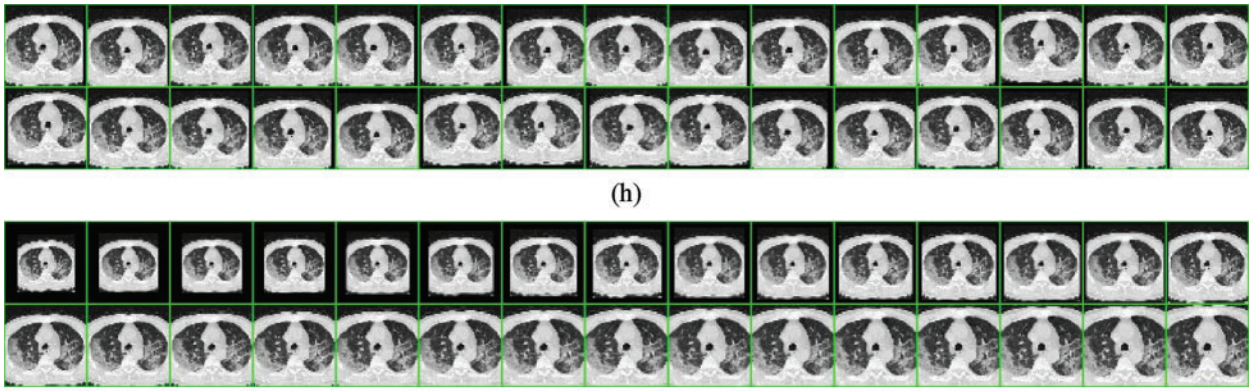
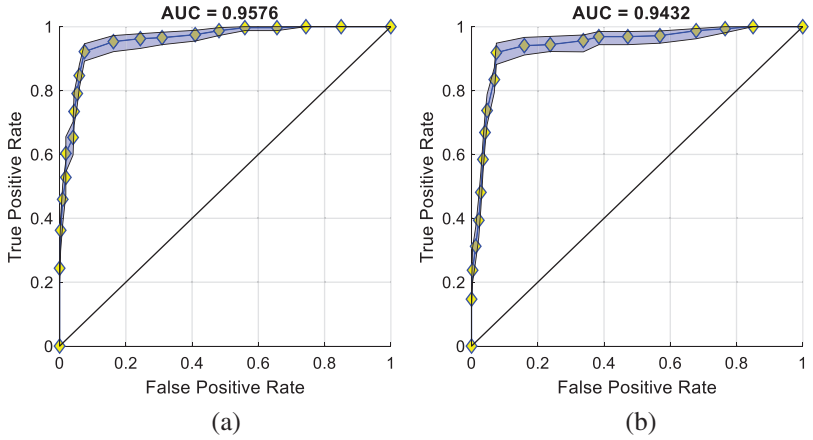


Figure 9: Q^D-way DA results of raw image (a) Gaussian noise (b) SAPN (c) SN (d) Horizontal shear (e) Vertical shear (f) Rotation (g) Gamma correction (h) Random translation (i) Scaling

Table 4: 10 Runs of statistical analysis of proposed PZM-DSSAE method

Run	Sen	Spc	Prc	Acc	F1	MCC	FMI
1	90.94	91.88	91.80	91.41	91.37	82.82	91.37
2	91.56	91.56	91.56	91.56	91.56	83.13	91.56
3	91.88	92.50	92.45	92.19	92.16	84.38	92.16
4	92.19	91.56	91.61	91.88	91.90	83.75	91.90
5	91.25	92.19	92.11	91.72	91.68	83.44	91.68
6	90.63	92.19	92.06	91.41	91.34	82.82	91.34
7	93.75	93.13	93.17	93.44	93.46	86.88	93.46
8	95.31	94.38	94.43	94.84	94.87	89.69	94.87
9	92.81	91.88	91.95	92.34	92.38	84.69	92.38
10	90.31	94.38	94.14	92.34	92.19	84.76	92.20
MSD	92.06 ± 1.54	92.56 ± 1.06	92.53 ± 1.03	92.31 ± 1.08	92.29 ± 1.10	84.64 ± 2.15	92.29 ± 1.10



(a) (b)

Figure 10: ROC curves of two settings (a) IP over UC (b) IP inside UC

Table 5: Comparison to state-of-the-art methods

Method	Sen	Spc	Prc	Acc	F1	MCC	FMI
ResNet-18 [11]	78.88 ± 2.57	89.28 ± 0.90	88.05 ± 0.79	84.08 ± 1.14	83.19 ± 1.44	68.55 ± 2.10	83.32 ± 1.36
ELM-BA [12]	56.91 ± 1.21	71.94 ± 2.17	67.01 ± 1.52	64.42 ± 0.88	61.53 ± 0.77	29.19 ± 1.88	61.74 ± 0.77
WEBBO [13]	72.94 ± 0.96	73.97 ± 1.02	73.70 ± 0.79	73.45 ± 0.69	73.31 ± 0.71	46.91 ± 1.38	73.32 ± 0.71
3SBBO [14]	85.94 ± 1.68	84.75 ± 2.42	84.96 ± 2.16	85.34 ± 1.81	85.44 ± 1.74	70.71 ± 3.61	85.44 ± 1.73
DeCovNet [15]	90.03 ± 1.22	90.34 ± 1.25	90.33 ± 1.07	90.19 ± 0.68	90.17 ± 0.69	80.39 ± 1.35	90.18 ± 0.68
FSVC [16]	90.25 ± 1.27	90.03 ± 0.80	90.06 ± 0.72	90.14 ± 0.70	90.15 ± 0.73	80.29 ± 1.41	90.15 ± 0.74
GoogleNet-COD [17]	89.44 ± 1.59	82.91 ± 1.64	83.98 ± 1.16	86.17 ± 0.67	86.61 ± 0.68	72.53 ± 1.32	86.66 ± 0.68
GLCMSVM [18]	72.38 ± 2.68	77.38 ± 1.96	76.22 ± 1.21	74.88 ± 0.86	74.21 ± 1.25	49.85 ± 1.70	74.25 ± 1.21
PZM-DSSAE (Ours)	92.06 ± 1.54	92.56 ± 1.06	92.53 ± 1.03	92.31 ± 1.08	92.29 ± 1.10	84.64 ± 2.15	92.29 ± 1.10

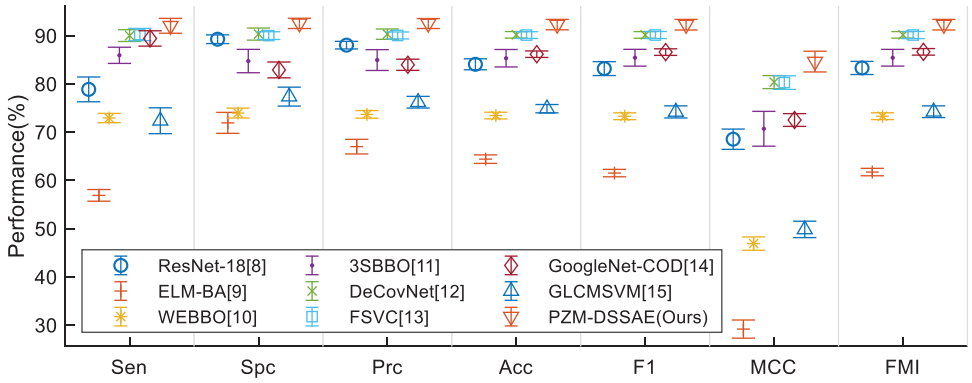


Figure 11: Error bar plot of method comparison

5 Conclusion

This study proposed a novel PZM-DSSAE system for COVID-19 diagnosis. As far as the authors' best known, we are the first to apply PZM to COVID-19 image analysis. Also, two other improvements are carried out: (i) DSSAE is used as the classifier, and (ii) multiple-way data augmentation is employed to generalize the classifier. Our model yields a sensitivity of 92.06% ± 1.54%, a specificity of 92.56% ± 1.06%, an accuracy of 92.31% ± 1.08%, and an AUC of 0.9576.

In the future, we shall collect more COVID-19 images from more patients and multiple modalities. Also, other advanced AI models will be tested, such as graph neural networks and attention networks.

Funding Statement: This study was supported by Royal Society International Exchanges Cost Share Award, UK (RP202G0230); Medical Research Council Confidence in Concept Award, UK (MC_PC_17171); Hope Foundation for Cancer Research, UK (RM60G0680); Global Challenges Research Fund (GCRF), UK (P202PF11)

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] B. Zemrani, M. Gehri, E. Masserey, C. Knob and R. Pellaton, "A hidden side of the COVID-19 pandemic in children: The double burden of undernutrition and overnutrition," *International Journal for Equity in Health*, vol. 20, no. 1, pp. 4, 2021.

[2] M. Bassi, L. Negri, A. Delle Fave and R. Accardi, "The relationship between post-traumatic stress and positive mental health symptoms among health workers during COVID-19 pandemic in Lombardy," *Italy Journal of Affective Disorders*, vol. 280, pp. 1–6, 2021.

[3] D. Simiies, A. R. Stengaard, L. Combs, D. Raben and T. C. I. A. Euro, "Impact of the COVID-19 pandemic on testing services for HIV, viral hepatitis and sexually transmitted infections in the WHO European Region, march to august 2020," *Eurosurveillance*, vol. 25, pp. 7, 2020.

[4] F. Salahshour, M.-M. Mehrabinejad, M. N. Toosi, M. Gity, H. Ghanaati *et al.,* "Clinical and chest CT features as a predictive tool for COVID-19 clinical progress: Introducing a novel semi-quantitative scoring system," *European Radiology*, vol. 11, pp. 1–11, 2021.

[5] A. M. Ismael and A. Sengur, "The investigation of multiresolution approaches for chest X-ray image based COVID-19 detection," *Health Information Science and Systems*, vol. 8, no. 1, pp. 1–13, 2020.

[6] A. Hata, M. Yanagawa, Y. Yoshida, T. Miyata, N. Kikuchi *et al.,* "The image quality of deep-learning image reconstruction of chest CT images on a mediastinal window setting," *Clinical Radiology*, vol. 76, no. 2, pp. 9, 2021.

[7] Y. Satoh, U. Motosugi, M. Imai, Y. Omiya and H. Onishi, "Evaluation of image quality at the detector's edge of dedicated breast positron emission tomography," *Ejnmmi Physics*, vol. 8, no. 1, pp. 14, 2021.

[8] M. A. Khan, S. Kadry, Y.-D. Zhang, T. Akram, M. Sharif *et al.,* "Prediction of COVID-19-pneumonia based on selected deep features and one class kernel extreme learning machine," *Computers & Electrical Engineering*, vol. 90, pp. 106960, 2021.

[9] T. Akram, M. Attique, S. Gul, A. Shahzad, M. Altaf *et al.,* "A novel framework for rapid diagnosis of COVID-19 on computed tomography scans," *Pattern Analysis and Applications*, pp. 1–14, 2021. https://doi.org/10.1007/s10044-020-00950-0.

[10] H. T. Rauf, M. I. U. Lali, M. A. Khan, S. Kadry, H. Alolaiyan *et al.,* "Time series forecasting of COVID-19 transmission in Asia Pacific countries using deep neural networks," *Personal and Ubiquitous Computing*, vol. 8, pp. 1–18, 2021.

[11] M. H. Guo and Y. Z. Du, "Classification of thyroid ultrasound standard plane images using ResNet-18 networks," in *IEEE 13th Int. Conf. on Anti-Counterfeiting, Security, and Identification*, Xiamen, China, pp. 324–328, 2019.

[12] S. Lu, "A pathological brain detection system based on extreme learning machine optimized by Bat algorithm," *CNS & Neurological Disorders-Drug Targets*, vol. 16, no. 1, pp. 23–29, 2017.

[13] X. Yao, "COVID-19 detection via wavelet entropy and biogeography-based optimization," in *COVID-19: Prediction, Decision-Making, and Its Impacts*, K. C. Santosh, A. Joshi (Eds.), Cham: Springer, pp. 69–76, 2020.

[14] X. Wu, "Diagnosis of COVID-19 by wavelet renyi entropy and three-segment biogeography-based optimization," *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, pp. 1332–1344, 2020.

[15] X. G. Wang, X. B. Deng, Q. Fu, Q. Zhou, J. P. Feng *et al.,* "A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2615–2625, 2020.

[16] E. S. M. El-kenawy, A. Ibrahim, S. Mirjalili, M. M. Eid and S. E. Hussein, "Novel feature selection and voting classifier algorithms for COVID-19 classification in CT images," *IEEE Access*, vol. 8, pp. 179317–179335, 2020.

[17] X. Yu, S.-H. Wang, X. Zhang and Y.-D. Zhang, "Detection of COVID-19 by GoogLeNet-COD," in *Int. Conf. on Intelligent Computing*, Bari, Italy, pp. 499–509, 2020.

[18] Y. Chen, "Covid-19 classification based on gray-level co-occurrence matrix and support vector machine," in *COVID-19: Prediction, Decision-Making, and Its Impacts*, K. C. Santosh, A. Joshi (Eds.), Singapore: Springer, Cham, pp. 47–55, 2020.
[19] M. A. Khan, N. Hussain, A. Majid, M. Alhaisoni, S. A. C. Bukhari *et al.,* "Classification of positive COVID-19 CT scans using deep learning," *Computers, Materials and Continua*, vol. 66, no. 2, pp. 1–15, 2021.
[20] Y. D. Zhang, "A seven-layer convolutional neural network for chest CT based COVID-19 diagnosis using stochastic pooling," *IEEE Sensors Journal*, vol. 1, pp. 1, 2021.
[21] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, pp. 179–187, 1962.
[22] S. S. Gornale, P. U. Patravali and P. S. Hiremath, "Automatic detection and classification of knee osteoarthritis using Hu's invariant moments," *Frontiers in Robotics and AI*, vol. 7, pp. 8, 2020.
[23] S. Ouchtati, A. Chergui, S. Mavromatis, B. Aissa and J. Sequeira, "Novel method for brain tumor classification based on use of image entropy and seven Hu's invariant moments," *Traitement Du Signal*, vol. 36, no. 6, pp. 483–491, 2019.
[24] M. R. Teague, "Image analysis via the general theory of moments," *Journal Optical Soceity America*, vol. 70, no. 8, pp. 920–930, 1980.
[25] A. Grosso and T. Scharf, "Scalar analytical expressions for the field dependence of Zernike polynomials in asymmetric optical systems with circular symmetric surfaces," *OSA Continuum*, vol. 3, no. 10, pp. 2749–2765, 2020.
[26] C. W. Chong, P. Raveendran and R. Mukundan, "The scale invariants of pseudo-Zernike moments," *Pattern Analysis and Applications*, vol. 6, no. 3, pp. 176–184, 2003.
[27] N. Borzue, K. Faez and IEEE, "Object contour detecting using pseudo zernike moment and multi-layer perceptron," in *22nd Iranian Conf. on Biomedical Engineering*, Tehran, Iran, pp. 304–308, 2015.
[28] S. P. Singh and S. Urooj, "An improved CAD system for breast cancer diagnosis based on generalized pseudo zernike moment and Ada-DEWNN classifier," *Journal of Medical Systems*, vol. 40, no. 4, pp. 1–23, 2016.
[29] S. Du, "Alzheimer's disease detection by pseudo zernike moment and linear regression classification," *CNS & Neurological Disorders-Drug Targets*, vol. 16, no. 1, pp. 11–15, 2017.
[30] Y. Jiang, "Exploring a smart pathological brain detection method on pseudo Zernike moment," *Multimedia Tools and Applications*, vol. 77, no. 17, pp. 22589–22604, 2018.
[31] D. K. Mohanty, A. K. Parida and S. Suman, "Financial market prediction under deep learning framework using auto encoder and kernel extreme learning machine," *Applied Soft Computing*, vol. 99, no. 3, pp. 14, 2021.
[32] J. Lequesne and P. Regnault, "vsgoftest: An R package for goodness-of-fit testing based on Kullback–Leibler divergence," *Journal of Statistical Software*, vol. 96, no. Code Snippet 1, pp. 26, 2020.
[33] S.-H. Wang, "Covid-19 classification by FGCNet with deep feature fusion from graph convolutional network and convolutional neural network," *Information Fusion*, vol. 67, pp. 208–229, 2021.
[34] C. Buitrago-Duque, R. Castaneda and J. Garcia-Sucerquia, "Pointwise phasor tuning for single-shot speckle noise reduction in phase wave fields," *Optics and Lasers in Engineering*, vol. 137, pp. 5, 2021.
[35] K. Vasanth and R. Varatharajan, "An adaptive content based closer proximity pixel replacement algorithm for high density salt and pepper noise removal in images," *Journal of Ambient Intelligence and Humanized Computing*, vol. 3, pp. 1–15, 2020.
[36] D. Muller and F. Kramer, "MIScnn: A framework for medical image segmentation with convolutional neural networks and deep learning," *BMC Medical Imaging*, vol. 21, pp. 11, 2021.
[37] J. Radoux and P. Bogaert, "About the pitfall of erroneous validation data in the estimation of confusion matrices," *Remote Sensing*, vol. 12, pp. 23, 2020.
[38] D. Dreizin, F. Goldmann, C. LeBedis, A. Boscak, M. Dattwyler *et al.,* "An automated deep learning method for tile AO/OTA pelvic fracture severity grading from trauma whole-body CT," *Journal of Digital Imaging*, vol. 21, pp. 1–13, 2021.

[39] R. Jena, B. Pradhan and A. M. Alamri, "Susceptibility to seismic amplification and earthquake probability estimation using recurrent neural network model in Odisha," *India Applied Sciences-Basel*, vol. 10, pp. 18, 2020.

[40] A. T. Young, K. Fernandez, J. Pfau, R. Reddy, N. A. Cao *et al.,* "Stress testing reveals gaps in clinic readiness of image-based diagnostic artificial intelligence models," *NPJ Digital Medicine*, vol. 4, pp. 8, 2021.