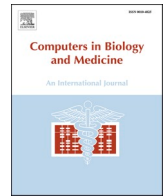




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

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.



Learning-to-augment strategy using noisy and denoised data: Improving generalizability of deep CNN for the detection of COVID-19 in X-ray images

Mohammad Momeny^{a,*}, Ali Asghar Neshat^{b,**}, Mohammad Arafat Hussain^c, Solmaz Kia^d, Mahmoud Marhamati^e, Ahmad Jahanbakhshi^f, Ghassan Hamarneh^c

^a Department of Computer Engineering, Yazd University, Yazd, Iran

^b Department of Environmental Engineering, Esfarayen Faculty of Medical Science, Esfarayen, Iran

^c School of Computing Science, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada

^d Department of Engineering Science, Faculty of Advanced Technologies, University of Mohaghegh Ardabili, Namin, Iran

^e Department of Medical-Surgical Nursing, Esfarayen Faculty of Medical Science, Esfarayen, Iran

^f Department of Biosystems Engineering, University of Mohaghegh Ardabili, Ardabil, Iran

ARTICLE INFO

Keywords:

Learning-to-augment
Data augmentation
Noise
X-ray images
Classification
COVID-19
Deep learning

ABSTRACT

Chest X-ray images are used in deep convolutional neural networks for the detection of COVID-19, the greatest human challenge of the 21st century. Robustness to noise and improvement of generalization are the major challenges in designing these networks. In this paper, we introduce a strategy for data augmentation using the determination of the type and value of noise density to improve the robustness and generalization of deep CNNs for COVID-19 detection. Firstly, we present a learning-to-augment approach that generates new noisy variants of the original image data with optimized noise density. We apply a Bayesian optimization technique to control and choose the optimal noise type and its parameters. Secondly, we propose a novel data augmentation strategy, based on denoised X-ray images, that uses the distance between denoised and original pixels to generate new data. We develop an autoencoder model to create new data using denoised images corrupted by the Gaussian and impulse noise. A database of chest X-ray images, containing COVID-19 positive, healthy, and non-COVID pneumonia cases, is used to fine-tune the pre-trained networks (AlexNet, ShuffleNet, ResNet18, and GoogleNet). The proposed method performs better results compared to the state-of-the-art learning to augment strategies in terms of sensitivity (0.808), specificity (0.915), and F-Measure (0.737). The source code of the proposed method is available at <https://github.com/mohadmomeny/Learning-to-augment-strategy>.

1. Introduction

1.1. COVID-19

As a worldwide pandemic, Coronavirus disease 2019 (COVID-19) has brought about a health crisis affecting all aspects of human life, and much effort has been being made to contain the virus since its inception. In the beginning, there were not many people who had contracted the disease, and it was not considered a great threat as most cases were treated in a short time. After a while, the World Health Organization (WHO) declared that the virus had an extreme potential to affect millions of individuals all around the world, especially in countries that had weaker healthcare systems. The disease is easily transmitted through

direct or indirect contact to the affected person [1]. The coronavirus statistics are horrifying [2].

The United States (US) has recorded one of the largest numbers of COVID-19 victims, though it is one of the leading countries in healthcare facilities. Brazil, India, Russia, South Africa, and 215 other countries around the world follow the US on the list. Many governments and administrative authorities across the globe are still imposing non-compromising lockdown restriction to ensure social distancing for the containment of the disease due to the ever-increasing number of new cases every single day [2,3]. According to the WHO and the US Center for Disease Control (CDC), fever, dry cough, vomiting, diarrhea, and myalgia are the most common symptoms of COVID-19 infection. To reduce morbidity rates, the general population in all countries has been

* Corresponding author.

** Corresponding author.

E-mail addresses: mohamad.momeny@gmail.com (M. Momeny), en.neshat@gmail.com (A.A. Neshat).

Table 1
Summary of patients' age, sex, and diagnosis.

Items	Value	
Number of images	1248	
Sex	Male	512
	Female	702
	Not available	34
Age	Available	1205
	Not available	43
	Years (Mean \pm Standard deviation)	48.1 \pm 17.5
Diagnosis	COVID-19	215
	Non-COVID pneumonia	533
	Healthy	500

made conscious of the symptoms so that they can seek treatment as soon as possible. Governments have begun to invest in COVID-19 vaccines and related research, and thus many studies and development activities are being conducted about the COVID-19 pandemic.

1.2. Detection of COVID-19 via X-ray images

Chest X-ray imaging has been playing a vital role in the rapid diagnosis and early management of COVID-19 [4,5]. It is reportedly used for COVID-19 detection in countries with a shortage of testing kits [6–8]. Recent studies [9–14] using machine learning (ML) and deep learning (DL) have shown promising results in the diagnosis of COVID-19. For example, convolutional neural networks (CNNs) have been applied for the classification of X-ray images [9–11] among COVID-19, non-COVID pneumonia (e.g., bacterial and viral pneumonia) and healthy cases.

1.3. Data augmentation using noise

The quality of a chest X-ray image may be deteriorated due to different types of noise generated by a malfunction in X-ray receiver sensors, bit errors in transmission, and faulty memory locations in the hardware [15]. Typically, noise-based data augmentation in DL is performed when there is a possibility of image data being corrupted by noise [16]. Data augmentation by noise addition is a strategy that improves the robustness and generalization of CNNs [17–24]. Moreno-Barea et al. [21] tested the noise injection to images from a Gaussian distribution, and showed it to be useful for improving CNN-based classification performance [23,24]. Sezer and Sezer [25] proposed a data augmentation approach, where CNN-based speckle-noise reduction is used in the neonatal hip ultrasound image classification task. Their work introduced a method employing an optimized Bayesian non-local mean filter to reduce speckle noise for data augmentation. This data augmentation strategy improved the performance of the CNN from 92.29% to 97.70%. Ofori-Oduro and Amer [20] proposed a noise-robust CNN via data augmentation using Artificial Immune System. Their noise-robust model was tested under noise and shown to be useful for improving the performance of a CNN. In our previous work [23], we proposed a noise-based adaptive data augmentation method to increase the CNN accuracy.

1.4. Current learning-to-augment strategies

Unlike conventional data augmentation approaches that use predefined rules and procedures for a specific target task, learning-to-augment strategies dynamically refine augmentation rules based on feedback network [26–28]. For example, Wang et al. [26] introduced an end-to-end compositional generative adversarial network architecture to generate natural and accurate face images. This model generates images of desired expressions, and edits the poses of faces. A reconstruction learning process was employed to re-generate the input data. The generators of the model encourage for preserving the important features of face. The augmented face images were used to train a robust

expression recognition model. Cai et al. [27] introduced a fully data-driven and learnable framework to change the data distribution near reliable samples. This data augmentation method selects efficient learning samples and reduces the impact of ineffective samples. Feng et al. [28] introduced an approach that generates new data obtained from a stationary distribution near the target data and implements a reinforced selector to automatically improve the augmentation strategy. However, these state-of-the-art methods [24–26] lack weak generalization ability under noisy conditions, which can cause overfitting.

1.5. Proposed data augmentation strategy

This paper focuses on the generation of noisy and denoised chest X-ray images as augmented images to improve the generalization of deep CNNs for COVID-19 detection. We propose a noise-based novel data augmentation approach, where our method optimizes parameters of different noise types to generate new data. We summarize the contributions of this paper below:

1. We introduce (i) noising- and (ii) denoising-based data augmenters to improve the generalization of a deep CNN. The denoising-based approach further uses an autoencoder to generate new augmented data.
2. We propose a “learning-to-augment” strategy to generate noisy images. The learning-to-augment approach employs a Bayesian optimizer to determine the optimal noise parameters for new augmented images.
3. We show the effectiveness of our proposed approach on a challenging task of COVID-19 detection in chest X-ray images and outperform state-of-the-art data augmentation methods.

2. Data

We accumulated a dataset of 1248 chest X-ray images of posterior–anterior view (666 images) and anterior–posterior view (582 images) from two public repositories [29–31]. The first repository contains chest X-ray images of 215 COVID-19 patients and 33 non-COVID pneumonia patients [30]. The second repository contains chest X-ray images of 500 healthy subjects and 500 non-COVID pneumonia patients [31]. We carefully eliminated the X-ray images of lateral view and CT images from the data cohort of first repository [28]. Table 1 summarizes the patients' demographics (age and sex). All the X-ray images were either in Portable Network Graphics (PNG) or Joint Photographic Experts Group (JPEG) file format. All the X-ray images were resized to a common input size for the pre-trained CNNs (i.e., AlexNet, GoogleNet, ResNet18, and ShuffleNet). This is a standard practice of preprocessing computer vision data for CNN training, and similar preprocessing for X-ray data resizing is also widely used in recent literature (e.g. Refs. [12–14]). On the other hand, X-ray images are stored and transmitted in the form of compressed data [32]. If a deep network is trained only using original images, the image distortion caused by lossy compression could deteriorate the test results. We therefore used lossless (i.e., PNG) and lossy (i.e., JPG) image compression as the inputs to deep convolutional neural networks.

Also note that we accumulated our dataset from public repositories, and the data do not contain the exposure time parameter associated with those. However, as the radiopacity variation is the key to visualize COVID-19 infection in a chest X-ray image and its quality is dependent on the exposure time, we assume that these X-ray images are acquired with an exposure more than 6 mA-seconds (mAs) to ensure good quality of image.

3. Methods

X-ray images often get degraded by impulse noise [15], Gaussian noise [33–35], speckle noise [36–38] and Poisson noise [39–42] at the

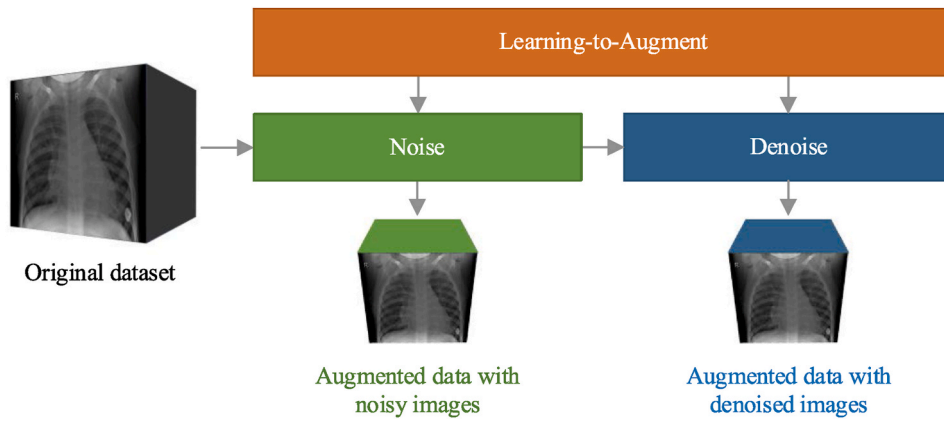


Fig. 1. Schematic pipeline of the proposed learning-to-augment strategy using noisy and denoised data.

Table 2
Parameters of different noise types.

Notation	Definition
μ	The mean of Gaussian and speckle noise
v	The variance of Gaussian and speckle noise
d	The noise density of impulse noise

time of acquisition, transmission, or storage. Noisy images can be used as inputs to a CNN in two ways [39]: (i) using unprocessed noisy images as inputs to the networks, (ii) using denoised images as inputs to the network. When the noisy images are fed as inputs to the network, data augmentation using noise may improve the robustness of the classifier. On the other hand, if preprocessing is used to denoise images, then data augmentation using restored (denoised) images in training can improve the generalizability of the network. Fig. 1 illustrates the pipeline of the

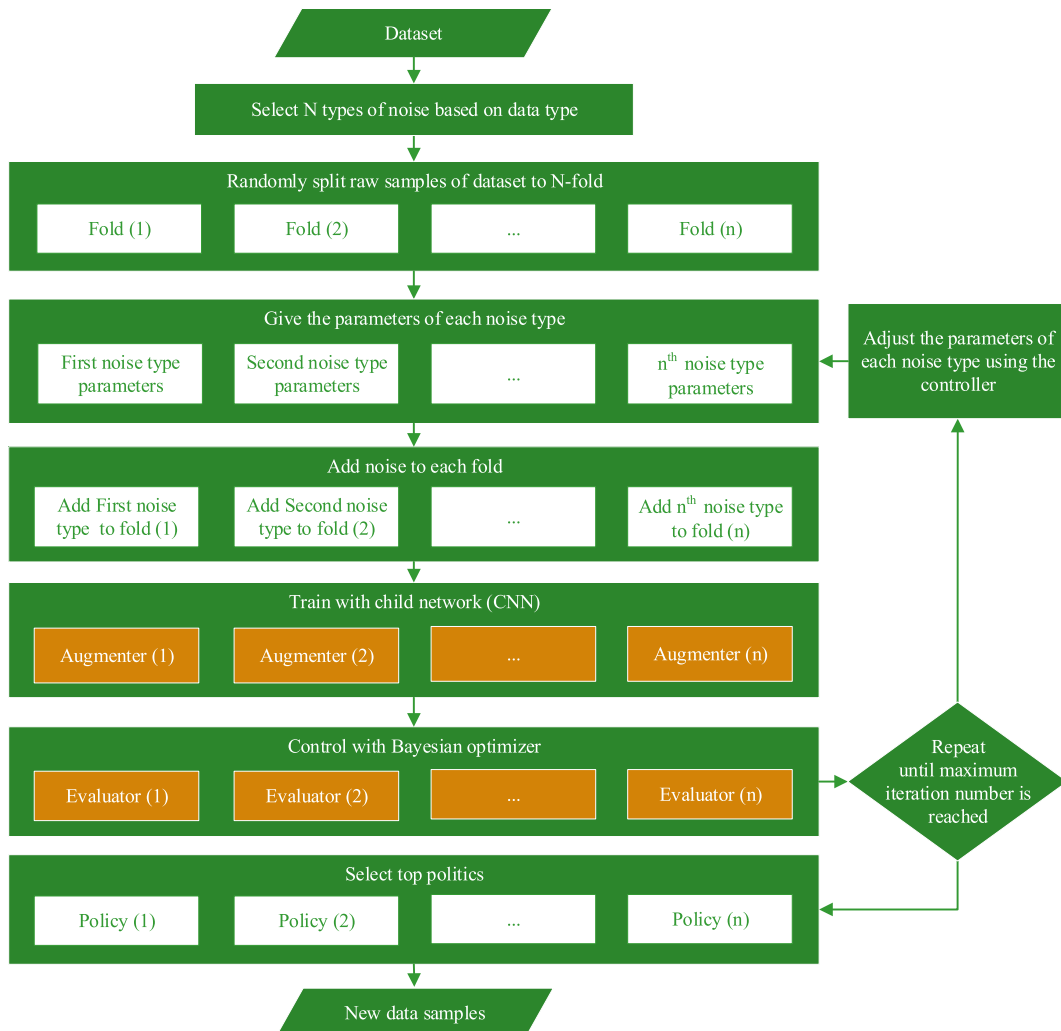


Fig. 2. Flowchart of our noise-based image data augmentation approach.

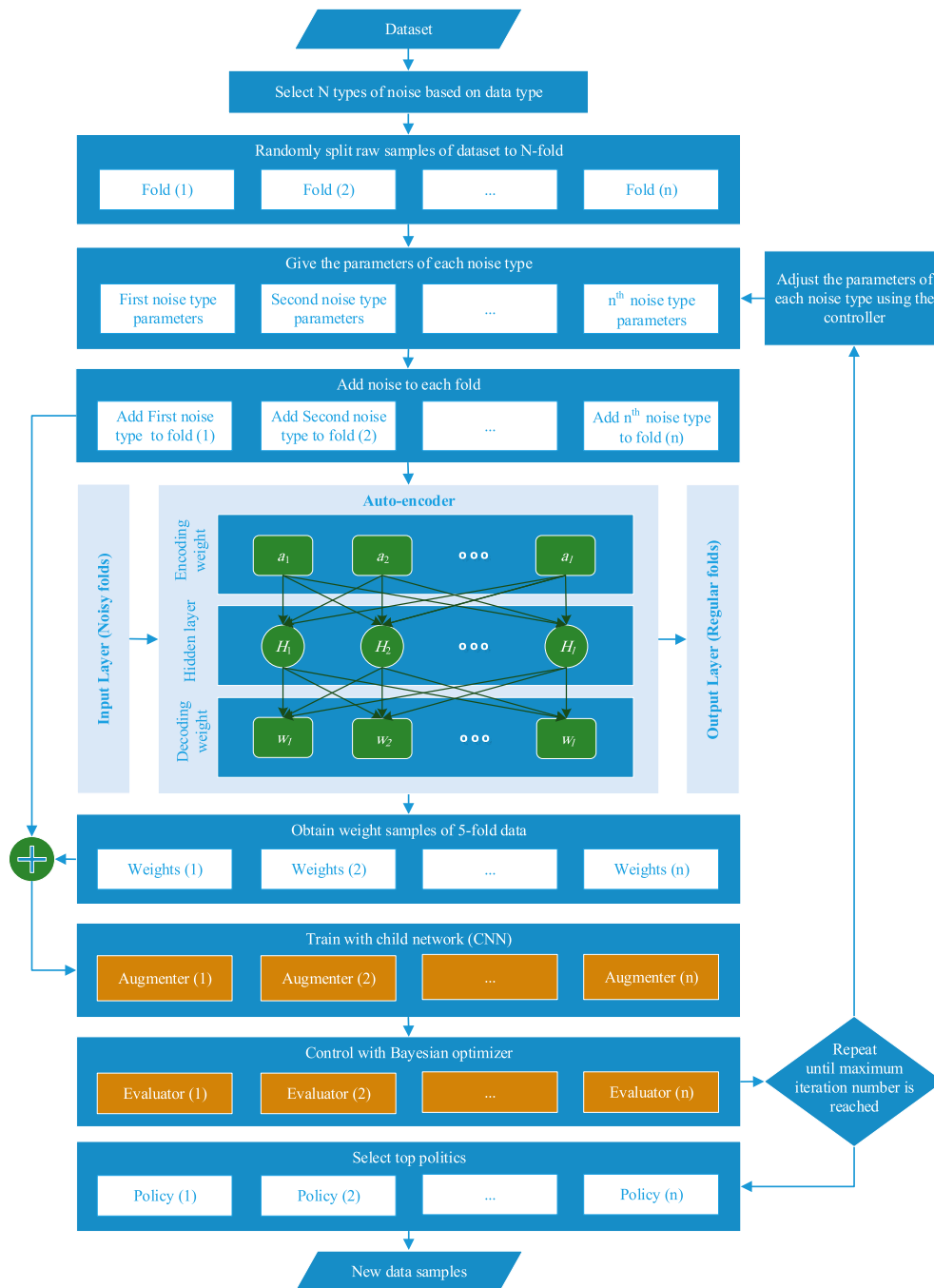


Fig. 3. Flowchart of the autoencoder-restored image-based data augmentation approach.

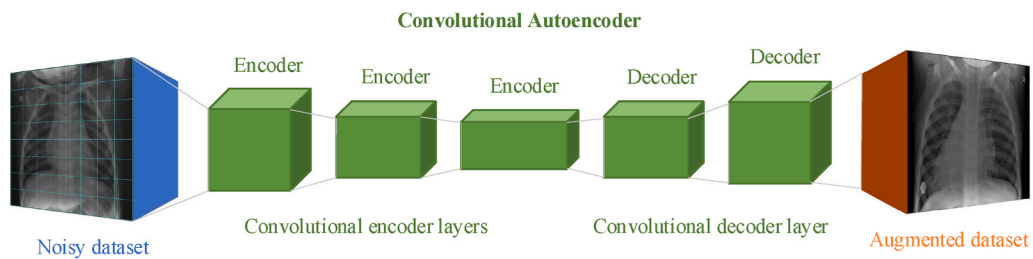


Fig. 4. The proposed convolutional autoencoder learns to denoise images.

Table 3
The configuration of the proposed autoencoder.

Type	Layer	Output Shape	# of Parameters	Padding
Input	Input Layer	(28, 28, 1)	0	–
Encoder	Convolution layer	(28, 28, 256)	2560	Same
Encoder	ReLU	–	–	–
Encoder	Max pooling	(14, 14, 256)	0	–
Encoder	Convolution layer	(14, 14, 512)	1,180,160	Same
Encoder	ReLU	–	–	–
Encoder	Max pooling	(7, 7, 512)	0	–
Encoder	Convolution layer	(7, 7, 512)	2,359,808	Same
Encoder	ReLU	–	–	–
Decoder	Upsampling	(14, 14, 512)	0	–
Decoder	Convolution layer	(14, 14, 256)	1,179,904	Same
Decoder	ReLU	–	–	–
Decoder	Upsampling	(28, 28, 256)	0	–
Decoder	Convolution layer	(28, 28, 1)	2305	Same
Decoder	Sigmoid	–	–	–

proposed learning-to-augment strategy using noisy and denoised data.

3.1. Data augmentation with noisy images

Adding noise to data is one approach to data augmentation [18]. We present a method for learning-to-augment via noisy input images. The learning-to-augment finds optimal noise parameters to generate the new data. The mean and variance are parameters of the Gaussian and speckle noise types [35,43]. The impulse (salt-and-pepper) noise is, on the other hand, specified with density parameter [44–46]. The parameters of different noise models are shown in Table 2.

As shown in Fig. 2, the proposed noisy image-based augmenter is composed of a noisy data generator, a controller, an augmenter, and child models. The steps of the noise-based data augmentation strategy are following:

Step 1. Noisy data generator adds noise to the original images with specific noise parameters. In the first iteration, the parameters of noise (i.e., μ , ν , and d) are set randomly. Then, the controller determines the parameters of each noise type as a new policy.

Step 2. The augmenter produces new data by applying the noise to the images.

Step 3. The child CNN models are trained using new augmented data to evaluate the performance of data augmentation policies.

Step 4. The controller, a Bayesian optimizer-based search algorithm [47,48], substitutes existing weak policies with new data augmentation policies by exploring the search space of the parameters for each noise type.

Step 5. The above steps are repeated until the best policies, i.e., the parameters of each noise type, are found.

As shown in Fig. 1, to augment the data, the input data pool is first divided into N equal folds. Then a noisy data generator adds noise (e.g., impulse, Gaussian, speckle, and Poisson noise) to each fold separately. The raw samples of the dataset were randomly split into N -folds to accelerate finding the best policies. Decreasing the number of samples using the N -fold split reduces the CNN training time for each fold [49].

The augmenters create new data based on new parameters that the Bayesian optimizer has found. In the next step, each fold is processed by the child CNN model. Use of child networks, instead of very deep CNN, in policy evaluation speeds up the execution of the proposed method. Based on the results of child CNNs, the controller improves weak policies and maintains strong policies. The controller uses the Bayesian optimizer to find the optimum set of augmentation policies (the parameters of each noise type) in a search space. Let \mathcal{S} be the search space and f be the loss function of a classifier, then the Bayesian optimizer can be represented as:

$$y = \arg \min_{G \in \mathcal{S}} f(G). \quad (1)$$

The optimization problem in Eq. (1) aims to find y that minimizes $f(G)$ for G in a bounded domain \mathcal{S} . The loss values of the child CNNs are used to calculate the loss function for the Bayesian optimizer. This process continues until the maximum iteration number is reached.

3.2. Data augmentation approach with denoised data

The proposed method provides an automatic augmentation policy search method using the generation of restored images that had been corrupted by noise. Noisy images can be restored by enhancement algorithms such as autoencoder networks [50]. However, depending on the noise type and density, the pixel values in the restored image and the original noise-free image are not exactly equal [51]. We aim to leverage the dissimilarity between restored and original pixels as a data augmentation strategy. First, noise of specific type and density is added to the image. Then, the noise is partially removed from the image by using the proposed autoencoder. The denoising autoencoder aims to produce the output from the noisy input, where the target is set as the original images. Finally, the restored images are used as augmented data. In the proposed noise-based data augmentation algorithm, the type

Table 4
Characteristics of pre-trained models.

Networks	Depth	Size	# of Parameters (millions)	Input Image Size (pixels)
AlexNet	8	227 MB	61	227 × 227
GooleNet	22	27 MB	7	224 × 224
ResNet18	18	44 MB	11.7	224 × 224
ShuffleNet	50	5.4 MB	1.4	224 × 224

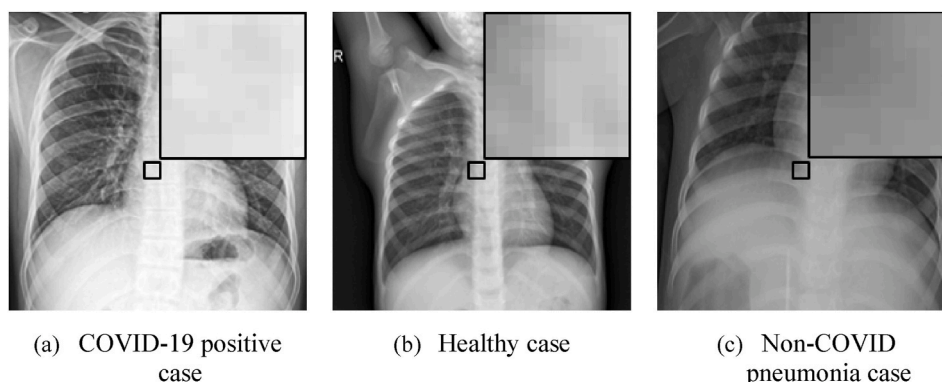


Fig. 5. Sample chest X-ray images for (a) COVID-19, (b) healthy, and (c) non-COVID pneumonia cases [56].

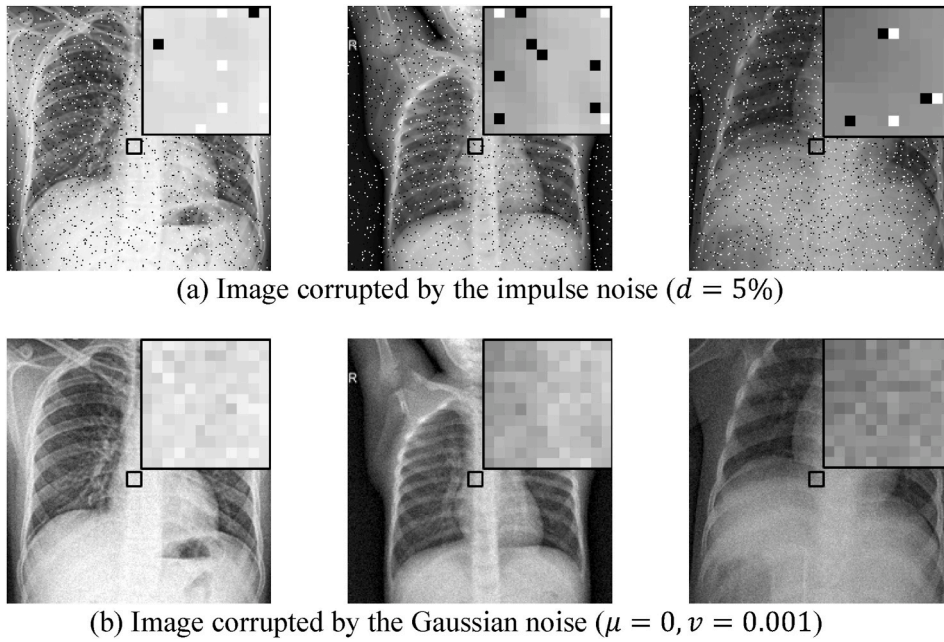


Fig. 6. Noise added to the chest X-ray images.

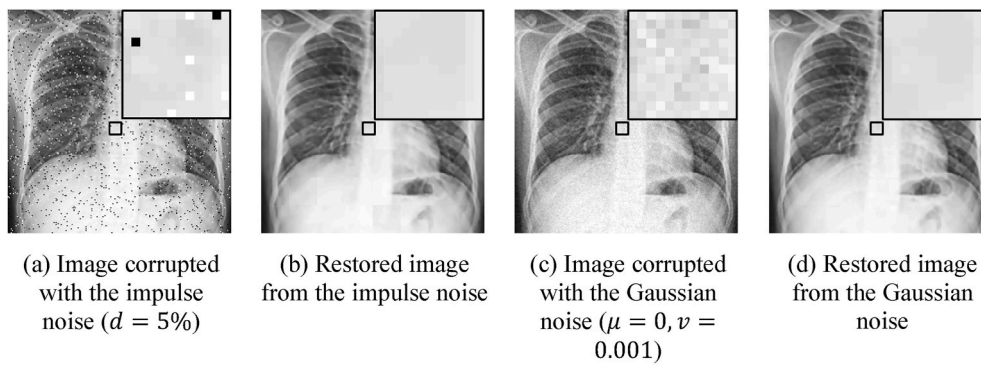


Fig. 7. Noise-added and restored COVID-19 positive chest X-ray images.

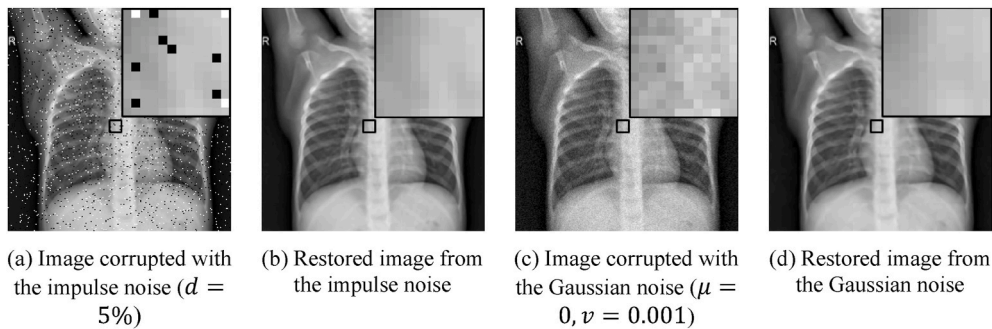


Fig. 8. Noise-added and restored healthy chest X-ray images.

and density of noise are important. Depending on the accuracy of the noise removal algorithm, the restored image could be very similar to the original image, especially when the noise magnitude is low, which could thus result in an ineffective augmented image. On the other hand, if the noise magnitude is high and the denoising is imperfect (as expected), the pixel values of the restored images would be more dissimilar from the original ones. As shown in Fig. 3, noise of specific parameters is added to the original images by the noise generator (section 3.1). The noisy

samples are then fed as inputs to the proposed autoencoder-based denoising model.

The decoding weights of the trained autoencoder can be compared to a conventional image denoising filter parameter. Once the autoencoder is trained, it produces new augmented image data from the noisy input data during inference. Like our noisy image-based augmentation approach (section 3.1), here also, we feed the new augmented data as input to the child CNN models. The Bayesian optimizer finds optimal

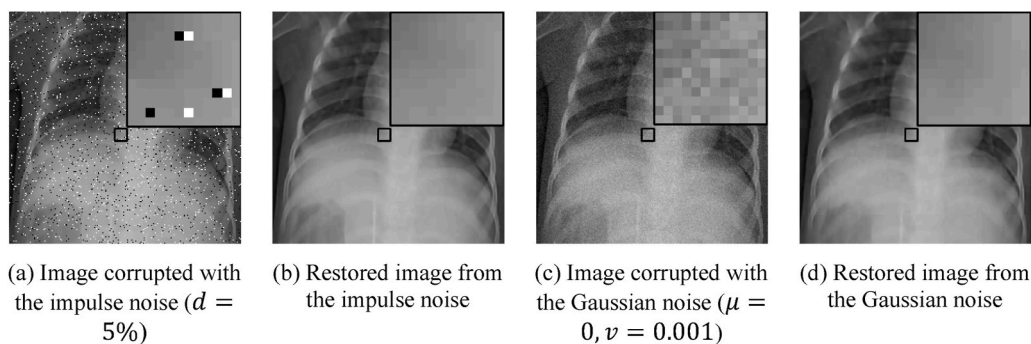


Fig. 9. Noise-added and restored chest X-ray images of non-COVID pneumonia types.

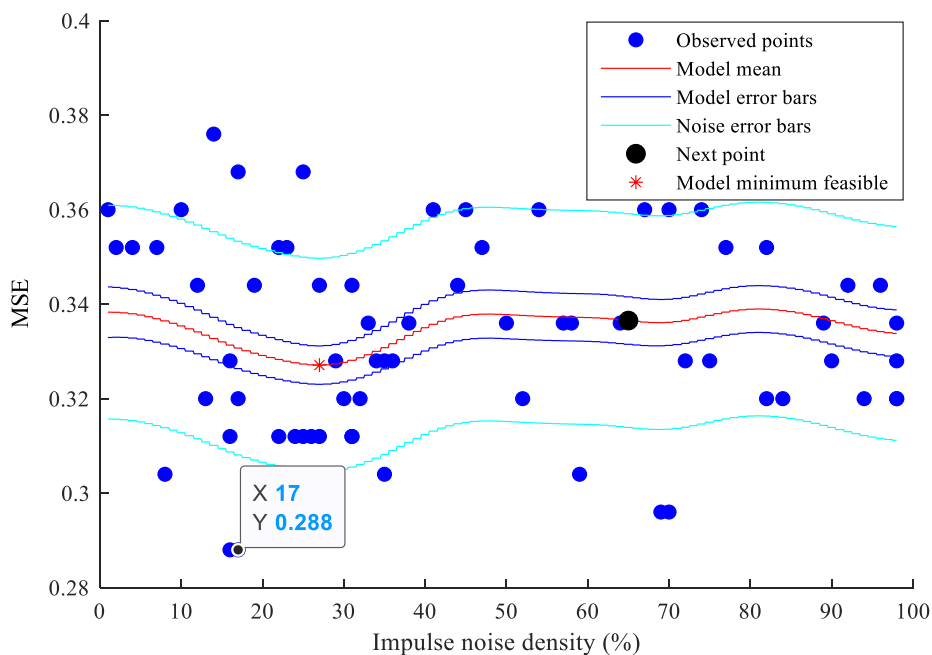


Fig. 10. The Bayesian optimizer evaluated the data augmentation policies for restored images, initially corrupted by impulse noise (impulse noisy density $d = 17\%$, MSE = 0.288).

noise parameters. After finding the optimal policies, we use the whole dataset and deep convolutional autoencoder for the noise-based data augmentation (Fig. 4).

4. Implementation details

4.1. Computational platform and training details

The proposed algorithm is trained using the deep learning toolbox of MATLAB 2020b in an Intel(R) Core(TM) i7-7700HQ CPU 2.81 GHz with 32 GB of RAM, and Nvidia GTX 1070 GPU with 8 GB VRAM. We used stochastic gradient descent with learning rate of 0.001 to train our networks (except for the autoencoder module) and used cross-entropy as the loss function. To train the autoencoder, we used Adam optimizer with learning rate of 0.001 and used mean squared error (MSE) as the loss function. We show the configuration of the autoencoder in Table 3. To cope with the limitation of our computation power, we split each original image of size 224×224 pixels into 64 patches of size 28×28 pixels. After training the autoencoder with 250 epochs, we combined the 64 patches as the restored image. We also used $N = 2$ in this study. We used 80% of the data (998 X-ray images) for training, 10% (125 X-ray images) of the data for validation, and 10% of the data (125 X-ray images) for testing. Fig. 5 shows sample chest X-ray images for COVID-19

positive, healthy, and non-COVID pneumonia cases.

We used pre-trained CNNs and finetuned them on the COVID-19 data for the classification task. We evaluated the performance of four pre-trained models, AlexNet [52], ShuffleNet [53], ResNet18 [54], and GoogleNet [55]. All of these networks were pre-trained on samples from the ImageNet Challenge database. Finetuning all these models on the COVID-19 dataset converge faster than training from scratch. Table 4 summarizes the ImageNet pre-trained networks used in the proposed framework.

4.2. Adding noise to images and denoising

The noisy data generators separately add impulse noise and Gaussian noise to the chest X-ray images of COVID-19 positive, healthy, and non-COVID pneumonia cases in each fold (Fig. 2). Since chest X-ray images are often corrupted by noise that is like the impulse and Gaussian noise, we used these two types of noise for simulation (Fig. 6).

Learning-to-augment with the restored or denoised images starts with dividing the dataset into N -folds (Fig. 3). The controllers initialize the noise parameters randomly. The noisy data generators create noisy images. The input to the proposed convolutional autoencoder is the noisy chest X-ray images and the target is set to the original images.

Table 5
The optimal values of the data augmentation policies.

Data augmentation method	Parameters of the method	Range of parameters	Optimal value of parameters	Error by AlexNet (%)
Changing brightness of the image	Value of brightness	[0,1]	0.215	0.256
Changing contrast of the image	Value of contrast	[0,255]	1.184	0.280
Adjusting hue of the image	Value of hue	[0,1]	0.164	0.264
Changing saturation of the image	Value of saturation	[0,1]	0.777	0.264
Rotation of image	Angle	0°–360°	21°	0.256
Adding impulse noise to the image	Noise density	[0,1]	0.025	0.288
Adding Gaussian noise to the image	Variance	[0,1]	0.714	0.256
Denoising image corrupted by impulse noise	Noise density	[0,1]	0.170	0.288
Denoising image corrupted by Gaussian noise	Variance	[0,1]	0.280	0.256

5. Results and discussion

We first present qualitative results after adding noise and denoising the images (section 5.1). We then quantify the proposed learning-to-augment strategy after finding the optimal data augmentation policies (section 5.2.1), and training with selected policies (section 5.2.2). Finally, we present our discussion on the results (section 5.2.3). It is worth noting that the proposed learning-to-augment strategy using

noisy and denoised data significantly increases the diversity in the training data. Unlike conventional data augmentation approaches that use predefined rules and procedures for a specific target task, the proposed learning-to-augment strategy dynamically refines augmentation rules based on feedback networks, and thus, reduces the negative effect of the small size of the training dataset.

5.1. Qualitative results

The denoised output images by the proposed convolutional auto-encoder for impulse and Gaussian noise-corrupted X-ray images of healthy cases are shown in Fig. 7. We show the noisy and restored chest X-ray images of the COVID-19 cases in Fig. 8 and Fig. 9, respectively. The augmenters use the outputs of the proposed convolutional autoencoder (restored images) to create new data.

5.2. Quantitative results

5.2.1. Optimal data augmentation policies

The data augmentation policies, i.e., the optimal parameters of noise, are determined by the Bayesian optimization algorithm. Fig. 10 shows the results of the Bayesian optimizer to evaluate the data augmentation policies using the restored images, which were initially corrupted by the impulse noise.

Learning-to-augment strategies using changing brightness, contrast, hue, saturation, and rotation of images are used to compare to the proposed approach. The optimal values of the parameters of the mentioned methods were chosen by the Bayesian optimizer for a fair comparison. The pre-trained networks (AlexNet, ShuffleNet, GoogleNet, and ResNet18) typically take input images of 3 channels (red, green, blue). Therefore, we use the same X-ray image and stack them three times to make the input image 3-channel. Then we use the augmentation operations, e.g., change of hue and saturation. As shown in Table 5, the

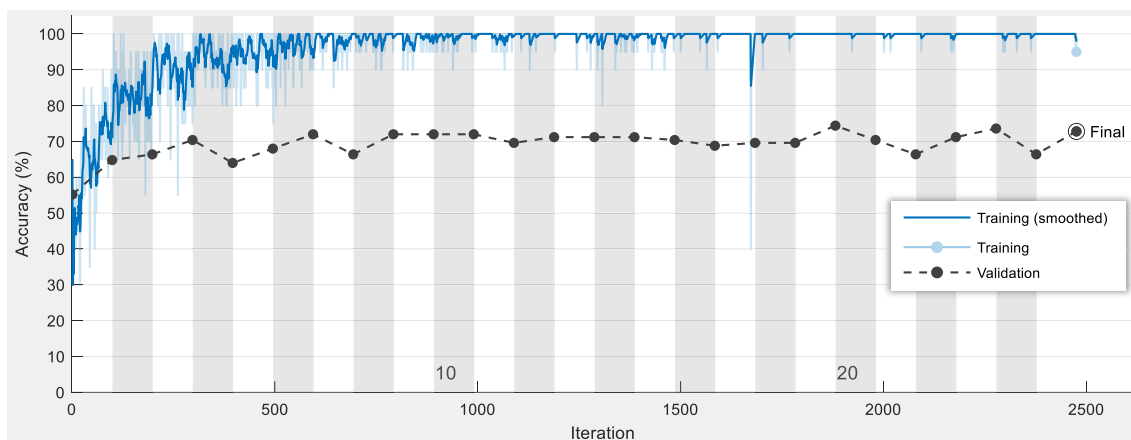


Fig. 11. Training and validation accuracy curves of ResNet18 in the X-ray image classification task.

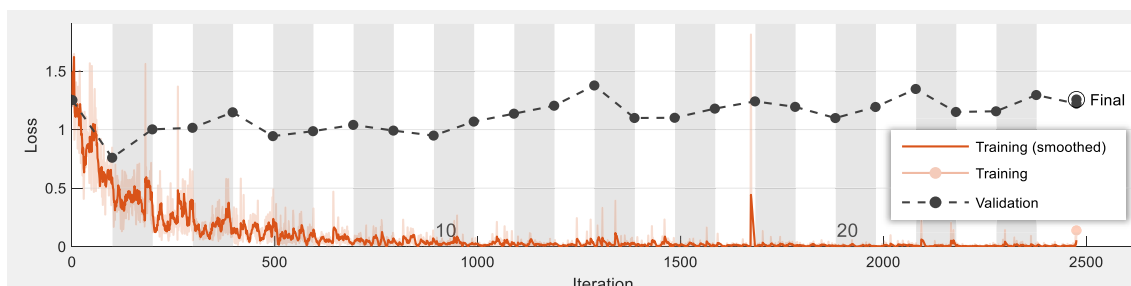


Fig. 12. Training and validation loss curves of ResNet18 in the X-ray image classification task.



Fig. 13. Confusion matrices for X-ray image classification by the proposed learning-to-augment approach using restored images corrupted by noise. Here, 'Normal' represents the 'Healthy' subjects and 'Other_Pneumonia' represents the 'non-COVID pneumonia' patients.

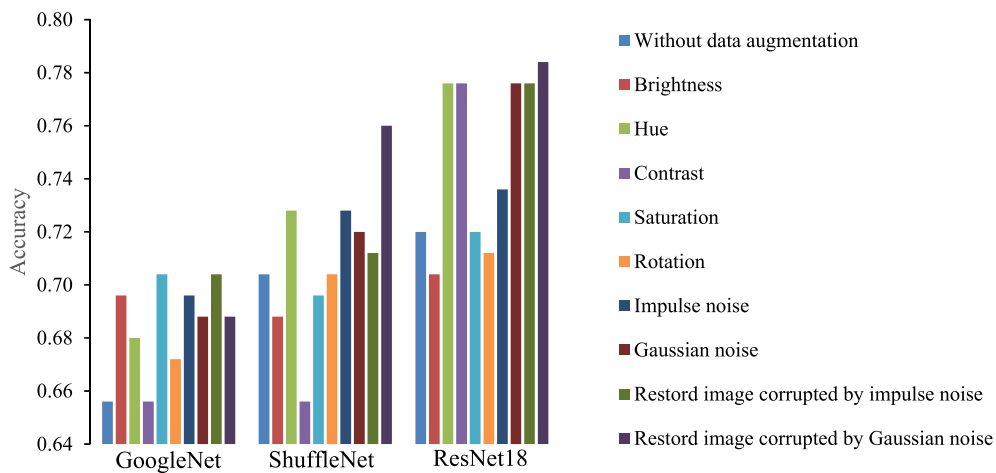


Fig. 14. Bar plot showing the accuracies of COVID-19 classification for different augmentation approaches.

optimal values of the parameters of data augmentation methods were chosen by the Bayesian optimizer for the AlexNet classifier.

5.2.2. Training with the selected policies

ShuffleNet with 50 layers, ResNet18 with 18 layers, and GoogleNet

with 22 layers were used for the classification of X-ray images. The COVID-19 classification accuracy curves during ResNet18 training and validation with the Gaussian noise corrupted and restored images are shown in Fig. 11. We also illustrate the ResNet18 training and validation loss curves in Fig. 12. The confusion matrices of COVID-19 classification

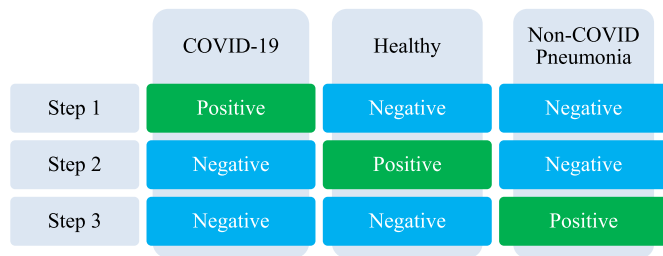


Fig. 15. Evaluation of the classification performance by identifying the positive and negative classes in three steps.

by the proposed data augmentation approach using restored images are shown in Fig. 13. The results show that the generalization of ResNet18 has been improved using the proposed data augmentation method based on restored images from the Gaussian noise corruption. The results, shown in Fig. 12, indicate that the data augmentation using the proposed noising- and denoising-based data augmenter performs overall better than other approaches. The ResNet18, trained with the restored images from the Gaussian noise corruption, performed the best among all the techniques.

5.2.3. Test results and discussion

To evaluate the pre-trained deep CNN models, four metrics are employed for verifying the quality of the COVID-19 classification

results, including accuracy, sensitivity, specificity, and F-Measure criteria [20,21]. We show the quantitative comparison of different augmentation strategies in Fig. 14, including changing brightness of the image, changing the contrast of the image, adjusting the hue of the image, changing the saturation of the image, rotation of the image, adding the impulse noise to the image, adding the Gaussian noise to the image, restoring the image corrupted by the impulse noise, and restoring image corrupted by the Gaussian noise. The pre-trained models were finetuned separately using optimal parameters of data augmentation methods (Table 5).

In Fig. 15, we show our 3-steps strategy to evaluate the X-ray image classification task in terms of sensitivity, specificity, and F-Measure. At each step of the evaluation, one class is considered positive and the others are considered negative classes. At first, we considered ‘COVID-19’ as the positive class and the other two (i.e., healthy and non-COVID pneumonia) as negatives. Table 6 demonstrates the efficacy of the proposed data augmentation strategy using noisy and denoised images with respect to the best performance by the state-of-the-art learning-to-augment method that uses the modification of the brightness, hue, contrast, saturation, and rotation. As shown in Table 6, ShuffleNet performance for restored images (our approach) is improved by 5.7% in terms of sensitivity from 54.3% to 60%, compared to saturation-based augmentation. Similarly, specificity of the ResNet18 method for the restored images is improved, from 89% to 90.9%, over the hue-based augmentation. Similarly, F-Measure is improved from 54.8% to 62.0% for the restored images compared to contrast-based augmentation. Also

Table 6

Comparing sensitivity, specificity, and F-measure of X-ray image classification with the learning-to-augment strategies.

Augmentation Method	Positive class	GoogleNet			ShuffleNet			ResNet18		
		Sensitivity	Specificity	F-Measure	Sensitivity	Specificity	F-Measure	Sensitivity	Specificity	F-Measure
Without data augmentation	COVID-19	0.410	0.767	0.427	0.500	0.793	0.507	0.441	0.824	0.462
	Healthy	0.625	0.681	0.619	0.698	0.708	0.667	0.667	0.765	0.685
	Non-COVID pneumonia	0.574	0.718	0.590	0.604	0.778	0.634	0.725	0.716	0.679
Brightness	COVID-19	0.500	0.763	0.457	0.390	0.833	0.451	0.410	0.837	0.464
	Healthy	0.607	0.768	0.642	0.741	0.648	0.672	0.709	0.700	0.678
	Non-COVID pneumonia	0.685	0.704	0.661	0.566	0.778	0.606	0.647	0.743	0.641
Hue	COVID-19	0.459	0.773	0.459	0.486	0.830	0.514	0.471	0.890	0.533
	Healthy	0.636	0.714	0.636	0.722	0.732	0.696	0.741	0.806	0.754
	Non-COVID pneumonia	0.611	0.732	0.623	0.642	0.792	0.667	0.745	0.797	0.731
Contrast	COVID-19	0.364	0.815	0.427	0.371	0.767	0.377	0.515	0.870	0.548
	Healthy	0.679	0.638	0.639	0.542	0.758	0.598	0.732	0.812	0.745
	Non-COVID pneumonia	0.519	0.761	0.566	0.712	0.616	0.632	0.765	0.784	0.736
Saturation	COVID-19	0.436	0.826	0.479	0.543	0.756	0.500	0.485	0.804	0.478
	Healthy	0.732	0.681	0.689	0.589	0.783	0.635	0.625	0.797	0.667
	Non-COVID pneumonia	0.566	0.806	0.619	0.648	0.732	0.648	0.750	0.699	0.690
Rotation	COVID-19	0.375	0.812	0.423	0.500	0.793	0.507	0.421	0.839	0.471
	Healthy	0.667	0.676	0.650	0.704	0.704	0.673	0.709	0.714	0.684
	Non-COVID pneumonia	0.574	0.746	0.602	0.574	0.803	0.626	0.667	0.743	0.654
Impulse noise	COVID-19	0.457	0.789	0.457	0.469	0.817	0.469	0.500	0.839	0.535
	Healthy	0.643	0.739	0.655	0.673	0.771	0.685	0.722	0.746	0.703
	Non-COVID pneumonia	0.660	0.722	0.648	0.750	0.712	0.696	0.618	0.829	0.673
Gaussian noise	COVID-19	0.432	0.795	0.451	0.500	0.802	0.493	0.529	0.868	0.563
	Healthy	0.679	0.696	0.661	0.667	0.761	0.673	0.778	0.775	0.750
	Non-COVID pneumonia	0.615	0.740	0.621	0.685	0.746	0.679	0.698	0.833	0.725
Restored images corrupted by the impulse noise	COVID-19	0.425	0.835	0.479	0.410	0.849	0.471	0.459	0.909	0.548
	Healthy	0.778	0.648	0.694	0.774	0.667	0.695	0.804	0.754	0.763
	Non-COVID pneumonia	0.547	0.819	0.611	0.582	0.814	0.640	0.686	0.838	0.714
Restored images corrupted by the Gaussian noise	COVID-19	0.486	0.773	0.480	0.600	0.811	0.545	0.579	0.874	0.620
	Healthy	0.636	0.729	0.642	0.636	0.857	0.700	0.811	0.764	0.761
	Non-COVID pneumonia	0.600	0.757	0.629	0.808	0.726	0.737	0.611	0.915	0.710

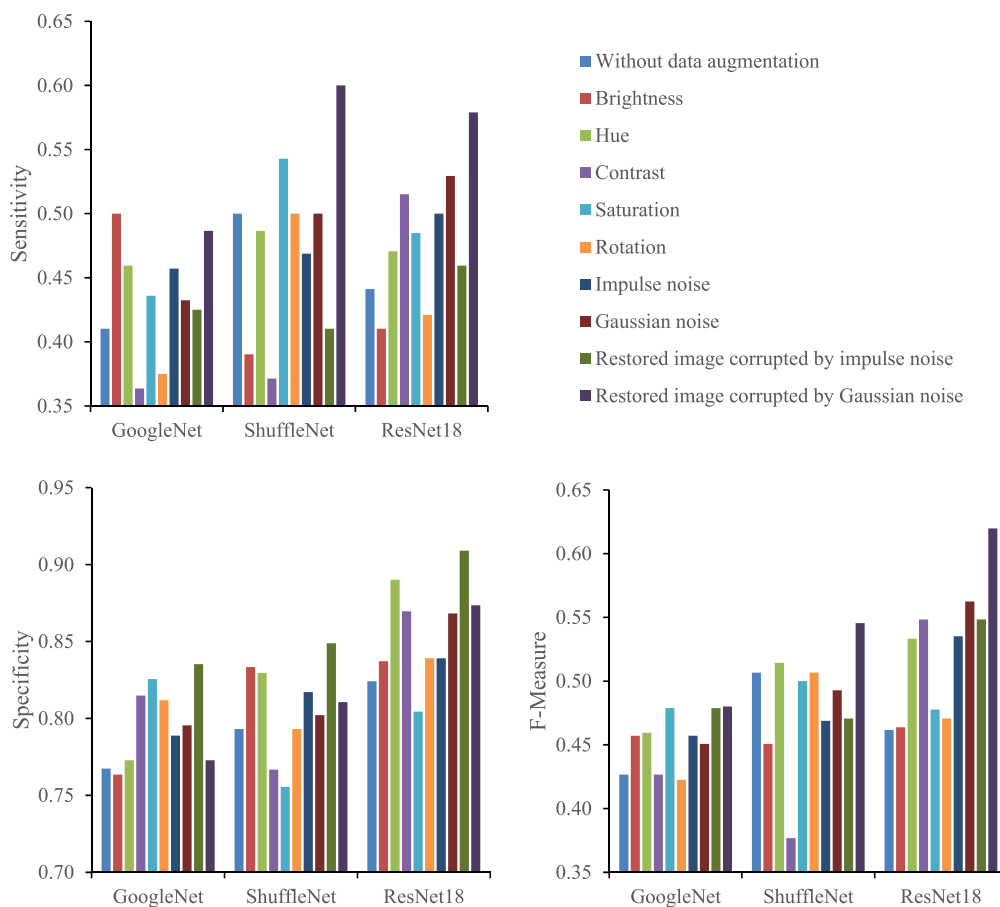


Fig. 16. Comparison of the sensitivity, specificity, and F-Measure of COVID-19 classification for different augmentation strategies.

Table 7

Comparison of proposed data augmentation strategy using restored images corrupted by the Gaussian noise to the state-of-the-art method in term of accuracy.

Strategy	Networks' architecture			Output of data augmentation		Accuracy (%)
	Depth (layer)	Size (MB)	# of parameters (million)	# of data augmentation methods	# of augmented data	
Nishio et al. [29] (DenseNet201)	201	77	20.0	6	5998	78.24 ± 2.23
Nishio et al. [29] (ResNet50)	50	96	25.6	6	5998	77.76 ± 1.18
Proposed (ResNet18)	19	44	11.7	1	998	77.60 ± 1.20

considering ‘healthy’ as the positive class, in ResNet18, the proposed strategy using restored images corrupted by Gaussian noise outperformed the results of hue-based augmentation in terms of sensitivity (from 74.1% to 81.1%). The performance of ShuffleNet increased in terms of specificity from 78.3% to 85.7%, when Gaussian noise-corrupted were restored, compared to saturation-based augmentation, and with ResNet18, F-Measure improved from 75.4% to 76.3% when images corrupted by impulse noise were restored rather than hue-based augmentation. Similar performance trends for the methods can be seen in Table 6 when the ‘non-COVID pneumonia’ cases are considered as the positive class in the third step. The proposed learning-to-augment strategy using denoised images corrupted by Gaussian noise yields the best results compared to contrast-based augmentation, for ShuffleNet the sensitivity was improved from 71.2% to 80.8%. Similarly, comparing with hue-based augmentation, the specificity of ResNet18 went from 79.7% to 91.5%, and the F-Measure of ShuffleNet was improved from 66.7% to 73.7%.

We show the evaluation results in terms of sensitivity, specificity, and F-Measure metrics in Fig. 16. The sensitivity and specificity values

for the classification of COVID-19 by the proposed approach are better than those by other augmentation techniques. Thus, the noising- and denoising-based data augmenter is applicable to improve the generalization of deep CNNs for image classification. The performance of the COVID-19 classification task was improved by our methods capability to choose noise parameters. In addition, using restored images by the proposed autoencoder model helps to generalize CNNs.

In a state-of-the-art method [29], Nishio et al. applied six data augmentation strategies on image data, consisting of ±15° rotation, ±15% x-axis shift, ±15% y-axis shift, horizontal flipping, and 85–115% scaling and shear transformation. They employed DenseNet201 (201 layers deep, 77 MB size, and 20.0 million parameters) and ResNet50 (50 layers deep, 96 MB size, and 25.6 million parameters) for the detection of COVID-19 in chest X-ray images. The accuracy on the test set for DenseNet201 and ResNet50 were 78.24 ± 2.23% and 77.76 ± 1.18%, respectively. On the other hand, according to results shown in Table 7, using the proposed data augmentation approach, ResNet18 (18 layers deep, 44 MB size, and 11.7 million parameters) showed an accuracy of 77.6 ± 1.20%. With a shallower network and less “augmented” data, the

proposed data augmentation strategy using restored image is almost as accurate as the state-of-the-art method.

6. Conclusion

In this paper, we proposed a learning-to-augment strategy using noisy and restored images to improve the generalizability of deep CNNs. Using a novel noise-based data augmentation approach, we tackled the overfitting problem of deep CNNs for automatic identification of COVID-19 in the chest X-ray images. A noisy data generator, a Bayesian optimizer-based controller, an autoencoder network, child augmenters, and child CNN models are key components of our proposed noising- and denoising-based data augments that increase the accuracy of the image classification task. Learning-to-augment strategies, including changing brightness of the image, changing the contrast of the image, adjusting the hue of the image, changing the saturation of the image, and the rotation of the image have been compared to the proposed method (adding impulse noise to an image, adding Gaussian noise to an image, restoring image corrupted by impulse noise, and restoring image corrupted by Gaussian noise). The proposed data augments also achieved the best performance in COVID-19, healthy, and non-COVID pneumonia classification in terms of sensitivity, specificity, and F-Measure. The learning-to-augment strategy on the restored Gaussian noise-corrupted images in a pretrained ResNet18 adapts properly on the new and previously unseen data (test set) and showed better classification accuracy compared to the state-of-the-art data augmentation approach. So, we can conclude that the proposed strategy improves the generalization of deep CNN.

Conflicts of interest

The authors have declared no conflict of interest.

References

- [1] S. Varela-Santos, P. Melin, "A new approach for classifying coronavirus COVID-19 based on its manifestation on chest X-rays using texture features and neural networks, *Information sciences* 545 (2021) 403–414.
- [2] S. Bhattacharya, P.K.R. Maddikunta, Q.V. Pham, T.R. Gadekallu, C.L. Chowdhary, M. Alazab, M.J. Piran, "Deep learning and medical image processing for coronavirus (COVID-19) pandemic: a survey, *Sustainable cities and society* 65 (2021) 102589.
- [3] I. Ahmed, M. Ahmad, J.J. Rodrigues, G. Jeon, S. Din, "A deep learning-based social distance monitoring framework for COVID-19, *Sustainable Cities and Society* 65 (2021) 102571.
- [4] E. Hussain, M. Hasan, M.A. Rahman, I. Lee, T. Tamanna, M.Z. Parvez, "CoroDet: a deep learning based classification for COVID-19 detection using chest X-ray images, *Chaos, Solitons & Fractals* (2021), <https://doi.org/10.1016/j.chaos.2020.110495>, 142, p. 110495, 2021/01/01.
- [5] E.Y. Lee, M.-Y. Ng, P.-L. Khong, "COVID-19 pneumonia: what has CT taught us?," *Lancet Infect. Dis.* 20 (4) (2020) 384–385.
- [6] H. Shi, X. Han, N. Jiang, Y. Cao, O. Alwalid, J. Gu, C. Zheng, "Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study," *the Lancet infectious diseases*, vol. *Lancet Infect. Dis.* 4 (2020) 425–434.
- [7] Y. Li, L. Xia, "Coronavirus disease 2019 (COVID-19): role of chest CT in diagnosis and management, *American Journal of Roentgenology* 6 (214) (2020) 1280–1286.
- [8] W. Kong, P.P. Agarwal, "Chest imaging appearance of COVID-19 infection, *radiology: cardiothoracic imaging* 2 (1) (2020) p. e200028.
- [9] J. Somasekar, P.P. Kumar, A. Sharma, G. Ramesh, "Machine learning and image analysis applications in the fight against COVID-19 pandemic: datasets, research directions, challenges and opportunities, *Materials Today: Proceedings* (2020).
- [10] H. Panwar, P. Gupta, M.K. Siddiqui, R. Morales-Mendez, P. Bhardwaj, V. Singh, "A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images, *Chaos, Solitons & Fractals* 140 (2020) 110190.
- [11] A.A. Ardakani, A.R. Kanafi, U.R. Acharya, N. Khadem, A. Mohammadi, "Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: results of 10 convolutional neural networks, *Computers in Biology and Medicine* 121 (2020) 103795.
- [12] E. Çallı, E. Sogancioglu, B. van Ginneken, K.G. van Leeuwen, K. Murphy, "Deep learning for chest X-ray analysis: a survey," *medical image analysis*, 102125, 2021/06/05, <https://doi.org/10.1016/j.media.2021.102125>, 2021.
- [13] A.M. Ismael, A. Şengür, "Deep learning approaches for COVID-19 detection based on chest X-ray images," *Expert Systems with Applications*, 164, p. 114054, 2021/02/01, <https://doi.org/10.1016/j.eswa.2020.114054>, 2021.
- [14] S. Thakur, A. Kumar, "X-ray and CT-scan-based automated detection and classification of covid-19 using convolutional neural networks (CNN)," *Biomedical Signal Processing and Control*, 69, p. 102920, 2021/08/01, <https://doi.org/10.1016/j.bspc.2021.102920>, 2021.
- [15] C.-T. Lu, M.-Y. Chen, J.-H. Shen, L.-L. Wang, C.-C. Hsu, "Removal of salt-and-pepper noise for X-ray bio-images using pixel-variation gain factors," *Computers & Electrical Engineering*, 71, pp. 862–876, 2018/10/01, <https://doi.org/10.1016/j.compeleceng.2017.08.012>, 2018.
- [16] M.D. Bloice, P.M. Roth, A. Holzinger, "Biomedical image augmentation using Augmentor," *Bioinformatics*, vol 21 (35) (2019) 4522–4524, <https://doi.org/10.1093/bioinformatics/btz259>.
- [17] P. Panda, K. Roy, "Implicit adversarial data augmentation and robustness with noise-based learning neural networks," 141, pp. 120–132, 2021/09/01, <https://doi.org/10.1016/j.neunet.2021.04.008>, 2021.
- [18] H.-J. Bae, C.W. Kim, N. Kim, B. Park, N. Kim, J.B. Seo, S.M. Lee, "A perlin noise-based augmentation strategy for deep learning with small data samples of HRCT images, *Scientific reports* 1 (8) (2018) 17687, <https://doi.org/10.1038/s41598-018-36047-2>, 12/06 2018.
- [19] M. Momeny, M.A. Sarram, A.M. Latif, R. Sheikhpour, Y.D. Zhang, "A noise robust convolutional neural network for image classification results in engineering, 100225, 2021/05/17, <https://doi.org/10.1016/j.rineng.2021.100225>, 2021.
- [20] M. Ofori-Oduro, M.A. Amer, "Data augmentation using artificial Immune systems for noise-robust CNN models," *IEEE Int. Conf. Image Process.* (2020) 833–837, <https://doi.org/10.1109/ICIP40778.2020.9191116>, 25–28 Oct. 2020 2020.
- [21] A.E. Bilali, A. Taleb, M.A. Bahlaoui, Y. Brouziyne, "An integrated approach based on Gaussian noises-based data augmentation method and AdaBoost model to predict faecal coliforms in rivers with small dataset," *Journal of Hydrology*, p. 126510, 2021/05/29, <https://doi.org/10.1016/j.jhydrol.2021.126510>, 2021.
- [22] M.J. Sáiz-Abajo, B.H. Mevik, V.H. Segtnan, T. Næs, "Ensemble methods and data augmentation by noise addition applied to the analysis of spectroscopic data," *Analytica Chimica Acta*, 533, no. 2, pp. 147–159, 2005/03/28, <https://doi.org/10.1016/j.aca.2004.10.086>, 2005.
- [23] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *journal of big data*, vol. 6, no. 1, p. 60, 2019/07/06 2019, doi: 10.1186/s40537-019-0197-0.
- [24] F. J. Moreno-Barea, F. Strazzera, J. M. Jerez, D. Urda, and L. Franco, "Forward noise adjustment scheme for data augmentation," in *2018 IEEE Symposium Series on computational intelligence (SSCI)*, 18–21 Nov. 2018 2018, pp. 728–734, doi: 10.1109/SSCI.2018.8628917.
- [25] A. Sezer, H.B. Sezer, "Deep convolutional neural network-based automatic classification of neonatal hip ultrasound images: a novel data augmentation approach with speckle noise reduction," *ultrasound in medicine & biology*, 46, no. 3, pp. 735–749, 2020/03/01, <https://doi.org/10.1016/j.ultrasmedbio.2019.09.018>, 2020.
- [26] W. Wang, Y. Fu, Q. Sun, T. Chen, C. Cao, Z. Zheng, G. Xu, H. Qiu, Y.G. Jiang, X. Xue, "Learning to augment expressions for few-shot fine-grained facial expression recognition, *Computer Vision and Pattern Recognition*, 2020 arXiv preprint arXiv:2001.06144.
- [27] H. Cai, H. Chen, Y. Song, C. Zhang, X. Zhao, D. Yin, "Data manipulation: towards effective instance learning for neural dialogue generation via learning to augment and reweight," arXiv preprint arXiv (2020), 2004.02594.
- [28] L. Feng, M. Qiu, Y. Li, H.-T. Zheng, and Y. Shen, "Learning to augment for data-scarce domain BERT knowledge distillation," arXiv preprint arXiv:2101.08106, 2021.
- [29] M. Nishio, S. Noguchi, H. Matsuo, and T. Murakami, "Automatic classification between COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy on chest X-ray image: combination of data augmentation methods," *Scientific Reports*, vol. 10, no. 1, p. 17532, 2020/10/16 2020, doi: 10.1038/s41598-020-74539-2.
- [30] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "Covid-19 image data collection: prospective predictions are the future," arXiv preprint arXiv: 2006.11988, 2020.
- [31] RSNA pneumonia detection challenge | kaggle." <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge> (accessed).
- [32] Y. Dong, Z. Jiang, H. Shen, W.D. Pan, "Classification accuracies of malaria infected cells using deep convolutional neural networks based on decompressed images, *SouthEastCon* (2017) 1–6, <https://doi.org/10.1109/SECON.2017.7925268>, 30 March–2 April 2017 2017.
- [33] A. Roy, P. Maity, "A Comparative Analysis of Various Filters to Denoise Medical X-Ray Images," in *2020 4th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, IEEE, 2020, pp. 1–5.
- [34] P. Gravel, G. Beaudoin, J.A.D. Guise, "A method for modeling noise in medical images," *IEEE Transactions on Medical Imaging*, vol 23 (10) (2004) 1221–1232, <https://doi.org/10.1109/TMI.2004.832656>.
- [35] Y. Zhu, W. Shen, F. Cheng, C. Jin, G. Cao, "Removal of high density Gaussian noise in compressed sensing MRI reconstruction through modified total variation image denoising method," *Heliyon*, 6, no. 3, p. e03680, 2020/03/01, <https://doi.org/10.1016/j.heliyon.2020.e03680>, 2020.
- [36] B. Goyal, S. Agrawal, B. Sohi, "Noise issues prevailing in various types of medical images," *Biomedical & Pharmacology Journal*, vol 11 (3) (2018) 1227.
- [37] K.-J. Xia, H.-S. Yin, G.-S. Rong, J.-Q. Wang, Y. Jin, "X-ray image enhancement based on the improved adaptive low-pass filtering," *Journal of Medical Imaging and Health Informatics* 7 (8) (2018) 1342–1348.
- [38] S. Bharati, T.Z. Khan, P. Podder, N.Q. Hung, "A comparative analysis of image denoising problem: noise models, denoising filters and applications, *Cognit. Internet.Med. Things.Smart.Healthc.:* Spring (2020) 49–66.

- [39] T.B. Chandra, K. Verma, "Analysis of quantum noise-reducing filters on chest X-ray images: a review," *Measurement*, 153, p. 107426, 2020/03/01, <https://doi.org/10.1016/j.measurement.2019.107426>, 2020.
- [40] M. Rafati, F. Farnia, M.E. Taghvaei, A.M. Nickfarjam, "Fuzzy genetic-based noise removal filter for digital panoramic X-ray images," *Biocybernetics and Biomedical Engineering*, 38, no. 4, pp. 941–965, 2018/01/01, <https://doi.org/10.1016/j.bbe.2018.08.005>, 2018.
- [41] D.A. Góes, N.D.A. Mascarenhas, "Poisson denoising under a Bayesian nonlocal approach using geodesic distances with low-dose CT applications," *Digital Signal Processing*, 106, p. 102835, 2020/11/01, <https://doi.org/10.1016/j.dsp.2020.102835>, 2020.
- [42] T. Kirti, K. Jitendra, S. Ashok, "Poisson noise reduction from X-ray images by region classification and response median filtering," *Sādhanā* 6 (42) (2017) 855–863.
- [43] N. Karimi, M.R. Taban, "A convex variational method for super resolution of SAR image with speckle noise," *Signal Processing: image Communication*, 90, p. 116061, 2021/01/01, <https://doi.org/10.1016/j.image.2020.116061>, 2021.
- [44] N. Zarmehi, F. Marvasti, "Removal of sparse noise from sparse signals," *Signal Processing*, 158, pp. 91–99, 2019/05/01, <https://doi.org/10.1016/j.sigpro.2019.01.002>, 2019.
- [45] L. Wang, D. Xiao, W.S. Hou, X.Y. Wu, L. Chen, "Weighted Schatten p-norm minimization for impulse noise removal with TV regularization and its application to medical images," *Biomedical Signal Processing and Control*, 66, p. 102123, 2021/04/01, <https://doi.org/10.1016/j.bspc.2020.102123>, 2021.
- [46] M. Nooshyar, M. Momeny, "Removal of High Density Impulse Noise Using a Novel Decision Based Adaptive Weighted and Trimmed Median Filter," in: *2013 8th Iranian Conference on machine Vision and image processing (MVIP)*, IEEE, 2013, pp. 387–391.
- [47] P. I. Frazier, "A tutorial on Bayesian optimization," arXiv preprint arXiv: 1807.02811, 2018.
- [48] M. Pelikan, D.E. Goldberg, E. Cantú-Paz, BOA: the Bayesian optimization algorithm," in: *Proceedings of the genetic and evolutionary computation conference GECCO-99 1 (1999)* 525–532. Citeseer.
- [49] S. Lim, I. Kim, T. Kim, C. Kim, S. Kim, Fast AutoAugment," in *NeurIPS*, 2019.
- [50] C. Zhang, L. Zhou, Y. Zhao, S. Zhu, F. Liu, Y. He, "Noise reduction in the spectral domain of hyperspectral images using denoising autoencoder methods," *Chemometrics and Intelligent Laboratory Systems*, 203, p. 104063, 2020/08/15, <https://doi.org/10.1016/j.chemolab.2020.104063>, 2020.
- [51] B. Guo, K. Song, H. Dong, Y. Yan, Z. Tu, L. Zhu, "NERNet: noise estimation and removal network for image denoising," *Journal of visual communication and image representation*, 71, p. 102851, 2020/08/01, <https://doi.org/10.1016/j.jvcir.2020.102851>, 2020.
- [52] A. Krizhevsky, I. Sutskever, G.E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, 25 (2012) 1097–1105.
- [53] X. Zhang, X. Zhou, M. Lin, J. Sun, Shufflenet: an extremely efficient convolutional neural network for mobile devices, *Proceedings of the IEEE conference on computer vision and pattern recognition (2018)* 6848–6856.
- [54] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition (2016)* 770–778.
- [55] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, *Proceedings of the IEEE conference on computer vision and pattern recognition (2015)* 1–9.
- [56] M.E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al Emadi, M.B.I. Reaz, "Can AI help in screening viral and COVID-19 pneumonia?," *IEEE Access* 8 (2020) 132665–132676.