



Artificial intelligence-assisted multistrategy image enhancement of chest X-rays for COVID-19 classification

Hongfei Sun^{1,2}, Ge Ren¹, Xinzhi Teng¹, Liming Song¹, Kang Li¹, Jianhua Yang², Xiaofei Hu³, Yuefu Zhan⁴, Shiu Bun Nelson Wan⁵, Man Fung Esther Wong⁵, King Kwong Chan⁶, Hoi Ching Hailey Tsang⁶, Lu Xu⁶, Tak Chiu Wu⁷, Feng-Ming (Spring) Kong⁸, Yi Xiang J. Wang⁹, Jing Qin¹⁰, Wing Chi Lawrence Chan¹, Michael Ying¹, Jing Cai¹

¹Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong, China; ²School of Automation, Northwestern Polytechnical University, Xi'an, China; ³Department of Radiology, Southwest Hospital, Third Military Medical University (Army Medical University), Chongqing, China; ⁴Department of Radiology, Hainan Women and Children's Medical Center, Hainan, China; ⁵Department of Radiology, Pamela Youde Nethersole Eastern Hospital, Hong Kong, China; ⁶Department of Radiology and Imaging, Queen Elizabeth Hospital, Hong Kong, China; ⁷Department of Medicine, Queen Elizabeth Hospital, Hong Kong, China; ⁸Department of Clinical Oncology, The University of Hong Kong, Hong Kong, China; ⁹Department of Imaging and Interventional Radiology, The Chinese University of Hong Kong, Hong Kong, China; ¹⁰School of Nursing, The Hong Kong Polytechnic University, Hong Kong, China

Contributions: (I) Conception and design: H Sun, J Cai; (II) Administrative support: J Cai; (III) Provision of study materials or patients: M Ying; (IV) Collection and assembly of data: X Hu, Y Zhan, SBN Wan, MFE Wong, KK Chan, HCH Tsang, L Xu, TC Wu; (V) Data analysis and interpretation: H Sun, G Ren, X Teng, L Song, K Li, J Yang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jing Cai, PhD; Michael Ying, PhD. Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, 11 Yuk Choi Rd., Hong Kong, China. Email: jing.cai@polyu.edu.hk; michael.ying@polyu.edu.hk

Background: The coronavirus disease 2019 (COVID-19) led to a dramatic increase in the number of cases of patients with pneumonia worldwide. In this study, we aimed to develop an AI-assisted multistrategy image enhancement technique for chest X-ray (CXR) images to improve the accuracy of COVID-19 classification.

Methods: Our new classification strategy consisted of 3 parts. First, the improved U-Net model with a variational encoder segmented the lung region in the CXR images processed by histogram equalization. Second, the residual net (ResNet) model with multidilated-rate convolution layers was used to suppress the bone signals in the 217 lung-only CXR images. A total of 80% of the available data were allocated for training and validation. The other 20% of the remaining data were used for testing. The enhanced CXR images containing only soft tissue information were obtained. Third, the neural network model with a residual cascade was used for the super-resolution reconstruction of low-resolution bone-suppressed CXR images. The training and testing data consisted of 1,200 and 100 CXR images, respectively. To evaluate the new strategy, improved visual geometry group (VGG)-16 and ResNet-18 models were used for the COVID-19 classification task of 2,767 CXR images. The accuracy of the multistrategy enhanced CXR images was verified through comparative experiments with various enhancement images. In terms of quantitative verification, 8-fold cross-validation was performed on the bone suppression model. In terms of evaluating the COVID-19 classification, the CXR images obtained by the improved method were used to train 2 classification models.

Results: Compared with other methods, the CXR images obtained based on the proposed model had better performance in the metrics of peak signal-to-noise ratio and root mean square error. The super-resolution CXR images of bone suppression obtained based on the neural network model were also anatomically close to the real CXR images. Compared with the initial CXR images, the classification accuracy rates of the internal and external testing data on the VGG-16 model increased by 5.09% and 12.81%, respectively, while the values increased by 3.51% and 18.20%, respectively, for the ResNet-18 model. The numerical results were better than those of the single-enhancement, double-enhancement, and no-enhancement CXR images.

Conclusions: The multistrategy enhanced CXR images can help to classify COVID-19 more accurately than the other existing methods.

Keywords: Coronavirus disease 2019 (COVID-19); chest X-ray (CXR); bone suppression; super-resolution

Submitted Jun 15, 2022. Accepted for publication Sep 17, 2022. Published online Nov 10, 2022.

doi: 10.21037/qims-22-610

View this article at: <https://dx.doi.org/10.21037/qims-22-610>

Introduction

The coronavirus disease 2019 (COVID-19), initially caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is one of the most severe diseases threatening lung health worldwide. COVID-19 is highly infectious and significantly impacts patients with poor physical immunity or a history of lung disease (1,2). In addition, fear of COVID-19 can also lead to a decline in mental health, further exacerbating the deterioration of the existing health conditions (3,4). Chest X-ray (CXR) is a commonly used first-line medical imaging modality that can be used to diagnose lung disease. Compared with computed tomography (CT), a CXR has the advantages of a low radiation dose, a low inspection price, and high universality. It has become one of the most commonly used methods of COVID-19 detection (5,6).

However, the conventional CXR diagnostic method has some limitations in detecting and assessing COVID-19 lung disease. In clinical applications, doctors need to read the images manually. The workload is enormous, and the diagnosis results often depend on each doctor's clinical experience. The overly mechanical and repetitive work makes doctors prone to fatigue, which leads to misdiagnosis and missed diagnoses (7). Based on the increase of medical image data and the development of artificial intelligence (AI) technology, AI-powered enhanced images can effectively improve the efficiency of doctors' diagnoses (8-12). However, CXR is a projected 2-dimensional image obtained by X-ray irradiation of human tissue structures. The overlapping clavicle and rib in CXR have strong signals in the imaging region, which interferes with the imaging of the lung's soft tissue structures and increases the misdiagnosis rate of manual diagnosis or deep learning-based diagnosis models (13). Therefore, enhancing the imaging quality of images by suppressing bone signals in the CXR imaging region has important clinical significance because it can improve the accuracy of COVID-19 diagnosis.

Currently, the bone-suppression function in a CXR

mainly includes hardware-based and software-based methods. Hardware-based methods primarily use dual-energy subtraction equipment to simultaneously expose patients to different energies to generate soft tissue CXR images that eliminate bone signals. However, this method increases the patient's radiation dose, and the patient's respiration and heartbeat produce motion artifacts during the second exposure process. The quality of the acquired bone-suppressed CXR images is also unstable (14,15). Software-based methods use image processing technology to suppress bone signals. The various algorithms are categorized as traditional machine learning and deep learning methods (16-20). The more commonly used machine learning bone signal suppression methods include artificial neural networks (21-23), K-nearest neighbor regression (24), principal component analysis (PCA) (25), and segmentation-based unsupervised methods (26,27). Although these methods can suppress the clavicle and rib signals to a certain extent, the resultant CXR image still has prominent residual bone images. The effectiveness of the process depends on the accuracy of the regression or segmentation algorithm. In recent years, deep learning models have made significant progress in medical image processing tasks. They have also been gradually applied to the study of bone signal suppression in CXR (28-33). Gusarev *et al.* (34) used a multilayer convolution network model with a self-encoder to suppress bone signals as noise signals to obtain CXR images without bone signals. Oh *et al.* proposed using the wavelet domain information as a guide. A conditional generative adversarial network model was used to learn the mapping relationship between the original and dual-energy subtraction CXR (35). CXR images containing only soft tissue structures could be directly obtained through this model. Rajaraman *et al.* (36) used a residual net (ResNet) model with 16 layers of residual blocks and introduced a short-circuit connection of residual cascades to eliminate the model convergence problem caused by the disappearance of gradients. Doing so enabled

the model to thoroughly learn the feature association between the original domain (CXR images) and the target domain (bone-suppressed CXR images), effectively suppressing bone signals in authentic CXR images.

The above AI-based image processing methods can suppress the CXR bone signal. However, they are all processed based on the global CXR images, and the resolution of the CXR images used for classification tasks is relatively low (37,38). A low-resolution (LR) CXR loses part of the image feature information, impacting the accuracy with which COVID-19 can be detected. Currently, the methods of super-resolution reconstruction mainly include interpolation (39), reconstruction (40), and deep learning (41). Interpolation is the process of estimating unknown intermediate pixel values using known discrete pixel points. The commonly used interpolation methods include the nearest neighbor and bicubic types. These methods are relatively simple, less robust, and inaccurate in reconstructing details. The reconstruction-based method extracts high-frequency information from multiple frames of LR images in the same scene and fuses them to generate high-resolution (HR) images. It mainly includes an iterative back projection, projection onto convex sets, and the maximum posterior probability method (42,43). However, these methods are relatively complicated to calculate, and the complexity of the CXR imaging content affects the accuracy, which is difficult to guarantee. The learning-based method uses the convolutional neural network (CNN) to establish a nonlinear mapping relationship between LR and HR images. In the past 2 years, super-resolution reconstruction methods based on deep learning models have also been gradually applied to the task of HR CXR image acquisition (44-46). The models include deep recursive neural network (DRCN) (47), HR Network (HRNet) (48), and super-resolution generative adversarial network (SRGAN) (49). The super-resolution CXR image reconstructed based on the model can obtain more accurate detailed information on CXR images and has a lower radiation dose exposure to patients than do these other methods. Considering the complexity of network parameters, the computational efficiency, and the robustness of the model, we decided to use a relatively lightweight CNN model to reconstruct HR CXR images.

When classifying COVID-19, a CXR has a certain application proportion. As the primary imaging mode of diagnosis, CXR is mainly combined with deep learning models to detect COVID-19. COVID-19 detection research can be divided into binary and multiclass

classification tasks (50-52). However, most scholars in these fields focus on improving the classification model to enhance the classification accuracy of COVID-19 or only focus on the analysis of a single field in the image processing task (53,54). Without multiple image enhancement mechanisms, the classification accuracy of COVID-19 CXR images could be affected. There are differences in the scanning parameters used by CXR machines in different hospitals or medical institutions, which are reflected in the differences in the signal intensity distribution of various organs and tissues in CXR images. There is a limited CXR image data set; therefore, the performance of training accuracy and robustness of the classification model in external testing is worthy of further research. In addition, there is a contradiction between bone suppression and super-resolution tasks. For the bone suppression task, it is better to use a skip connection to fuse the interpolated input and the convolved feature map so that the obtained HR CXR image can have better imaging quality (55). However, for the super-resolution task, the purpose is to suppress the bone signals in the original CXR images; retention of shallow feature information affects the final performance of bone suppression. The comparative experiments have shown that a single deep learning network model is insufficient for effectively processing CXR images enhanced by multitasking.

Considering the above problems and referring to the idea of Stacked Generative Adversarial Networks (stackGAN), we propose a multistrategy enhanced method based on the combination of segmentation, bone suppression, and super-resolution to obtain lung-only bone-suppressed CXR images with HR (56). The enhanced CXR images can be used for the classification of COVID-19 and normal cases. This study makes 3 major contributions. (I) A bone-suppressed model based on a multidilated-rate strategy is proposed to improve the texture feature performance of the lung region in CXR images. (II) A step-by-step model training strategy is used to simplify the image processing mode of complex tasks. (III) The new method is used in the internal and external classification tasks to verify the application value of the enhanced CXR images in COVID-19 detection. This study sheds light on the prospect of further improving the performance of deep learning-based classifiers through multienhanced CXR images. We present the following article in accordance with the TRIPOD reporting checklist (available at <https://qims.amegroups.com/article/view/10.21037/qims-22-610/rc>).

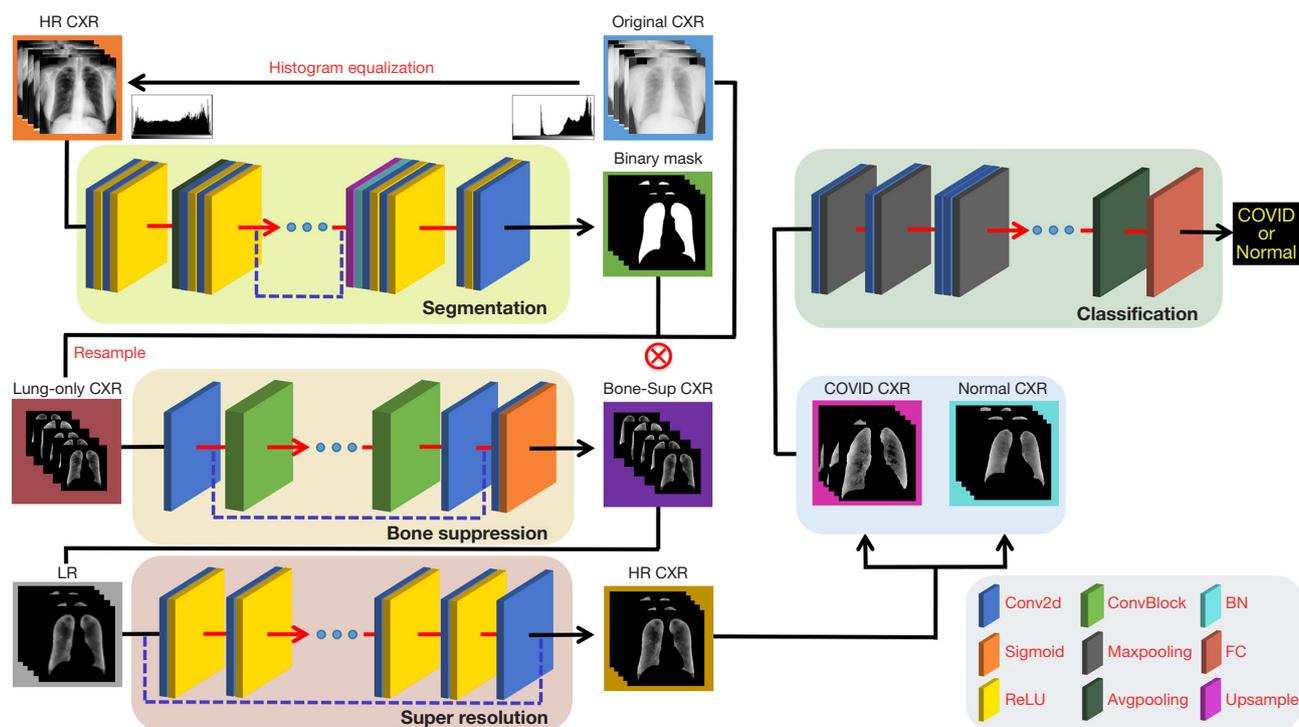


Figure 1 An overview of the acquisition of multistrategy enhanced CXR images and their application to COVID-19 detection. LR CXR, low-resolution chest X-ray; HR CXR, high-resolution chest X-ray; COVID-19, coronavirus disease 2019; Conv2d, convolutional layers contained two-dimensional convolution; ReLU, rectified linear units; ConvBlock, convolutional block; BN, batch normalization; FC, fully connected.

Methods

Overview

This study proposes an AI-assisted multistrategy enhancement technique to obtain lung-only bone-suppressed CXR images with HR. The process contained 3 stages. In the first stage, histogram equalization is performed on the CXR images to enhance the contrast of the images, and the improved U-Net model with variable encoder (VAE) is used to segment the lung region in the CXR images. In the second stage, the nonlinear mapping relationship between the CXR images with and without the bone region is established through the ResNet model with the multidilated rates to eliminate the bone signal in the original CXR images. In the third stage, the CNN model with residual cascade is used to improve the resolution of bone suppression CXR images containing only lung regions. In addition, the improved visual geometry group (VGG)-16 and ResNet-18 models are adopted to classify the CXR images based on the above multienhancement operations. The accuracy of the enhanced CXR images obtained based on this study used for COVID-19 detection was verified by the comparative experiments of various enhancement schemes.

The specific acquisition steps of multistrategy enhanced CXR images are shown in *Figure 1*.

Data and image preprocessing

The CXR images included in the experiment belonged to public data sets or clinical retrospective cases. The public data set for the bone suppression task was the Japanese Society of Radiological Technology (JSRT) (57), and the size of the images in the data set was $2,048 \times 2,048$. The pixel size was $0.175 \times 0.175 \text{ mm}^2$. Bone-suppressed CXR images acquired using the method proposed by Juhász *et al.* (58), were used as the ground truth for this task. For use in subsequent image processing tasks, the CXR and ground truth images were downsampled using ImageJ software (*National Institutes of Health*, Bethesda), the size of the processed images was 256×256 , and the pixel size was $1 \times 1 \text{ mm}^2$. This study randomly selected the data of 1,300 CXRs of normal and viral pneumonia classes in the public data set of the Radiological Society of North America (RSNA) for the super-resolution task (59). Similarly, ImageJ software was used to reduce the resolution of the original data, and the resulting LR image

Table 1 The data sets used for training, validation, and testing based on different image tasks

Image task	Training and validation stage	Testing stage
Bone suppression	JSRT [177]	JSRT [40]
Super-resolution	RSNA [1,200]	RSNA [100]
Classification	RSNA [Normal] [1,047]	PYNEH [Normal] [364]
	COVID-QU-Ex [COVID-19] [1,000]	MIDRC-RICORD-1C [COVID-19] [356]

The numbers in parentheses represent the size of the data sets. JSRT, Japanese Society of Radiological Technology; RSNA, Radiological Society of North America; PYNEH, Pamela Youde Nethersole Eastern Hospital; MIDRC-RICORD-1C, Medical Imaging Data Resource Center (MIDRC)-RSNA International COVID-19 Open Radiology Database (RICORD) Release 1c-chest X-ray COVID+; COVID-19, coronavirus disease 2019.

size was 256×256. The original 1,024×1,024 size data were used as the ground truth for this task. In the COVID-19 classification task, the COVID-19 and normal CXR image data in the training and testing stages came from different public data sets or medical institutions. We also selected 1,047 normal CXR images in the RSNA data set as one class of data in the training set of the classification task, and we selected the data of 1,000 cases of COVID-19 in the COVID-QU-Ex data set as another class (60). Data from 356 COVID-19 cases from the Medical Imaging Data Resource Center (MIDRC)-RSNA International COVID-19 Open Radiology Database (RICORD) Release 1c-CXR COVID+ (MIDRC-RICORD-1C) and 364 normal CXRs from Pamela Youde Nethersole Eastern Hospital (PYNEH) in Hong Kong were used as the 2 classes of data in the testing stage for the classification task (61). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics committee of the Hong Kong East Cluster Research (No. HKECREC-2020-119), and informed consent was obtained from all the patients. A summary of the various types of data sets used in this study is shown in *Table 1*.

Before the above models were trained to solve the problem of limited data, this study adopted the data augmentation method of geometric transformation, which increased the size of the training data set and enabled the enhanced data to build better deep learning models. In addition, since the lesions of COVID-19 are mainly concentrated in the lung regions of the human body, to avoid the interference of redundant pixel information on image processing tasks, the lung VAE model was used to segment the lung regions in the data set. Segmented lung-only CXR images were used in the subsequent tasks of bone suppression, super-resolution, and COVID-19 classification. Interested readers can refer to the paper by Selvan *et al.* (62) on the detailed design of the segmentation model with a variational autoencoder.

Bone suppression

CXR images containing only lung imaging were acquired in the preprocessing stage, which was followed by bone suppression. The bone suppression task aimed to use a deep learning model to fit the feature distribution between normal and no-bone CXR images to eliminate the overlapping bone components in the CXR images and ensure that the lung structural information was precise and accurate. *Figure 2* shows the acquisition of bone-suppressed CXR images. The mainframe of the proposed network consisted of 16 convolutional blocks, each of which contained 3 convolutional layers with different dilation convolution rates and 2 rectified linear units (ReLU). The output of each convolution block was the element-wise residual concatenation result between the current input and the feature map obtained after convolution according to the weight ratio. The sigmoid layer activated the final feature map to obtain the bone-suppressed CXR images. The size of the convolution kernel of each convolutional layer in the network was 3×3, the number of filters was 64, the stride was 1, and the padding was SAME. *Figure 3* shows the schematic diagram of the distribution of the receptive fields of different network layers based on the multidilated-rate strategy. This study selected the “3-2-1” dilation rate variation mode to fill the convolution block’s receptive field for the dilated convolutional layer group. *Figure 3* shows that an element in the feature map of the Nth layer corresponds to the 13×13 feature map in the N-3 layer; that is, it contains the shallow feature information in the above feature map, avoiding the spacing effect caused by the traditional “2-2-2” dilation rate variation mode. The feature information between network layers can be fully used in the new model, and the model has better robustness than the traditional dilation rate variation model.

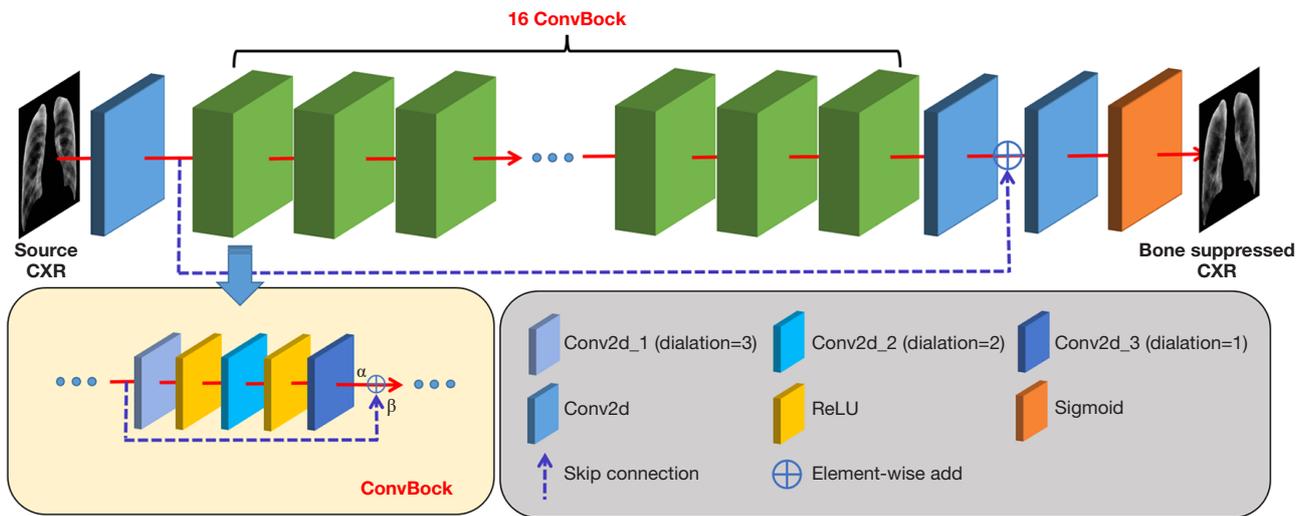


Figure 2 Flowchart of bone-suppressed CXR image acquisition. CXR, chest X-ray; Conv2d, convolutional layers contained two-dimensional convolution; ReLU, rectified linear units; ConvBlock, convolutional blkok.

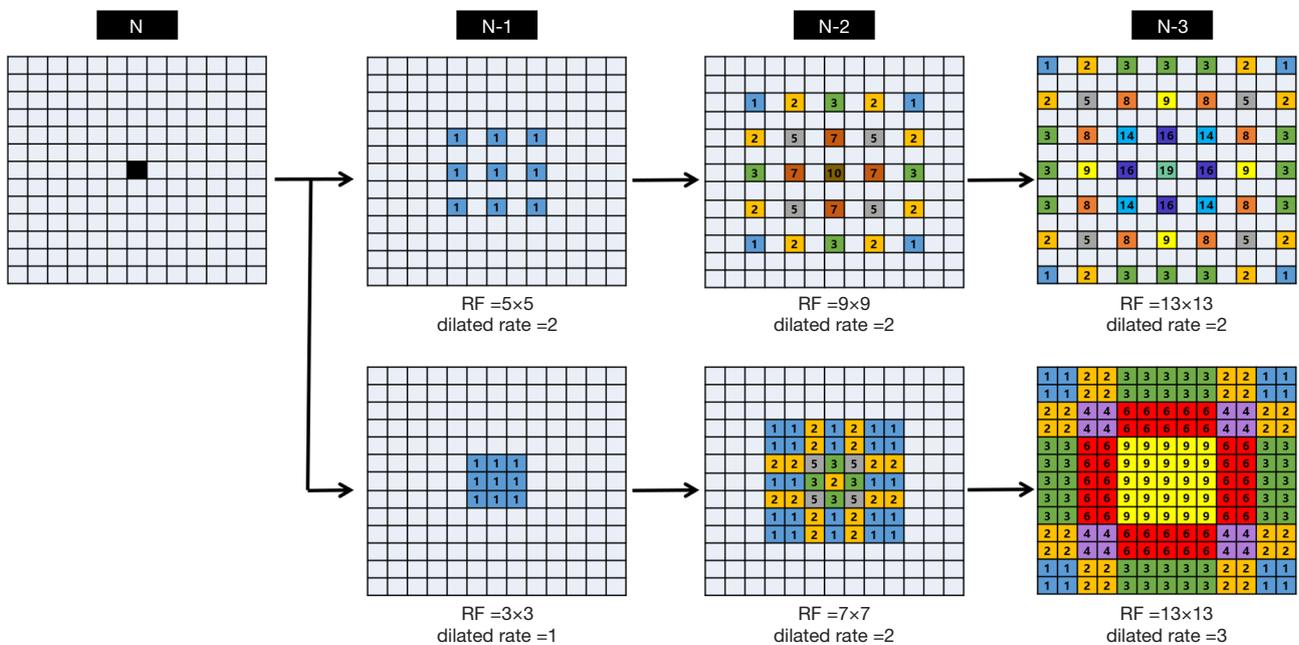


Figure 3 The schematic diagram of the distribution of the receptive fields of different network layers. The first row is the receptive field distribution based on the traditional dilation rate variation model. The second row is the receptive field distribution based on the multidilated-rate strategy. The numbers in the image indicate how many times the field of view used that position element. N means the Nth network layer, and RF denotes the receptive field.

Super-resolution

The acquisition of super-resolution CXR images involves upgrading LR CXR images to HR images through a

particular algorithm or model. After super-resolution processing, the CXR image has more pixels and anatomical structure details than it had before processing, which is conducive to the extraction of image feature information

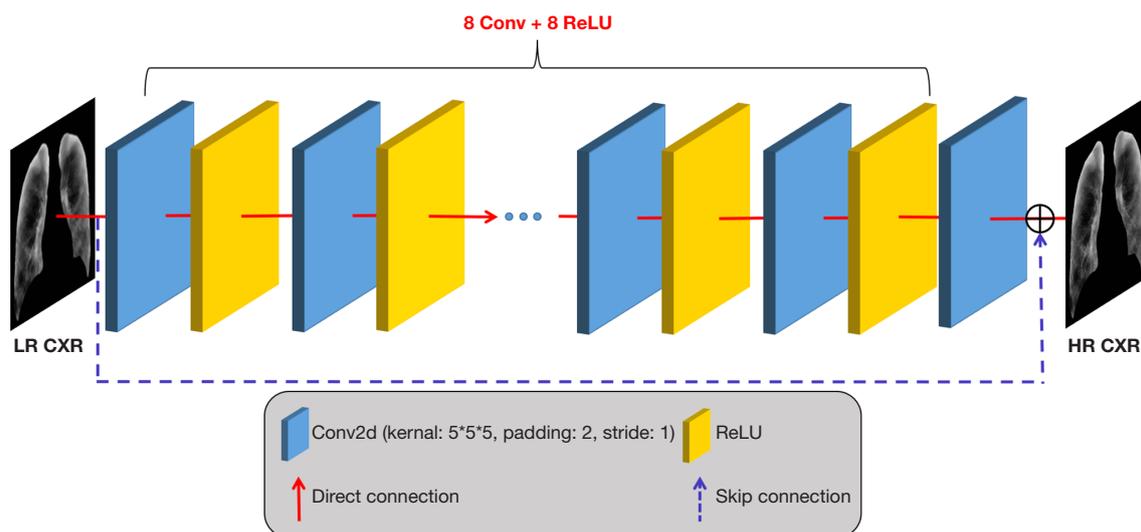


Figure 4 Flowchart of super-resolution CXR image acquisition. LR CXR, low-resolution chest X-ray; HR CXR, high-resolution chest X-ray; Conv, convolution; ReLU, rectified linear units; Conv2d, convolutional layers contained two-dimensional convolution.

and facilitates the diagnosis of COVID-19 by clinicians (63). This study chose to use a CNN model with 8 convolutional layers for super-resolution reconstruction of bone-suppressed CXR images. The LR CXR images with a size of 256×256 were used as the input for training. The first 7 convolutional layers contained two-dimensional convolution (Conv2d) and ReLU operations, and the last layer only had Conv2d operations. To reduce the training difficulty of the model, we performed bicubic interpolation on LR CXR images to change their dimensions to $1,024 \times 1,024$. The interpolated CXR images were connected with the feature maps after the eighth convolution layer by a long skip connection on element-wise. Finally, the HR CXR images with the size of $1,024 \times 1,024$ were output. The kernel size of each convolutional layer in the model was 5×5 , the stride was 1, the padding was 2, and the number of feature maps was 16, 32, 64, 64, 64, 64, and 1, respectively. The specific model architecture of the super-resolution task is shown in *Figure 4*.

Classification

After acquiring the multienhanced CXR image data, to verify the accuracy of the new data for COVID-19 detection, this study applied the data to the 2 classification tasks for COVID-19. The selected classification models were the VGG-16 and ResNet-18 models, with high universality in computer vision. Since the data sets used for

classification task training consisted of CXR images with a size of $1,024 \times 1,024$, they could not be directly used as the input of the classical network model for training. To solve this problem, we improved the architecture of the 2 classification models. In terms of preprocessing, for the control group of CXR images obtained without the super-resolution model, we performed traditional cubic spline interpolation operations on them, and the up-sampled images had the same size as the other HR CXR image data. To ensure the uniformity of the feature maps in each layer of the network layer of the classic model and reduce the interference of redundant pixels in the image background, the resize operation was performed on the HR CXR image, and the new image data 896×896 in size that was obtained was used as the input of the classification model. In terms of model architecture improvement, for the VGG-16 model, 2 additional convolution and pooling layers were added to the beginning of the original model, and the rest of the architecture remained unchanged. The convolution layers included conv, batch normalization (BN), and ReLU operations. The kernel size was 3×3 , the stride and padding were both 1, and the pooling layer was maximum pooling. We also replaced the fully connected layers in VGG-16 with an adaptive average pooling layer to avoid the overfitting problem by reducing the model training parameters. For the ResNet-18 model, this study modified the convolutional layer with a kernel size of 7×7 in the original model into 3 convolutional layers with a kernel size

of 3×3. The convolutional layers included conv, BN, and ReLU operations, and the maximum pooling layer was used for dimension reduction after each convolutional layer. To ensure the regular training of the model, we set the stride of the first residual block in ResNet-18 to 2, and the rest of the network layers of the model remained unchanged. Please refer to the relevant literature for a more detailed description of the 2 classical model architectures applied to classification tasks, the details of which are not repeated here (64,65).

The loss functions of different models

The loss functions in the image enhancement tasks involved in this study were optimized based on the difference in pixel information between the original domain and target domain images. For the bone suppression task, the loss function L_{BS} of the model included 2 parts: L_{MSE} and L_{SSIM} . L_{MSE} was the ratio of the squared sum of the pixel deviations between the ground truth CXR image X_g and the model-based bone-suppressed CXR image X_p to the number of pixels N of the total input image, and its function expression is shown in Eq. [1]:

$$L_{MSE}(X_g, X_p) = \frac{1}{N} \sum_{i=1}^N (X_{p,i} - X_{g,i})^2 \tag{1}$$

L_{SSIM} was based on the original structure similarity index (SSIM) function by calculating the SSIM values in the local patches under the corresponding M sliding windows in the X_g and X_p images and then averaging the numerical calculation results L_{MSSIM} . The specific mathematical formula is as follows:

$$L_{MSSIM}(X_g, X_p) = \frac{1}{N} \sum_{j=1}^M L_{SSIM}(X_{g,j}, X_{p,j}) \tag{2}$$

$$L_{SSIM}(X_g, X_p) = \frac{(2\mu X_g X_p + c_1)(2\sigma X_g X_p + c_2)}{(\mu X_g^2 X_p^2 + c_1)(\sigma X_g^2 + \sigma X_p^2)} \tag{3}$$

In summary, the total loss function of the bone suppression task is shown in Eq. [4], and the value is 0.84.

$$L_{BS}(X_g, X_p) = (1-\alpha)L_{MSE}(X_g, X_p) + \alpha L_{MSSIM}(X_g, X_p) \tag{4}$$

For super-resolution tasks, the loss functions L_{SR} of the model also includes L_{MSE} and L_{SSIM} . In addition, in this study, the Huber loss function was added to ensure the stability of HR image data training. The loss function has the advantages of using the mean absolute error and mean square error, which weakens the oversensitivity problem

of discrete group points in training and is derivable everywhere. While having a faster convergence speed, the change of the model training gradient was relatively more minor, avoiding the problem of gradient explosion or disappearance. The specific expressions of L_{SR} and L_{Huber} are shown in Eq. [5] and Eq. [6], where β is 0.7.

$$L_{SR}(X_g, X_p) = (1-\beta)[L_{MSE}(X_g, X_p) + L_{Huber}(X_g, X_p)] + \beta L_{MSSIM}(X_g, X_p) \tag{5}$$

$$L_{Huber}(X_g, X_p) = \begin{cases} \frac{1}{2}(X_p - X_g)^2, & |X_p - X_g| < 1 \\ |X_p - X_g| - \frac{1}{2}, & |X_p - X_g| \geq 1 \end{cases} \tag{6}$$

For the COVID-19 classification task, the loss functions of both classification models involved in the study were binary cross-entropy. The cross-entropy describes the distance between the predicted probability and the actual label. The smaller the numerical result of the cross-entropy, the higher the accuracy of the classification model prediction. The cross-entropy mainly uses negative log-likelihood to solve the classification problem. The value of the loss item L_c of the current category can be calculated through the softmax function, and the logarithmic evaluation and its specific expression are as follows:

$$L_c(x, class) = -\log\left(\frac{e^{x[class]}}{\sum_k e^{x[k]}}\right) \tag{7}$$

The target category adopts one-hot encoding. The class represents the index corresponding to the current sample category in one-hot encoding, in which $x[k]$ is the k th output predicted by the model.

Training details

The model network frameworks involved in this study were built by the Pytorch library. The computing platform used for training was NVIDIA GeForce RTX 3090 GPU with 24 GB video memory, AMD Ryzen 9 5900X 12-Core CPU, and 32 GB RAM, which could meet the requirements for the calculation of network models. In terms of the details of hyperparameter selection, for the network model of the bone suppression task, the batch size was 10, the epochs were 200, the initial learning rate was set to 10^{-3} , and the learning rate of the model was 0.75 times the original every 50 epochs. The proportions of the deep and shallow feature maps in the residual cascade of ResNet were 0.9 and 0.1, respectively, and the optimizer of the network model

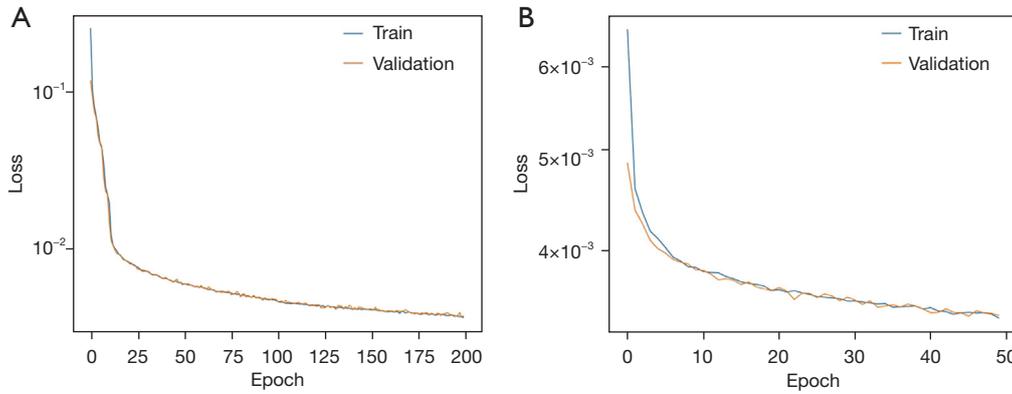


Figure 5 Training curves of the bone suppression and super-resolution task. (A) Bone suppression. (B) Super-resolution.

was Adam. For the super-resolution task, the batch size of the network model was set to 5, the epochs were 50, the learning rate was 10⁻⁴, and the optimizer also used Adam. The training curves of the models are shown in *Figure 5*. *Figure 5A* and *Figure 5B* are the optimized results of bone suppression and super-resolution, respectively. The loss value of the model gradually decreased and tended to be stable within the specified epoch. For the improved VGG-16 and ResNet-18 models in the classification task, the batch size was set to 10, the epochs were 80 and 50, and the learning rates were 10⁻⁵ and 5x10⁻⁶, respectively. The optimizers used in the 2 models were stochastic gradient descent (SGD) and Adam. The weight decay of the former was set to 0.01, and the momentum decay factors of the latter were the default values of 0.9 and 0.999, respectively.

Evaluation

This study mainly used peak signal-to-noise ratio (PSNR), SSIM, and root mean square error (RMSE) to evaluate the similarity between the CXR images synthesized by deep learning models and ground truth CXR images (66,67).

For the quantization index PSNR, the formula is as follows:

$$PSNR = 10 \log \left(\frac{MAX_I^2}{\sqrt{\frac{1}{XY} \sum_{x=0}^{Y-1} \sum_{y=0}^{X-1} |I_{gr}(x,y) - I_{ps}(x,y)|^2}} \right) \quad [8]$$

In Eq. [8], I_{gr} and I_{ps} represent the ground truth and the pseudo-CXR images obtained based on the deep learning model, respectively. X and Y represent the size of the image, and MAX_I represents the maximum gray value in

the CXR images. The higher the PSNR value is, the closer the synthesized pseudo CXR image is to the ground truth. The remaining 2 quantitative indicators, SSIM and RMSE, are described in the previous section. In addition, compared with the super-resolution task, the original image of the data set used in the bone suppression task was more minor. To avoid the randomness of training, we used 8-fold cross-validation to train the bone suppression model. The CXR image data were randomly divided into 8 groups: 7 groups were selected for training in each experiment, and the remaining group of data was used for interesting.

For the COVID-19 classification task, a visual confusion matrix was often used to represent the prediction results of the classification model. In terms of quantitative evaluation, the indicators mainly included accuracy, precision, recall, F1-score value, and area under the curve (AUC). Their corresponding mathematical expressions are as follows:

$$accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad [9]$$

$$F1 - score = \frac{2 \cdot precision \cdot recall}{precision + recall} \quad [10]$$

$$precision = \frac{TP}{TP + FP} \quad [11]$$

$$recall = \frac{TP}{TP + FN} \quad [12]$$

TP is the number of normal data samples that the model predicts as normal, TN is the number of CXR data samples predicted to be COVID-19, and FP represents the number of COVID-19 data samples falsely predicted to be normal by the model. FN represents the number of normal CXR data samples falsely predicted to be COVID-19. Therefore,

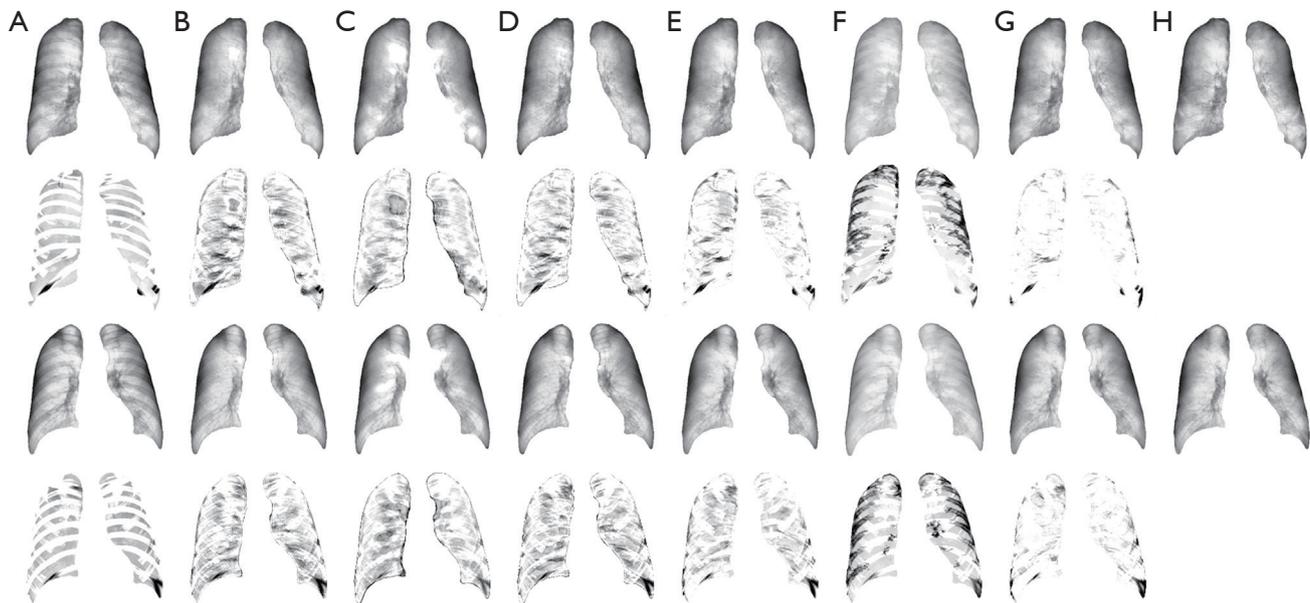


Figure 6 The random-selected bone-suppressed lung-only CXR images obtained based on different deep learning models. The second and fourth rows show the difference comparison maps. The closer the color is to black, the more significant the difference. (A) Original CXR. (B) CXR obtained based on the 5-layer CNN model. (C) CXR obtained based on the 6-layer CNN model. (D) CXR obtained based on the autoencoder model. (E) CXR obtained based on the ResNet model. (F) CXR obtained based on the DecGAN model. (G) CXR obtained based on the proposed model. (H) Ground truth bone-suppressed CXR. CXR, chest X-ray; CNN, convolutional neural network; DecGAN, decoupling generative adversarial network.

based on the above 4 values, it can be determined that $TP + FN + FP + TN$ is the total number of all samples, $TP + FN$ is the number of CXR data samples of all true normal classes, $FP + TN$ is the number of CXR data samples of all true COVID-19 classes, $TP + FP$ is the number of samples with all predicted values of the normal class, and $TN + FN$ is the number of samples with all predicted values of the COVID-19 class.

For the above indicators, the larger the value is, the better the model's performance. Suppose the recall obtained by the classification model is low. In that case, there are more false-negative patients in the predicted value, which affects the diagnosis of COVID-19 cases and causes missed diagnoses.

Gradient-weighted class activation mapping (Grad-CAM++) was also used in this study to provide a visual display of the predictive performance of the classification models (68). Grad-CAM++ fuses the class discriminative property of the class activation mapping (CAM) with existing pixel-space gradient visualization techniques, which highlight fine-grained details in the image. Grad-CAM++ can localize the predicted class more accurately

than can Grad-CAM. The generated heatmaps better explain the performance of the classification model when multiple instances of a single class are present in the image. The redder the region in the heatmap is, the greater the contribution of the corresponding region in the CXR image to the classification model. The detected COVID-19 regions are framed by the bounding boxes.

Results

Qualitative evaluation results of multistrategy enhanced CXR images

Figure 6 shows the visualization results of bone-suppressed CXR images obtained based on different deep learning models. Figure 6 shows that the image results obtained based on the model proposed in this study could better suppress the bone regions in the CXR images in terms of details. The rib and clavicle signals in the lung imaging region could be removed well. The contrastive difference map shows relatively small anatomical differences between the bone-suppressed CXR images proposed in this study

Table 2 The average numerical results of PSNR (dB) and RMSE under different deep learning models in the bone suppression task

Methods	PSNR (dB)	RMSE
DecGAN	32.89±0.87	0.0228±0.0023
5-layer CNN	38.46±2.13	0.0124±0.0033
6-layer CNN	40.27±2.42	0.0101±0.0030
Autoencoder	37.17±1.33	0.0139±0.0023
ResNet	41.55±2.68	0.0080±0.0030
Propose	43.21±3.14*	0.0074±0.0029*

The values are presented as mean ± standard deviation. *, the best performance. PSNR, peak signal-to-noise ratio; RMSE, root mean square error; DecGAN, decoupling generative adversarial network; CNN, convolutional neural network; ResNet, residual net.

and the ground truth. *Table 2* shows the quantitative measurement results between the target CXR images and the bone-suppressed CXR images obtained by different deep learning models. The numerical results in the *Table 2* show that the similarity between the CXR images obtained based on the model proposed in this study and the target CXR was relatively high. Compared with the ResNet (36), Autoencoder (34), 6-layer CNN, 5-layer CNN, and decoupling generative adversarial network (DecGAN) (30), the proposed model had an average increase of about 1.7, 6.0, 2.9, 4.7, and 10.3 dB for the PSNR metric. The values for the RMSE metric were reduced by an average of 0.0006, 0.0065, 0.0027, 0.0050, and 0.154, respectively. After addition of a multidilated-rate convolutional layer to the ResNet model, the visualization and numerical results showed that the nonlinear mapping relationship between the original and the enhanced CXR image could be better established. The bone signals in the enhanced CXR images obtained by the proposed model were suppressed while retaining more accurate feature information of soft tissues. The fold with the best performance in the 8-fold cross-validation was selected as the model for the bone suppression task, which had better numerical measurement results of PSNR and RMSE.

Figure 7 shows the visualization results of the synthesized super-resolution CXR image based on the improved CNN model and the bicubic interpolation algorithm. *Figure 7* shows that the CNN model used in this study obtained not only HR CXR images with bone signals but was also suitable for super-resolution reconstruction of CXR images after bone suppression. Further, the numerical results of PSNR, SSIM, and RMSE of CXR images obtained by the CNN

model were better than those of traditional interpolation algorithms, and HR CXR images met the image quality requirements of the next-order classification task. *Table 3* shows the quantitative results of the super-resolution CXR images obtained based on the traditional cubic spline interpolation and CNN model. For the CXR images with bone, compared with the interpolation algorithm, the super-resolution method based on the CNN model increased the PSNR and SSIM values by about 9.5 dB and 2.1%, respectively, and reduced the RMSE values by about 0.0196. For the bone-suppressed CXR images, the PSNR and SSIM values increased by about 11.0 dB and 2.2%, respectively, and the RMSE value decreased by about 0.0265.

Accuracy evaluation of COVID-19 classification

Figure 8 shows the example explanation heatmaps and COVID-19 detection results for original and multienhanced CXR images. The first 2 lines are the original CXR images and their corresponding visual analysis results, and the last 2 lines are the results of multienhanced CXR images. *Figure 8A* shows COVID-19 CXR images, and the first and third rows correspond to the original and multienhanced CXR images of the same COVID-19 patient, respectively, as do the second and fourth rows. *Figure 8B, 8D* are the heatmaps generated by Grad-CAM++ based on the improved VGG-16 classification model and the corresponding COVID-19 region detection maps. *Figure 8C, 8E* are the visual results based on the improved ResNet-18 classification model. As shown in the *Figure 8*, the performance of the classification models could be improved based on the multienhanced CXR so that the models could more accurately predict COVID-19 than could the other models.

Figure 9A, 9B are the confusion matrix results of internal and external testing of the improved VGG-16 model based on CXR image data obtained by different enhancement methods. *Figure 10A, 10B* are the confusion matrices of the internal and external testing of the improved ResNet-18 model, respectively. The closer the color in the confusion matrix image is to the darker color, the greater the amount of data the model predicts for this class, and vice versa. It can be seen from the *Figures 9, 10* that, for the internal and external test data sets, the classification results trained on the CXR image data based on any enhancement method had an improved ability to recognize the 2 classes of data. This shows that the training model based on the enhanced CXR image proposed in this study had relatively good accuracy in recognizing the 2 categories of images.

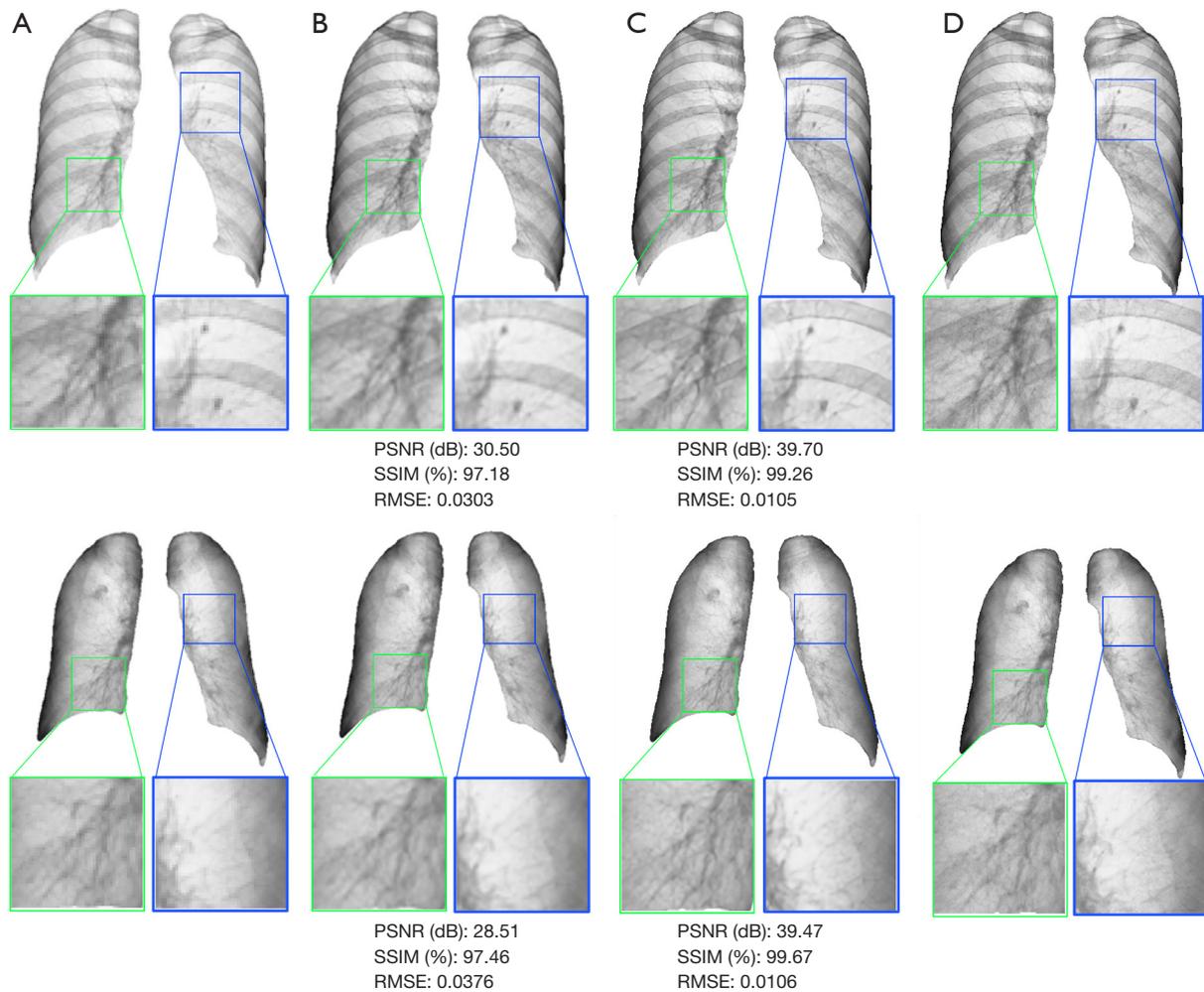


Figure 7 The visual and quantitative results between the synthesized super-resolution CXR image based on the improved CNN model and the bicubic interpolation algorithm. (A) LR CXR images. (B) HR CXR images obtained by bicubic interpolation. (C) HR CXR images obtained based on an improved CNN model. (D) Ground truth CXR images. PSNR, peak signal-to-noise ratio; SSIM, structure similarity index; RMSE, root mean square error; CNN, convolutional neural network; LR CXR, low-resolution chest X-ray; HR CXR, high-resolution chest X-ray.

Table 3 The quantitative results of the super-resolution task with bone or bone suppression obtained based on the bicubic interpolation and improved CNN model

CXR images	Methods	PSNR (dB)	SSIM (%)	RMSE
With bone	Bicubic	30.25±1.00	97.13±0.49	0.0300±0.0034
	CNN	39.71±0.96	99.21±0.24	0.0104±0.0012
Bone-suppressed	Bicubic	28.50±0.47	97.48±0.27	0.0371±0.0020
	CNN	39.51±0.43	99.66±0.07	0.0105±0.0005

The values are presented as mean ± standard deviation. CNN, convolutional neural network; CXR, chest X-ray; PSNR, peak signal-to-noise ratio; SSIM, structure similarity index; RMSE, root mean square error.

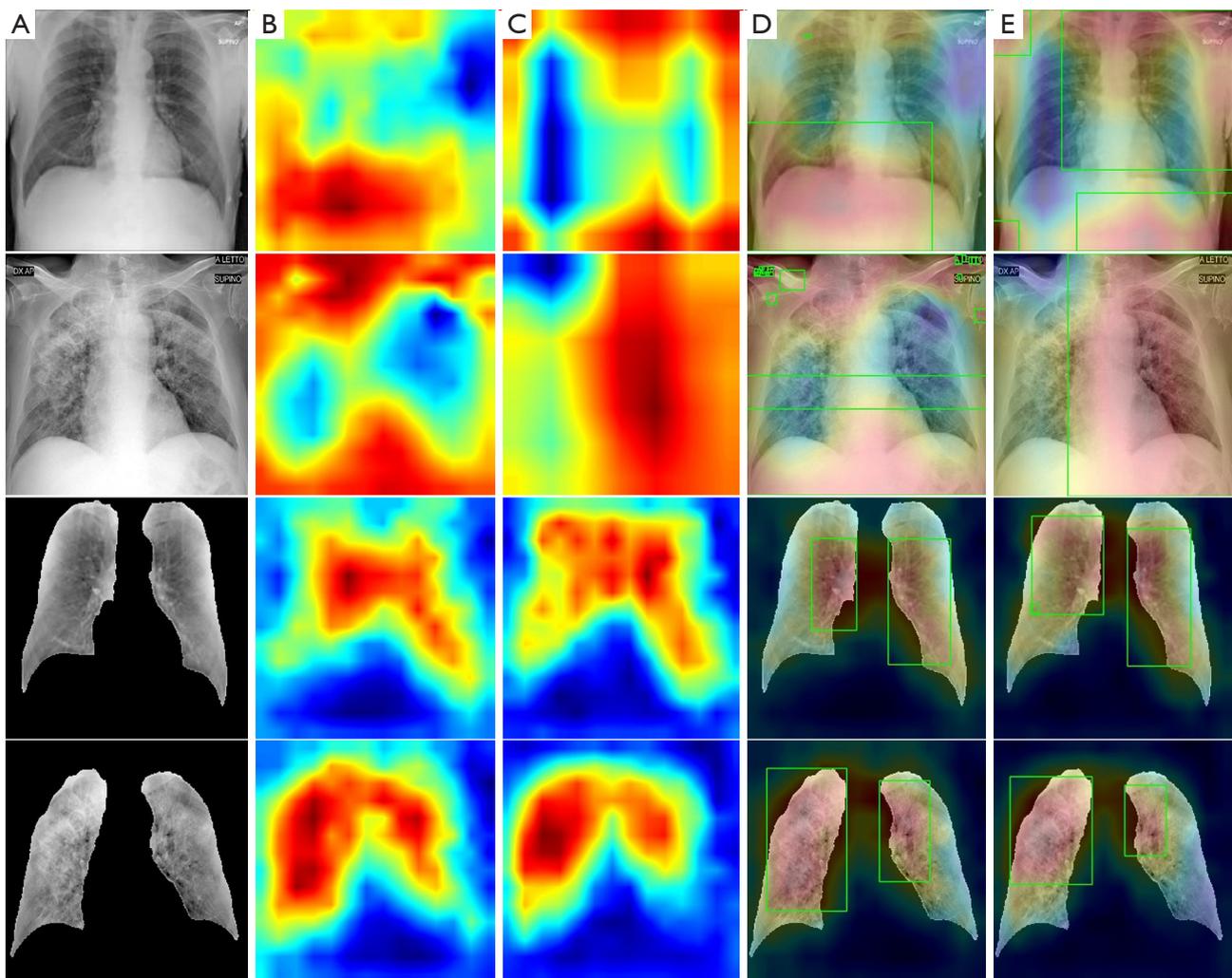


Figure 8 Example explanation heatmaps and COVID-19 detection results for CXR images. The first 2 lines are the original CXR images and their corresponding visual analysis results, and the last 2 lines are the results of multienhanced CXR images. (A) COVID-19 images. (B) The heatmaps generated by Grad-CAM++ based on the improved VGG-16 classification model. (C) The heatmaps generated by Grad-CAM++ based on the improved ResNet-18 classification model. (D) COVID-19 region detection maps based on the improved VGG-16 model. (E) COVID-19 region detection maps based on the improved ResNet-18 model. COVID-19, coronavirus disease 2019; CXR, chest X-ray; Grad-CAM++, gradient-weighted class activation mapping; VGG, visual geometry group.

Figure 11 show the different metrics of the VGG-16 model for classifying 2 CXR image data categories on the internal and external testing data sets, respectively. Figure 12 show the various metrics of the ResNet-18 model for classifying 2 CXR image data categories on the internal and external test data sets. The quantitative analysis showed that the accuracy of the 2 classification models after training based on the multistrategy enhanced CXR image data obtained in this study was the highest. The accuracy rates of internal and external testing based on the VGG-16 model

were 92.97% and 83.06%, respectively. The accuracy rates of the internal and external tests based on the ResNet-18 model were 90.23% and 78.89%, respectively. Based on the CXR images obtained by the enhancement method proposed in this study, the paired *t*-test was performed with the comparison experiment. The difference between the accuracy rates was statistically significant ($P < 0.05$). In addition, for the VGG-16 model, compared with several other comparison methods, the F1-score values based on the multienhanced CXR classification were the highest

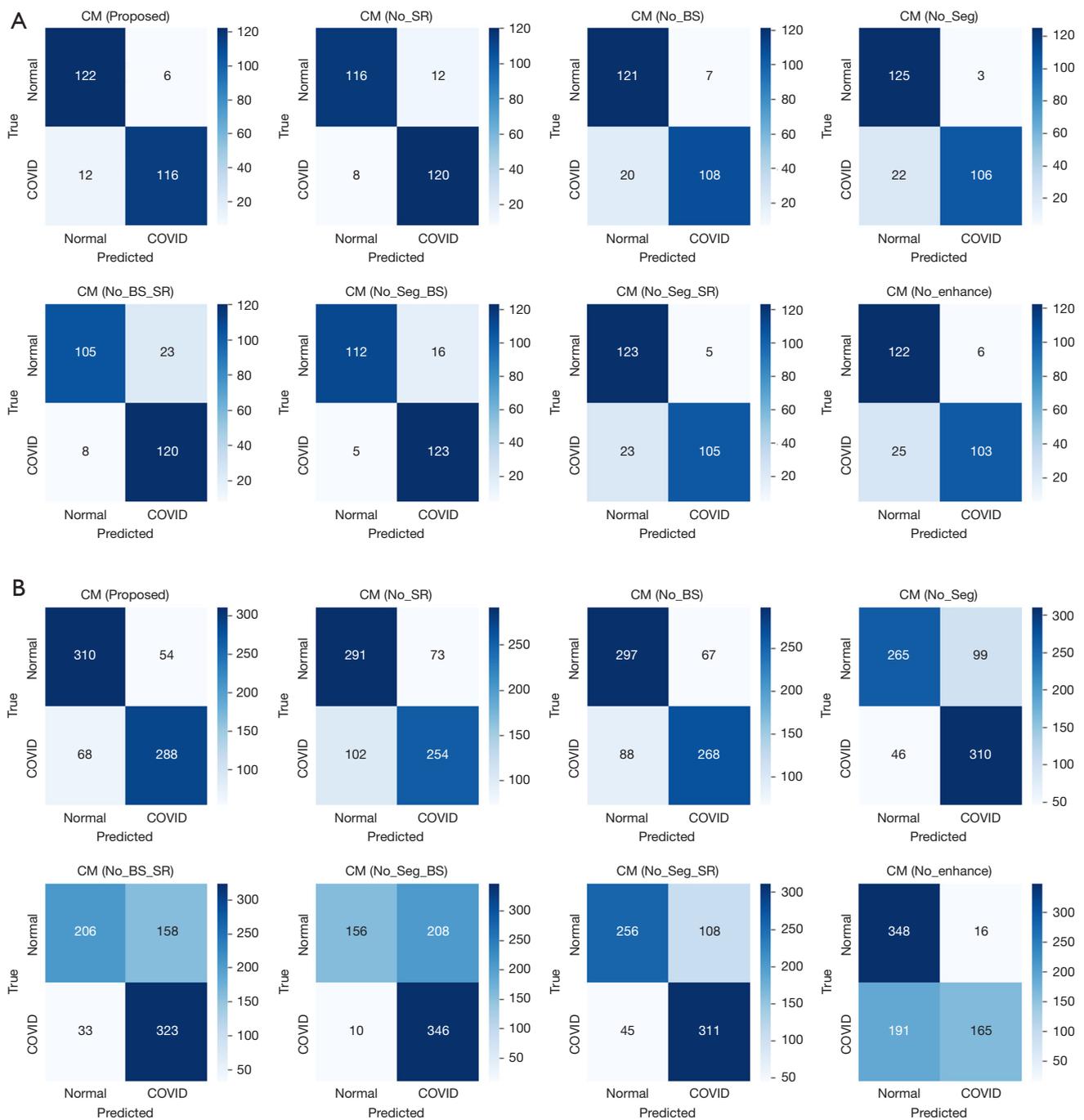


Figure 9 The confusion matrix results of the improved VGG-16 model based on different CXR images. (A) The confusion matrix results of internal testing of the improved VGG-16 model. (B) The confusion matrix results of external testing of the improved VGG-16 model. Proposed, CXR images obtained based on the multienhancement for the classification; No_SR, CXR images without super-resolution enhancement for the classification task; No_BS, CXR images without bone suppression enhancement for the classification task; No_Seg, CXR images without segmentation enhancement for the classification task; No_BS_SR, CXR images without bone suppression and super-resolution enhancement for the classification task; No_Seg_BS, CXR images without segmentation and bone suppression enhancement for the classification task; No_Seg_SR, CXR images without segmentation and super-resolution enhancement for the classification task; No_enhance, original CXR images for the classification task; CXR, chest X-ray; COVID, coronavirus disease; CM, confusion matrix; VGG, visual geometry group.

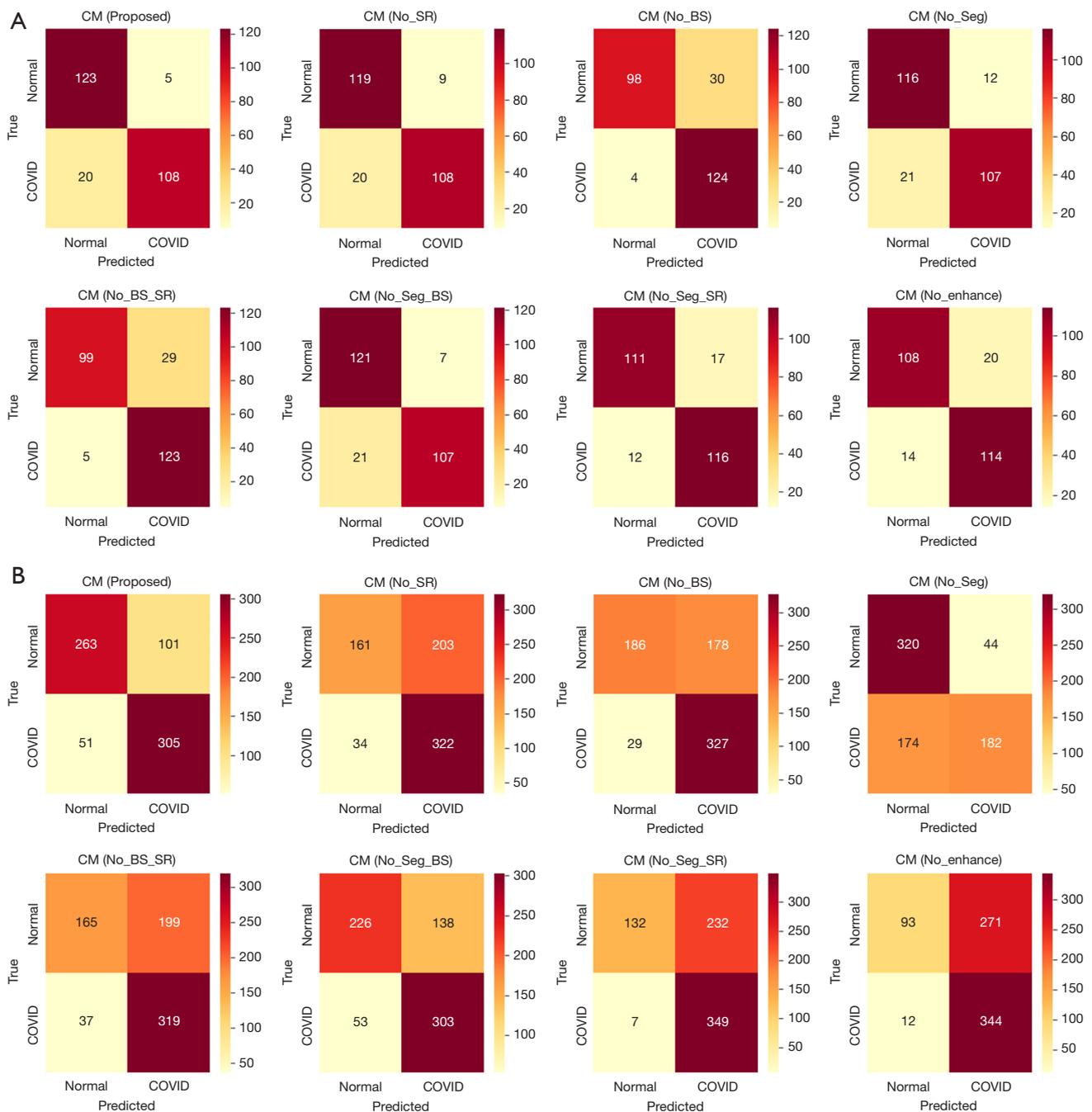


Figure 10 The confusion matrix results of the improved ResNet-18 model based on different CXR images. (A) The confusion matrix results of internal testing of the improved ResNet-18 model based on CXR image data obtained by different enhancement methods. (B) The confusion matrix results of external testing of the improved ResNet-18 model based on CXR image data obtained by different enhancement methods. Proposed, CXR images obtained based on the multienhancement for the classification; No_SR, CXR images without super-resolution enhancement for the classification task; No_BS, CXR images without bone suppression enhancement for the classification task; No_Seg, CXR images without segmentation enhancement for the classification task; No_BS_SR, CXR images without bone suppression and super-resolution enhancement for the classification task; No_Seg_BS, CXR images without segmentation and bone suppression enhancement for the classification task; No_Seg_SR, CXR images without segmentation and super-resolution enhancement for the classification task; No_enhance, original CXR images for the classification task; CXR, chest X-ray; COVID, coronavirus disease; ResNet, residual net; CM, confusion matrix.

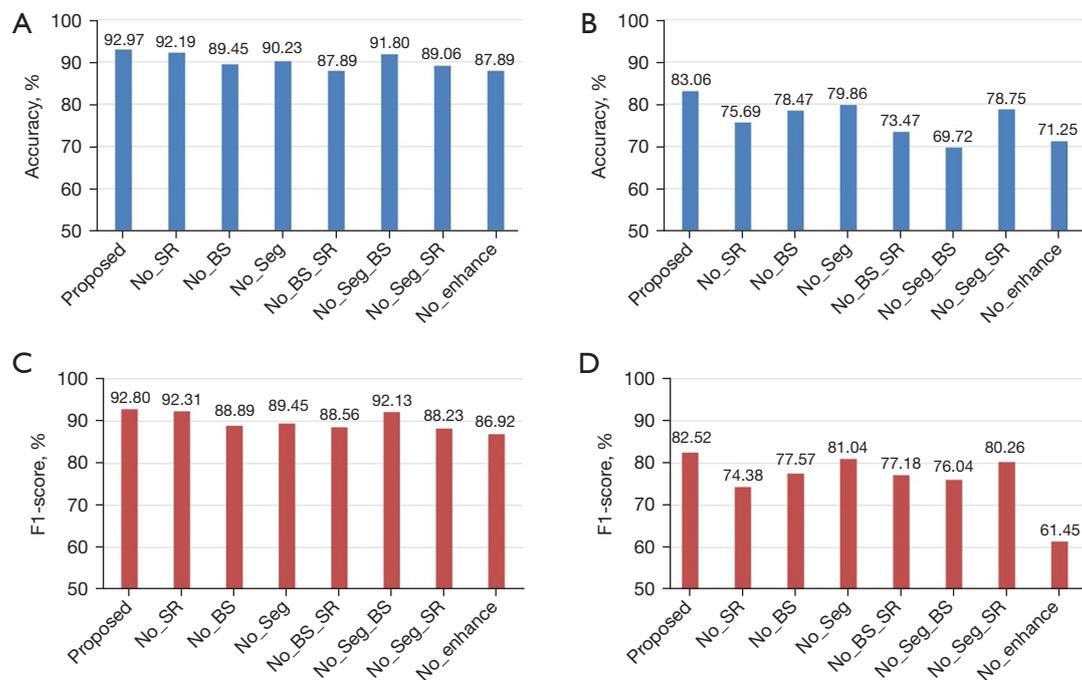


Figure 11 The performance metrics of the VGG-16 model for classifying 2 CXR image data categories on the internal and external testing data sets. (A) The accuracy values for the internal testing data sets. (B) The accuracy values for the external testing data sets. (C) The F1-score values for the internal testing data sets. (D) The F1-score values for the external testing data sets. Proposed, CXR images obtained based on the multienhancement for the classification; No_SR, CXR images without super-resolution enhancement for the classification task; No_BS, CXR images without bone suppression enhancement for the classification task; No_Seg, CXR images without segmentation enhancement for the classification task; No_BS_SR, CXR images without bone suppression and super-resolution enhancement for the classification task; No_Seg_BS, CXR images without segmentation and bone suppression enhancement for the classification task; No_Seg_SR, CXR images without segmentation and super-resolution enhancement for the classification task; No_enhance, original CXR images for the classification task; CXR, chest X-ray; VGG, visual geometry group.

in both the internal and external tests, and their values were 92.80% and 82.52%, respectively. Similarly, the F1-score values based on multienhanced CXR classification were also the highest in the internal and external tests of the ResNet-18 model, with values of 89.63% and 80.05%, respectively. The classification stability and robustness of the models were also relatively good. For the evaluation index of AUC, the numerical results of the internal and external tests of the VGG-16 model were 0.930 and 0.830, and the numerical results of the ResNet18 model were 0.902 and 0.790, respectively. The AUC numerical results of the 2 classification models based on different enhanced CXR images are shown in *Table 4*.

Discussion

In this study, multistrategy enhanced CXR images were

used to improve the accuracy of classifying COVID-19. It can be seen from the experimental results that the multistrategy enhanced CXR images could better represent the characteristic information of COVID-19 than the other enhanced CXR images. Compared with the baseline CXR images, the multistrategy enhanced CXR images used in the internal COVID-19 classification testing increased accuracy by approximately 3–5%. The accuracy rates in the external testing e increased by approximately 12–18%. For the F1-score metric, the numerical results based on internal and external tests increased by about 2.5–6% and 9–21%, respectively. These findings have significance for the auxiliary diagnosis of patients with COVID-19 in clinical practice.

As mentioned above, this study adopted the ResNet network with dilated convolutional layers to build the framework of the bone suppression training model.

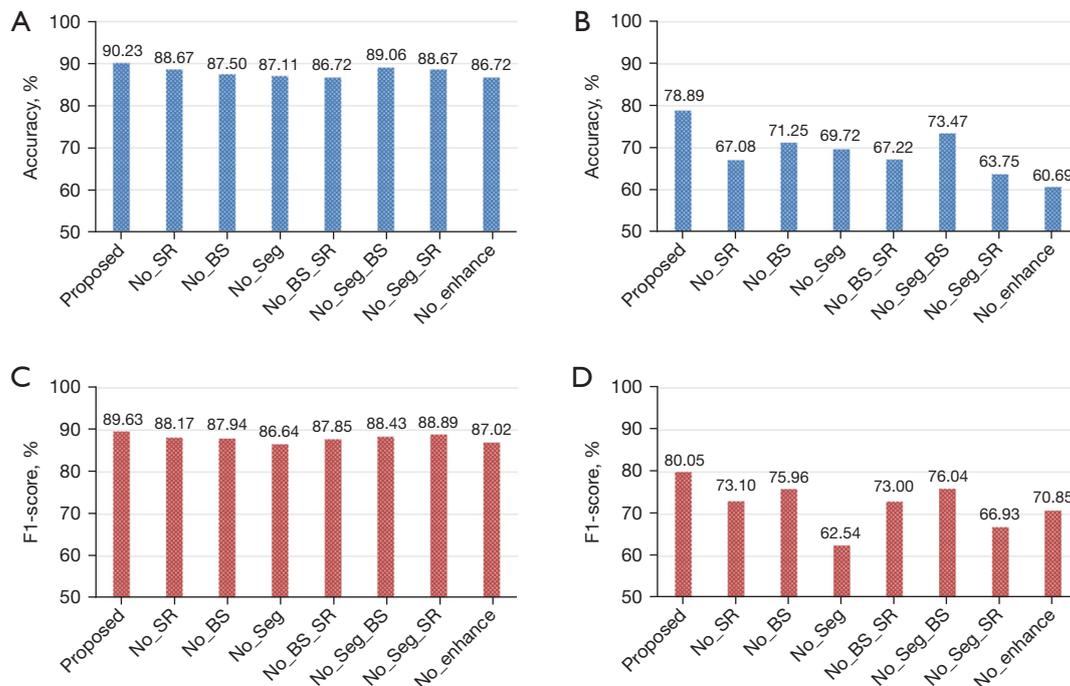


Figure 12 The performance metrics of the ResNet-18 model for classifying 2 CXR image data categories on the external testing data sets. (A) The accuracy values for the internal testing data sets. (B) The accuracy values for the external testing data sets. (C) The F1-score values for the internal testing data sets. (D) The F1-score values for the external testing data sets. Proposed, CXR images obtained based on the multienhancement for the classification; No_SR, CXR images without super-resolution enhancement for the classification task; No_BS, CXR images without bone suppression enhancement for the classification task; No_Seg, CXR images without segmentation enhancement for the classification task; No_BS_SR, CXR images without bone suppression and super-resolution enhancement for the classification task; No_Seg_BS, CXR images without segmentation and bone suppression enhancement for the classification task; No_Seg_SR, CXR images without segmentation and super-resolution enhancement for the classification task; No_enhance, original CXR images for the classification task; CXR, chest X-ray images; ResNet, residual net.

Table 4 The numerical AUC results of classification models of the VGG-16 and ResNet-18 models based on different enhanced CXR images

Model	Testing type	Proposed CXR	No_SR CXR	No_BS CXR	No_Seg CXR	No_BS_SR CXR	No_Seg_BS CXR	No_Seg_SR CXR	No_enhance CXR
VGG-16	Internal	0.930	0.922	0.895	0.902	0.879	0.918	0.891	0.879
	External	0.830	0.756	0.784	0.799	0.737	0.700	0.788	0.710
ResNet-18	Internal	0.902	0.887	0.875	0.871	0.867	0.891	0.887	0.867
	External	0.790	0.673	0.715	0.695	0.675	0.736	0.639	0.611

The values are presented as mean \pm standard deviation. AUC, area under curve; VGG, visual geometry group; ResNet, residual net; CXR, chest X-ray; Proposed, multistrategy enhanced; No_SR, no super resolution; No_BS, no bone suppression; No_Seg, no segmentation; No_BS_SR, no bone suppression and super resolution; No_Seg_BS, no segmentation and bone suppression; No_SR, no super resolution; No_enhance, no enhancement.

Compared with traditional convolution layers, the dilated convolution has 2 advantages. First, it increases the receptive field so that each convolution output contains a

more extensive range of information. The receptive field is the region in the input space at which a particular CNN feature looks. Its central location and size can describe a

receptive field of an element. In a deep net, downsampling (pooling or Conv2d with 2 strides) is always carried out to increase the receptive field and reduce the amount of calculation. Although the receptive field can be improved, the spatial resolution is reduced. In order not to lose resolution and still expand the receptive field, dilated convolution can be used. Second, dilated convolution can capture multiscale context information. Dilated convolution has a parameter that can set the dilation rate, and the specific meaning is how many zeros are filled in the convolution kernel. Therefore, the receptive fields differ when different dilation rates are set. That is, multiscale information is obtained. However, the model cannot fully use the pixel information if the same dilated rate-based kernel is adopted. Some features are used many times, and some feature information is not used at all. Here, in the improved model, the receptive fields of convolutional layers with different dilated rates are inconsistent. All feature points in the feature map can be obtained in crosswise fashion, which avoids the loss of feature information and better distinguishes the foreground (soft tissue) and background (bone) regions to suppress the bone signal in the CXR images, making the lung texture imaging clearer. Therefore, its quantitative and qualitative results for the bone suppression task of CXR were better than those of ResNet. For the autoencoder and CNN models, the results of bone suppression were inferior to those of the proposed method. Both models do not contain skip connections; therefore, they cannot fully combine the shallow and deep feature information of the CXR images, and the fitting ability of the models is insufficient, resulting in the loss of the detailed structural information of the CXR images. For the DecGAN model, the numerical results obtained in the bone suppression task were relatively stable. However, the model was trained based on unsupervised learning. The obtained bone-suppressed images still retained the edge contours of some bone regions. The enhanced images still retained the edge contours of some bone regions, resulting in a mutation in the grayscale information of the soft tissues. In terms of quantitative and qualitative results, the DecGAN model also exhibited relatively significant differences with the ground truth images.

In acquiring HR CXR images of bone suppression, this study used a step-by-step strategy to divide the total task into the LR CXR image bone suppression task and HR CXR image reconstruction task to replace the direct bone suppression processing of HR CXR images with bone signals. The reason for this was that the training of deep

learning models based on HR image data often requires a long time for iteration and optimization, which depends on the computer's video memory, virtual memory storage space, and GPU computing power. The iterative training cycles are also not conducive to timely adjustment of the hyperparameters in the network model, and the model training efficiency is significantly reduced. In addition, LR CXR images with bone signals could not directly establish a nonlinear mapping relationship with HR bone-suppressed CXR images. For the bone suppression task, the aim of the network training was to eliminate the bone signal on the original input CXR image data as much as possible and preserve the anatomical structure information of the soft tissue. The LR image first went through an interpolation process for the super-resolution task. The skip connection established the relationship between the interpolated image before inputting the network and the convoluted feature map. Throughout the literature consulted, this is an essential step in super-resolution. Omitting this step makes obtaining HR images difficult. If the super-resolution reconstruction task is performed, the shallow feature information of the image should be preserved as much as possible. In contrast, the bone suppression task must suppress the bone signal component in the shallow feature information of the CXR image. The 2 training tasks are incompatible. In summary, the step-by-step training strategy can adjust the hyperparameters in the network model in time to reduce the computer operating load and avoid conflict between the multitask training models to obtain accurate lung-only bone-suppressed CXR images with HR. In addition, through experiments, we also found that the super-resolution model trained on images with the bone signal can also be applied well to LR bone-suppressed CXR images. Since the network model learns the nonlinear mapping relationship based on the image feature information and whether bones or soft tissues are reflected by pixel information, the strength of different signals in the image has little impact on the reconstruction task.

In the COVID-19 classification task, this study adopted the improved VGG-16 and ResNet-18 models to detect the categories of HR bone-suppressed CXR image data that only contained lung imaging regions. To our knowledge, this study is the first to examine the application of CXR images to a COVID-19 classification task. Compared with the original CXR images without enhancement, the feature information in the image-enhanced CXR data based on deep learning can be more easily learned by the classification model. The predicted results are more precise than the

other enhanced CXR images. Moreover, in this study, the CXR images obtained based on the proposed technique are more suitable than the compared deep learning methods for the task of detecting COVID-19. It can also be seen from the AUC values that, compared with the nonenhanced CXR images for the VGG-16 and ResNet-18 models, the numerical results based on the proposed technique increased by 0.051 and 0.035 in the internal test, and 0.120 and 0.179 in the external test, respectively. This is because the nonenhanced CXR images have more interfering pixel information, which makes the classification model unable to judge the focus regions of COVID-19 accurately. For the overall imaging region, the proportion of the lung generally does not exceed 50%. A large amount of redundant information makes the loss function of the classification model unable to converge well during the training process. In contrast, the enhanced CXR images can provide higher quality and accurate pixel information for the classification model. In addition, compared with single- or multienhanced CXR images, the new technique can obtain stable detection results. The model's accuracy on the internal and external testing data sets was relatively higher, and the model's generalization ability was somewhat better. The synergy of multiple enhancement methods can leverage the advantages of each component. For the enhanced CXR image after lung segmentation, since the COVID-19 virus mainly invades the patient's lung tissue, it has no apparent impact on the rest of the tissues and organs, and the imaging of COVID-19 features on the CXR image is also mainly concentrated in the lungs. The segmentation of the lungs can prevent the classification model from additionally learning the pixel information features of other regions, resulting in a decrease in the accuracy of the classification task. After the segmentation operation, the whole CXR image only contains the lung imaging region, which not only improves the accuracy of subsequent bone suppression and super-resolution tasks but also ensures that the filtered pixel information can make it easier for the classification model to distinguish COVID-19 and normal CXR images in the training process. After bone suppression, the enhanced CXR image can effectively reduce the imaging interference of the bone signals in the lungs to the COVID-19 lesion region. The classification model can focus more accurately on the image features of COVID-19. In addition, compared with the LR CXR image, the enhanced HR image has more pixel information so that the classification network model can fully extract the image feature information in the lung imaging area under

different CXR categories. The robustness and accuracy of the classification model have been improved to a certain extent. This study mainly used the measures of accuracy and F1-score instead of precision and recall. In the classification task, the higher the precision of a certain category is, the lower the misdiagnosis rate of the category, while the higher the recall of a category is, the lower the missed diagnosis rate of the category. Usually, precision and recall contradict each other, leading to a higher value of one measure and a lower value of the other. The clinical reference significance of a separate measure of the two is questionable. The F1-score is the comprehensive quantity of precision and recall. It is a commonly used indicator for evaluating classification models, which can solve the limitations of precision and recall indicators in evaluation.

There are some limitations in comparing the classification accuracy of the recently existing methods with the proposed method. The former is based on the relatively small size of CXR images for classification tasks, while the latter is based on images with a size of 1,024×1,024 (69,70). For comparison, 1 of the following 2 options is required. The first method is to reduce the dimensionality of the 1,024×1,024 image to a suitable size for the input without changing the network structure of the classification model. The second method is to modify the network structure of the classification model to meet the input data size of 1,024×1,024. However, both methods would change the originality of their experiments, and the accuracy of the classification based on the existing methods can be affected when they are modified. Therefore, only the improved VGG-16 and ResNet-18 models were used in this study to verify the accuracy of COVID-19 classification based on the multienhanced CXR images.

In addition, there is no clear rule for improving COVID-19 classification tasks based on single-enhanced or double-enhanced CXR images. It is impossible to determine which of the single or double enhancement method has the most significant impact on classification accuracy. Since this is a preliminary study on applying CXR image under multistrategy enhancement in COVID-19 classification, we will continue to search for the correlation between different enhancement schemes and COVID-19 classification accuracy in subsequent experiments. By building a more accurate image enhancement model, the robustness of the classification model can be further improved, and increasingly accurate classification results can be obtained. In addition, this study was only based on COVID-19 binary classification data for training. We will

continue to add other pneumonia categories in the future to verify the ability of multistrategy enhanced CXR images to detect various categories of pneumonia.

Conclusions

This study developed a novel COVID-19 classification technique based on the multistrategy enhanced step-by-step CXR image acquisition method. The new technique was applied to the COVID-19 classification task. In terms of quantitative verification, the bone-suppressed model with multidilated-rate convolution could better eliminate bone signals in CXR images than the other existing methods. The HR CXR images of bone suppression based on the CNN model were also less different from the ground truth CXR images. In terms of classification accuracy evaluation, the 2 classification models trained based on the multistrategy enhanced CXR images could accurately predict the normal and COVID-19 data categories at the internal and external testing stage. The CXR images synthesized by the new technique based on deep learning have good application prospects in diagnosing COVID-19.

Acknowledgments

Funding: This work was supported by the Health and Medical Research Fund (No. HMRF COVID190211), the Food and Health Bureau, and the Government of the Hong Kong Special Administrative Region.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://qims.amegroups.com/article/view/10.21037/qims-22-610/rc>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://qims.amegroups.com/article/view/10.21037/qims-22-610/coif>). YXJW serves as the editor-in-chief of *Quantitative Imaging in Medicine and Surgery*. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was

conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics committee of the Hong Kong East Cluster Research (No. HKECREC-2020-119), and informed consent was obtained from all patients.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *Int J Antimicrob Agents* 2020;55:105924.
- Chen Y, Klein SL, Garibaldi BT, Li H, Wu C, Osevala NM, Li T, Margolick JB, Pawelec G, Leng SX. Aging in COVID-19: Vulnerability, immunity and intervention. *Ageing Res Rev* 2021;65:101205.
- Shahi TB, Sitaula C, Paudel N. A Hybrid Feature Extraction Method for Nepali COVID-19-Related Tweets Classification. *Comput Intell Neurosci* 2022;2022:5681574.
- Sitaula C, Basnet A, Mainali A, Shahi TB. Deep Learning-Based Methods for Sentiment Analysis on Nepali COVID-19-Related Tweets. *Comput Intell Neurosci* 2021;2021:2158184.
- Sethy PK, Behera SK, Anitha K, Pandey C, Khan MR. Computer aid screening of COVID-19 using X-ray and CT scan images: An inner comparison. *J Xray Sci Technol* 2021;29:197-210.
- Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra Acharya U. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med* 2020;121:103792.
- Nayak SR, Nayak DR, Sinha U, Arora V, Pachori RB. Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: A comprehensive study. *Biomed Signal Process Control* 2021;64:102365.

8. Oh Y, Park S, Ye JC. Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets. *IEEE Trans Med Imaging* 2020;39:2688-700.
9. Zebin T, Rezvy S. COVID-19 detection and disease progression visualization: Deep learning on chest X-rays for classification and coarse localization. *Appl Intell (Dordr)* 2021;51:1010-21.
10. Khan AI, Shah JL, Bhat MM. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. *Comput Methods Programs Biomed* 2020;196:105581.
11. Xing L, Krupinski EA, Cai J. Artificial intelligence will soon change the landscape of medical physics research and practice. *Med Phys* 2018;45:1791-3.
12. Jia X, Ren L, Cai J. Clinical implementation of AI technologies will require interpretable AI models. *Med Phys* 2020;47:1-4.
13. Bae K, Oh DY, Yun ID, Jeon KN. Bone Suppression on Chest Radiographs for Pulmonary Nodule Detection: Comparison between a Generative Adversarial Network and Dual-Energy Subtraction. *Korean J Radiol* 2022;23:139-49.
14. van der Heyden B. The potential application of dual-energy subtraction radiography for COVID-19 pneumonia imaging. *Br J Radiol* 2021;94:20201384.
15. Baltruschat IM, Steinmeister L, Ittrich H, Adam G, Nickisch H, Saalbach A, Berg J, Grass M, Knopp T. When does bone suppression and lung field segmentation improve chest x-ray disease classification? 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) 2019:1362-1366. doi: 10.1109/ISBI.2019.8759510.
16. Erdaw Y, Tachbele E. Machine Learning Model Applied on Chest X-Ray Images Enables Automatic Detection of COVID-19 Cases with High Accuracy. *Int J Gen Med* 2021;14:4923-31.
17. Yousefi B, Kawakita S, Amini A, Akbari H, Advani SM, Akhloufi M, Maldague XPV, Ahadian S. Impartially Validated Multiple Deep-Chain Models to Detect COVID-19 in Chest X-ray Using Latent Space Radiomics. *J Clin Med* 2021;10:3100.
18. Li F, Lu X, Yuan J. MHA-CoroCapsule: Multi-Head Attention Routing-Based Capsule Network for COVID-19 Chest X-Ray Image Classification. *IEEE Trans Med Imaging* 2022;41:1208-18.
19. Han L, Lyu Y, Peng C, Zhou SK. GAN-based disentanglement learning for chest X-ray rib suppression. *Med Image Anal* 2022;77:102369.
20. Lam NFD, Sun H, Song L, Yang D, Zhi S, Ren G, Chou PH, Wan SBN, Wong MFE, Chan KK, Tsang HCH, Kong FS, Wáng YXJ, Qin J, Chan LWC, Ying M, Cai J. Development and validation of bone-suppressed deep learning classification of COVID-19 presentation in chest radiographs. *Quant Imaging Med Surg* 2022;12:3917-31.
21. Sheng Chen, Suzuki K. Separation of bones from chest radiographs by means of anatomically specific multiple massive-training ANNs combined with total variation minimization smoothing. *IEEE Trans Med Imaging* 2014;33:246-57.
22. Chen S, Zhong S, Yao L, Shang Y, Suzuki K. Enhancement of chest radiographs obtained in the intensive care unit through bone suppression and consistent processing. *Phys Med Biol* 2016;61:2283-301.
23. Suzuki K, Abe H, MacMahon H, Doi K. Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN). *IEEE Trans Med Imaging* 2006;25:406-16.
24. Loog M, van Ginneken B, Schilham AM. Filter learning: application to suppression of bony structures from chest radiographs. *Med Image Anal* 2006;10:826-40.
25. Li X, Luo S, Hu Q, Li J, Wang D. Rib suppression in chest radiographs for lung nodule enhancement. 2015 IEEE International Conference on Information and Automation; 08-10 August 2015; Lijiang, China. IEEE, 2015:50-5.
26. Simkó G, Orbán G, Máday P, Horváth G. Elimination of clavicle shadows to help automatic lung nodule detection on chest radiographs. 4th European Conference of the International Federation for Medical and Biological Engineering 2009:488-91. doi:10.1007/978-3-540-89208-3_116.
27. Hogeweg L, Sanchez CI, van Ginneken B. Suppression of translucent elongated structures: applications in chest radiography. *IEEE Trans Med Imaging* 2013;32:2099-113.
28. Mamalakis M, Swift AJ, Vorselaars B, Ray S, Weeks S, Ding W, Clayton RH, Mackenzie LS, Banerjee A. DenResCov-19: A deep transfer learning network for robust automatic classification of COVID-19, pneumonia, and tuberculosis from X-rays. *Comput Med Imaging Graph* 2021;94:102008.
29. Ren G, Xiao H, Lam SK, Yang D, Li T, Teng X, Qin J, Cai J. Deep learning-based bone suppression in chest radiographs using CT-derived features: a feasibility study. *Quant Imaging Med Surg* 2021;11:4807-19.
30. Li H, Han H, Li Z, Wang L, Wu Z, Lu J, Zhou SK. High-Resolution Chest X-Ray Bone Suppression Using Unpaired CT Structural Priors. *IEEE Trans Med Imaging*

- 2020;39:3053-63.
31. Cho K, Seo J, Kyung S, Kim M, Hong GS, Kim N. Bone suppression on pediatric chest radiographs via a deep learning-based cascade model. *Comput Methods Programs Biomed* 2022;215:106627.
 32. Lin Z, He Z, Xie S, Wang X, Tan J, Lu J, Tan B. AANet: Adaptive Attention Network for COVID-19 Detection From Chest X-Ray Images. *IEEE Trans Neural Netw Learn Syst* 2021;32:4781-92.
 33. Loey M, El-Sappagh S, Mirjalili S. Bayesian-based optimized deep learning model to detect COVID-19 patients using chest X-ray image data. *Comput Biol Med* 2022;142:105213.
 34. Gusarev M, Kuleev R, Khan A, Rivera AR, Khattak AM. Deep learning models for bone suppression in chest radiographs. 2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB); 23-25 August 2017; Manchester, UK. IEEE, 2017:1-7.
 35. Oh DY, Yun ID. Learning bone suppression from dual energy chest X-rays using adversarial networks 2018. arXiv preprint arXiv:1811.02628.
 36. Rajaraman S, Zamzmi G, Folio L, Alderson P, Antani S. Chest X-ray Bone Suppression for Improving Classification of Tuberculosis-Consistent Findings. *Diagnostics (Basel)* 2021;11:840.
 37. Rajaraman S, Cohen G, Spear L, Folio L, Antani S. DeBoNet: A deep bone suppression model ensemble to improve disease detection in chest radiographs. *PLoS One* 2022;17:e0265691.
 38. Chetoui M, Akhloufi MA. Explainable Vision Transformers and Radiomics for COVID-19 Detection in Chest X-rays. *J Clin Med* 2022;11:3013.
 39. Bätz M, Eichenseer A, Seiler J, Jonscher M, Kaup A. Hybrid super-resolution combining example-based single-image and interpolation-based multi-image reconstruction approaches. 2015 IEEE international conference on image processing (ICIP); 27-30 September; Quebec City, QC, Canada. IEEE, 2015:58-62.
 40. Zhou J, Zhou C, Zhu J, Fan D. A method of super-resolution reconstruction for remote sensing image based on non-subsampled contourlet transform. *Acta Optica Sinica* 2015;35:0110001.
 41. Xiao J, Liu E, Zhu L, Lei J. Improved image super-resolution algorithm based on convolutional neural network. *Acta Optica Sinica* 2017;37:0318011.
 42. Shen W, Fang L, Chen X, Xu H. Projection onto Convex Sets Method in Space-frequency Domain for Super Resolution. *J Comput* 2014;9:1959-66.
 43. Bareja MN, Modi CK. An effective iterative back projection based single image super resolution approach. 2012 International Conference on Communication Systems and Network Technologies; 11-13 May 2012; Rajkot, Gujarat, India. IEEE, 2012:95-99.
 44. Du YB, Jia RS, Cui Z, Yu JT, Sun HM, Zheng YG. X-ray image super-resolution reconstruction based on a multiple distillation feedback network. *Applied Intelligence* 2021;51:5081-94.
 45. Monday HN, Li J, Nneji GU, Nahar S, Hossin MA, Jackson J, Ejayi CJ. COVID-19 Diagnosis from Chest X-ray Images Using a Robust Multi-Resolution Analysis Siamese Neural Network with Super-Resolution Convolutional Neural Network. *Diagnostics (Basel)* 2022;12:741.
 46. Yu Y, She K, Liu J. Wavelet Frequency Separation Attention Network for Chest X-ray Image Super-Resolution. *Micromachines (Basel)* 2021;12:1418.
 47. Zhao CY, Jia RS, Liu QM, Liu XY, Sun HM, Zhang XL. Chest X-ray images super-resolution reconstruction via recursive neural network. *Multimedia Tools and Applications* 2021;80:263-77.
 48. Ahmed S, Hossain T, Hoque OB, Sarker S, Rahman S, Shah FM. Automated COVID-19 Detection from Chest X-Ray Images: A High-Resolution Network (HRNet) Approach. *SN Comput Sci* 2021;2:294.
 49. Moran MBH, Faria MDB, Giraldo GA, Bastos LF, Conci A. Using super-resolution generative adversarial network models and transfer learning to obtain high resolution digital periapical radiographs. *Comput Biol Med* 2021;129:104139.
 50. Mahmud T, Rahman MA, Fattah SA. CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. *Comput Biol Med* 2020;122:103869.
 51. Umair M, Khan MS, Ahmed F, Baothman F, Alqahtani F, Alian M, Ahmad J. Detection of COVID-19 Using Transfer Learning and Grad-CAM Visualization on Indigenously Collected X-ray Dataset. *Sensors (Basel)* 2021;21:5813.
 52. Hussain E, Hasan M, Rahman MA, Lee I, Tamanna T, Parvez MZ. CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images. *Chaos Solitons Fractals* 2021;142:110495.
 53. Singh D, Kumar V, Yadav V, Kaur M. Deep neural network-based screening model for COVID-19-infected

- patients using chest X-ray images. *International Journal of Pattern Recognition and Artificial Intelligence* 2021;35:2151004.
54. Jia G, Lam HK, Xu Y. Classification of COVID-19 chest X-Ray and CT images using a type of dynamic CNN modification method. *Comput Biol Med* 2021;134:104425.
 55. Zamzmi G, Rajaraman S, Antani SK. Accelerating Super-Resolution and Visual Task Analysis in Medical Images. *Appl Sci* 2020;10:4282.
 56. Zhang H, Xu T, Li H, Zhang S, Wang X, Huang X, Metaxas DN. StackGAN++: Realistic Image Synthesis with Stacked Generative Adversarial Networks. *IEEE Trans Pattern Anal Mach Intell* 2019;41:1947-62.
 57. Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu K, Matsui M, Fujita H, Kodera Y, Doi K. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. *AJR Am J Roentgenol* 2000;174:71-4.
 58. Juhász S, Horváth Á, Niházy L, Horváth G. Segmentation of anatomical structures on chest radiographs. *XII Mediterranean Conference on Medical and Biological Engineering and Computing* 2010:359-62. doi: 10.1007/978-3-642-13039-7_90.
 59. Radiological Society of North America. RSNA Pneumonia Detection Challenge, 2019. Available online: <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge>
 60. Tahir AM, Chowdhury MEH, Khandakar A, Rahman T, Qiblawey Y, Khurshid U, Kiranyaz S, Ibtehad N, Rahman MS, Al-Maadeed S, Mahmud S, Ezeddin M, Hameed K, Hamid T. COVID-19 infection localization and severity grading from chest X-ray images. *Comput Biol Med* 2021;139:105002.
 61. Tsai EB, Simpson S, Lungren MP, Hershman M, Roshkovan L, Colak E, et al. The RSNA International COVID-19 Open Radiology Database (RICORD). *Radiology* 2021;299:E204-13.
 62. Selvan R, Dam EB, Detlefsen NS, Rischel S, Sheng K, Nielsen M, Pai A. Lung segmentation from chest X-rays using variational data imputation. 2020;arXiv preprint arXiv:2005.10052.
 63. Nneji GU, Cai J, Monday HN, Hossin MA, Nahar S, Mgbejime GT, Deng J. Fine-Tuned Siamese Network with Modified Enhanced Super-Resolution GAN Plus Based on Low-Quality Chest X-ray Images for COVID-19 Identification. *Diagnostics (Basel)* 2022;12:717.
 64. Almezghwi K, Serte S, Al-Turjman F. Convolutional neural networks for the classification of chest X-rays in the IoT era. *Multimed Tools Appl* 2021;80:29051-65.
 65. Shelke A, Inamdar M, Shah V, Tiwari A, Hussain A, Chafekar T, Mehendale N. Chest X-ray Classification Using Deep Learning for Automated COVID-19 Screening. *SN Comput Sci* 2021;2:300.
 66. Zarshenas A, Liu J, Forti P, Suzuki K. Separation of bones from soft tissue in chest radiographs: Anatomy-specific orientation-frequency-specific deep neural network convolution. *Med Phys* 2019;46:2232-42.
 67. Sundhari RP. Enhanced histogram equalization based nodule enhancement and neural network based detection for chest x-ray radiographs. *Journal of Ambient Intelligence and Humanized Computing* 2021;12:3831-9.
 68. Chattopadhyay A, Sarkar A, Howlader P, Balasubramanian VN. Grad-CAM++: Improved Visual Explanations for Deep Convolutional Networks. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV); 12-15 March 2018; Lake Tahoe, NV, USA. IEEE, 2018:839-47.
 69. Sitaula C, Hossain MB. Attention-based VGG-16 model for COVID-19 chest X-ray image classification. *Appl Intell (Dordr)* 2021;51:2850-63.
 70. Sitaula C, Shahi TB, Aryal S, Marzbanrad F. Fusion of multi-scale bag of deep visual words features of chest X-ray images to detect COVID-19 infection. *Sci Rep* 2021;11:23914.

Cite this article as: Sun H, Ren G, Teng X, Song L, Li K, Yang J, Hu X, Zhan Y, Wan SBN, Wong MFE, Chan KK, Tsang HCH, Xu L, Wu TC, Kong FMS, Wang YXJ, Qin J, Chan WCL, Ying M, Cai J. Artificial intelligence-assisted multistrategy image enhancement of chest X-rays for COVID-19 classification. *Quant Imaging Med Surg* 2023;13(1):394-416. doi: 10.21037/qims-22-610