

Bone Suppression of Chest Radiographs With Cascaded Convolutional Networks in Wavelet Domain

YINGYIN CHEN¹, XIAOFANG GOU¹, XIUXIA FENG¹, YUNBI LIU¹, GENGGENG QIN²,
QIANJIN FENG¹, WEI YANG¹, AND WUFAN CHEN¹

¹Guangdong Provincial Key Laboratory of Medical Image Processing, School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China

²Radiology Department, Nanfang Hospital, Southern Medical University, Guangzhou 510515, China

Corresponding authors: Wei Yang (weiyanggm@gmail.com) and Wufan Chen (chenwf@fimmu.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 81771916 and Grant 61471187, in part by the Guangdong Provincial Key Laboratory of Medical Image Processing under Grant 2014B030301042, and in part by the State's Key Project of Research and Development Plan under Grant 2018YFC2001203.

ABSTRACT Bone suppression of chest radiographs (CXRs) is potentially useful for diagnosing lung diseases for radiologists and computer-aided diagnosis. This paper presents a cascaded convolutional network model in wavelet domain (Wavelet-CCN) for bone suppression in single conventional CXR. Wavelet coefficients are sparse and suitable as the output of convolutional network. The convolutional networks are trained to predict the wavelet coefficients of bone images from the wavelet coefficients of CXRs, using real two-exposure dual energy subtraction (DES) CXRs as training data. By combining the multi-level wavelet decomposition and a cascaded refinement framework, the Wavelet-CCN model can work automatically with a multi-scale approach and progressively refine the prediction in terms of accuracy and spatial resolution. Compared with previous work of CamsNet model which preforms bone prediction in gradient domain, the Wavelet-CCN model predicts the wavelet coefficients to reconstruct bone images and can avoid the inconsistent background intensity caused by 2D integration of gradients. The predicted bone image is subtracted from the original CXR to produce a soft-tissue image. The Wavelet-CCN model and its variants with different wavelet basis are evaluated on a dataset that consists of 504 cases of real two-exposure DES CXRs (404 cases for training and 100 cases for test). Experimental results show that among all the variants and different wavelet bases, the Wavelet-CCN model with Haar wavelet performs best. The average peak signal-to-noise ratio and structural similarity index of the soft-tissue images produced by the proposed Wavelet-CCN model are both higher than those of the previous CamsNet model in gradient domain, reaching values of 39.4 (± 0.94) dB and 0.977 (± 0.004), respectively. The results also demonstrate that the Wavelet-CCN model can process the CXRs acquired by four types of X-ray machines.

INDEX TERMS Bone suppression, convolutional networks, chest radiographs, wavelet transform.

I. INTRODUCTION

Chest radiography (chest X-ray or CXR) is one of the most important tools for lung disease diagnosis. However, lesion areas (such as lung nodules), especially those located in the occlusion area of ribs and clavicle are difficult to detect by both radiologists and computers owing to the anatomical structure overlap in the CXR [1]. Therefore, identifying and removing the effect of superimposed structures in CXRs are necessary to reduce manual and computer-based interpretation errors [2]. The soft-tissue image can be obtained by bone suppression, which subtracts the bone image from

the original CXR. Bone suppression of CXRs is potentially useful for radiologists and computer-aided diagnosis [3], [4].

The bone suppression methods of CXRs can be divided into two classes: hardware implementation and image processing method. Hardware implementation can be implemented by two technical solutions of dual energy subtraction (DES) [5]. One is single-shot technique using a single exposure by a thin copper filter between two storage phosphor plates to produce low- and high-energy images; the other is double-shot technique, which produces the low- and high-energy radiographs by two exposures consecutively.

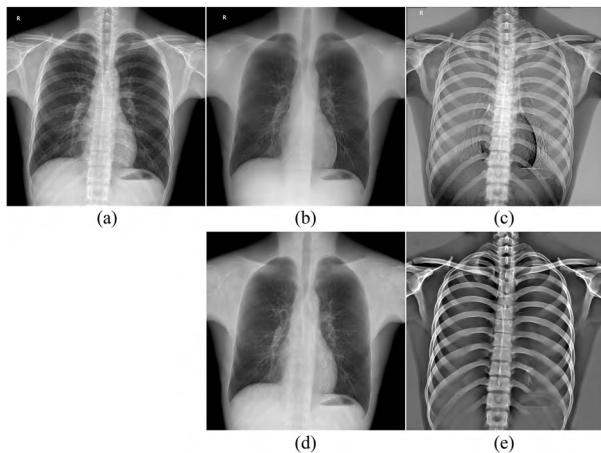


FIGURE 1. Examples of bone images and soft-tissue images produced by a real DES unit and our Wavelet-CCN model. (a) is a standard CXR; (b) and (c) are the corresponding soft-tissue image and bone image produced by a real DES unit, respectively; (d) and (e) are soft-tissue image and bone images produced by our Wavelet-CCN model, respectively.

However, DES has some disadvantages. Single-shot DES produces noisy images with limited energy separation. The bone and soft-tissue images produced by double-shot DES may contain the motion artifacts caused by patients' breath and motion, as shown in Fig. 1. In addition, both DES techniques require special equipment and long operation time. The quality of images derived from DES can be unstable and imaging procession leads to extra radiation.

Image processing methods for bone suppression can be divided into two categories: unsupervised and supervised methods. The unsupervised method finds bone targets and then removes them from CXR [6], [7]. Clavicle and ribs can be modeled as the elongated structures and high gradient edges [2], [12]. Blind source separation techniques or gradient modification are used to remove the bone structures from CXRs [2], [6], [12]. The detected bone boundaries or segmented bone regions are used as the intermediate results to remove bone structures [8]–[11]. However, visible bone shadow remains after these tedious unsupervised processes. For supervised bone suppression methods, DES image has become an appealing research object for supervised bone suppression methods. Massive-training artificial neural networks (MTANN) was proposed by Suzuki *et al.* [1], [13] is one of the earliest works which used the DES images as training data to address bone suppression problem. In these works, MTANNs were trained under a multi-resolution framework, in which each resolution of bone prediction model was trained independently. Prediction results from the multi-resolution models of MTANN were combined to provide a complete high-resolution bone image. In Chen and Suzuki's [15] work, MTANN was further improved by training in different anatomic segments and combining total variation minimization. The k-nearest neighbor regression method with optimized local features has been also proposed to perform the separation of bone and tissue components in

chest radiographs [14]. Deep learning models transform standard CXR images into soft-tissue images by treating bones as noise through autoencoder-like model and multi-layer neural model for bone suppression [16]. An improved performance of bone suppression method was obtained by learning the mapping between the gradients of the CXRs and the corresponding bone images [4].

Our previous work proposed a cascade of multi-scale convolutional network (CamsNet) to progressively refine the predicted gradients of bone images by increasing resolutions successively in each convolutional network (ConvNet) [4]. The sparsity of gradient domain makes the bone and soft-tissue much easier to separate [17]. The CamsNet model predicts bone gradient images and the bone image is reconstructed through 2D integration. However, 2D integration of the predicted bone gradients may lead to inconsistent background intensity. The inconsistent background intensity can be corrected through post-processing by guide filtering [18]. However, the correction of guided filtering for CamsNet model in gradient domain is only an ad hoc solution, and the inconsistent background intensity caused by the non-integrable gradient field cannot be addressed.

Wavelet decomposition is commonly used for the multi-resolution image analysis. Knapp *et al.*'s patent also mentioned the wavelet transform for bone suppression using a multi-layer perceptron (MLP) neural network architecture, the inputs of which were well-designed image features such as multi-scale harmonics derivatives and shape indices [19]. Chen and Suzuki's [15] work decomposed a CXR into seven sub-band images by wavelet decomposition to train MTANNs. The MLP neural network used to construct MTANN is a traditional neural network model, which has relatively small receptive field and cannot extract rich information. Many approaches for low-dose X-ray CT reconstruction [20], [21], super-resolution [22], [23], and classification work [24] are proposed through fusing wavelet transform and ConvNet.

In this work, a model is developed for bone suppression, which combines ConvNets and wavelet decomposition in a cascaded refinement framework. The proposed model is called Wavelet-CCN (cascaded convolutional networks in wavelet domain). The Wavelet-CCN model makes use of sparsity and multi-resolution analysis of wavelet transform with ConvNet. The Wavelet-CCN model offers three advantages:

- 1) The combination of all wavelet sub-bands of the input CXR can help improve the predicted wavelet sub-bands of bone images.
- 2) The predicted approximation sub-band of bone images combined with high-frequency sub-band of the input CXR can help improve the prediction of finer structures.
- 3) Reconstruction of bone images from wavelet coefficients can avoid the inconsistent background intensity in the CamsNet models of gradient domain.

The remainder of this paper is organized as follows. Section II describes the framework and details of our method.

The experimental results are provided in Section III. The results and future work are discussed in Section IV. Finally, a summary of the proposed bone suppression method is presented in Section V.

II. METHOD

In this section, the Wavelet-CNN model for bone suppression is introduced and used to predict a series of wavelet sub-band images of bone image.

A. WAVELET DECOMPOSITION OF IMAGE

The 2D discrete wavelet transform (*dwt*) is one of the basic operations for multi-resolution image analysis. An image I (with a size of $M \times N$) can be decomposed into L ($0 < L < \min(\log_2 M, \log_2 N)$) levels by *dwt*. The procedure of multi-level wavelet decomposition of sub-band image A using in this manuscript is denoted by:

$$[A^l, H^l, V^l, D^l] = dwt(A^{l-1}) \quad (1)$$

The set of these decomposed sub-band images is called wavelet representations in 2D [25]. The sub-band images are reconstructed by inverse discrete wavelet transform (*idwt*). This procedure is denoted by:

$$A^{l-1} = idwt(A^l, H^l, V^l, D^l) \quad (2)$$

where $l \in \{1, 2, \dots, L\}$, when $l = 1$, $A_0 = I$. *dwt* of an image produces a non-redundant image representation. The results of wavelet decomposition can be interpreted as a set of independent, spatially oriented frequency channels [26]. Compared with other multi-scale representations such as Gaussian and Laplacian pyramid decomposition, *dwt* of image provides more powerful insight into spatial and spectral localization of image formation.

The wavelet coefficients are sparse and follow the heavy-tailed distribution. An example of wavelet decomposition results of a CXR, the corresponding DES bone image and soft-tissue image are shown in Fig. 2. The wavelet coefficients are distributed near zero, and large absolute value of wavelet coefficients indicate the position of edges shown in Fig. 2.

B. ARCHITECTURE OF WAVELET-CCN

In this section, the architecture of Wavelet-CCN model is introduced, which predicts bone image from a single CXR. The bone image predicted by a Wavelet-CCN model is subtracted from the input CXR to obtain a bone-suppressed image (soft-tissue image). Bone image can be refined progressively by several consecutive ConvNets in a Wavelet-CCN model.

Fig. 3 shows an example of two-level Wavelet-CCN model which starts training from the coarsest level L where $L = 2$. $\mathbf{T} = \{(I_n; B_n); n = 1, \dots, N\}$ denotes the training data set where I_n is an input CXR, and B_n denotes the corresponding ground truth of bone image. I_n and B_n are all decomposed

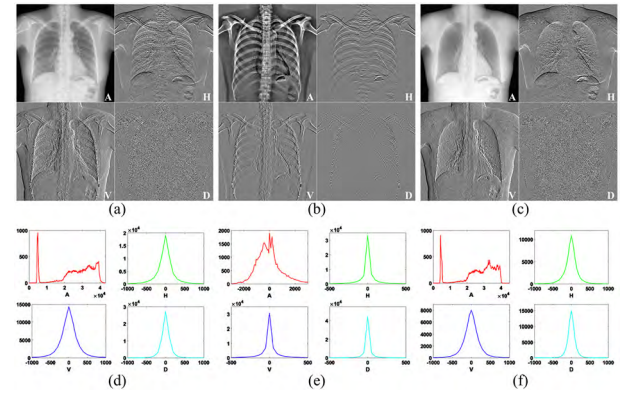


FIGURE 2. Wavelet decomposition results of (a) CXR image, (b) DES bone image, and (c) DES soft-tissue image. (d), (e), and (f) are the corresponding histograms of each sub-band images in (a), (b), and (c), respectively.

by L -level *dwt* for training of the models. For each level of Wavelet decomposition, a level-specific ConvNet is trained for predicting the Wavelet sub-bands of bone images. In the sequel, the subscript n for notational convenience is omitted.

1) BASIC UNIT OF WAVELET-CCN

The basic unit of Wavelet-CCN is used to learn the mapping from the four sub-band images $\hat{A}_B^l, \hat{H}_B^l, \hat{V}_B^l, \hat{D}_B^l$ to the four sub-band images $A_B^l, H_B^l, V_B^l, D_B^l$ of the corresponding bone image B at level l , and \hat{A}_B^{l-1} is obtained through *idwt*. *idwt* is the reverse operation of wavelet transform. As shown in Fig. 3, $\hat{A}_B^l, \hat{H}_B^l, \hat{V}_B^l, \hat{D}_B^l$ are the predicted results from $\hat{A}_B^l, \hat{H}_B^l, \hat{V}_B^l, \hat{D}_B^l$ with level l prediction unit of Wavelet-CCN. \hat{A}_B^{l-1} is the intermediate result from the *idwt* of level l , and $\hat{A}_B^{l-1} = idwt(\hat{A}_B^l, \hat{H}_B^l, \hat{V}_B^l, \hat{D}_B^l)$. The four sub-band images $A_B^l, H_B^l, V_B^l, D_B^l$ which only appear in the lowest level L are initial inputs, and $\hat{A}_B^l, \hat{H}_B^l, \hat{V}_B^l, \hat{D}_B^l$ are the input sub-band images for other level l . The corresponding prediction procedures of sub-band bone images can be formulated into:

$$\begin{cases} \hat{A}_B^l, \hat{H}_B^l, \hat{V}_B^l, \hat{D}_B^l = \mathcal{F}_l(\hat{A}_B^{l-1}, H_B^{l-1}, V_B^{l-1}, D_B^{l-1}), & l < L \\ \hat{A}_B^l, \hat{H}_B^l, \hat{V}_B^l, \hat{D}_B^l = \mathcal{F}_L(A_B^l, H_B^l, V_B^l, D_B^l), & l = L \end{cases} \quad (3)$$

where $\hat{A}_B^l, \hat{H}_B^l, \hat{V}_B^l, \hat{D}_B^l$ are the predicted results and $\mathcal{F}(\cdot)$ is the network mapping between the inputs and outputs.

2) WAVELET-CCN CASCADE

The features of image at a coarse resolution provide the “context” (structural information) of the image, whereas the finer details correspond to the particular local information [25]. Wavelet-CCN continues with the coarse-to-fine processing strategy as the CamsNet model [4]. The predicted results of each level are preserved and passed onto the next level. *idwt* is a key operation, which connects the basic prediction units of level l and level $l-1$ which means it is used to link (current resolution) to the following resolution and make full use of the predicted bone image information.

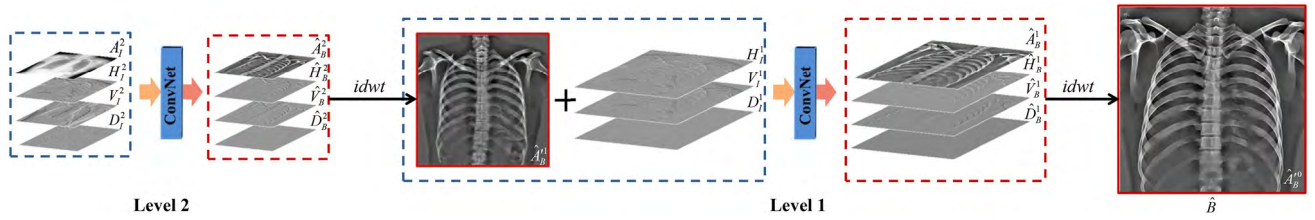


FIGURE 3. Flowchart of bone image prediction with a two-level Wavelet-CCN model. The inputs of each level are with dashed gray-blue rectangle; the outputs of each level are with dashed red rectangle; the prediction result of bone image for each level is with red rectangle, and the right is the predicted bone image.

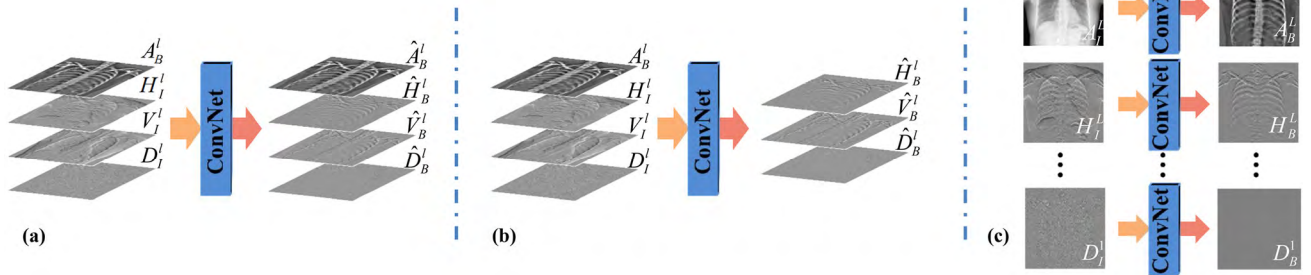


FIGURE 4. Different architectures in wavelet domain: (a) Wavelet-CCN, (b) Pyramidal architecture, and (c) Channel-wise architecture.

The *idwt* operation plays a connection role between ConvNets of adjacent scales. The outputs of each ConvNet are upsampled by an upscaling factor 2 via *idwt* operation and then become the input of the next scale ConvNet. The intermediate predicted sub-band bone image \hat{A}_B^{l-1} is passed to the following level. For the norm preservation properties, the real approximation sub-band image A_B^{l-1} is replaced with the intermediate predicted sub-band image \hat{A}_B^{l-1} , which is close to the real sub-band image of bone A_B^{l-1} ; the formula can be expressed as: $A_B^{l-1} = \hat{A}_B^{l-1}$. Three high-frequency sub-band images H_B^{l-1} , V_B^{l-1} and D_B^{l-1} with more detailed information are combined with \hat{A}_B^{l-1} for following prediction in level $l-1$. The final prediction result of bone image \hat{B} can be obtained through the outputs of last level with $\hat{B} = \hat{A}_B^0 = idwt(\hat{A}_B^1, \hat{H}_B^1, \hat{V}_B^1, \hat{D}_B^1)$. The whole model is trained from the coarsest scale step by step, where *idwt* is used as the connection of cascaded ConvNets.

C. VARIANTS OF CONVNET ARCHITECTURE IN WAVELET DOMAIN

To investigate the performance of different architectures, two different non-cascade architectures are designed in wavelet domain for bone suppression, as shown in Fig. 4.

1) PYRAMIDAL ARCHITECTURE

The prediction unit in pyramidal architecture predicts three high-frequency sub-bands of bone images using three high-frequency sub-bands of CXRs and the predicted bone images of previous level as the input for each level of wavelet decomposition, as shown in Fig. 4(b). The predicted

approximation sub-band of bone image for each level is obtained by $\hat{A}_B^l = idwt(\hat{A}_B^{l+1}, \hat{H}_B^{l+1}, \hat{V}_B^{l+1}, \hat{D}_B^{l+1})$. The predicted bone image is directly obtained through the *idwt* from the outputs of each level. The difference between the pyramidal architecture and the cascade architecture of Wavelet-CCN is that the outputs of ConvNet for each level do not include the approximation sub-band of bone images, and the approximation sub-bands of bone images are not refined by the ConvNets.

2) CHANNEL-WISE ARCHITECTURE

For each sub-band in each level of wavelet decomposition, one ConvNet is trained to predict the sub-band of bone image using the corresponding sub-band of CXR as the input in the channel-wise architecture, as shown in Fig. 4(c). This architecture was also studied by Chen and Suzuki [15] with the traditional ANN models.

D. IMPLEMENTATION DETAILS

1) CONVNETS IN WAVELET-CCN

For each level of Wavelet-CCN, a simple ConvNet with four convolutional layers (without padding) is trained to predict the wavelet coefficients of bone images, as shown in Fig. 5. Each convolutional layer, except the last one, is followed by a batch normalization (BN) layer and a ReLU non-linearity activation layer. The filter sizes of convolutional layers are set as $f_1 \times f_1$, $f_2 \times f_2$, $f_3 \times f_3$, and $f_4 \times f_4$, respectively, and the network architecture is denoted as f_1 - f_2 - f_3 - f_4 . The filter numbers of all convolutional layers are set to be the same and vary from 64 to 256. The filter weights of ConvNets are

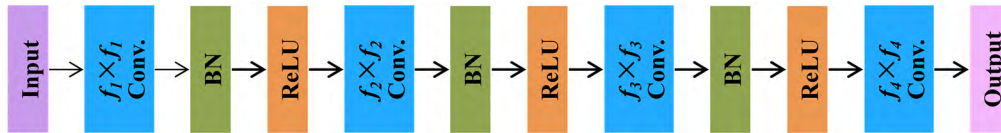


FIGURE 5. Architecture of a ConvNet with convolution layers (short as Conv.).

initialized by the improved “Xavier” method proposed by He *et al.* [27].

2) TRAINING

Driven by Zhao *et al.*’s [28] work, L_1 loss function is employed for training the ConvNets. The parameters Θ of ConvNets can be optimized by minimizing the loss between the predicted sub-bands and the corresponding ground truth sub-bands of bone images. Given a set of training pairs $\{(A_{B,i}, \hat{A}_{B,i}), (H_{B,i}, \hat{H}_{B,i}), (V_{B,i}, \hat{V}_{B,i}), (D_{B,i}, \hat{D}_{B,i})\}$; ($i = 1, 2, \dots, n$), the L_1 loss function is defined as follows:

$$L_1(\Theta) = \frac{1}{n} \sum_{i=1}^n |A_{B,i} - \hat{A}_{B,i}| + |H_{B,i} - \hat{H}_{B,i}| + |V_{B,i} - \hat{V}_{B,i}| + |D_{B,i} - \hat{D}_{B,i}| \quad (4)$$

where n is the number of training samples. L_1 loss function is good at handling noise, and may be smoothing image much better than L_2 does. Thus, L_1 obtaining a good local minimum is more likely than using L_2 loss function, which treats noise independent of the local characteristics.

The patch pairs of wavelet sub-bands of CXRs and the bone images are randomly sampled from the augmented training image pairs as the training samples for ConvNets. The augmentation transformations on training image pairs include translations, horizontal reflections, and scale. Before patch sampling, each wavelet sub-band of CXR is normalized by zero mean unit variance normalization, and the normalization parameters are used to normalize the corresponding wavelet sub-band of bone image. The input and output patch sizes are set to 60×60 and 40×40 , respectively. For each ConvNet in Wavelet-CNN, 0.3 million pairs of patches of wavelet sub-bands are collected. ConvNets are then optimized by stochastic gradient descent (SGD) algorithm with batch updating and momentum from scratch. The batch size is set to 256. The momentum for SGD is set to 0.8. The learning rate starts from 10^{-3} and lower 10 times after 10 epochs. All models are trained and tested with MatConvNet toolbox [29] in Matlab 2016a using a NVIDIA Titan X graphics processing unit (GPU) with 12-GB memory. The GPU computing platform contains 2 3.00GHz Intel(R) Xeon(R) E5-2623 CPUs and 128 GB RAM. Training a single ConvNet took approximately 12 hours within 40 epochs.

The whole wavelet sub-bands of a CXR cannot be fed directly to the ConvNets to obtain the predicted outputs due to the limited memory of GPU and the large size of CXR. For testing, the wavelet sub-bands of a CXR are first padded

by 20 pixels and divided into overlapped 60×60 patches with step size of 40 pixels. Then, the patches are fed to the ConvNets, and the predicted 40×40 patches are rearranged as the wavelet sub-bands of bone image.

3) PREPROCESSING OF CXR

Our dataset is acquired with a two-exposure DES DR (digital radiography) machine. Each case in our collected dataset includes a DES soft-tissue image I^S , a DES bone image I^B , and a standard CXR I^0 . The soft-tissue image I^S , bone image I^B , and standard CXR I^0 are all referred to as the DES CXR. Due to the sophisticated nonlinear post-processing of the raw image data, the relationship $I^0 = I^S + I^B$ is not eventually satisfied, and I^S/I^B is not exactly the soft-tissue/bone component in the CXR I^0 .

As in our previous work [4], the standard CXR I^0 , the original soft-tissue image I^S , and the original bone image I^B are used to generate the bone component B and the soft-tissue component S in I^0 as ground truth for training and evaluation. First, the noise in the three images was reduced by BM3D [30]. Then, the intensity ranges of I^S were calibrated by multiplying a coefficient α such that the correlation between $I - \alpha I^S$ and I^S was minimized. Secondly, an intermediate result is got through $I_1 = I - \alpha I^S$, and only the details of the soft-tissue components and the bone components are kept in I_1 . Third, the gradients \mathbf{G} of the bone components in I were obtained as the transformed gradient field of I_1 using cross projection tensors [31] obtained from I^B that contains the structure information of the bone component. Finally, the bone component B in I was ultimately reconstructed from \mathbf{G} through 2D integration. The corresponding soft-tissue image S is then obtained as $I - B$.

III. EXPERIMENTAL RESULTS

A. EXPERIMENTAL DATA

A total of 646 posterior–anterior DES CXRs are acquired with GE machine of Discovery XR656 at Nanfang Hospital, Guangzhou, China. The X-ray tube voltages of DES devices use the typical value of 120 and 60 kV for the consecutively exposed twice at two different energy levels. There are 142 DES radiographs which contain serious motion artifacts in the soft-tissue and bone images. As in [4], these 142 radiographs are excluded from the dataset. A total of 404 cases are used as training set, and the remaining 100 cases are used as test set. The sizes of DES radiographs ranged from 1800×1800 to 2022×2022 pixels. The pixel sizes range from

0.1931 to 0.1943 mm and are restored in DICOM format with a 14-bit depth.

B. EVALUATION METRICS

The performance of bone suppression models is evaluated in terms of four metrics including relative mean absolute error (RMAE), peak signal-to-noise ratio (PSNR), structural similarity (SSIM) index, and bone suppression ratio (BSR).

These metrics are defined and computed as follows. Denote \mathbf{Z} as a ground truth image and $\hat{\mathbf{Z}}$ as a prediction of \mathbf{Z} . Ω denotes the valid regions in the image \mathbf{Z} , N is the number of pixels in Ω , (x, y) denotes the pixel locations in \mathbf{Z} , and Z_{\max} and Z_{\min} are the maximum value and the minimum value of pixels in Ω , respectively. 0.1 and 99.9 percentiles in the ground truth image \mathbf{Z} are set as the values of Z_{\max} and Z_{\min} , respectively, to reduce the effect of outliers on the quantitative metrics.

RMAE is defined as:

$$RMAE = \sqrt{\frac{1}{N} \sum_{(x,y) \in \Omega} |\hat{\mathbf{Z}}(x, y) - \mathbf{Z}(x, y)| / (Z_{\max} - Z_{\min})} \quad (5)$$

PSNR is defined as:

$$PSNR = 20 \log_{10} \left(\frac{(Z_{\max} - Z_{\min})}{\sqrt{\sum_{(x,y) \in \Omega} (\hat{\mathbf{Z}}(x, y) - \mathbf{Z}(x, y))^2 / N}} \right) \quad (6)$$

SSIM, is a well-known denoising performance metric, can be used to evaluate the quality of the predicted soft-tissue images. The intensity ranges of bone or soft-tissue images are rescaled into the range of [0, 255], and the default setting parameters in the implementation of SSIM are used to compute the values of SSIM indices.

BSR is defined as:

$$BSR = 1 - \sum_{(x,y) \in \Omega} (\hat{\mathbf{S}}(x, y) - \mathbf{S}(x, y))^2 / \sum_{(x,y) \in \Omega} \mathbf{B}(x, y)^2 \quad (7)$$

where $\hat{\mathbf{S}}$ is an estimation of a ground truth soft-tissue image \mathbf{S} , and \mathbf{B} is the corresponding ground truth bone image. $BSR = 1$ indicates perfect performance.

In addition, to evaluate prediction performance in each scale, RMAE and relative root mean square error (RRMSE) are calculated between the original sub-band bone image and the predicted sub-band bone image in each scale.

C. EFFECTS OF PARAMETER SETTINGS

The RMAE of the predicted bone images (denoted as RMAE-B), the PSNR and SSIM of the produced soft-tissue images (denoted as PSNR-S and SSIM-S, respectively), and the BSR are determined to evaluate the performances of different models.

1) EFFECT OF CONVNET ARCHITECTURES

The ConvNet architecture adopted in our work is largely inspired by super-resolution convolutional neural network for image super-resolution in [31]. In general, the performance can be improved if the filter sizes and the depth of network are increased moderately. To be consistent with the basic Wavelet-CCN model, the filter number and the wavelet decomposition levels are fixed to 128 and 3, respectively. The results suggest that a reasonably larger filter size and deeper structures could grasp richer structural information, which in turn lead to better results (see Table 1).

TABLE 1. Performance of Wavelet-CCN models with different ConvNet architectures.

ConvNet architecture	RMAE-B (%)	PSNR-S (dB)	SSIM-S	BSR (%)
9-1-7	4.25±0.58	37.54±0.97	0.969±0.005	77.97±14.02
5-3-1-3	4.14±0.49	37.89±0.75	0.970±0.004	80.36±6.31
9-5-1-5	4.03±0.69	38.13±1.09	0.973±0.004	80.47±14.45
9-7-1-7	3.44±0.51	39.40±0.94	0.977±0.004	85.85±6.57

The large values of standard deviation of BSR for some Wavelet-CCN models are caused by a case with serious motion artifacts in test set. The running time of the full bone suppression procedure with a Wavelet-CCN model on GPU is approximately 6-8 s for an input CXR of 2000×2000 pixels.

2) EFFECT OF WAVELET DECOMPOSITION LEVELS AND FILTER NUMBERS

Table 2 shows that the more levels number and filters number of Wavelet-CCN model indicates better quantitative indices. Compared with three-level Wavelet-CCN, four-level Wavelet-CCN model can achieve better prediction performance. But, the prediction performance of Wavelet-CNN models was only marginally improved when the wavelet decomposition levels and the convolutional filter numbers were increased.

TABLE 2. Performance of Wavelet-CCN models with different wavelet decomposition levels and filter numbers.

L	Filter Number	RMAE-B (%)	PSNR-S (dB)	SSIM-S	BSR (%)
4	64	3.40±0.63	39.43±1.36	0.978±0.004	84.7±15.6
	128	3.48±0.47	39.33±0.90	0.978±0.004	85.7±5.80
	256	3.38±0.51	39.53±1.00	0.979±0.004	86.1±7.50
3	64	3.57±0.60	39.07±1.03	0.976±0.004	84.5±8.80
	128	3.44±0.51	39.40±0.94	0.977±0.004	85.85±6.57
	256	3.54±0.60	39.23±1.01	0.976±0.005	85.10±7.76
2	64	4.09±0.54	37.99±0.83	0.970±0.004	80.61±7.84
	128	4.05±0.44	38.12±0.74	0.971±0.004	81.48±5.28
	256	4.07±0.44	38.08±0.73	0.971±0.004	81.31±5.16

The sizes of CXRs are about 2000×2000 and the number of training images is 404. When a CXR is decomposed into the fourth level by *dwt*, the size of the sub-band images of coarsest level (level L) is about only 125×125 . When the decomposition level is more than 4, very limited number

of image patches can be collected for training the ConvNets. The experiment results also show that the performance of 4-level Wavelet-CNN model is not much better than the 3-level Wavelet-CNN model. In the following experiments, the wavelet decomposition levels and the convolutional filter numbers are set to 3 and 128, respectively, with the trade-off between prediction accuracy and computation time. The more ConvNets in the model and the more convolution filters in each layer of ConvNet, the higher prediction accuracy of bone image can be achieved and the more time will be spent for processing. In order to evaluate the performance of other designs of Wavelet-CCN model, the number of wavelet decomposition levels and convolutional filters are set to 3 and 128 in the following experiments, respectively.

3) EFFECT OF LOSS FUNCTIONS

The experiments on Wavelet-CCN models with different loss functions are conducted. The results confirm L_1 loss function is better than L_2 loss function in terms of four evaluation metrics, as listed in Table 3. The conclusion is remarkably consistent with Zhao *et al.* [28] who reported that when the L_2 or PSNR metrics were used to evaluate image quality, the network trained with the L_1 loss function outperformed the one trained with the L_2 loss function.

TABLE 3. Performance of Wavelet-CCN models trained with different loss functions.

Loss	RMAE-B (%)	PSNR-S (dB)	SSIM-S	BSR (%)
L_1	3.44±0.51	39.40±0.94	0.977±0.004	85.85±6.57
L_2	3.93±0.56	38.30±0.89	0.972±0.004	81.84±7.92

D. EFFECTS OF MODEL ARCHITECTURES

1) COMPARISON WITH PYRAMIDAL ARCHITECTURE AND CASCADE ARCHITECTURE

Fig. 6 illustrates the quantitative results of the model of pyramidal architecture and Wavelet-CCN. The average

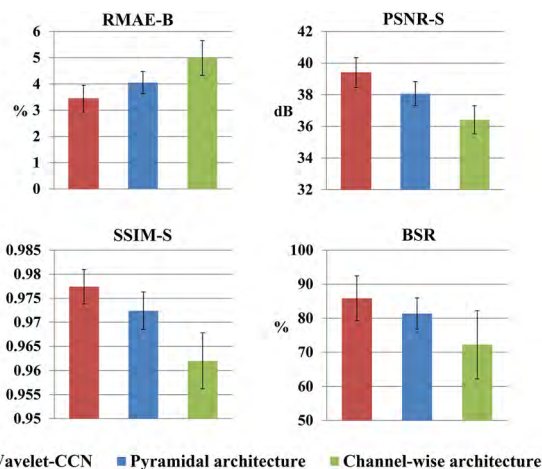


FIGURE 6. Performance comparison of models of different architectures. The number of filters in each convolutional layer is fixed as 128, and the level of wavelet decomposition is set to 3.

RMAE of the bone images predicted by the pyramidal model is significantly higher than that of the Wavelet-CCN (4.05% to 3.57%). The average PSNR of the soft-tissue images produced by the pyramidal model is lower than 1 dB the PSNR of Wavelet-CCN. The average BSR of the Wavelet-CCN model is higher than 3.15% that of the pyramidal model.

To further verify the effectiveness of Wavelet-CCN model, the quality metrics of the approximate sub-bands of bone images of each level predicted by the Wavelet-CCN model and pyramidal model are evaluated. We assume that after each forecast of a higher resolution and adding the high-frequency information into the inputs of subsequent level, the prediction accuracy of approximation sub-bands should improve accordingly. The results are consistent with our expectation, as shown in Fig. 7.

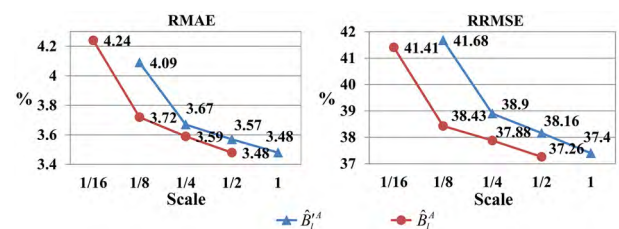


FIGURE 7. Prediction performance of Wavelet-CCN models for approximate sub-bands of different levels.

2) COMPARISON BETWEEN CHANNEL-WISE MODEL AND WAVELET-CCN MODEL

In this section, the channel-wise model is compared with the Wavelet-CCN model. 10 individual ConvNets are trained as the channel-wise model to predict the approximate sub-band and nine detail sub-bands of bone images for three-level wavelet decomposition. The final prediction result can be obtained by *idwt* from the ten wavelet sub-bands. The prediction performance of Wavelet-CCN and the channel-wise model are reported in Fig. 6. The performance of the Wavelet-CCN model is better than that of the channel-wise model in terms of RMAE-B, PSNR-S, SSIM-S, and BSR.

The examples of bone suppression results produced by the models of different architectures are shown in Fig. 8. Less rib shadow remains in soft-tissue image of Fig. 8(b), (d), and (e). The Wavelet-CCN model produces cleaner and sharper bone images than the channel-wise model. The bone image produced by channel-wise model has some artifacts beside spinal column, as shown in Fig. 8(d). The prediction results of each level are shown in Fig. 9. The top half of Fig. 9(d) shows that the channel-wise model cannot predict the details clearly of channel D at all levels. The sub-bands of bone images predicted by the Wavelet-CCN model are more similar to the ground truths.

E. HAAR VS DAUBECHIES-N WAVELET BASIS

The effect of different wavelet basis for wavelet decomposition on the performance of Wavelet-CCN models

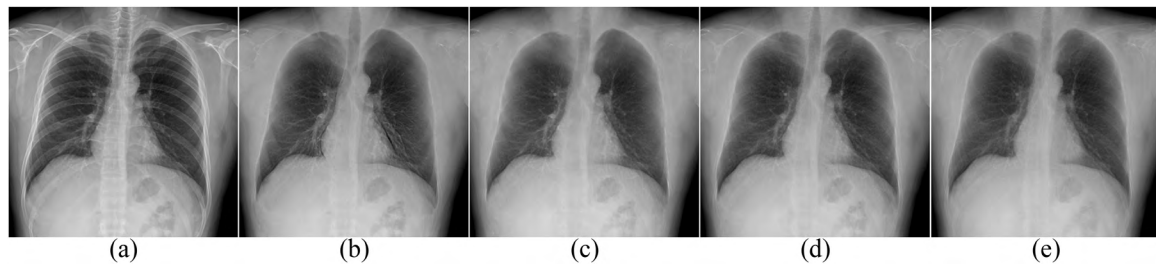


FIGURE 8. An original CXR and soft-tissue images produced by different methods: (a) a CXR, (b) the corresponding ground truth of soft-tissue image, (c) the soft-tissue image produced by a Wavelet-CCN model, (d) the soft-tissue image produced by a model of pyramidal architecture, and (e) a soft-tissue image produced by a model of channel-wise architecture.

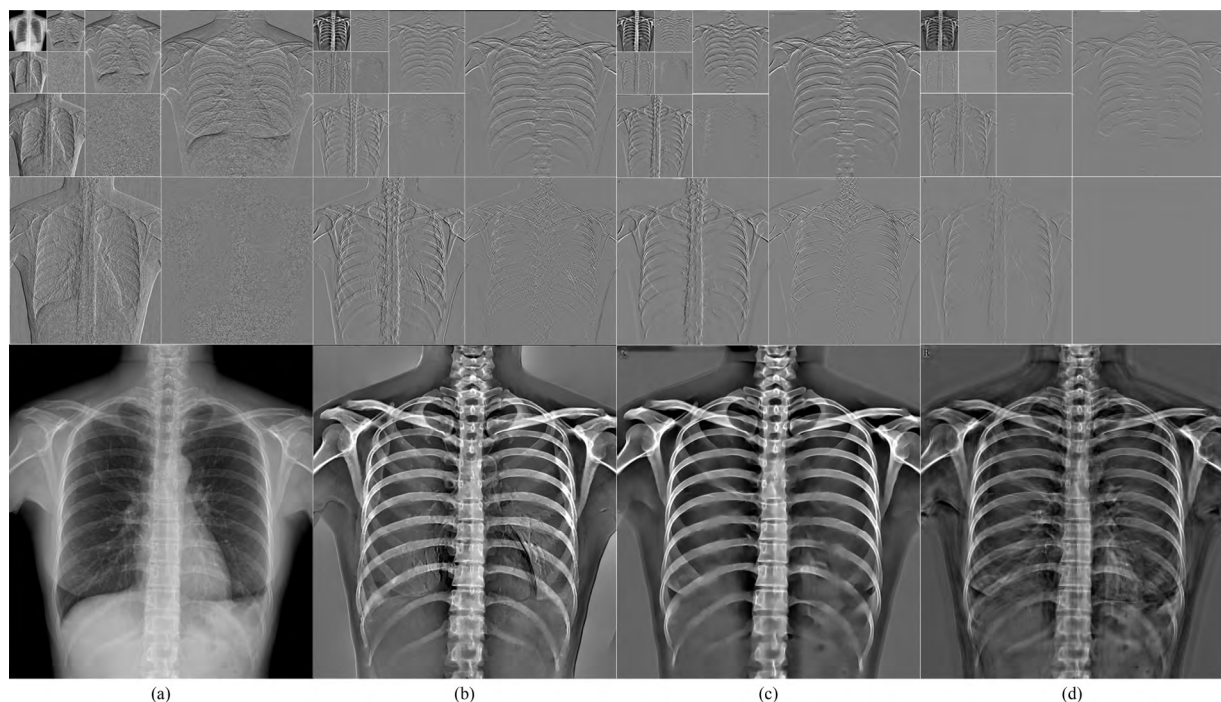


FIGURE 9. (a) Visualization of wavelet coefficients of an input CXR, (b) corresponding ground truth of bone image, (c) bone image produced by our Wavelet-CCN model, and (d) bone image produced by channel-wise wavelet ConvNet model. Top: wavelet coefficients. Bottom: images.

is evaluated. To maintain the norm preservation property of the entire procedure, only the orthogonal db wavelets are considered for evaluation. Among db wavelets, Haar wavelet is the only orthogonal wavelet which has both linear phase and symmetrically compact support. Meanwhile, the Haar wavelet [33] has the shortest support among all orthogonal wavelets, and Haar dwt can perform efficiently with a computational overhead linear. In terms of all evaluation metrics of our experiments as shown in Fig. 10, the Wavelet-CCN model with Haar wavelet shows better performance than others with db wavelets as shown in Fig. 10.

F. COMPARISON WITH CAMSNET IN GRADIENT DOMAIN

The experiment settings of the CamsNet models are same as in our previous work [4]. The performance gains profit in the wavelets domain against gradient domain. Compared with the CamsNet model, RMAE and SSIM of

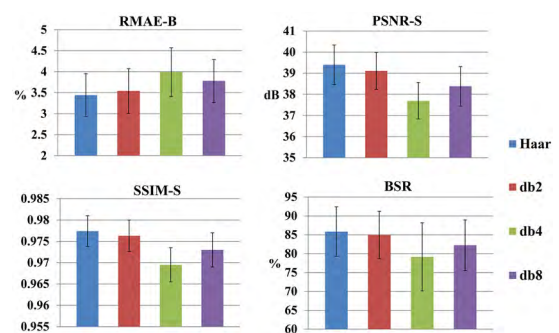


FIGURE 10. Performance comparisons of the Wavelet-CCN models with different wavelet bases.

soft-tissue image and BSR of Wavelet-CCN are improved by 0.35, 0.18, and 1.88 percentage points, respectively, and PSNR of soft-tissue image is also improved by 0.53 dB, as listed in Table 4. Standard deviation of each measuring

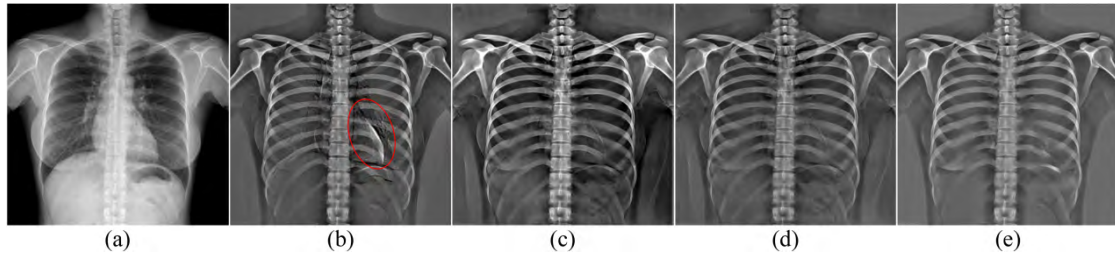


FIGURE 11. The bone images produced by different models from an input CXR (a). (b) is the corresponding ground truth of bone image. (c), (d), and (e) are the bone images produced by CamsNet, CamsNet with post-processing, and Wavelet-CCN, respectively.

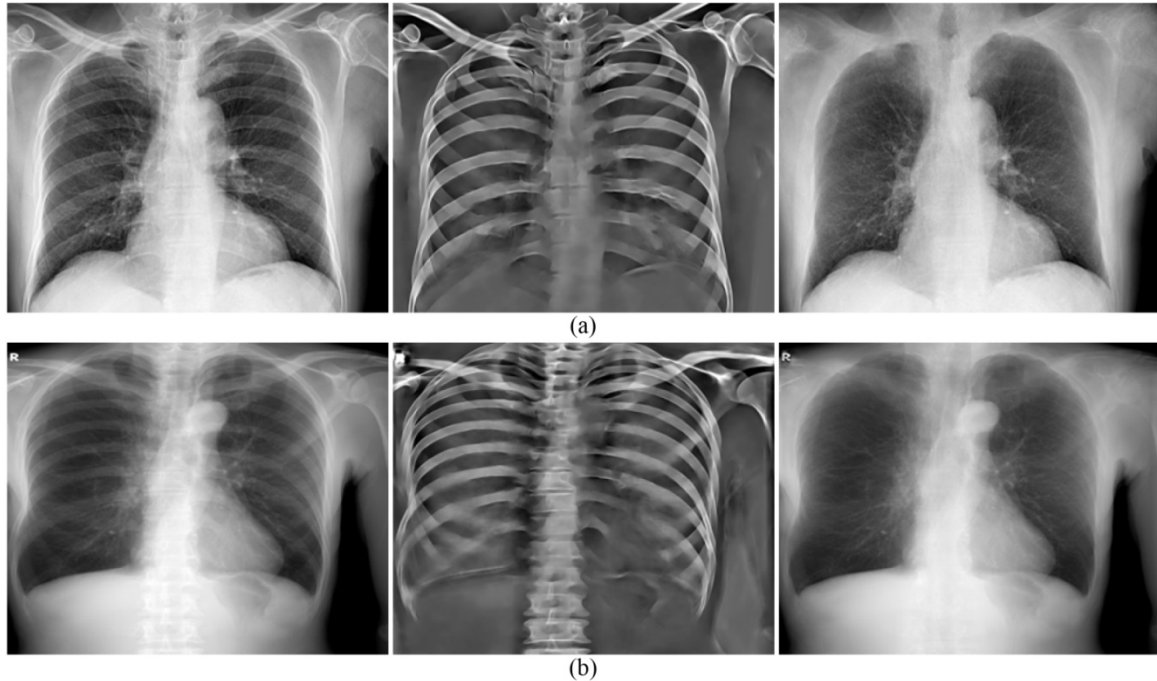


FIGURE 12. Illustration of cross-dataset generalization of a Wavelet-CCN model. From left to right: the input CXRs, the corresponding bone images, and the soft-tissue images produced by a Wavelet-CCN model. The input CXRs in (a) and (b) are acquired with a DR machine (Siemens FD-X system) and a CR machine (FUJIFILM system), respectively.

TABLE 4. Performance comparison of three-level Wavelet-CCN and four-level CamsNet models for bone suppression.

	RMAE-B (%)	PSNR-S (dB)	SSIM-S	BSR (%)
CamsNet	3.83±0.60	38.70±1.10	0.976±0.004	83.80±8.00
Wavelet-CCN	3.44±0.51	39.40±0.94	0.977±0.004	85.85±6.57

metric has different degrees of reduction, and the prediction results of Wavelet-CCN model are more stable. Our wavelet-CCN achieves the highest value in PSNR, SSIM, and BSR for produced soft tissue image, and the lowest value in RMAE for bone image. Compared with CamsNet, the bone structure produced by the Wavelet-CCN model is smoother and has no intensity inhomogeneous problem. Meanwhile, the motion artifacts are almost eliminated visually, as shown in Fig. 11.

The speed of the proposed Wavelet-CCN model is six times faster than the CamsNet model with the multi-scale fusion

procedure which takes 66 s for the full bone suppression procedure on a CXR of about 2000×2000 pixels.

G. GENERALIZATION CAPABILITY OF WAVELET-CCN ON DIFFERENT DATASET

Different X-ray machines, different PA CXRs including DR images acquired with Siemens FD-X (Siemens Healthcare) and CR (computed radiograph) acquired with FUJIFILM and GE Healthcare Definium 6000 are collected to verify whether or not Wavelet-CCN can work well on CXRs acquired from different X-ray machines. These CXRs were resized to a fixed spatial resolution of $0.194 \text{ mm} \times 0.194 \text{ mm}$ by bi-cubic interpolation and fed to the Wavelet-CCN model as the inputs.

The examples of bone suppression produced by a Wavelet-CCN model are shown in Fig. 12. The Wavelet-CCN model is applied on a CXR, which contains a lung nodule overlapped by a rib, as shown in Fig. 13, to observe the effect

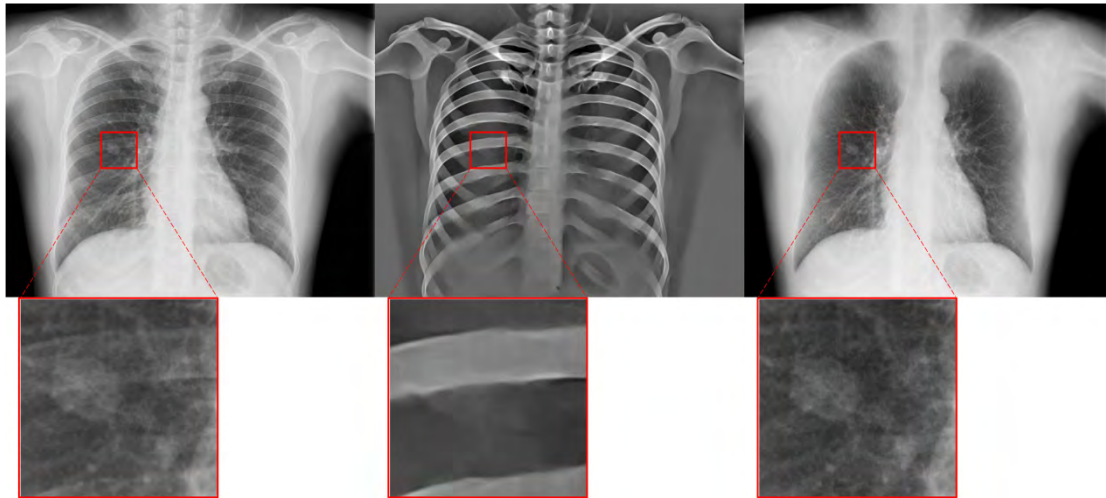


FIGURE 13. Illustration of cross-dataset generalization of Wavelet-CCN model. From left to right: the input CXR with a nodule, the corresponding bone image and soft-tissue image produced with our method. The input CXR is acquired using a CR machine.

of bone suppression on lung nodule. The contrast and shape of the lung nodule are preserved in the soft-tissue image.

IV. DISCUSSION

From the experimental results, it can be seen that the proposed cascaded convolutional network model in wavelet domain can perform bone suppression effectively. However, there are still some limitations in this work. Considering that the ground truth for training is not the real DES soft-tissue and bone images, the trained Wavelet-CCN models cannot produce the same result as DES soft-tissue images. It is preferable that the raw data of DES CXRs be used as the ground truth. Unfortunately, the raw data of DES chest radiographs are not currently available to us from commercial DES systems. And, the quality of soft-tissue images produced with the Wavelet-CCN model in some regions (e.g., trachea) should be further improved. Furthermore, the usefulness of our method have not been proved for clinical diagnosis and computerized analysis.

The Wavelet-CCN model can be further investigated to improve the performance in several aspects. One is that the loss function for training can be replaced with the mix of multi-scale SSIM and L_1 loss function which may remove some of the local minima and converge to a better local minimum [28]. Another is to train multiple Wavelet-CCN models for the different anatomical regions of CXRs. Wavelet transform for Wavelet-CCN model can be also replaced with other transforms such as contourlet [34], ridgelet [35], and curvelet transform [36]. These transforms which can provide more powerful and sparse representations may lead to better predication of bone and soft-tissue images.

V. CONCLUSION

This paper has presented a Wavelet-CCN model for bone suppression which combines the benefits of wavelet transform with the flexibility of convolutional networks. It has

made a considerable improvement for bone suppression over the existing CamsNet model both quantitatively and qualitatively. Our proposed Wavelet-CCN model is superior over the CamsNet model in gradient domain in terms of several aspects. First, the sparsity of wavelet coefficients inherits the advantages of the gradient domain used in our previous work [4]. Second, wavelet reconstruction can avoid the inconsistent background intensity, which appears in 2D integration of the predicted gradients. Third, *idwt* completes up-sampling, perfectly cascades the wavelet convolution network between different scales simultaneously, and combines all levels of the enriched contextual information.

REFERENCES

- [1] K. Suzuki, H. Abe, F. Li, and K. Doi, "Suppression of the contrast of ribs in chest radiographs by means of massive training artificial neural network," *Proc. SPIE*, May 2004, pp. 1109–1119.
- [2] L. Hogeweg, C. I. Sanchez, and B. van Ginneken, "Suppression of translucent elongated structures: Applications in chest radiography," *IEEE Trans. Med. Imag.*, vol. 32, no. 11, pp. 2099–2113, Nov. 2013.
- [3] S. Chen and K. Suzuki, "Computerized detection of lung nodules by means of 'virtual dual-energy' radiography," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 2, pp. 369–378, Feb. 2013.
- [4] W. Yang et al., "Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain," *Med. Image Anal.*, vol. 35, pp. 421–433, Jan. 2016.
- [5] P. Vock and Z. Szucs-Farkas, "Dual energy subtraction: Principles and clinical applications," *Eur. J. Radiol.*, vol. 72, no. 2, pp. 231–237, Nov. 2009.
- [6] G. Simkó, G. Orbán, P. Máday, and G. Horváth, "Elimination of clavicle shadows to help automatic lung nodule detection on chest radiographs," in *Proc. 4th Eur. Conf. Int. Fed. Med. Biol. Eng. (MBEC)*. Antwerp, Belgium, 2008, pp. 488–491.
- [7] H. Oğul, B. B. Oğul, A. M. Ağlıdere, T. Bayrak, and E. Sümer, "Eliminating rib shadows in chest radiographic images providing diagnostic assistance," *Comput. Methods Programs Biomed.*, vol. 127, pp. 174–184, Apr. 2016.
- [8] M. Loog, B. van Ginneken, and M. A. Viergever, "Segmenting the posterior ribs in chest radiographs by iterated contextual pixel classification," *Proc. SPIE*, vol. 5032, pp. 609–618, May 2003.
- [9] L. E. Hogeweg, "Automatic detection of tuberculosis in chest radiographs," Ph.D. dissertation, Dept. Radboud Univ. Med. Center, Radboud Univ., Nijmegen, The Netherlands, 2013.

- [10] M. Loog and B. Ginneken, "Segmentation of the posterior ribs in chest radiographs using iterated contextual pixel classification," *IEEE Trans. Med. Imag.*, vol. 25, no. 5, pp. 602–611, May 2006.
- [11] H. X. Nguyen and T. T. Dang, "Ribs suppression in chest X-ray images by using ICA method," in *Proc. 5th Int. Conf. Biomed. Eng. Vietnam*, 2015, pp. 194–197.
- [12] T. Rasheed, B. Ahmed, M. A. U. Khan, M. Bettayeb, S. Lee, and T.-S. Kim, "Rib suppression in frontal chest radiographs: A blind source separation approach," in *Proc. 9th Int. Symp. Signal Process. Appl. (ISSPA)*, Feb. 2007, pp. 1–4.
- [13] K. Suzuki, H. Abe, H. MacMahon, and K. Doi, "Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN)," *IEEE Trans. Med. Imag.*, vol. 25, no. 4, pp. 406–416, Apr. 2006.
- [14] M. Loog, B. van Ginneken, and A. M. R. Schilham, "Filter learning: Application to suppression of bony structures from chest radiographs," *Med. Image Anal.*, vol. 10, no. 6, pp. 826–840, Dec. 2006.
- [15] S. Chen and K. Suzuki, "Separation of bones from chest radiographs by means of anatomically specific multiple massive-training ANNs combined with total variation minimization smoothing," *IEEE Trans. Med. Imag.*, vol. 33, no. 2, pp. 246–257, Feb. 2014.
- [16] M. Gusarev, R. Kuleev, A. Khan, A. R. Rivera, and A. M. Khattak, "Deep learning models for bone suppression in chest radiographs," in *Proc. IEEE Conf. Comput. Intell. Bioinf. Comput. Biol. (CIBCB)*, Aug. 2017, pp. 1–7.
- [17] Y. Chen, T.-C. Chang, C. Zhou, and T. Fang, "Gradient domain layer separation under independent motion," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 694–701.
- [18] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.
- [19] J. Knapp and S. Worrell, "Feature based neural network regression for feature suppression," U.S. Patent 8 204 292 B2, Jun. 19, 2012.
- [20] E. Kang, J. Min, and J. C. Ye. (2017). "Wavelet domain residual network (WavResNet) for Low-dose X-ray CT reconstruction." [Online]. Available: <https://arxiv.org/abs/1703.01383>
- [21] J. Gu and J. C. Ye. (2017). "Multi-scale wavelet domain residual learning for limited-angle CT reconstruction." [Online]. Available: <https://arxiv.org/abs/1703.01382>
- [22] N. Kumar, R. Verma, and A. Sethi, "Convolutional neural networks for wavelet domain super resolution," *Pattern Recognit. Lett.*, vol. 90, pp. 65–71, Apr. 2017.
- [23] H. Huang, R. He, Z. Sun, and T. Tan, "Wavelet-SRNet: A wavelet-based CNN for multi-scale face super resolution," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1689–1697.
- [24] T. Williams and R. Li, "Advanced image classification using wavelets and convolutional neural networks," in *Proc. 15th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Anaheim, CA, USA, Dec. 2016, pp. 233–239.
- [25] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 37, no. 12, pp. 2091–2110, Dec. 1989.
- [26] R. Sihag, R. Sharma, and V. Setia, "Wavelet thresholding for image denoising," in *Proc. Int. Conf. VLSI, Commun. Instrum. (ICVCI)*, vol. 2011, pp. 20–23.
- [27] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2015, pp. 1026–1034.
- [28] H. Zhao, O. Gallo, I. Frosio, and J. Kautz. (2015). "Loss functions for neural networks for image processing." [Online]. Available: <https://arxiv.org/abs/1511.08861>
- [29] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in *Proc. Proc. 23rd ACM Int. Conf. Multimedia*, Oct. 2015, pp. 689–692.
- [30] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [31] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2015.
- [32] S. Mallat, *A Wavelet Tour of Signal Processing*. London, U.K.: Academic, 1999.
- [33] K. Zhang and L. Zhang. (2017). "Rectifier neural network with a dual-pathway architecture for image denoising." [Online]. Available: <https://arxiv.org/abs/1609.03024>
- [34] M. N. Do and M. Vetterli, "Contourlets: A directional multiresolution image representation," in *Proc. Int. Conf. Image Process.*, Rochester, NY, USA, Sep. 2002, p. 1.
- [35] E. J. Candès, "Ridgelets: Estimating with ridge functions," *Ann. Statist.*, vol. 31, no. 5, pp. 1561–1599, 2003.
- [36] J.-L. Starck, E. J. Candès, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Trans. Image Process.*, vol. 11, no. 6, pp. 670–684, Jun. 2002.



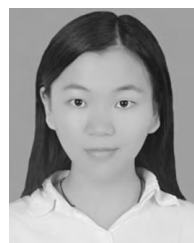
YINGYIN CHEN was born in China in 1988. She received the B.S. degree from the Department of Biomedical Engineering, South Medical University, Guangzhou, China, in 2010, where she is currently pursuing the Ph.D. degree with the Guangdong Provincial Key Laboratory of Medical Image Processing. Her research interests include medical image analysis, machine learning, deep learning, computerized-aid diagnosis, and medical image retrieval.



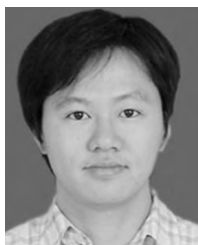
XIAOFANG GOU was born in China in 1994. She received the bachelor's degree in biomedical engineering from Southern Medical University, Guangzhou, China, in 2016, where she is currently pursuing the M.E. degree with the Department of Biomedical Engineering. Her research interests include the medical images analysis and computer vision.



XIUXIA FENG was born in China in 1993. She received the B.S. degree in biomedical engineering from Southern Medical University, Guangzhou, China, in 2016, where she is currently pursuing the M.S. degree in engineering with the Department of Biomedical Engineering. Her research interests include the medical image analysis and computer vision.



YUNBI LIU was born in China in 1994. She received the bachelor's degree in biomedical engineering from Southern Medical University, Guangzhou, China, in 2015, where she is currently pursuing the Ph.D. degree with the Department of Biomedical Engineering. Her research interests include medical image analysis and computerized-aid diagnosis.



USA. He has published 26 journal papers. His current research interests include radiomics, medical image analysis, deep learning, machine learning, and artificial intelligence.

GENGGENG QIN received the B.Sc. degree in medicine from Sun Yat-sen University, Guangzhou, China, in 2007, and the M.Sc. degree in medical imageology from Southern Medical University, Guangzhou, in 2014. He is currently an Associate Chief Physician with the Department of Radiology, Nanfang Hospital, Southern Medical University, and also a Visiting Scholar with the Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, TX,



research interests include medical image analysis, machine learning, and computerized-aid diagnosis.

WEI YANG received the B.Sc. degree in automation from the Wuhan University of Science and Technology, Wuhan, China, in 2001, the M.Sc. degree in control theory and control engineering from Xiamen University, Xiamen, China, in 2005, and the Ph.D. degree in biomedical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009. He is currently a Professor with the School of Biomedical Engineering, Southern Medical University, Guangzhou, China. His



His research interests include medical image analysis, pattern recognition, and computerized-aided diagnosis.

QIANJIN FENG received the M.S. and Ph.D. degrees in biomedical engineering from First Military Medical University, China, in 2000 and 2003, respectively. From 2003 to 2004, he was a Faculty Member with the School of Biomedical Engineering, First Military Medical University. Since 2004, he has been with Southern Medical University, China, where he is currently a Professor and the Dean of the School of Biomedical Engineering.



His research interests include the medical imaging and medical image analysis.

WUFAN CHEN received the B.S. and M.S. degrees in applied mathematics and computational fluid dynamics from the Peking University of Aeronautics and Astronautics, China, in 1975 and 1981, respectively. From 1981 to 1987, he was with the School of Aerospace, National University of Defense Technology, China. From 1987 to 2004, he was with the Department of Training, First Military Medical University, China. Since 2004, he has been with Southern Medical University, China, where he is currently a Professor with the School of Biomedical Engineering and the Director of the Guangdong Provincial Key Laboratory for Medical Image Processing.

...