LOGO

Multi-scale reconstruction of undersampled spectral-spatial OCT data for coronary imaging using deep learning

Xueshen Li, Shengting Cao, Hongshan Liu, Xinwen Yao, Brigitta C. Brott, Silvio H. Litovsky, Xiaoyu Song, Yuye Ling, Yu Gan

Abstract—**Coronary artery disease (CAD) is a cardio-vascular condition with high morbidity and mortality. Intravascular optical coherence tomography (IVOCT) has been considered as an optimal imagining system for the diagnosis and treatment of CAD. Constrained by Nyquist theorem, dense sampling in IVOCT attains high resolving power to delineate cellular structures/features. There is a trade-off between high spatial resolution and fast scanning rate for coronary imaging. In this paper, we propose a viable spectral-spatial acquisition method that downscales the sampling process in both spectral and spatial domain while maintaining high quality in image reconstruction. The down-scaling schedule boosts data acquisition speed without any hardware modifications. Additionally, we propose a unified multi-scale reconstruction framework, namely Multiscale-Spectral-Spatial-Magnification Network (MSSMN), to resolve highly down-scaled (compressed) OCT images with flexible magnification factors. We incorporate the proposed methods into Spectral Domain OCT (SD-OCT) imaging of human coronary samples with clinical features such as stent and calcified lesions. Our experimental results demonstrate that spectral-spatial down-scaled data can be better reconstructed than data that are down-scaled solely in either spectral or spatial domain. Moreover, we observe better reconstruction performance using MSSMN than using existing reconstruction methods. Our acquisition method and multi-scale reconstruction framework, in combination, may allow faster SD-OCT inspection with high resolution during coronary intervention.**

Index Terms—**Optical coherence tomography; Coronary**

imaging; Image reconstruction; Deep learning

I. INTRODUCTION

Coronary Artery Disease (CAD) is the narrowing of coronary arteries caused by build-up of atherosclerotic plaques. It is the most common type of heart disease, leading to 1 in 7 deaths in the U.S [1]. Percutaneous Coronary Intervention (PCI) is one of the most common non-surgical procedures to open clogged coronary arteries. Approximately one million PCI are performed annually [2]. To assess PCI performance, imaging tools are necessary before and after the procedure, among which Intravascular Optical Coherence Tomography (IVOCT) offers a real-time three-dimensional evaluation with highest resolution [3]–[5].

Most commercial IVOCT is based on Swept Source Optical Coherence Tomography (SS-OCT), which has a high scanning rate ($>$100 frames/s) with a spatial resolution around 10 μm [6]. This resolution, however, is insufficient to delineate key features such as endothelial lining, stent coverage, and cholesterol crystals, at a much finer level of 2 μm for accurate assessment of stent implantation. In contrast, Spectral Domain Optical Coherence Tomography (SD-OCT) at a shorter wavelength outperforms its swept source counterpart in resolution and contrast, and has demonstrated its capability in visualizing critical structures [7]. Nonetheless, resolving at 2 μm requires a denser sampling that inevitably leads to a lower scanning rate of about 8 frames/s [7], insufficient for performing catheter pullback to visualize the length of arteries during PCI. Currently, there is no technology available to simultaneously maintain high spatial resolution at such a 2 μm level and a fast scanning rate. A real-time visualization with higher resolution will be desirable to better PCI outcome.

Real-time visualization requires both fast acquisition and display. A cost-effective approach to improve scanning rate is to sample and store less data. Spectral domain-OCT acquires data and processes data for visualization in spectral domain. Various methods have been proposed to down-scale OCT data from either spatial or spectral domain [8]–[14]. For example, low signal-to-noise ratio (SNR) OCT images from spatial undersampling were obtained and reconstructed via sparse representation [8]–[11]. Also, spectral raw data were undersampled and reconstructed using machine learning or deep
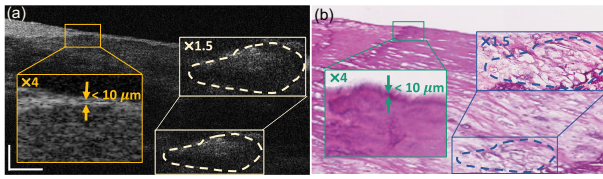
Fig. 1. Representative of a pairwise match between coronary OCT image (left) and its histology (right). Two regions are magnified with different scale factors for coronary analysis. scale bar: 200 μm.

learning (DL) methods [12]–[14]. However, these methods shared downsampling factors up to 4-fold, still insufficient to booster acquisition and display from the level of 8 frames/s to the level of 100 frames/s for coronary intervention [6].

Moreover, multi-scale magnification is needed during reconstruction of down-scaled images. An example is shown in Fig. 1, where one region of fibrous cap in human coronary is shown in an OCT cross-sectional image with its matched histology. There are two regions of clinical interest: the endothelial layer (gold box in (a) and green box in (b)) and the lipid pool (white box in (a) and blue box in (b)). Based on different needs, the two regions are magnified at different scales. As gatekeepers for passage of lipoprotein and leukocytes into intima, the endothelial layer needs to be magnified with a larger scale for accurate thickness assessment. For lipid pools, however, thickness assessment is of less interest thus a lower magnification suffices. A binary justification of whether the distance between pool edge to lumen is larger than 65 μm or not suffices the evaluation of risk of rupture [15]. Resolving images to the finest resolution, unfortunately, will increase the total number of pixels in display and add computational burden for visualization. As a result, a multi-scale magnification framework would be beneficial for real-time coronary visualization and analysis. At present, multi-scale magnification is illusive for OCT image reconstruction.

Existing reconstruction methods [13], [14] treat each scale factor as a single task and only work for scale factors that are integers. Whenever a new scale factor is needed, a new model is needed and requires re-training. As a result, multiple neural networks and individual training are necessary if following existing methods to maintain the flexibility with multi-scale magnification. A unified neural network model with high flexibility in magnification factors and computational cost-effectiveness is preferable.

In this paper, we devise an efficient spectral-spatial domain acquisition method for SD-OCT systems. Our method reduces data size in both spectral and spatial domains during imaging. Additionally, we propose a unified neural network, namely Multiscale-Spectral-Spatial-Magnification Network (MSSMN) for image reconstruction. Inspired by super-resolution technology, MSSMN magnifies compressed information in both spectral and spatial domains with any arbitrary magnification factors. With extensive experiments in human coronary OCT images, we demonstrate high accuracy, data efficiency, and magnification flexibility of our proposed method. Our contributions are as below:

- We propose a novel method to reduce data size by downscaling SD-OCT data in both spectral domain and spatial domain. The proposed method improves acquisition

speed. Our method does not require any additional optical setup and can be easily integrated into existing SD-OCT systems.

- We devise a unified network, MSSMN, to address the issue of multi-scale magnification. To the best of our knowledge, it is the first multi-scale OCT reconstruction framework in spectral-spatial domain and allows for arbitrary magnification factors.

- We propose the first super-resolution neural network for coronary OCT imaging with an aim to highlight pathological features such as calcified region, stent strut, and lipid region.

II. RELATED WORK

A. OCT image acquisition/compression

In SD-OCT, images can be compressed by removing part of their spectrum data [12], [16]. Lichtenegger et al. acquired discontinuous spectrum by modifying light source [16]; Zhang et al. studied 2-fold and 3-fold undersampled spectral data [12]. In spatial domain, previous research aimed to compress OCT images without modifying the spectrum. Algorithms based on transform coding, such as discrete cosine transform (DCT) and discrete wavelet transform (DWT), have been used in OCT image compression [17]. Also, algorithms based on sparse representation were also applied to OCT image compressing [8], [9]. The U-Net [18] structure has been used in compressing OCT images because of their unique bottleneck layers [19].

B. OCT image reconstruction

1) None-DL methods for OCT image reconstruction: Sparse representation algorithms have been applied to reconstruct OCT images [8], [9]. Recently, a nonlocal weighted sparse representation (NWSR) method was proposed [10]. The NWSR algorithm achieved better estimation of sparse representation by exploiting information from noisy and denoised patches. Very recently, a mixed low rank approximation and second-order tensor-based total variation (LRSOTTV) approach was presented for OCT [11]. Compared with other low-rank tensor-based methods, the LRSOTTV achieved better results and lower complexity by adopting a tensor prior.

2) DL-based methods for OCT image reconstruction: Variations of U-Net have been adopted in reconstructing OCT images [12], [13], [16], [20]–[22]. Hao et al. reconstructed high quality OCT images from their low bit-depth counterparts using a U-Net based architecture [21]. Also with the U-Net architecture, Liang et al. enhanced the resolution while recovering realistic speckles in the OCT images [13]. Further, Qiu et al. achieved simultaneously denoising and reconstruction of OCT images using a U-Net based network [22]. The U-Net based networks have been used in reconstructing OCT images compressed in spectral domain.

Apart from the U-Net based DL models, other DL architectures such as SRResNet, residual dense network, deep back-projection network, and residual-in-residual dense network have been used in different research of reconstructing OCT
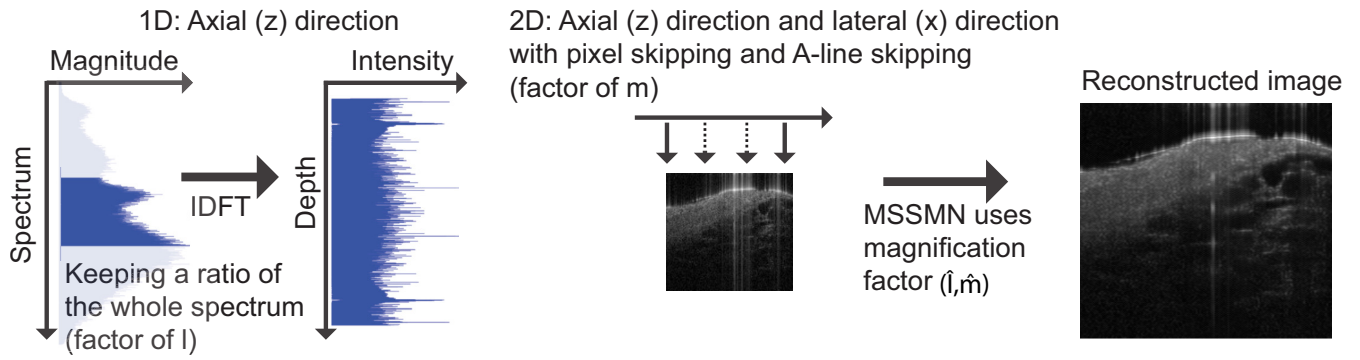
Fig. 2. Illustration of the proposed spectral-spatial acquisition method and multi-scale reconstruction framework. During scanning, the A-lines are skipped in lateral direction at factor of m. When acquiring depth information through A-lines, spectrum data are truncated at factor of l. After IDFT, we skip pixels in spatial domain at factor of m. The spectrum truncation and pixel skipping reduces data storage consumption; the A-line skipping boosts image acquisition speed. The structure of MSSMN is detailed in Fig. 3.

images [14], [23], [24]. Furthermore, the Generative adversarial network (GAN) based networks were also incorporated in many OCT reconstruction research [14], [21], [23]–[26]. Among these research, Yuan et al. proposed to super-resolve optical resolution of OCT images using DL algorithms [14]; Cao et al. increased both optical and digital resolutions of OCT [24]. However, none of previous research managed to develop a unified framework that was capable of reconstructing OCT images in both spectral and spatial domains by multiple magnification factors.

III. METHOD

In this part, we first discuss spectral-spatial OCT image acquisition in detail. Then we define the problem of multi-scale OCT reconstruction (magnification) in spectral-spatial domain. Finally, we introduce the framework and implementations.

A. OCT acquisition in spectral-spatial domain

We collaboratively reduce the amount of data in both spatial domain and spectral domain during OCT image acquisition with a viable scale factor of (l, m), where l is a acquisition factor in spectral domain and m is a acquisition factor in spatial domain. As shown in Fig. 2, in spectral domain, we only collect $\frac{1}{l}$ of the spectrum at the center of the bandwidth, since it carries most information spectrum. In spatial domain, we skip $m - 1$ A-lines per m A-lines during scanning. The A-line skipping is carried out in lateral direction; in order to further reduce data storage consumption and maintain the aspect ratio of compressed OCT images in spatial domain, we keep $\frac{1}{m}$ pixels in axial direction after Inverse Discrete Fourier Transform (IDFT). The proposed method can be integrated with existing workflow of SD-OCT without any modifications of hardware.

With less data sampled, such acquisition process could be considered as a degradation process. The data-acquisition and reconstruction process are degradation and super-resolution processes [27]. In SD-OCT images, the degradation process can be reformulated as:

$$I_{l,m}^C = \mathcal{F}^{-1}(S \cdot T_l) \downarrow_m \qquad (1)$$

where $I_{l,m}^C$ is a OCT image compressed with a acquisition factor (l, m) in spectral-spatial domain; $\mathcal{F}^{-1}$ stands for inverse Fourier transform operator; S is the raw spectrum acquired by SD-OCT systems; T_l is a truncating kernel which keeps $\frac{1}{l}$ of the spectrum in spectral domain; $\downarrow_m$ stands for skipping operation in both axial and lateral directions in spatial domain, keeping $\frac{1}{m}$ pixels in each direction. Note that the truncating operation in the spectral domain is equivalent to the convolution in the spatial domain because of the convolution theorem. T_l and $\downarrow_m$ are simple operations with time complexity of $O(1)$. We refer to OCT images acquired by (l, m) with $l > 1$ and $m > 1$ as low resolution (LR) image.

B. Multi-scale magnification of compressed OCT images

The acquisition process of compressed OCT images is deterministic; once the acquisition factors l and m are fixed, a unique $I_{l,m}^C$ is generated. Reconstruction of OCT images can be considered as a reversing process of compression, which provides high quality images given $I_{l,m}^C$. Ideally, the output after reconstruction should be $I^{Ref} = \mathcal{F}^{-1}(S)$. However, reconstruction is a ill-posed inverse problem since multiple solutions may exist for a single $I_{l,m}^C$. The DL-based algorithm learns a single non-linear mapping from $I_{l,m}^C$ to $\hat{I}^{Rec}$. In this paper, we rather propose a novel reconstruction framework for resolving compressed OCT images by multiple magnification factors:

$$\hat{I}_{\hat{l},\hat{m}}^{Rec} = G_{\hat{l},\hat{m}}(I_{l,m}^C; \theta_{\hat{l},\hat{m}}) \qquad (2)$$

here $G_{\hat{l},\hat{m}}$ is a unified neural network which has input of LR image and two magnification factors: magnification factor $\hat{l}$ in spectral domain and magnification factor $\hat{m}$ in spatial domains; $\theta_{\hat{l},\hat{m}}$ stands for parameters in $G_{\hat{l},\hat{m}}$.

C. Unified deep learning framework for OCT image magnification

Our framework of MSSMN consists of the following modules: feature extractor, meta-upscaling and meta-restoration modules [20], and loss function.

1) Feature extractor: The feature map F is extracted from $I_{l,m}^C$, using feature extractor E which contains parameters θ_E. We choose to use Residual Channel Attention Network (RCAN) for feature extraction, which adopts the Residual-In-Residual (RIR) structure [28]. The RCAN achieves much deeper network structure [29].
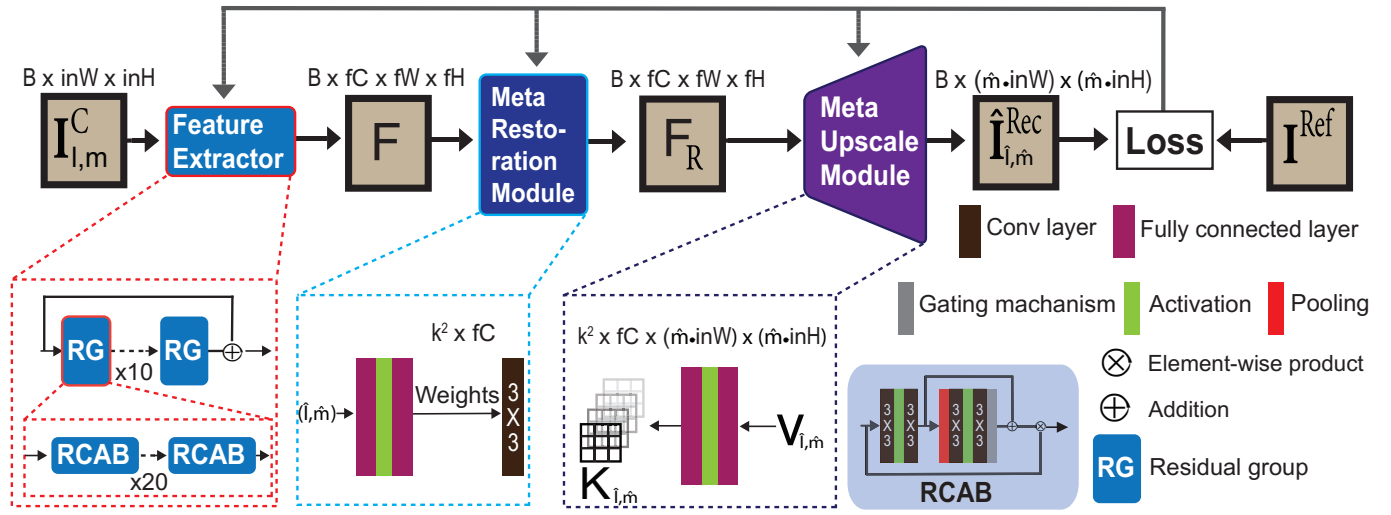
Fig. 3. Scheme of the proposed Multiscale-Spectral-Spatial-Magnification Network (MSSMN) using Residual Channel Attention Network (RCAN) as feature extractor. Residual Channel Attention Blocks (RCAB) are the building block of RCAN. The key components of our framework are feature extractor, meta-upscale module, and meta-restoration module. The feature map F is extracted by feature extractor, then processed by meta-restoration (generating F_R) and meta-upscaling modules (generating $\hat{I}_{\hat{l},\hat{m}}^{Rec}$). Magnification factors $(\hat{l}, \hat{m})$ determine the reconstruction weights. The width and height of reconstructed OCT images are magnified by spatial magnification factor $\hat{m}$. Number of feature channels fC $= 64$ and convolution kernel size $k = 3$.

2) Meta-upscaling and meta-restoration modules for multiple degradation parameters: The $\hat{I}_{\hat{l},\hat{m}}^{Rec}$ is reconstructed from the feature map F. We develop a unified framework that resolves compressed OCT images in multiple magnification factors. For this purpose, we adopt the meta-upscaling module [30] and meta-restoration module [31]. The magnification factors $(\hat{l}, \hat{m})$ are used to determine the weights of the convolution kernels in meta-learning modules. The meta-restoration module φ_R dynamically predicts weights W_R of a regular convolution layer based on magnification factors $(\hat{l}, \hat{m})$. W_R possesses a dimension $k^2 \times$ fC. The convolution layer restores F and generates restored feature map F_R. Two feature maps, F and F_R, have the same dimension of B$\times$fC$\times$fW$\times$fH, where B is the number of patches. θ_{φ_R} denotes parameters in φ_R. The meta-upscaling module φ_U dynamically predicts weights of kernels $K_{\hat{l},\hat{m}}$, given magnification factors $(\hat{l}, \hat{m})$. $K_{\hat{l},\hat{m}}$ has a shape of $k^2 \times$fC$\times(\hat{m}\cdot$inW$)\times(\hat{m}\cdot$inH$)$. With F_R and $K_{\hat{l},\hat{m}}$, the $\hat{I}_{\hat{l},\hat{m}}^{Rec}$ is reconstructed by a feature mapping function Φ:

$$K_{\hat{l},\hat{m}}(w, h) = \varphi_U(v_{\hat{l},\hat{m}}; \theta_{\varphi_U})$$
$$\hat{I}_{\hat{l},\hat{m}}^{Rec}(w, h) = \Phi(F_R(w', h'), K_{\hat{l},\hat{m}}(w, h)) \quad (3)$$
$$(w', h') = (\lfloor \frac{w}{\hat{m}} \rfloor, \lfloor \frac{h}{\hat{m}} \rfloor)$$

where θ_{φ_U} denotes parameters in φ_U, (w, h) denotes locations of pixels in $I_{l,m}^C$, (w', h') denotes locations of pixels in F_R, $v_{\hat{l},\hat{m}}$ is the input vector relating to magnification factors $(\hat{l}, \hat{m})$ and size of $I_{l,m}^C$, $\lfloor \cdot \rfloor$ stands for floor operator. The $K_{\hat{l},\hat{m}}(w, h)$ is a $k \times k$ kernel with its center at (w', h'). Note that for a single (w', h') in F_R, there will be multiple kernels $K_{\hat{l},\hat{m}}(w, h)$ for reconstructing the feature map with varying magnification factors $(\hat{l}, \hat{m})$. We set number of feature channels fC $= 64$ and convolution kernel size $k = 3$.

The $v_{\hat{l},\hat{m}}$ and Φ are defined as:

$$v_{\hat{l},\hat{m}}(w, h) = (\frac{w}{\hat{m}} - \lfloor \frac{w}{\hat{m}} \rfloor, \frac{h}{\hat{m}} - \lfloor \frac{h}{\hat{m}} \rfloor, \hat{l}) \quad (4)$$
$$\Phi(F_R(w', h'), K_{\hat{l},\hat{m}}(w, h)) = F_R(w', h') \otimes K_{\hat{l},\hat{m}}(w, h)$$

Based on this meta-learning scheme, we magnify compressed OCT images by multiple scale factors with a unified network.

3) Loss function: Mean squared error (MSE) is widely used loss function for general image restoration. However, the L1 (Mean absolute error) has demonstrated potential of achieving better convergence [32]. We adopted the L1 loss for training all models.

IV. EXPERIMENTS AND RESULTS

A. Experimental dataset

Human coronary samples were collected from the School of Medicine at the University of Alabama at Birmingham (UAB) and delivered to the University of Alabama for imaging. The autopsy specimens were de-identified and not considered as human subjects, in compliance with UAB's Institutional Review Board (IRB). Specimens were imaged via a commercial OCT system (Thorlabs Ganymede, Newton, NJ) [24]. We collected 2996 OCT images (B-scans) from 23 specimens. The depth of OCT images were 2.56 mm. The width of the images ranged from 2 mm to 4 mm depending on actual sample size. Among all images, the pixel size remained 2 μm $\times$ 2 μm within a B-scan. Samples were processed for Hematoxylin-Eosin (H&E) histology after imaging.

We acquired the OCT images using the proposed spectral-spatial acquisition method. In spectral domain, we used $\frac{1}{2}$, $\frac{1}{3}$ and $\frac{1}{4}$ of spectrum data (denoted as $l = 2, 3, 4$ respectively). In the spatial domain, we used $m = 2, 3, 4$ for A-line skipping
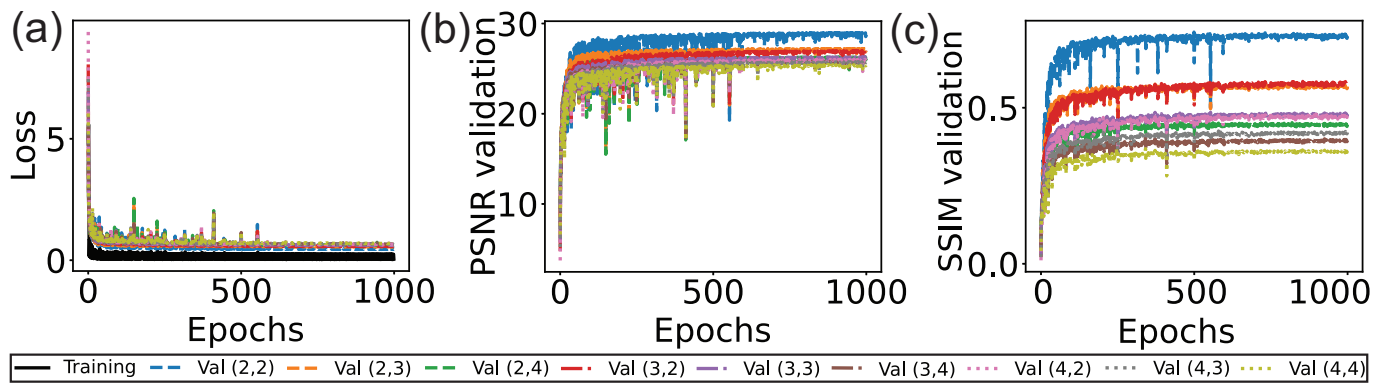
Fig. 4. Representative loss function plots and validation PSNR and SSIM scores. (a) Training and validation loss values over epochs. (b) PSNR scores over epochs in validation dataset. (c) SSIM scores over epochs in validation dataset.

TABLE I
RESULTS OF RECONSTRUCTION FROM OCT IMAGE OF ACQUISITION FACTOR (l, m) GIVEN A COMPRESSION RATE OF 25%. ALL RESULTS ARE AVERAGED BASED ON FIVE-FOLD CROSS-VALIDATION. RED INDICATES THE BEST PERFORMANCE.

(l, m) Metrics	(1,4)	(4,1)	(2,2)
PSNR	25.8183	26.6053	29.1390
SSIM	0.4860	0.4622	0.7570

during scanning and pixel skipping after IDFT. The experimental dataset was divided into 5 folds by ID of specimen for cross-validation.

B. Network implementation

We use RCAN as the feature extractor to implement MSSMN. As a comparative study, we implemented reconstruction networks with other feature extractors, including Multi-Scale Residual Network (MSRN) [33], Residual Dense Network (RDN) [34], and Residual Dense Unet (RDU) [35]. For the network implementations, we used 8 multi-scale residual blocks for MSRN; 10 residual dense blocks (RDB) for RDN, with each RDB containing 6 convolution layers; 7 levels for RDU, with each level containing 1 denoising block; 10 residual groups (RG) for RCAN, with each RG containing 20 residual channel attention blocks (RCAB). We implemented the sub-pixel upsampling module [36] for each features extractors. In comparison with the abovementioned network structure, RCAN achieves deep trainable network structure and is capable of adaptively learn useful channel-wise features [28].

C. Training details

The intensities were scaled to $[0, 1]$. For each network, the training procedure was repeated 5 times on different data folds. During the training, we randomly extract 16 non-overlapped patches with the size of 64×64. Those patches are the input of the model. We augment the patches by randomly flipping to prevent overfitting. Within each batch, magnification factors $(\hat{l}, \hat{m})$ for MSSMN remained the same. We used Adam optimizer [37] for training. The learning rate

was initialized as 10^{-4}, following by a half decay for every 200 epochs. In total the networks were trained 1000 epochs to ensure convergence. The experiments were carried out in parallel on three RTX A6000 GPUs. Training on each data fold take around 10 hours.

D. Evaluation metrics

We used peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [38] to measure the quality of image magnification. The PSNR calculates pixel-wise differences between the reconstructed image and the reference image; the SSIM focuses on structural similarly between reconstructed image and the reference image.

E. Network convergence

For each fold, we trained the MSSMN for 1000 epochs to ensure convergence. Fig. 4 shows representative plots of training loss and validation loss, PSNR and SSIM scores, all from one fold of validation dataset. The MSSMN is capable of multi-scale reconstruction. We calculated validation loss values and PSNR and SSIM scores for each magnification factor. Within 1000 epochs, the MSSMN converges for each magnification factor.

F. The efficiency of spectral-spatial compression

We fixed a compression ratio of 25% and compared the reconstruction performance from OCT images among three schemes: i) solely in spectral domain ($l = 4, m = 1$); ii) solely in spatial domain ($l = 1, m = 4$); iii) jointly in spectral and spatial domains ($l = 2, m = 2$). In Table I, we show the PNSR and SSIM of magnifying OCT images of acquisition factors $(1, 4)$, $(4, 1)$ and $(2, 2)$ using MSSMN. The joint acquisition, with a factor of $(2, 2)$, provides much higher scores after reconstruction, comparing to that of $(1, 4)$ and $(4, 1)$. Thus, acquiring OCT images in the spectral-spatial domain is a better strategy than pushing the compression limit in a single domain.

When acquiring spectrum data, there are two possible approaches: truncating the spectrum from center and dropping spectral data uniformly. We compared the two strategies by

TABLE II

RESULTS OF RECONSTRUCTION FROM OCT IMAGES OF ACQUISITION FACTOR (l, m) USING VARIOUS FEATURE EXTRACTORS (MSRN, RDN, RDU, RCAN) AND META LEARNING (MSSMN WITH RCAN). ALL RESULTS ARE AVERAGED BASED ON FIVE-FOLD CROSS-VALIDATION. MAGNIFICATION FACTORS $(\hat{l}, \hat{m})$ ARE THE SAME AS ACQUISTION FACTORS (l, m). RED AND BLUE INDICATE THE BEST AND THE SECOND BEST PERFORMANCE, RESPECTIVELY.

| Method	Metrics	(2,2)	(2,3)	(2,4)	(3,2)	(3,3)	(3,4)	(4,2)	(4,3)	(4,4)
MSRN	PSNR	29.1191	27.3316	26.3894	26.8260	26.2016	25.7575	25.8892	25.6713	25.3854
	SSIM	0.7566	0.6018	0.4866	0.5996	0.5061	0.4288	0.4964	0.4456	0.3941
RDN	PSNR	29.1173	27.3165	26.4040	26.8226	26.2199	25.7389	25.9015	25.6276	25.3776
	SSIM	0.7569	0.6023	0.4874	0.5967	0.5046	0.4302	0.4952	0.4477	0.3941
RDU	PSNR	28.9336	27.3247	26.4121	26.3479	26.2179	25.7630	25.6872	25.5999	25.4010
	SSIM	0.7526	0.5999	0.4870	0.5796	0.5065	0.4295	0.4827	0.4443	0.3932
RCAN	PSNR	29.1395	27.3534	26.4421	26.9176	26.2668	25.7503	25.9547	25.7287	25.4509
	SSIM	0.7570	0.6024	0.4876	0.6023	0.5082	0.4326	0.4973	0.4486	0.3947
MSSMN	PSNR	29.1789	27.3914	26.4562	27.0109	26.3555	25.8655	26.0577	25.7946	25.4995
	SSIM	0.7578	0.6040	0.4876	0.6031	0.5104	0.4333	0.4993	0.4507	0.3978
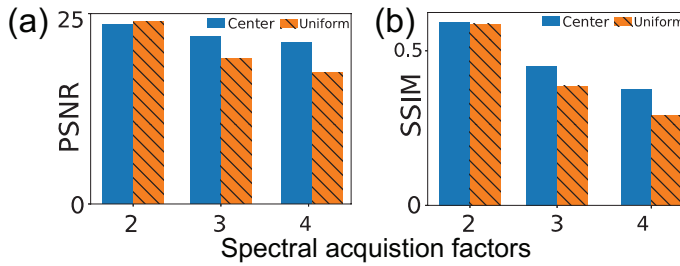
Fig. 5. Results of acquiring OCT images using different strategies: truncating the spectrum from center and dropping spectral data uniformly. (a): PSNR. (b): SSIM.

fixing the spatial acquisition factor m to 1 and varying spectral acquisition factor l. For the uniform acquisition, we drop every $l-1$ points in the A-line spectrum data to achieve an acquisition factor of l as the same sampling strategy in [12]. We calculated PSNR and SSIM scores for OCT images within the whole dataset acquired from both strategies. As shown in Fig. 5, truncating the spectrum from center provides higher PSNR and SSIM scores when $l = 3$ and $l = 4$.

G. Performance comparison among feature extractors and meta-learning

We compare the performance of RCAN as a feature extractor with that of MSRN, RDN and RDU, as shown in Table II. The feature extractors were separately-trained for OCT images of each acquisition factor (l, m). Among the separately-trained models, the RCAN method reports the best PSNR and SSIM scores comparing to other methods for most of acquisition factors. Additionally, the MSRN method reports comparatively high PSNR and SSIM scores against RDN and RDU, whereas the number of trainable parameters in MSRN is significantly less than other methods. The RDU model does not excel in our OCT dataset even if it has more trainable parameters than RCAN. One possible explanation is that variations of U-Net have bottleneck layers, which might restrain the information flow in reconstruction tasks for OCT images [39].

In Table II, we also report performance of meta-learning integrated with the best feature extractor RCAN (referred as MSSMN). To make a fair comparison, we consistently set the

magnification factors $(\hat{l}, \hat{m})$ equal to acquistion factors (l, m) for MSSMN. The MSSMN achieves the highest PSNR and SSIM scores among all acquisition factors. Using the same feature extractor, meta-learning brings an average improvement of 0.0674 in PSNR and an average improvement of 0.0015 in SSIM towards all acquisition factors from RCAN to MSSMN. For larger acquisition factors (less data acquired), the improvements are more significant. Overall, the combination of RCAN and meta-learning framework achieves the best performance for OCT image reconstruction in our dataset.

We observe that the SSIM scores for reconstructed OCT images are lower compared to that of nature images [30], [31]. This is consistent with existing reports on OCT image reconstruction ([12], [13], [14], [16]) due to the comparatively low image quality. The high resolution OCT has lower image quality than high resolution natural images in terms of signal to noise ratio and image contrast. As the high resolution OCT image is with lower image quality than natural image as a reference to calculate SSIM and PSNR, we expect the SSIM and PSNR are lower in OCT-related tasks than those in natural image-related tasks. In addition, our acquisition strategy on both spatial and spectral domain brings extra challenges for MSSMN compared to downsampling OCT data in a single domain.

H. Performance comparison using a unified network on different acquisition factors

We tested the flexibility of using a unified network to reconstruct images when magnification factors $(\hat{l}, \hat{m})$ were different from acquisition factors (l, m). We compared the performance of using MSSMN with other RCAN models that were trained for single magnification factors. The RCAN$(\hat{l}, \hat{m})$ stands for a RCAN model that is trained for magnification factor of $(\hat{l}, \hat{m})$. The output images of RCAN$(\hat{l}, \hat{m})$ were resized to the reference OCT images in order to calculate the PSNR and SSIM scores. The magnification factors $(\hat{l}, \hat{m})$ for MSSMN were equal to acquisition factors (l, m).

As shown in Table III, the MSSMN reports higher PSNR and SSIM scores for resolving OCT images of most acquisition factors. We find that the best performance of RCAN$(\hat{l}, \hat{m})$ generally occurs when ideally $l = \hat{l}$ and $m = \hat{m}$. Also, the second best PSNR and SSIM scores for a RCAN$(\hat{l}, \hat{m})$ model

TABLE III
PSNR AND SSIM RESULTS OF RECONSTRUCTION FROM OCT IMAGES OF ACQUISITION FACTOR (l, m) USING RCAN$(\hat{l}, \hat{m})$. ALL RESULTS ARE AVERAGED BASED ON FIVE-FOLD CROSS-VALIDATION. RED AND BLUE INDICATE THE BEST AND THE SECOND BEST PERFORMANCE, RESPECTIVELY.

Method (l, m)	Metrics	(2,2)	(2,3)	(2,4)	(3,2)	(3,3)	(3,4)	(4,2)	(4,3)	(4,4)
RCAN ($\hat{l}=2, \hat{m}=2$)	PSNR	29.1395	26.7070	25.9195	26.0177	25.2471	24.8725	24.1563	24.2241	24.1444
	SSIM	0.7570	0.5586	0.4510	0.6018	0.4829	0.4100	0.4858	0.4274	0.3794
RCAN ($\hat{l}=2, \hat{m}=3$)	PNSR	28.2387	27.3534	26.1379	25.4245	24.5441	25.1339	23.5747	23.6788	24.3677
	SSIM	0.6848	0.6024	0.4639	0.5404	0.4210	0.4203	0.4388	0.3771	0.3874
RCAN ($\hat{l}=2, \hat{m}=4$)	PNSR	28.1176	25.7341	26.4421	25.1130	24.5077	25.2193	23.4057	23.5806	24.3706
	SSIM	0.6579	0.4659	0.4876	0.5202	0.4099	0.4319	0.4294	0.3692	0.3952
RCAN ($\hat{l}=3, \hat{m}=2$)	PNSR	28.3228	26.2502	25.3176	26.9176	25.7484	25.1228	25.4813	25.2145	24.8207
	SSIM	0.7006	0.5066	0.4033	0.6023	0.4670	0.3879	0.5065	0.4334	0.3740
RCAN ($\hat{l}=3, \hat{m}=3$)	PNSR	27.9758	25.7518	25.8236	26.2365	26.2668	25.4906	24.5388	24.6297	25.0996
	SSIM	0.6534	0.4610	0.4345	0.5472	0.5082	0.4108	0.4591	0.3931	0.3903
RCAN ($\hat{l}=3, \hat{m}=4$)	PNSR	27.6449	25.8364	26.0717	25.5899	25.1103	25.7503	23.8753	24.3177	25.0535
	SSIM	0.6235	0.4552	0.4522	0.5239	0.4189	0.4326	0.4464	0.3864	0.3999
RCAN ($\hat{l}=4, \hat{m}=2$)	PNSR	26.3729	24.9097	24.2386	26.3620	25.0821	24.4809	25.9547	25.0280	24.5449
	SSIM	0.5671	0.4074	0.3392	0.5373	0.4005	0.3390	0.4973	0.3922	0.3400
RCAN ($\hat{l}=4, \hat{m}=3$)	PNSR	27.4361	25.6006	25.4481	26.4774	25.3258	25.3649	25.3811	25.7287	25.1836
	SSIM	0.6086	0.4377	0.4052	0.5318	0.4132	0.3913	0.4617	0.4486	0.3793
RCAN ($\hat{l}=4, \hat{m}=4$)	PNSR	27.3637	25.7250	25.7403	26.1471	25.4230	25.6337	24.6886	25.0076	25.4509
	SSIM	0.5875	0.4350	0.4216	0.5135	0.4120	0.4054	0.4489	0.3901	0.3947
MSSMN ($\hat{l}=l, \hat{m}=m$)	PNSR	29.1789	27.3914	26.4562	27.0109	26.3555	25.8655	26.0577	25.7946	25.4995
	SSIM	0.7578	0.6040	0.4876	0.6031	0.5104	0.4333	0.4993	0.4507	0.3978

TABLE IV
RESULTS OF RECONSTRUCTED OCT IMAGES BY MAGNIFICATION FACTOR $(\lfloor\hat{l}\rfloor, \lfloor\hat{m}\rfloor)$ USING RCAN AND MAGNIFICATION FACTOR $(\hat{l}, \hat{m})$ USING MSSMN. ALL RESULTS ARE AVERAGED BASED ON FIVE-FOLD CROSS-VALIDATION. RED AND BLUE INDICATE THE BEST AND THE SECOND BEST PERFORMANCE, RESPECTIVELY.

Method $(\hat{l}, \hat{m})$	Metrics	(2.5,2.5)	(2.5,3.5)	(2.5,4.5)	(3.5,2.5)	(3.5,3.5)	(3.5,4.5)	(4.5,2.5)	(4.5,3.5)	(4.5,4.5)
RCAN	PSNR	26.3234	25.4072	25.2823	25.6820	25.1737	25.0225	25.3419	25.0736	25.0012
	SSIM	0.5760	0.4441	0.4011	0.4899	0.4034	0.3735	0.4284	0.3797	0.3559
MSSMN	PSNR	27.2603	26.3926	25.8637	26.1884	25.7945	25.4855	25.6679	25.4557	25.2380
	SSIM	0.6123	0.5005	0.4236	0.5103	0.4398	0.3887	0.4469	0.4047	0.3677

are more likely to occur where at least one of the acquisition factors (either in spatial or spectral domain) matches $(\hat{l}, \hat{m})$. For acquisition factor $(4, 4)$, among RCAN$(\hat{l}, \hat{m})$ models, the best PSNR (25.4509) is achieved by RCAN$(4, 4)$, and the best SSIM (0.5095) is achieved by RCAN$(4, 3)$. For acquisition factor $(4, 2)$, among RCAN$(\hat{l}, \hat{m})$ models, the best PSNR (25.9547) is achieved by RCAN$(4, 2)$, and the best SSIM (0.5065) is achieved by RCAN$(3, 2)$. Overall, Table III also supports our argument that the output of RCAN$(\hat{l}, \hat{m})$ models will be deteriorated if applied to OCT images of different acquisition factors (l, m). In contrast, training a unified network, MSSMN, generally achieves better performance than training each RCAN-based model to handle various magnification factors due to the nature of multi-task learning. Reconstructing OCT using multiple magnification factors can be considered as solving multiple tasks. Useful information is exploited from related tasks during multi-task learning, which in turn improves learning procedure for each task [40]. The meta-upscaling and meta-restoration modules in MSSMN provide platforms for multi-task learning, which leads to the improvements in each magnification factor compared to baseline RCAN models trained for specific acquisition factors.

I. Performance comparison on OCT images of untrained acquisition factors

The choice of acquisition factors (l, m) should be adjustable based on needs in coronary analysis. However, it is technically impossible to train the reconstruction framework for OCT images of all possible acquisition factors, especially the non-integer factors. Thus, it is desirable to propose a framework that is robust and able to generalize for untrained magnification factors, especially the factors which are non-integers.

We used the MSSMN to reconstruct OCT images of untrained acquisition factors (l, m), here $l = 2.5, 3.5, 4.5$ and $m = 2.5, 3.5, 4.5$. The floor operation is used to achieve non-integer spatial acquisition factor m. Using floor operation, the A-lines indexes are adjusted while the desired spatial acquisition factor m can be achieved. To compare the performances, we also calculated the PSNR and SSIM scores of reconstruction using RCAN$(\lfloor l \rfloor, \lfloor m \rfloor)$. For example, we used the RCAN model trained for magnification factor $(2, 2)$ to resolve OCT images of acquisition factor $(2.5, 2.5)$ if $l = 2.5$ and $m = 2.5$.

With the meta-upscaling and meta-restoration modules, our MSSMN is capable of magnifying OCT images of untrained acquisition factors. In Table IV, MSSMN method reports higher PSNR and SSIM scores compared to the baseline RCAN$(\lfloor l \rfloor, \lfloor m \rfloor)$ models. The differences of scores between
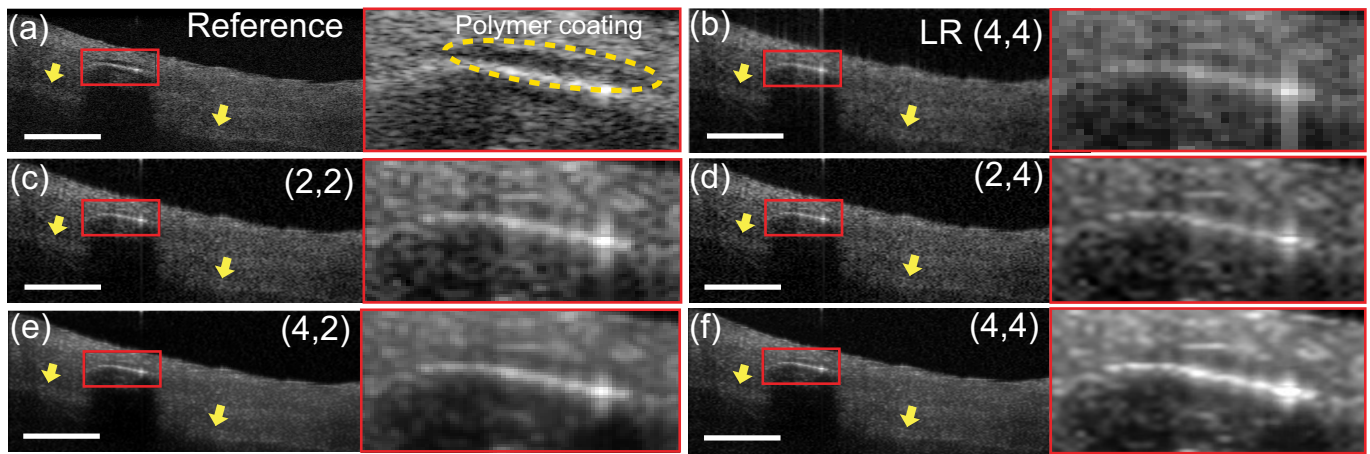
Fig. 6. Multi-scale image reconstruction of drug-eluting stents. (a) The reference OCT image. (b) The LR OCT image is down-scaled with an acquisition factor of $(4, 4)$. (c)-(f) The reconstructed images using different magnification factors $(2, 2)$, $(2, 4)$, $(4, 2)$ and $(4, 4)$. Scale bar: 200 μm. Stent alloy, polymer coating and boundary between media and adventitia can be resolved using different magnification factors.
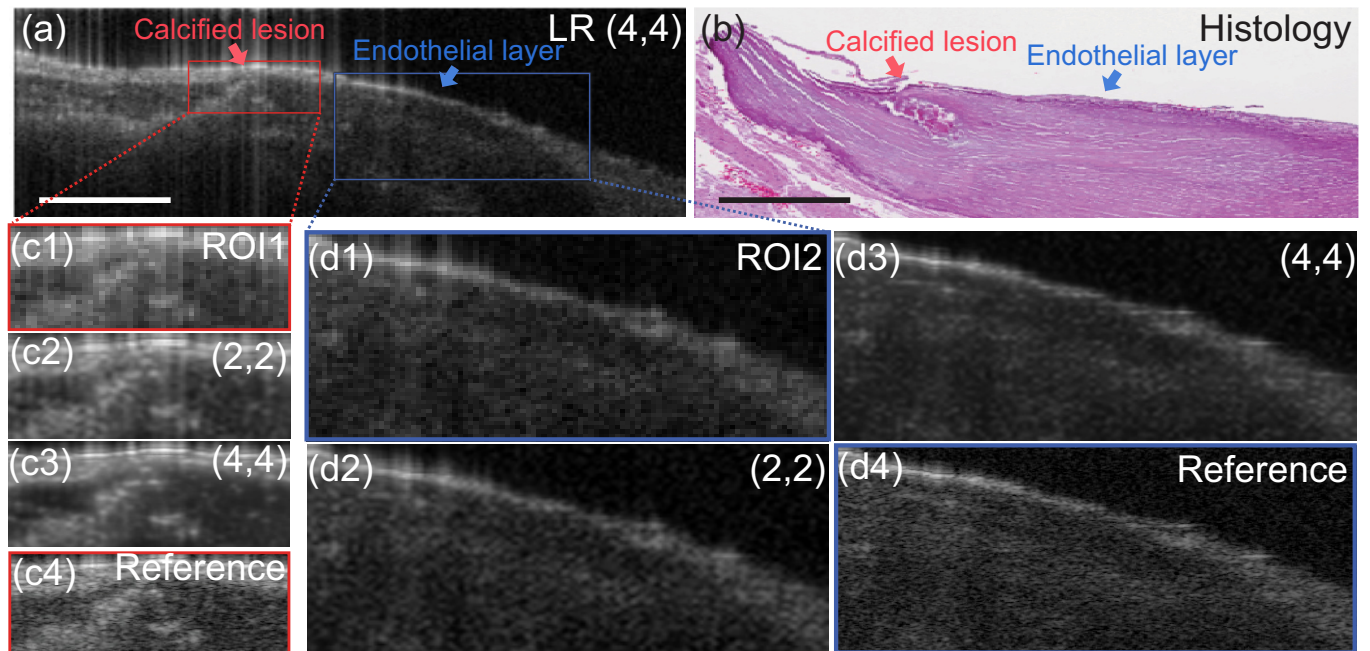
Fig. 7. Multi-scale image reconstruction of two ROIs. (a) The LR OCT image is down-scaled with an acquisition factor of $(4, 4)$. (b) Histology image, (c1-c3,d1-d3) ROIs. (c4,d4) The reference OCT images. Scale bar: 200 μm. ROIs are marked by red and blue rectangles respectively. ROIs are magnified by MSSMN using magnification factors of $(2, 2)$ and $(4, 4)$. For ROI1, a magnification factor of $(2, 2)$ resolves general shape of calcified lesion; For ROI2, a magnification factor of $(4, 4)$ provides better assessment of the thickness of endothelial layer.

RCAN and MSSMN are larger in resolving OCT images of untrained acquisition factors compared to that of trained factors (shown in Table II), which means our framework is generalizable.

J. Visualization of drug-eluting stents in coronary imaging

Imaging microstructures and tissues adjacent to stent struts is crucial in clinic. It is critical provides accurate morphological information of interactions between the stent and vessel wall. With drug-eluting stents (DESs), polymer coatings elute drugs, which prevents intimal proliferation as well as reduces the risk of restenosis after stent implantation. Polymer coating on DESs usually appears as clear rims of material around the metallic struts [7]. After image down-scaling, such information might be compromised in LR OCT images.

In Fig. 6(a), we demonstrate the appearance of polymer coating on a stent captured by OCT. However, the polymer coating and stent alloy are hardly visible in compressed OCT image with a acquisition factor of $(4, 4)$ as shown in Fig. 6(b). We applied our framework to the compressed OCT image, magnifying them by different factors in both spectral and spatial domains. With a magnification factor of $(2, 2)$, the structure of stent alloy was restored as shown in Fig. 6(c). However, polymer coating on the stent was still blurry. Thus, we used a higher magnification factor in spatial domain in order to mitigate the blurring artifacts, as shown in Fig. 6(d).
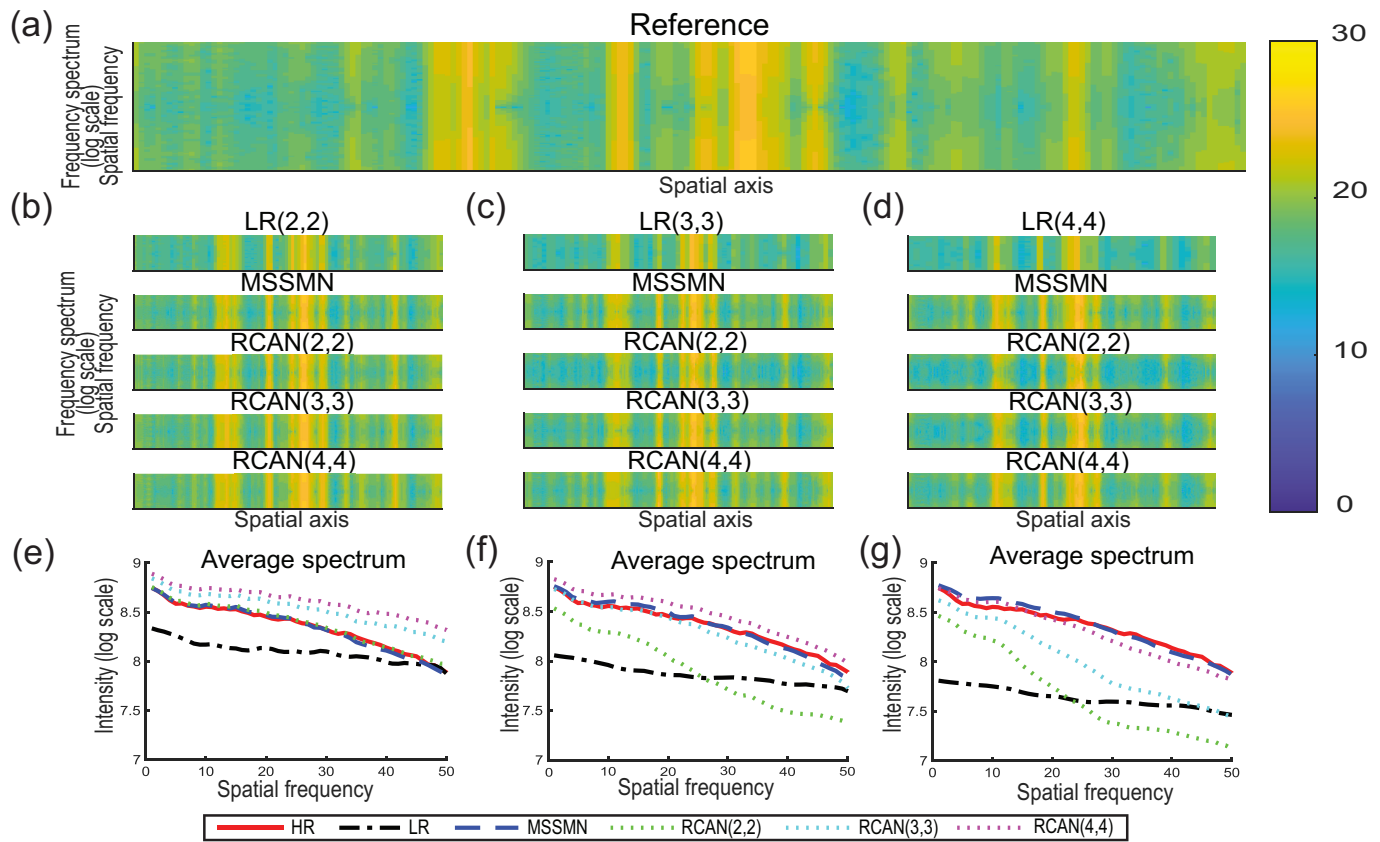
Fig. 8. Spatial frequency analysis of the reference, LR and reconstructed OCT images. (a) Log-scaled spatial frequency spectra of the reference zoomed ROI1 in Fig. 7. (b)-(d) First row: Log-scaled spatial frequency spectra of OCT images compressed by factors $(2,2)$, $(3,3)$ and $(4,4)$. Second row: outcome from MSSMN. Third row to Fifth row: outcome from RCAN$(2,2)$, RCAN$(3,3)$, and RCAN$(4,4)$. (e)-(g) Averaged intensity of spectral profiles in (b)-(d). Our framework, MSSMN, generates OCT images that better match the spectrum of the reference image.

With the magnification factor of $(2,4)$, blurry artifacts were removed thus the polymer coating can be better visualized. In this case, a magnification factor $(2,4)$ should be enough if we are interested in the stent and polymer coating. But for more detailed textures such as boundary between media and adventitia (highlighted by yellow arrows), a higher magnification factor in spatial domain is required. In Fig. 6(e), we used a magnification factor $(4,2)$ to resolve the compressed OCT image, which provided information about stent, polymer coating, and the boundary between media and adventitia. Further with a magnification factor of $(4,4)$, the detailed textures were better recovered with a higher magnification factor in terms of digital resolution as shown in Fig. 6(f).

In Fig. 6, we demonstrated performance of our framework using multiple magnification factors to resolve a compressed OCT image containing important features including stent, polymer coating, and boundary between media and adventitia. The result suggests that resolving compressed OCT images with multiple magnification factors in both spectral and spatial domain provides flexibility which suits structures with various sizes.

K. Visual inspection and assessment with H&E stained histological micrographs

To demonstrate flexibility of our framework in magnifying pathological features, we applied our framework to two

magnified regions of interests (ROIs), ROI1 and ROI2, in an example of low resolution (LR) OCT image acquired using a factor of $(4,4)$.

We applied magnification factors of $(2,2)$ and $(4,4)$ to magnify two ROIs. The visual inspections are shown in Fig. 7. In ROI1, a calcified lesion is highlighted. For ROI1, a magnification factor $(2,2)$ is clear enough to draw the general shape of calcified region. In ROI2, the thickness of endothelial layer needs to be precisely assessed. For ROI2, a magnification factor $(4,4)$ is necessary for an accurate assessment of the thickness of endothelium. In conventional setup, two networks need to be separately trained to magnify two scales. On the contrary, our framework can freely switch between magnification factors using a unified network. We thus provide flexibility for analyzing pathological features within coronary OCT images.

L. Comparison of output spectrum

To further quantify our network inference results against the ground truth images, we performed spatial frequency analysis to the reference OCT image, LR OCT images, reconstruction results from MSSMN and RCAN models. The reference image is a zoomed ROI, the same on as shown in the red box of Fig. 7(a). We magnified each LR image using MSSMN, RCAN$(2,2)$, RCAN$(3,3)$ and RCAN$(4,4)$. The magnification factor of MSSMN is equal to the acquisition factor of LR.

The LR OCT images were acquired using acquisition factors of $(2, 2)$, $(3, 3)$ and $(4, 4)$. We converted the images into the spatial frequency domain using 1D Fourier transform for each A-line (along the vertical axis). Results of spatial frequency analysis are shown in Fig. 8. The results of spatial frequency comparison for each A-line, as shown in Fig. 8(b)-(d), demonstrate that our framework is capable of providing OCT images that closely match the reference in the spectral domain. The detailed quantitative comparison of spectrum are given in Fig. 8(e)-(g). Moreover, the separately trained RCAN models with fixed magnification factor $(\hat{l}, \hat{m})$, if applied to compressed OCT image of different acquisition factors, outcomes diverged spectrum. The spatial frequency analysis supports our argument that the separately-trained RCAN models are less robust and can hardly be applied to compressed OCT images of different factors.

V. Discussion

We propose a new image acquisition method and a reconstruction framework to achieve fast image visualization at flexible magnification factors. We devise a new image acquisition method in spectral-spatial domain for SD-OCT systems, which improves acquisition speed and reduces data storage consumption while maintaining high image quality in reconstruction. To reconstruct the compressed OCT images of different acquisition factors, we develop a multi-scale framework which functions in both spectral and spatial domains. Our framework, MSSMN, achieves better PSNR and SSIM scores comparing to that of separately-trained MSRN, RDN, RDU and RCAN models. Experimental results suggest that separately-trained RCAN models report deteriorated PSNR and SSIM scores if applied to OCT images of different acquisition factors. We confirm that the proposed multi-scale reconstruction framework is beneficial for supporting our acquisition method. Moreover, our framework achieves higher PSNR and SSIM scores when adopting untrained magnification factors comparing to that of RCAN models. Our viable acquisition method and multi-scale reconstruction framework are particularly valuable to the development of catheter based IVOCT systems. The image acquisition speed of SD-OCT systems can be improved without any hardware modifications using our acquisition method, which makes it more suitable to be applied in intravascular applications. Meanwhile, the resolving power of SD-OCT systems is maintained by adopting our multi-scale reconstruction framework. We have demonstrated the flexibility of our multi-scale framework in resolving pathological features.

We compare our work with existing OCT reconstruction works from three aspects: data acquisition strategy, neuron network, and data size. We present a unique spectral-spatial OCT data acquisition method which modifies OCT images in both spectral and spatial domains. According to our experiment, our acquisition method outperforms uniform sampling when the spectral acquisition l factor is larger than 2. Moreover, we have performed extensive reconstruction experiments on four feature extractors for comparison: MSRN, RDN, RDU, and RCAN. We demonstrated RCAN is the optimal selection of

feature extractor for MSSMN. Lastly, our research is carried out on a reasonably large dataset. In terms of deep learning-based OCT super-resolution or enhancement, the training data size used in our work is larger than that of [14] (100 images), [20] (320 images), [21] (1388 images), [22] (128 images), [23] (25 images), [25] (280 images) and [26] (429 images).

There are a few directions to further extend our framework in the future. First, our data acquisition method and multi-scale reconstruction network can be easily adopted in SS-OCT systems. SS-OCT and SD-OCT share similar principle in acquiring spectral data and use Fourier Transform for image reconstruction. Second, the image reconstruction framework can be integrated in a GAN structure. There is a potential to further improve image quality by adding a discriminator in current neural network. Third, we plan to incorporate de-blurring and de-noising modules in MSSMN with special considerations on blurring and noise corruption. Finally, we will extend the study from benchtop OCT data acquisition to catheter-based OCT data acquisition.

VI. Conclusion

In this paper, we propose a spectral-spatial based sampling method to boost efficiency of image acquisition while maintaining high performance and flexibility in reconstruction. We develop a multi-scale reconstruction framework, MSSMN, for compressed OCT images. Our experimental results suggest that the framework outperforms separately-trained networks. We demonstrate that our framework is capable of reconstructing OCT images and magnifying pathological structures with flexible magnification factors as determined by clinical needs. With extensive experiments in human coronary OCT images, we demonstrate high accuracy, data efficiency, and magnification flexibility. Our work may allow for faster SD-OCT inspection with high resolution during coronary intervention.

Acknowledgment

References

[1] E. J. Benjamin *et al.*, "Heart disease and stroke statistics—2019 update: a report from the american heart association," *Circulation*, vol. 139, no. 10, pp. e56–e528, 2019.

[2] A. M. Benjo *et al.*, "High dose statin loading prior to percutaneous coronary intervention decreases cardiovascular events: a meta-analysis of randomized controlled trials," *Catheter. Cardiovasc. Interv*, vol. 85, no. 1, pp. 53–60, 2015.

[3] W. Wijns *et al.*, "Optical coherence tomography imaging during percutaneous coronary intervention impacts physician decision-making: ILUMIEN I study," *Eur. Heart J*, vol. 36, no. 47, pp. 3346–3355, 2015.

[4] N. Meneveau *et al.*, "Optical coherence tomography to optimize results of percutaneous coronary intervention in patients with non–st-elevation acute coronary syndrome: results of the multicenter, randomized doctors study (does optical coherence tomography optimize results of stenting)," *Circulation*, vol. 134, no. 13, pp. 906–917, 2016.

[5] D. A. Jones *et al.*, "Angiography alone versus angiography plus optical coherence tomography to guide percutaneous coronary intervention: outcomes from the pan-london pci cohort," *JACC Cardiovasc. Interv*, vol. 11, no. 14, pp. 1313–1321, 2018.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2022.3175670, IEEE Transactions on Biomedical Engineering

LI *et al.*: BARE DEMO OF IEEETRAN.CLS FOR IEEE JOURNALS
11

[6] G. J. Tearney, I.-K. Jang, and B. Bouma, "Imaging coronary atherosclerosis and vulnerable plaques with optical coherence tomography," in *Optical Coherence Tomography*. New York, NY, USA: Springer International Publishing, 2008, pp. 1083–1101.

[7] L. Liu *et al.*, "Imaging the subcellular structure of human coronary atherosclerosis using micro–optical coherence tomography," *Nat. Med.*, vol. 17, no. 8, pp. 1010–1014, 2011.

[8] L. Fang *et al.*, "Fast acquisition and reconstruction of optical coherence tomography images via sparse representation," *IEEE Trans. Med. Imag.*, vol. 32, no. 11, pp. 2034–2049, 2013.

[9] L. Fang *et al.*, "Segmentation based sparse reconstruction of optical coherence tomography images," *IEEE Trans. Med. Imag.*, vol. 36, no. 2, pp. 407–421, 2017.

[10] A. Abbasi *et al.*, "Optical coherence tomography retinal image reconstruction via nonlocal weighted sparse representation," *J. Biomed. Opt.*, vol. 23, no. 3, pp. 1–11, 2018.

[11] P. G. Daneshmand, A. Mehridehnavi, and H. Rabbani, "Reconstruction of optical coherence tomography images using mixed low rank approximation and second order tensor based total variation method," *IEEE Trans. Med. Imag.*, vol. 40, no. 3, pp. 865–878, 2021.

[12] Y. Zhang *et al.*, "Neural network-based image reconstruction in swept-source optical coherence tomography using undersampled spectral data," *Light. Sci. Appl.*, vol. 10, no. 1, p. 155, 2021.

[13] K. Liang *et al.*, "Resolution enhancement and realistic speckle recovery with generative adversarial modeling of micro-optical coherence tomography," *Biomed. Opt. Express*, vol. 11, no. 12, pp. 7236–7252, 2020.

[14] Z. Yuan *et al.*, "Axial super-resolution study for optical coherence tomography images via deep learning," *IEEE Access*, vol. 8, pp. 204 941–204 950, 2020.

[15] G. J. Tearney *et al.*, "Consensus standards for acquisition, measurement, and reporting of intravascular optical coherence tomography studies: a report from the international working group for intravascular optical coherence tomography standardization and validation," *J. Am. Coll. Cardiol.*, vol. 59, no. 12, pp. 1058–1072, 2012.

[16] A. Lichtenegger *et al.*, "Reconstruction of visible light optical coherence tomography images retrieved from discontinuous spectral data using a conditional generative adversarial network," *Biomed. Opt. Express.*, vol. 12, no. 11, pp. 6780–6795, 2021.

[17] A. Kurniawan, T. W. Purboyo, and A. L. Prasasti, "Implementation of image compression using discrete cosine transform (DCT) and discrete wavelet transform (DWT)," *Int. J. Appl. Eng. Res.*, vol. 12, no. 23, pp. 13 951–13 958, 2017.

[18] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *MICCAI*, 2015, pp. 234–241.

[19] P. Guo, D. Li, and X. Li, "Deep OCT image compression with convolutional neural networks," *Biomed. Opt. Express.*, vol. 11, no. 7, pp. 3543–3554, 2020.

[20] Y. Xu *et al.*, "Improving the resolution of retinal oct with deep learning," in *Medical Image Understanding and Analysis*. New York, NY, USA: Springer International Publishing, 2018, pp. 325–332.

[21] Q. Hao *et al.*, "High signal-to-noise ratio reconstruction of low bit-depth optical coherence tomography using deep learning," *J. Biomed. Opt.*, vol. 25, no. 12, pp. 1–16, 2020.

[22] B. Qiu *et al.*, "N2NSR-OCT: Simultaneous denoising and super-resolution in optical coherence tomography images using semisupervised deep learning," *J. Biophotonics*, vol. 14, no. 1, p. e202000282, 2021.

[23] H. Pan *et al.*, "More realistic low-resolution OCT image generation approach for training deep neural networks," *OSA Contin.*, vol. 3, no. 11, pp. 3197–3205, 2020.

[24] S. Cao *et al.*, "Super-resolution technology to simultaneously improve optical & digital resolution of opticalsuper-resolution technology to simultaneously improve optical & digital resolution of opticalhe coherence tomography via deep learning," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, 2020, pp. 1879–1882.

[25] T. Zhou *et al.*, "Digital resolution enhancement in low transverse sampling optical coherence tomography angiography using deep learning," *OSA Contin.*, vol. 3, no. 6, pp. 1664–1678, 2020.

[26] V. Das, S. Dandapat, and P. Bora, "Unsupervised super-resolution of oct images using generative adversarial network for improved age-related macular degeneration diagnosis," *IEEE Sens. J.*, vol. PP, p. 1, 2020.

[27] W. Yang *et al.*, "Deep learning for single image super-resolution: A brief review," *IEEE Trans. Multimedia.*, vol. 21, no. 12, pp. 3106–3121, 2019.

[28] Y. Zhang *et al.*, "Image super-resolution using very deep residual channel attention networks," in *ECCV*, 2018, pp. 294–310.

[29] X. Wang *et al.*, "ESRGAN: Enhanced super-resolution generative adversarial networks," in *ECCV*, 2019, pp. 63–79.

[30] X. Hu *et al.*, "Meta-SR: A magnification-arbitrary network for super-resolution," in *CVPR*, 2019, pp. 1575–1584.

[31] X. Hu *et al.*, "Meta-USR: A unified super-resolution network for multiple degradation parameters," *IEEE Trans. Neural. Netw. Learn. Syst.*, pp. 1–15, 2020.

[32] H. Zhao *et al.*, "Loss functions for image restoration with neural networks," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 47–57, 2017.

[33] J. Li *et al.*, "Multi-scale residual network for image super-resolution," in *ECCV*, 2018, pp. 527–542.

[34] Y. Zhang *et al.*, "Residual dense network for image super-resolution," in *CVPR*, 2018, pp. 2472–2481.

[35] J. Gurrola-Ramos, O. Dalmau, and T. E. Alarcón, "A residual dense U-Net neural network for image denoising," *IEEE Access*, vol. 9, pp. 31 742–31 754, 2021.

[36] W. Shi *et al.*, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *CVPR*, 2016, pp. 1874–1883.

[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[38] Z. Wang *et al.*, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image. Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[39] S. Lee and I. V. Bajić, "Information flow through U-Nets," in *Proc. IEEE Int. Symp. Biomed. Imaging*, 2021, pp. 812–816.

[40] Y. Zhang and Q. Yang, "A survey on multi-task learning," *IEEE Trans. Knowl. Data. Eng.*, pp. 1–1, 2021.