



A low-cost pathological image digitalization method based on 5 times magnification scanning

Kai Sun^{1^}, Yanhua Gao^{2^}, Ting Xie^{1^}, Xun Wang^{1^}, Qingqing Yang^{1^}, Le Chen^{1^}, Kuansong Wang^{3^}, Gang Yu^{1^}

¹Department of Biomedical Engineering, School of Basic Medical Science, Central South University, Changsha, China; ²Department of Ultrasound, Shaanxi Provincial People's Hospital, Xi'an, China; ³Department of Pathology, School of Basic Medical Science, Central South University, Changsha, China

Contributions: (I) Conception and design: G Yu, K Sun, Y Gao, T Xie; (II) Administrative support: G Yu; (III) Provision of study materials or patients: G Yu, Y Gao, K Wang; (IV) Collection and assembly of data: K Sun, Y Gao, T Xie, X Wang, Q Yang, L Chen; (V) Data analysis and interpretation: G Yu, K Sun, T Xie, X Wang, Q Yang, L Chen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Gang Yu, PhD. School of Basic Medical Science, Central South University, #172 Tongzipo Road, Changsha 410013, China. Email: yugang@mail.csu.edu.cn.

Background: Digital pathology has aroused widespread interest in modern pathology. The key to digitalization is to scan the whole slide image (WSI) at high magnification. The file size of each WSI at 40 times magnification (40×) may range from 1 gigabyte (GB) to 5 GB depending on the size of the specimen, which leads to huge storage capacity, very slow scanning and network exchange, seriously increasing time and storage costs for digital pathology.

Methods: We design a strategy to scan slides with low resolution (LR) (5×), and a superresolution (SR) method is proposed to restore the image details during diagnosis. The method is based on a multiscale generative adversarial network, which can sequentially generate three high-resolution (HR) images: 10×, 20×, and 40×. A dataset consisting of 100,000 pathological images from 10 types of human body systems is used for training and testing. The differences between the generated images and the real images have been extensively evaluated using quantitative evaluation, visual inspection, medical scoring, and diagnosis.

Results: The file size of each 5× WSI is approximately 15 Megabytes. The peak-signal-to-noise ratios (PSNRs) of 10× to 40× generated images are 24.167±3.734 dB, 22.272±4.272 dB, and 20.436±3.845 dB, and the structural similarity (SSIM) index values are 0.845±0.089, 0.680±0.150, and 0.559±0.179, which are better than those of other SR networks and conventional digital zoom methods. Visual inspections show that the generated images have details similar to the real images. Visual scoring average with 0.95 confidence interval from three pathologists are 3.630±1.024, 3.700±1.126, and 3.740±1.095, respectively, and the P value of analysis of variance is 0.367, indicating the pathologists confirm that generated images include sufficient information for diagnosis. The average value of the Kappa test of the diagnoses of paired generated and real images is 0.990, meaning the diagnosis of generated images is highly consistent with that of the real images.

Conclusions: The proposed method can generate high-quality 10×, 20×, 40× images from 5× images, which can effectively reduce the time and storage costs of digitalization up to 1/64 of the previous costs, which shows the potential for clinical applications and is expected to be an alternative digitalization method after large-scale evaluation.

Keywords: Digital pathology; superresolution (SR); low-resolution (LR) scanning; low cost

^ ORCID: Kai Sun, 0000-0002-6232-2094; Yanhua Gao, 0000-0002-7930-8301; Ting Xie, 0000-0001-8742-3855; Xun Wang, 0000-0003-2823-4688; Qingqing Yang, 0000-0001-8277-6854; Le Chen, 0000-0003-3644-4589; Kuansong Wang, 0000-0002-7828-2648; Gang Yu, 0000-0003-3599-8985.

Submitted Jul 24, 2021. Accepted for publication Jan 06, 2022.

doi: 10.21037/qims-21-749

View this article at: <https://dx.doi.org/10.21037/qims-21-749>

Introduction

In pathological examinations, technicians make glass slides of human tissue, such as routine hematoxylin and eosin (HE) (1), frozen sections (2,3), and special staining (4). The pathologist magnifies the slides with a microscope to diagnose the disease. In recent years, the digitalization of pathological slides has begun to be promoted because of convenient storage, network transmission, and effective computer-aided diagnosis (5-7).

A high-resolution (HR) scanner with 20 times magnification (20×) or even a 40× lens is used to scan the glass slide, and the whole slide image (WSI) is obtained (8). Because of multiple magnifications of image, one WSI contains up to 50,000×50,000 pixels (9), and the file size per 40× WSI may range from 1 gigabyte (GB) to 5 GB, depending on the size of the specimen in the slide (10), which significantly increases the storage costs and takes a few minutes for network exchange (11). The development of digital pathology requires the preservation of a large number of WSIs for a long time. Taking China as an example, there are as many as 50 million new pathological examinations every year, which requires 50,000,000 to 250,000,000 GB storage capacity. Moreover, the HR scanning of each WSI needs a few minutes or more (12), so the digitization of the WSIs requires more than 200 years of scanning time. In addition, because the depth of field of the HR lens is very small, the tissue on the slide must be very flat or parts of the WSI may be blurred (13). Therefore, the HR digitalization of pathological glass slides is inefficient and costly.

We assume that low-resolution (LR) scanning, such as 5× magnification, can provide an alternative strategy. Compared with 40× scanning, the speed of 5× scanning increases by 64 times, and the size of WSI reduces to 1/64. Second, the depth of field of the LR lens is larger, which could reduce requirements about the flatness of human tissues. However, the 5× images have lost many details for diagnosis, so a method must be proposed to restore the missing details and generate a series of HR images from 5× images.

The conventional methods of obtaining HR images from LR images are algorithms such as bilinear interpolation and bicubic interpolation (14). However, it is difficult for

them to recover the lost details. Recently, to obtain high-quality images from LR images, superresolution (SR) methods based on deep learning have been widely studied in natural images. There are many excellent SR algorithms, such as deep recursive residual networks (DRRN) (15), SR generative adversarial networks (SRGAN) (16), progressive generative adversarial networks (P-GAN) (17), and multiscale information cross-fusion networks (MSICF) (18). SR methods are also used in medical images such as magnetic resonance (19,20), computer tomography (21), and ultrasound images (22). For example, the SR convolutional neural network (SRCNN) is used for computer tomography images (23), and the pathological SR model is used for cytopathological images (24).

Although SR methods have been successfully applied in many domains, pathological images are different from natural and other medical images. There are a large number of cells and rich microtissue structures in them, but general SR methods pay attention to the larger image texture and structure. Moreover, the maximum magnification of almost all SR methods is 4 times (25), but the magnification of the pathological image is up to 8 times, e.g., from 5× to 40×. For the convenience of diagnosis, a series of HR images should be continuously obtained from 5× images; however, the magnification of almost all SR methods is fixed. We propose a multiscale SR method (MSR), which continuously generates 10×, 20×, and 40× SR images from 5× images. The highlights of the proposed method are as follows:

The size of the 5× scanned image is small enough, approximately 15 Megabytes (MB), and only one-64th of the image file is scanned by 40×. The presented method can generate 10×, 20×, and 40× SR images simultaneously. The optimization is applied on three resolutions, and the generated images are of realistic and rich details, better than the existing methods.

The differences between generated and real images are evaluated based on the pathological diagnosis of 10 types of human body systems, which shows the potential for clinical applications. Therefore, MSR is promising as a low-cost method for pathological digitalization.

We present the following article in accordance with the MDAR checklist (available at <https://qims.amegroups.com/article/view/10.21037/qims-21-749/rc>).

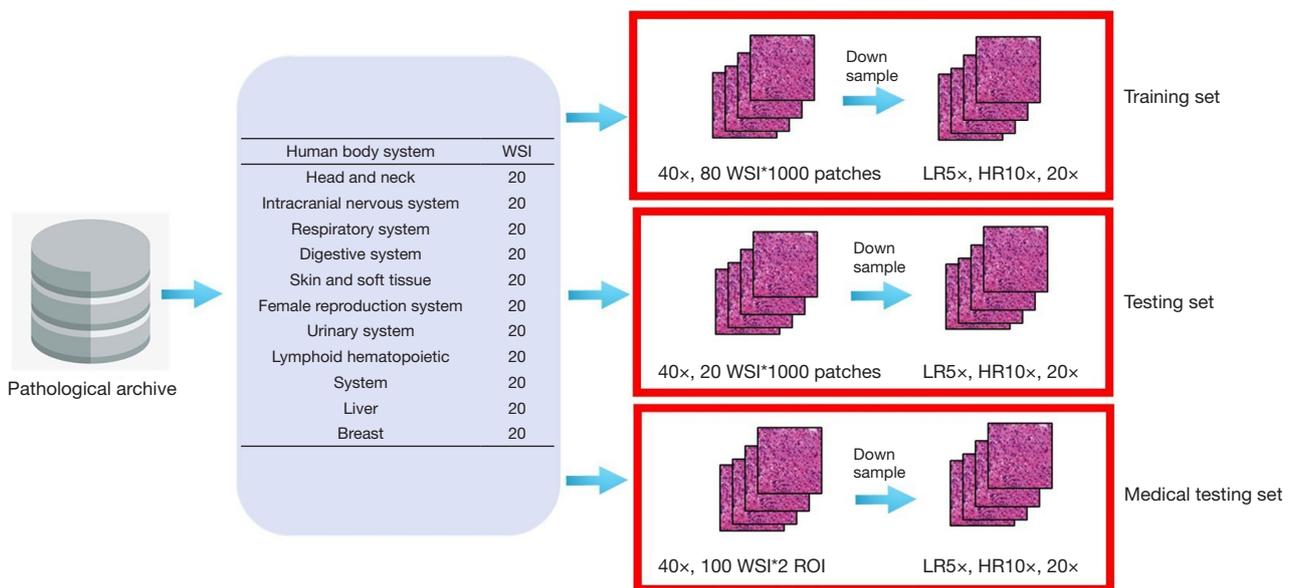


Figure 1 Datasets. A total of 40× WSIs of 200 (20 of each human body system) were randomly selected from the pathological archive. The 100 WSIs are used for the training and testing sets. Another 100 WSIs are used for the medical testing set. The images were made using the FFPE method and stained with HE. WSI, whole slide image; LR, low-resolution; HR, high-resolution; FFPE, formalin-fixed and paraffin-embedded; HE, hematoxylin-eosin; ROI, regions of interest.

Methods

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Institutional Review Board of Xiangya Hospital, Central South University (No. 2019030596), and individual consent for this retrospective analysis was waived.

Dataset collection

Pathological glass slides were obtained from Xiang-ya Hospital of Central South University, which is one of the most famous hospitals in China. The technicians randomly selected 200 whole slides, including 10 types of human body systems, from the pathological archive of the hospital (Table S1). Slides from 10 types of human body systems were selected to ensure the generalization of the dataset because the pathological structures of different systems are different. The number of whole slides of each human body system was set to 20, and the WSIs of some typical diseases were randomly selected by technicians using the pathological archive search tool. The slides were then scanned using a KF-PRO-005 digital pathology scanner (KFBIO company, Ningbo City, China) to obtain WSIs at 40× magnification. After WSIs were obtained, the

technicians no longer participated in the follow-up studies.

Dataset review

Two pathologists with “associate chief physician” titles and at least 10 years in the clinic reviewed the image quality of WSIs to ensure that the WSIs were suitable for pathological diagnosis. If the image quality of WSI was too poor to be suitable for diagnosis, it was excluded. They then independently reviewed the pathological tissue structures of WSIs for the diagnosis of diseases. If a consensus was reached, the WSI was included. No WSI was excluded due to image quality or inconsistent diagnosis. Finally, they discussed selecting two nonoverlapping regions of interest (ROIs) from each WSI for subsequent medical testing. The two ROIs were considered to include sufficient information on disease diagnosis. After reviewing WSIs, the two pathologists would no longer participate in follow-up studies.

Dataset division and preprocessing

The WSIs were randomly divided into training, testing, and extended medical testing sets, as detailed in Figure 1. The training set consists of 80 WSIs, which come from 10

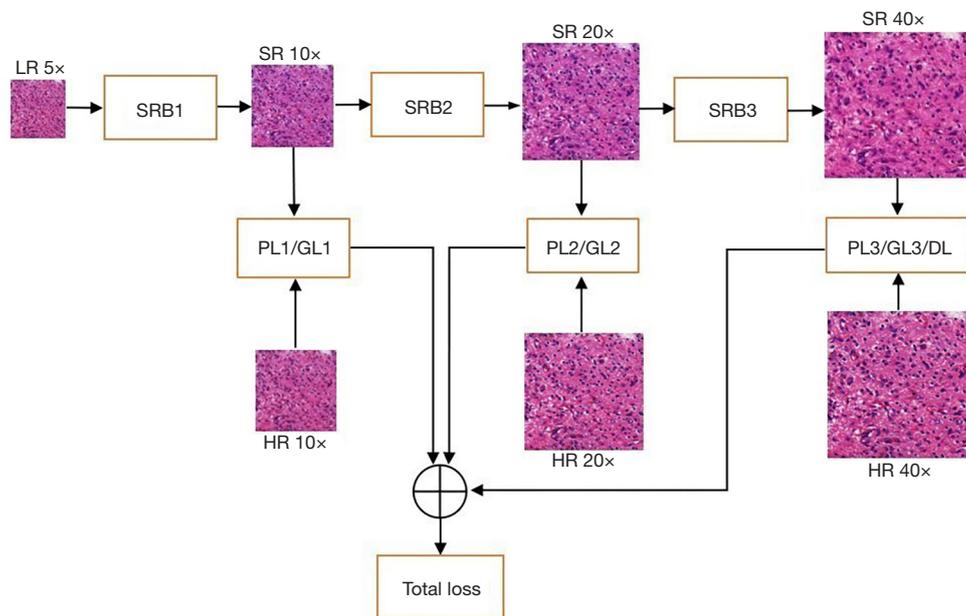


Figure 2 The flowchart of network pipeline. The LR 5 \times images in the training set are input into the network, where the three SRBs are connected together, and each block generates one SR image. The three generated images are compared with real HR images, and the loss is calculated for training the network. The images were made using the FFPE method and stained with HE. LR, low-resolution; SR, superresolution; PL, perceptual loss; GL, generator loss; DL, discriminator loss; HR, high-resolution; SRB, superresolution block; FFPE, formalin-fixed and paraffin-embedded; HE, hematoxylin-eosin.

types of human body systems, and each system has 8 WSIs. The testing and medical testing sets consist of 20 and 100 WSIs, respectively. Because the size of one WSI is too large to be inputted into the convolutional neural network, the 1,000 nonoverlapping images (tiles or patches) from every WSI in the training and testing sets are randomly taken out, and the size of every image is 1,024 \times 1,024 pixels (40 \times). There are a total of 100,000 images in the training and testing sets. Among the 100 WSIs of the medical testing set, two ROIs of each WSI, selected by the two pathologists mentioned above, are used for medical evaluation.

Because the pathological scanner has only two lenses, 20 \times and 40 \times , it is impossible to directly obtain LR images. To obtain LR images from the 40 \times images, the bicubic interpolation algorithm is used, where 3 continuous downsamples from each 40 \times image are used to obtain LR images of 20 \times , 10 \times , and 5 \times (26,27), which are 512 \times 512, 256 \times 256, and 128 \times 128 pixels, respectively. Therefore, the 5 \times –20 \times images in the training and testing sets are obtained, which are treated as the real LR images. Meanwhile, the size of each pixel in the 40 \times image is 0.25 micrometers (KFBIO scanner by KFBIO company, Ningbo City, China), while the pixel sizes of 20 \times , 10 \times , and 5 \times images are 0.5, 1,

and 2 mm, respectively.

The minimum dataset can be found at <https://doi.org/10.6084/m9.figshare.15173634>. The whole dataset can be obtained by contacting the author for research but cannot be used for commercial purposes.

Network architecture

The proposed network comprises three superresolution blocks (SBRs), and each SRB can magnify the input image by 2 times. Therefore, by combining three SRBs, the network can magnify the input image by 2 times, 4 times, and 8 times. The 5 \times LR image is inputted into the network, and the 10 \times , 20 \times , and 40 \times SR images are continuously generated by the three SRBs, as shown in *Figure 2*. The generated images are compared with real images at three resolutions, and the pixel differences (generator loss), feature differences (perceptual loss), and image differences (discriminator loss) are weighted together to guide the training of the network.

The structures of these SRBs are the same, as shown in *Figure 3*, where the input LR image is first convolved by 64 kernels with a size of 3 \times 3, and the 10 repeated basic blocks are concatenated to predict the details in the SR image. The

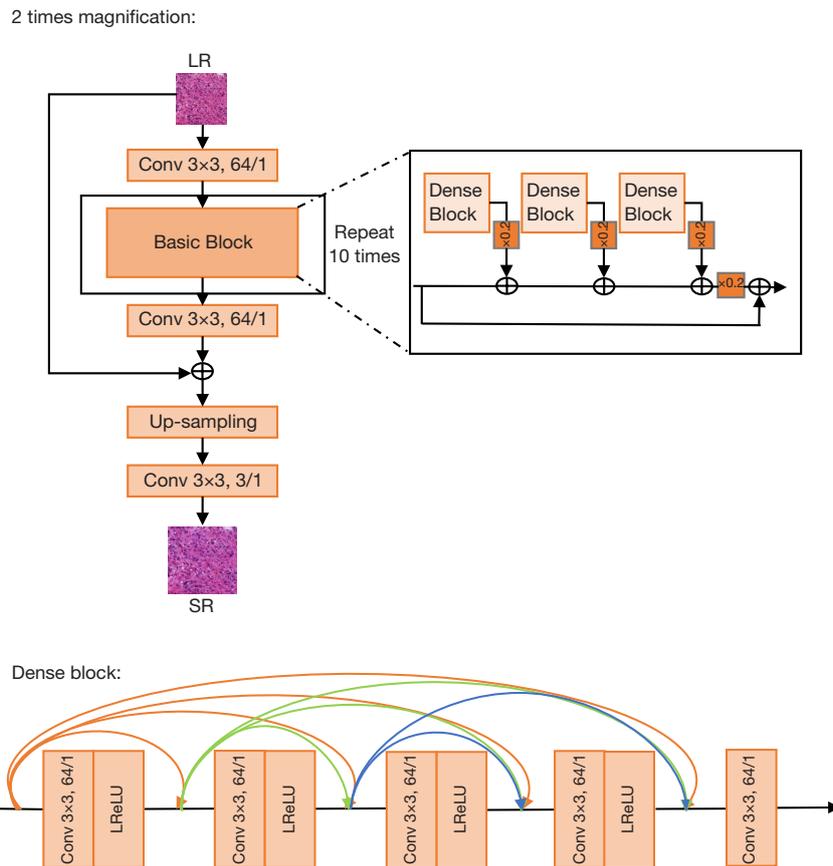


Figure 3 The structure of the superresolution block, where the convolutional layer is denoted as conv, the kernel size, and the number of channels/stride size, LReLU is an activation function, and the upsampling layer is used for pixel shuffling to enlarge the image. The 0.2 is the residual scaling parameter. The images were made using the FFPE method and stained with HE. LR, low-resolution; SR, superresolution; FFPE, formalin-fixed and paraffin-embedded; HE, hematoxylin-eosin.

predicted details are mixed by a convolutional layer and then added to the input LR image. The sum of predicted and original details is upsampled, which is performed by a subpixel layer, where four convolution kernels are learned to increase the number of channels by four times, and the channels are reshaped to three for two upsamplings. Meanwhile, every basic block includes three dense blocks, where five dense convolutional layers are concatenated to find multiscale features from the LR image to reconstruct the lost details.

The loss functions, including the generator, perceptual, and discriminator losses, are based on three levels: pixel, feature, and whole image. The generator loss compares the differences of the pixels one by one of the generated images and real images and is defined as the mean absolute error.

$$GL_{X_j} = \frac{1}{m} \sum_{i=1}^m |SR_{X_j,i} - HR_{X_j,i}| \tag{1}$$

$$X_j = 10X, 20X, 40X$$

where X_j is the magnification and m is the pixel number in the image. The SR is the generated image, and HR is the real image.

The high-level representations of generated SR images and real HR images are extracted, where a network of 19 layers from the Visual Geometry Group (VGG19) (28) is selected as the extractor because of its excellent performance on representation learning. The middle-level (layer 5) and high-level (layer 9) representations are compared here for the perceptual loss of the generated image and real image, defined as Eq. [2].

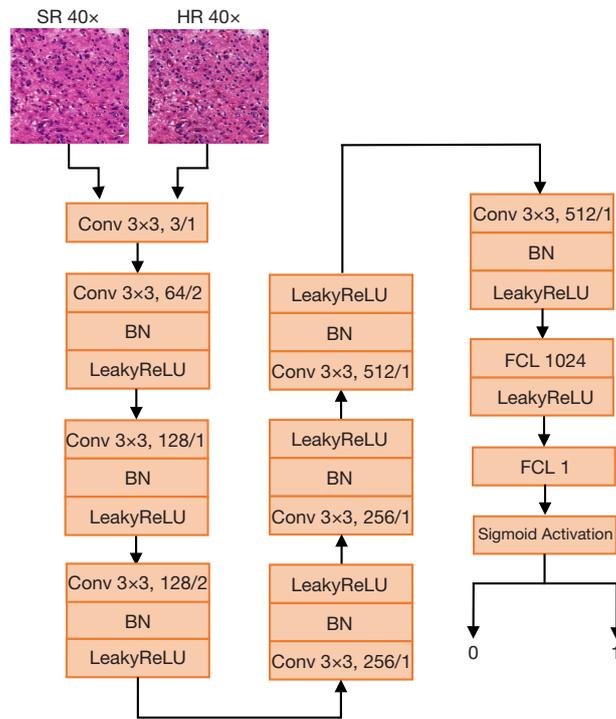


Figure 4 The network framework of the discriminator. The images were made using the FFPE method and stained with HE. SR, super-resolution; HR, high-resolution; Conv, convolutional layer; BN, batch normalization layer; FCL, fully connected layer; FFPE, formalin-fixed and paraffin-embedded; HE, hematoxylin-eosin.

$$\begin{aligned}
 PL_{X_j} = & \text{MSE} \left[\phi_s \left(SR_{X_j} \right) - \phi_s \left(HR_{X_j} \right) \right] \\
 & + \text{MSE} \left[\phi_g \left(SR_{X_j} \right) - \phi_g \left(HR_{X_j} \right) \right]
 \end{aligned}
 \tag{2}$$

where $X_j=10\times,20\times,40\times$, the MSE is the mean square error function, and ϕ is the extractor of VGG19.

To make better generated images that have a high degree of authenticity, the discriminator network is used to identify the true or false values of the real image and the generated image. Because the discriminator network will be trained with the generator, we only use one discriminator to compare the authenticity of the images with the maximum 40x magnification to reduce the training cost. Generally, the greater the magnification is, the more likely the generated image is to be distorted. Therefore, if the quality of the 40x image is good, 20x and 10x should also be good. The discriminator loss is defined as Eq. [3].

$$DL_{40x} = -E \left[\log \left(D \left(HR_{40x} \right) \right) \right] - E \left[\log \left(1 - D \left(SR_{40x} \right) \right) \right] \tag{3}$$

where D is the discriminator network. Because there are

so many details in 40x pathological images, a very deep discriminator network is used, where there are seven convolutional modules and two fully connected layers. Every convolutional module includes some convolution layers with 64–512 convolutional kernels of 3x3 size, and the stride size is 1–2. The sigmoid function is connected to the end, outputting True [1] and False [0], as shown in *Figure 4*.

The two or three losses at each magnification are added and then weighted together for the total loss function.

$$Loss = \sum_{X_j} w_{X_j} Loss_{X_j} \tag{4}$$

where $Loss_{X_j}$ is the sum of generator, perceptual, and discriminator losses at magnification X_j and w_{X_j} is the weighted value.

Network training

The proposed network is implemented using the TensorFlow framework (version 1.15.0) (29) and trained on a server consisting of two graphics processing units (GPUs)

Table 1 Hyperparameters used in our model (Optimizer: Adam)

Hyperparameters	Value
Epochs	120
Steps per epoch	500
Batch size	2
Learning rate	0.0004
Decay factor	0.5
Decay frequency	30
Min flatness	0
Max flatness	0.15
Increase factor	0.01
Increase frequency	5
Generator loss weight	0.06
Perceptual loss weight	0.083
Discriminator loss weight	0.04

of Tesla V100 32 GB (NVIDIA company, Santa Clara, California, USA). The pretrained model of VGG19 is downloaded from the model zoo of TensorFlow and remains fixed during training. The generator and discriminator are initialized randomly and trained alternately. The source code can be found: http://github.com/CSU-BME/pathology_MSR.

Ten percent of the images in the training set (8 WSIs, 8,000 images) are used as the validation set for hyperparameter selection. Different parameter values were tried, such as the optimizer (SGD, Adadelta, Adam), the initial learning rate (0.0005, 0.0004, 0.0001), and the batch size [2, 4, 6]. The optimal parameters with the best performance on the validation set were selected. The hyperparameters are listed in *Table 1*. The Adam optimizer is selected because of its good performance. The initial learning rate is 0.0004 and decays exponentially, where the rate is reduced by 0.5 times every 30 epochs. The total number of training epochs was 120, and 500 steps were performed every epoch. The batch size is fixed to 2, considering the GPU memory, patch size, and training costs. The 50 smaller patches with a size of 64×64 pixels are randomly selected from the input LR 128×128 image, and the small patches whose standard deviations are larger than the flatness value are kept and then input to the proposed network for training. The flatness is the parameter controlling the complexity of the input patches,

and only the patches with complex details should be used as the training data. The flatness increases 0.01 from 0 every 5 epochs, but not more than 0.15. The weights of the three losses are determined after hyperparameter selection, where the weight of the generator loss is 0.06, perceptual loss 0.083, and discriminator loss 0.04.

Experiment settings for quantitative comparison

The quantitative experiments are described here, and the 10×, 20×, and 40× SR images are generated from 5× images in the testing set by the proposed MSR method. While the 5× image is converted to 10×–40× images, the numbers of pixels are the same as the real 10×–40× images. Two quantitative indices are computed to evaluate the generation quality. The first is peak-signal-to-noise-ratios (PSNR), which is defined as the pixel-by-pixel similarity between two images (30). Its definition is as follows.

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad [5]$$

The mean square error (MSE) represents the difference in pixel-by-pixel value between two images, which is defined as Eq. 6.

$$MSE = \sqrt{\frac{\sum_{i=1, j=1}^{i=M, j=N} (SR(i, j) - Real(i, j))^2}{MN}} \quad [6]$$

where SR and Real indicate the generated image and real image, respectively. (i, j) are the coordinates of one pixel on the image, and M and N represent the width and height, in pixels, of the image, respectively. The image intensity ranges from 0 to 255, and the unit of PSNR is decibels (dB). From Eqs. [5,6], larger values of PSNR would indicate more similarity or less difference.

The second is the structural similarity (SSIM) index, which is defined as the similarity in human vision between two images (31). The definition is shown as follows:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad [7]$$

where x, y are two images; μ_x and μ_y represent the mean brightness of the images x, y ; and σ_x , σ_y , and σ_{xy} indicate the two variances and covariance, respectively. C_1 and C_2 are the two known constants. Larger values of SSIM would indicate more similarity or less difference between the two images x, y .

To compare our model and other recent SR models, such as deep back-projection networks (DBPN) (32), efficient

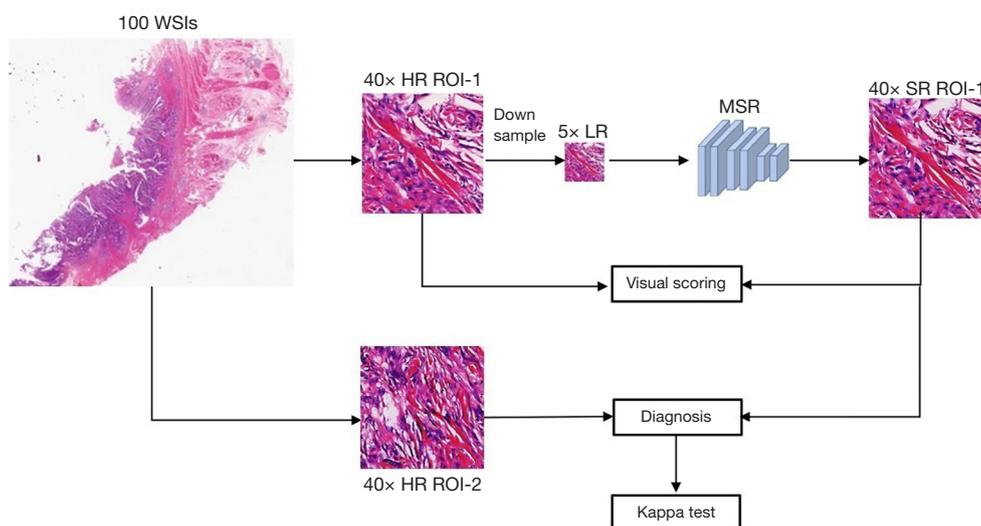


Figure 5 The flowchart of medical evaluation. Two ROIs from every WSI are selected, while ROI-1 is downsampled and then generated by our MSR method. The comparison of SR ROI-1 and HR ROI-1 by pathologist for visual scoring and HR ROI-2 and SR ROI-1 were paired for consistent diagnosis experiments. The images were made using the FFPE method and stained with HE. WSI, whole slide image; HR, high-resolution; ROI, regions of interest; LR, low-resolution; SR, superresolution; MSR, multiscale SR method; FFPE, formalin-fixed and paraffin-embedded; HE, hematoxylin-eosin.

sub-pixel convolutional neural network (ESPCN) (33), enhanced deep superresolution network (EDSR) (34), multiscale deep superresolution system (MDSR) (34), and residual dense network (RDN) (34), the existing models are retrained and retested in our training and testing sets. Because these models support only a fixed magnification, three versions with 2 times, 4 times, and 8 times magnification of each model are trained and tested, for example, DBPN-10 \times , DBPN-20 \times , and DBPN-40 \times , which, respectively realize the generation of 10 \times , 20 \times , and 40 \times images from 5 \times images in the testing set. Similarly, after the PSNR and SSIM of the 10 \times , 20 \times , and 40 \times images generated by each model and the real 10 \times , 20 \times , and 40 \times images in the testing set are computed, all PSNRs and SSIMs are averaged at different resolutions. The codes of these models are downloaded from the links given in their published papers and then retrained at 2 times, 4 times, and 8 times magnification according to the provided hyperparameters in the codes. The RDN code produced some errors when magnifying the 5 \times image by 8 times (34), and the author of the RDN model did not give a solution, so the PSNR and SSIM of RDN at 40 \times images cannot be provided. Two conventional digital zoom methods, BILINEAR (bilinear interpolation) and BICUBIC (bicubic interpolation), are also compared with our model.

Experiment setting for medical diagnosis

Another 100 40 \times WSIs in the medical testing set were used to evaluate the pathological diagnosis by pathologists, as shown in *Figure 5*. Two nonoverlapping ROIs of every WSI were defined as HR ROI-1 and ROI-2, which were selected by the two pathologists mentioned in the dataset review subsection. The two ROIs from the same WSI are paired, and a total of 100 paired ROIs are obtained. We invited six highly experienced pathologists (A-F) who had all been in pathological clinics for at least 10 years to participate in this medical evaluation. None of them had access to the images before this experiment.

The reason why a preselected ROI is diagnosed is to exclude some influencing factors. When we compare the diagnostic difference between real images and generated images, there are three factors to consider: the difficulty of the disease, the pathologist's diagnostic ability, and the difference in images. Among them, the first two differences must be ruled out first, and then the difference of diagnoses can show the difference between the two images. Therefore, we selected a pair of ROIs related to the disease from a WSI. Because these two ROIs come from the same patient, the same doctor diagnoses the same patient twice, so the two factors of the patient and the doctor's ability will not affect the results. Therefore, the significant difference

based on the ROI is only related to the image quality. In contrast, if the doctor is asked to see two different WSIs, one is real and the other is generated, then the differences of the patient and the doctor's ability are included. In future diagnosis based on the newly scanned 5× images, we will start with 5× WSI and generate 10×, 20×, and 40× WSI by the presented SR model. The HR ROI-1 is downsampled to 5× LR and then input into the presented method (MSR) to generate the SR ROI-1. Because each pair of HR ROI-1 and SR ROI-1 comes from the same WSI of the same patient, pathologists should make the same diagnosis, regardless of the diagnostic ability of pathologists.

Results

The comparison of quantitative comparison

The 10×–40× SR images are generated from 20,000 5× images in the testing set and compared with real 10–40× HR images. The PSNR and SSIM distributions of our model, five SR models (DBPN, ESPCN, EDSR, MDSR, RND), and two conventional digital zoom methods (BILINEAR, BICUBIC) at different resolutions are shown in *Figure 6*. Additionally, the PSNR and SSIM of each body system are listed in *Table S2*. The average PSNR with 0.95 confidence interval of our method on 10×, 20×, and 40× are 24.167 ± 3.734 dB, 22.272 ± 4.272 dB, and 20.436 ± 3.845 dB, respectively, and the average with 0.95 confidence interval are 0.845 ± 0.089 , 0.680 ± 0.150 , and 0.559 ± 0.179 .

From *Figure 6*, it can be seen that the SSIM and PSNR on the 10× images of all methods are the highest, which means that the images generated by 2 times magnification are the closest to the real images, and the image quality is the highest. Because 40× images are downsampled to 5× images and most details are lost, the quality of the generated 40× images is relatively poor. In contrast, our method has a higher mean PSNR and SSIM than the other methods at 10×, 20×, and 40× generated images, which shows that our method has a better restoration of image details (P value <0.05).

The comparison of visual inspections

Although the mean PSNR and SSIM can evaluate the similarity in signal and vision between the generated images and the real images, they are still far from human visual perception. In *Figure 7*, the real HR 10×, 20×, and 40× images are displayed in the left column, in which the ROIs are drawn by the green bounding box. The 10×, 20×,

and 40× images generated from 5× ROIs by our model, the other SR models and two conventional digital zoom methods, and real ROI images are shown in the right column. By visually inspecting the generated images, it can be found that for 10×, all the generated images are of high quality and the image details are quite realistic.

Additionally, it is evident that the 20× images from our model are relatively close to the real image, but other methods have different degrees of distortion. Moreover, the 40× images generated by other models appear blurred, while our model can preserve important image structures. It is worth noting that from 5× to 40×, some backgrounds are too small to be completely restored, which is consistent with the PSNR and SSIM.

The images generated by our method and real images of one typical sample are shown in *Figure 8*, which shows that the visual perception of generated and real images is very similar. Pixel-sized scale is also shown on the 40× image.

Visual comparison for medical evaluations

The first comparison is the visual scoring of HR ROI-1 and SR ROI-1. Three pathologists (A–C) carefully observed the two images and then gave a score on the diagnostic consistency of the two ROI images. The scoring rules are shown in *Table 2*, where the visual scoring results are set to five different levels denoted by 1–5 to represent the quality of generated images. A more significant value implies better quality. When pathologists compare any pair of real and generated ROIs, if they think that important information for a correct diagnosis is retained in the generated ROI, the score is 3. For example, the features of cancer cells in cancer images are completely preserved. A total of 100 pairs of HR ROI-1 and SR ROI-1 were scored by pathologists A–C. The distributions of visual scorings from three pathologists are shown in *Figure 9*, and the average score on each body system is shown in *Table S3*. *Figure 9* shows that the three pathologists agreed that all generated ROIs retained diagnostic features, so the ROIs scored at least 3. Regarding the closeness of the generated ROI to the real ROI, pathologists have inconsistent scores of 4–5, which may be due to differences in their subjective perceptions of the restoration of the generated image. The average visual scorings with 0.95 confidence intervals from pathologists A, B, and C were 3.630 ± 1.024 , 3.700 ± 1.126 , and 3.740 ± 1.095 , respectively, indicating that the pathologists believe that there is no visual difference in diagnosis between the generated image and the real image or that the visual

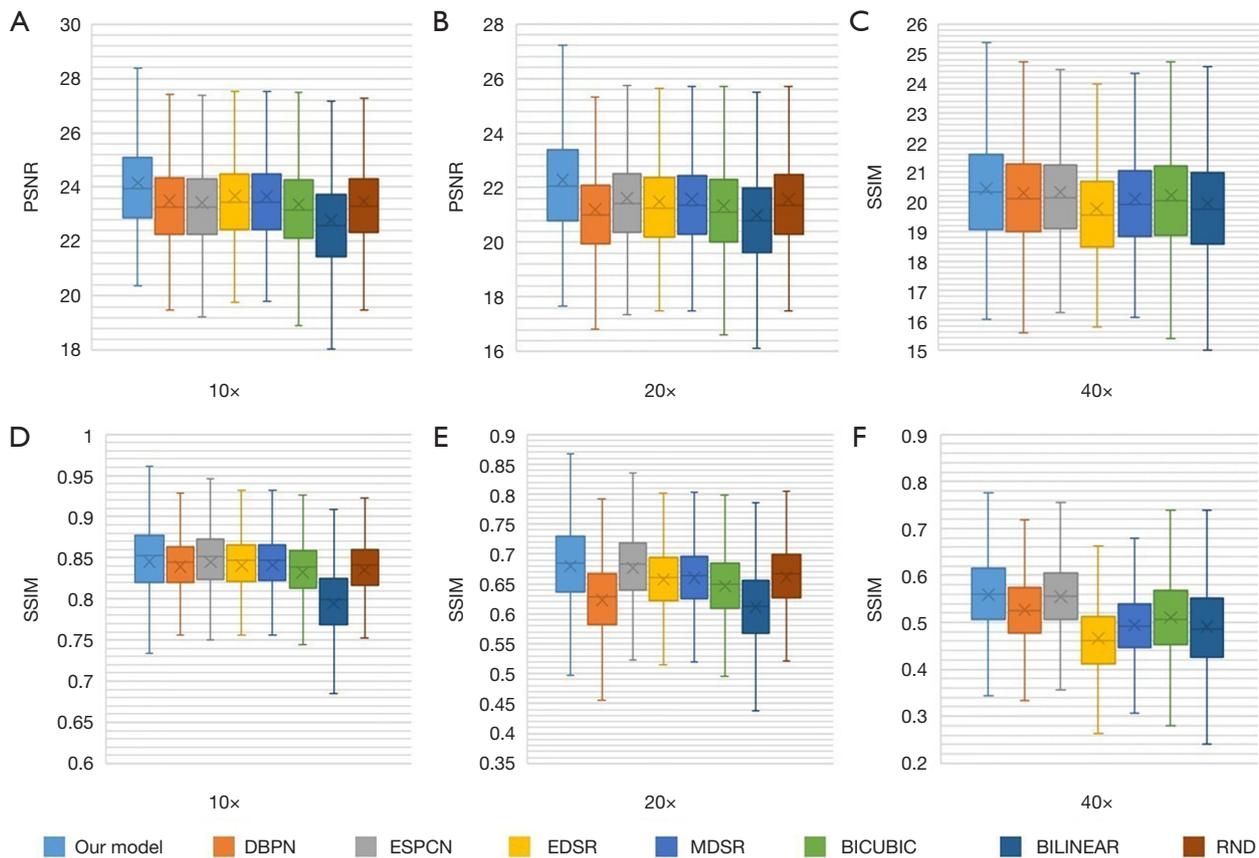


Figure 6 SSIM and PSNR comparison. The PSNR and SSIM distribution of our model, five SR models (DBPN, ESPCN, EDSR, MDSR, RND), and two conventional digital zoom methods (BILINEAR, BICUBIC) at the testing set and two-sided paired T test are performed. The boxes indicate the upper and lower quartile values, and the whiskers indicate the minima and maxima values. The horizontal bar in the box indicates the median, while the cross indicates the mean. (A), (B) and (C) are the distributions of PSNR on 10x, 20x and 40x images, respectively. (A) The average PSNR with 0.95 confidence interval on 10x of our method *vs.* DBPN, ESPCN, EDSR, MDSR, BICUBIC, BILINEAR, RND: 24.167 ± 3.734 *vs.* 23.502 ± 3.895 dB, P value = 1.978×10^{-156} ; 23.435 ± 3.564 dB, P value = 0; 23.668 ± 3.821 dB, P value = 2.054×10^{-89} ; 23.660 ± 3.822 dB, P value = 4.422×10^{-92} ; 23.368 ± 4.034 dB, P value = 3.592×10^{-212} ; 22.785 ± 4.314 dB, P value = 0; 23.486 ± 3.647 dB, P value = 1.041×10^{-169} . (B) 20x: 22.272 ± 4.272 *vs.* 21.191 ± 3.906 dB, P value = 0; 21.609 ± 3.827 dB, P value = 0; 21.495 ± 4.079 dB, P value = 2.019×10^{-174} ; 21.572 ± 4.062 dB, P value = 1.784×10^{-143} ; 21.339 ± 4.263 dB, P value = 2.656×10^{-235} ; 20.990 ± 4.432 dB, P value = 0; 21.584 ± 4.022 dB, P value = 4.409×10^{-140} ; (C) 40x: 20.436 ± 3.845 *vs.* 20.295 ± 4.172 dB, P value = 6.479×10^{-8} ; 20.334 ± 3.772 dB, P value = 5.471×10^{-99} ; 19.777 ± 4.106 dB, P value = 5.600×10^{-139} ; 20.096 ± 3.846 dB, P value = 1.103×10^{-41} ; 20.216 ± 4.264 dB, P value = 9.039×10^{-17} ; 19.938 ± 4.393 dB, P value = 9.573×10^{-76} . (D), (E) and (F) are the distributions of SSIM on 10x, 20x and 40x images, respectively. (D) The average SSIM with a 0.95 confidence interval on 10x of our method *vs.* DBPN, ESPCN, EDSR, MDSR, BICUBIC, BILINEAR, RND: 0.845 ± 0.089 *vs.* 0.840 ± 0.075 , P value = 3.850×10^{-255} ; 0.844 ± 0.084 , P value = 4.116×10^{-8} ; 0.841 ± 0.077 , P value = 1.034×10^{-15} ; 0.841 ± 0.077 , P value = 2.060×10^{-15} ; 0.832 ± 0.083 , P value = 0; 0.794 ± 0.099 , P value = 0; 0.835 ± 0.075 , P value = 2.361×10^{-74} . (E) 20x: 0.680 ± 0.150 *vs.* 0.623 ± 0.137 , P value = 0; 0.677 ± 0.139 , P value = 5.287×10^{-57} ; 0.657 ± 0.133 , P value = 2.048×10^{-126} ; 0.660 ± 0.132 , P value = 1.794×10^{-100} ; 0.646 ± 0.146 , P value = 0; 0.612 ± 0.162 , P value = 0; 0.662 ± 0.155 , P value = 5.523×10^{-83} . (F) 40x: 0.559 ± 0.179 *vs.* 0.526 ± 0.183 , P value = 3.155×10^{-164} ; 0.555 ± 0.181 , P value = 1.832×10^{-4} ; 0.466 ± 0.193 , P value = 0; 0.494 ± 0.177 , P value = 0; 0.510 ± 0.203 , P value = 0; 0.491 ± 0.213 , P value = 0. SSIM, structural similarity; PSNR, peak-signal-to-noise-ratio; DBPN, deep back-projection network; ESPCN, efficient sub-pixel convolutional neural network; EDSR, enhanced deep superresolution network; MDSR, multiscale deep superresolution system; RND, residual dense network.

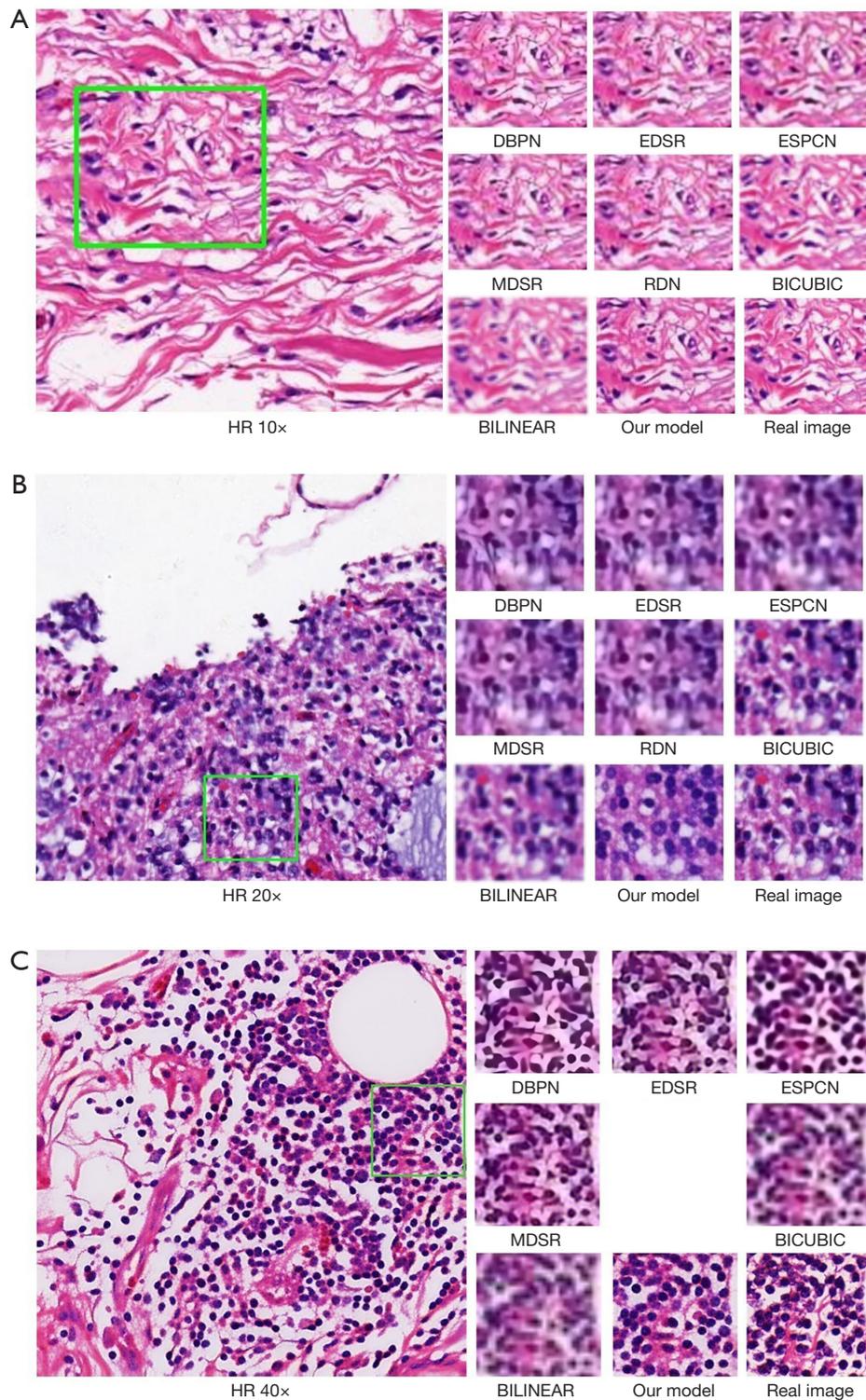


Figure 7 Visual inspections of 10×–40× images generated by different methods. On the far left are the real 10×–40× images, and for shapely layout, 10×, 20×, and 40× images are all adjusted to the same size. The right side of the real 10×–40× images are the same magnification images generated by different methods. The images were made using the FFPE method and stained with HE. HR, high-resolution; DBPN, deep back-projection network; EDSR, enhanced deep superresolution network; ESPCN, efficient sub-pixel convolutional neural network; MDSR, multiscale deep superresolution system; FFPE, formalin-fixed and paraffin-embedded; HE, hematoxylin-eosin.

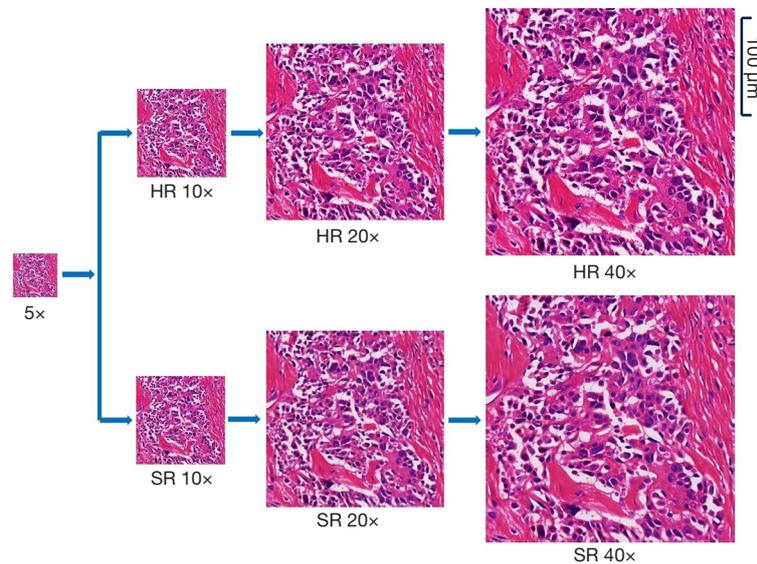


Figure 8 Visual comparison of generated and real images of one typical sample. The 10 \times , 20 \times , and 40 \times images generated by our model and 10 \times , 20 \times , and 40 \times real images are shown. The images were made using the FFPE method and stained with HE. HR, high-resolution; SR, superresolution; FFPE, formalin-fixed and paraffin-embedded; HE, hematoxylin-eosin.

Table 2 Visual scoring rules of generated images by pathologists

Scores	Descriptions
5	The differences between the generated and real images are so small that they are almost invisible
4	There is a slight difference, but it does not affect the diagnosis
3	There is a certain difference, but it does not affect the diagnosis
2	There are noticeable differences, which may affect the diagnosis
1	The difference is large enough not to be used for diagnosis

difference will not affect the diagnosis. The P value of analysis of variance was 0.367, meaning that the scores of the three pathologists were not significantly different.

Diagnosis comparison for medical evaluations

To obtain quantitative comparisons of the generated image and the real image diagnosis, the covariates of the difference between the diagnosis difficulty of the sample and the diagnosis ability of the pathologists should be removed. Therefore, the 100 paired SR ROI-1 and HR ROI-2 are randomly shuffled. Except for the human body system from which the images come, other information, including the WSIs to which the ROIs belong, has been eliminated. Another three pathologists (D-F) blindly reviewed the SR ROI-1 and HR ROI-2 independently and made diagnoses

for every ROI.

The diagnosis comparison is shown in *Figure 10*, where the ROIs of the generated images were misdiagnosed twice and the ROIs of the real images were misdiagnosed once. Compared to the total number of ROIs, the error ratio of the generated and real images is very close. The Kappa test shows that K_D (K value of pathologist D) about the 100 paired SR ROI-1 and HR ROI-2 from pathologist D is 0.990, K_E from pathologist E is 1.000, K_F from pathologist F is 0.980, showing no significant difference in the diagnosis of three pathologists based on the generated images and the real images.

Generation time and diagnosis time

A 5 \times ROI containing 128 \times 128 pixels can be converted

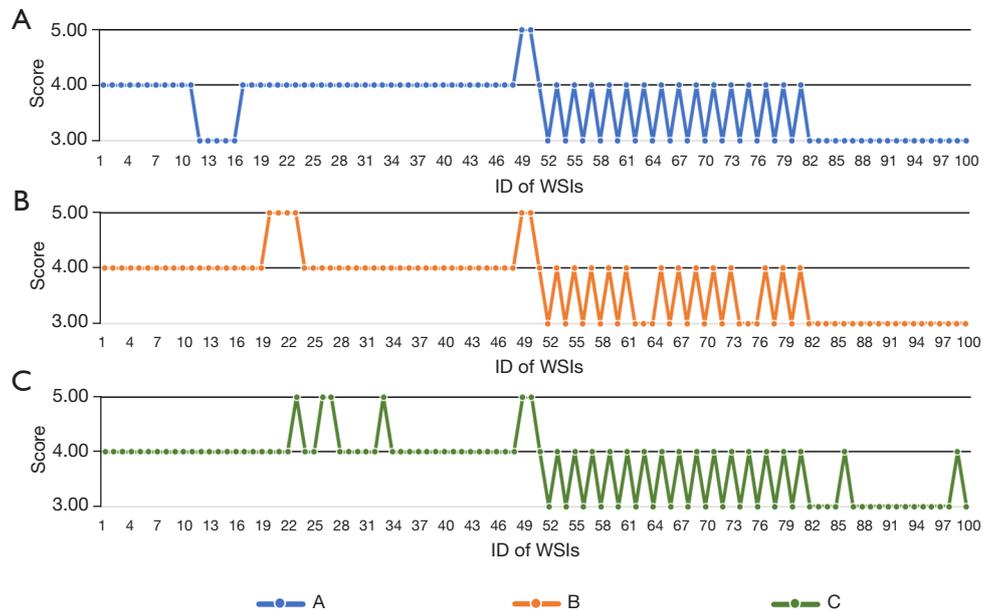


Figure 9 The line figure of visual scoring. (A), (B), and (C) represent the visual score distribution of pathologists (A), (B), and (C) on 100 WSIs, respectively. The average visual scores with 0.95 confidence intervals were 3.630 ± 1.024 (pathologist A), 3.700 ± 1.126 (pathologist B), and 3.740 ± 1.095 (pathologist C). Moreover, the median score of the three pathologists was 4.000. The P value of analysis of variance was 0.367, meaning there was no significant difference among the scores of the three pathologists. WSI, whole slide image.

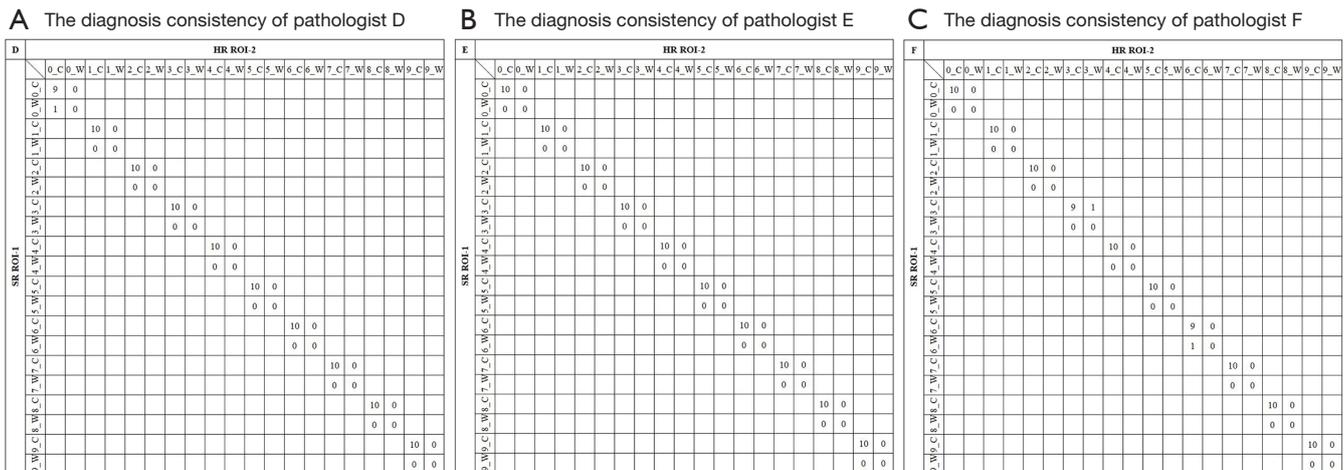


Figure 10 The diagnosis consistency of pathologists D-F. n ($n=0, 1, 2, \dots, 9$) represents the n human body system, which refers to Table S1. C and W represent correct and incorrect diagnoses, respectively. The number in the table represents the number of consistently correct, consistently wrong, inconsistent diagnoses on the n human body systems. HR, high-resolution; SR, superresolution.

to $10\times$, $20\times$, and $40\times$ ROIs in 200 ms using a graphics processing unit of Tesla V100. If the server has 4 GPUs and the WSI has 4,000 ROIs with human tissue, the conversion from a $5\times$ WSI to $10\times$ – $40\times$ WSI would take 3–4 min.

The average time with 0.95 confidence intervals while

diagnosing one generated ROI or one real ROI was compared by a two-sided paired T test. Pathologist A: 5.849 ± 2.200 vs. 5.951 ± 2.349 s, P value = 0.534; pathologist B: 6.497 ± 2.649 vs. 6.269 ± 2.944 s, P value = 0.285; pathologist C: 7.147 ± 3.691 vs. 6.893 ± 2.398 s, P value = 0.248. There

was no significant difference in the diagnosis time between the generated image and the real image.

Discussion

Pathological slides are not conducive to preservation and are inconvenient to search and exchange. With the advent of HR scanners, pathological images from glass slides have been applied in modern pathology. The glass slides are scanned into digital WSIs, which can be conveniently stored and transmitted over the network for remote diagnosis. Because there are many microscopic cells and tissues in pathological slides, HR scanning is necessary to retain these image details. However, the file size of HR scanning is too large, which severely affects the acceptance of digital pathology because of the slow scanning and data exchange and huge storage costs.

We design an alternative strategy for LR scanning, where the 5× images are only approximately 15 MB, which is promising to solve the storage and exchange problems of pathological images. However, LR images, such as 5× images, only show large-scale structural information of human tissue. In fact, pathologists first observe LR images on a large scale and then use a larger magnification to observe the cells on a small scale. In other words, 40× image diagnosis and 5× image diagnosis are based on the information at different scales. The 5× images cannot support a complete pathological diagnosis because of the lack of small-scale information provided by 40× images.

If the structure is larger than 2 microns (one-pixel size in the 5× image), it can be recovered by SR. It is worth noting that structures smaller than 2 microns may be lost, or they cannot be fully restored based on adjacent pixels. The smallest structure concerned in pathological diagnosis is the nucleus of small lymphocytes, whose diameter is approximately 7 microns (35). The size of each pixel in the 40× image is 0.25 microns, so the nucleus of small lymphocytes includes 28×28 pixels. In the 5× image, the nucleus would include 3.5×3.5 pixels. Therefore, the smallest structure is present on the 5× image and should be restored from the 5× image to the 40× image by a good SR method.

Although the existing SR models have been widely used in natural and medical images, the main problem is fixed magnification, and the magnification is not more than 4 times. According to the needs of pathological diagnosis, the SR method must be magnified continuously, and the

maximum magnification can reach 8 times. The proposed method meets the above requirements and can continuously generate 10×, 20×, and 40× images.

We performed paired T tests on SSIM and PSNR between our method and existing methods. PSNR is an index for evaluating the similarity of pixel-by-pixel values of two images, and no image structure, such as edge, region, and texture, is considered. A large PSNR value means that the pixel-by-pixel values of two images are close, but it does not mean that the human visual perceptions of two images are close. SSIM can compare the similarity of image structures; it is closer than PSNR but still different from human perception. In all experiments, the PSNR and SSIM of our method are higher than those of other methods.

Visual comparison is more effective in illustrating the advantages of our method, which is significantly better than those of the existing methods, especially the restoration of cell structure and clusters related to pathological diagnosis. However, it is worth noting that the PSNR and SSIM of all methods are lower than those of the natural images. Although our method can restore local details very well, the color of the generated image is slightly changed. This may be because the method pays attention to microstructures such as cells in the pathological image but not macrostructures such as the color and background of the image. The PSNR and SSIM mainly consider the differences in macrostructures, so the two indicators are not too high.

Because the existing SR models have only fixed magnifications, we trained three versions, i.e., the generations of 5×–10×, 5×–20×, 5×–40×. In theory, a model that considers a single magnification ratio should perform better than a model that generates multiple magnification ratios simultaneously because the learning task is easier. However, our method has achieved better results than those in quantitative evaluation and visual comparison experiments. The possible reason is that the image generated by our method is optimized with the real image at multiple resolutions/magnifications, but other methods only consider the optimization at a single magnification.

For application in pathological domains, whether the generated image can replace the real image, the most important factor is whether it changes the diagnosis result. Therefore, medical confirmation is more important. In the visual assessment, the pathologists believed that although there were visual differences between the generated image

and the real image, the information required for diagnosis was complete.

Our clinical evaluation first chose the most important task of pathological diagnosis for cancer, including multiple cancers involving multiple organ systems. In addition, some inflammation and other benign lesions, such as hippocampal sclerosis, endometrial polyps, renal medullary lipoma, and kidney atrophy, are also included. Because the pathological diagnosis is based on human tissue structure and cells on the image, the process for cancer and other diseases is similar. If SR can provide the information of cells and human tissue structure needed for these diseases, we have reason to believe that the information needed for other disease diagnoses should also be generated. The blind diagnosis from 10 body systems confirmed that there was no significant difference between the diagnosis of generated images and real images in the medical testing set.

The development of digital pathology requires the preservation of a large number of whole slide images for a long time. Whether these images are 5× or 40× is related to storage capacity and copying or network transmission. Our research suggests permanently keeping 5× images instead of 40× images to reduce storage and transmission costs. The practical application of pathological images can be roughly divided into two common scenarios, among which the images to be generated are well known. First, for new patients in clinical practice, we can generate 40× images from new scanned 5× images in advance for pathologists to diagnose. After diagnosis, the 40× images are then deleted. Second, for scientific research issues, researchers have designed clear experimental settings, looking for suitable slides (5× images) from pathological archives, quickly transmitting the 5× images to the local server via the network, and then generating 40× images for research. When the image to be generated cannot be known in advance, although this is not common, it is a wonderful alternative that generates in real-time the 40× region of interest that the pathologist is observing because the generation time of a region is only approximately 200 ms.

Quantitative and visual experiments have confirmed the excellent image quality generated by the proposed method, and medical experiments on some specific diseases preliminarily confirmed that the generated images have the potential for clinical pathological diagnosis. The main limitation of this study is that there were not enough types of diseases in the test set. However, it shows the possibility of LR digitization, and it is worthwhile to further evaluate

its feasibility on a wide range of datasets in further work.

Conclusions

Digital pathology can easily store and exchange WSIs, but the HR scan will produce large image files, high storage costs, and slow network transmission. The proposed multiscale method based on LR scanning can generate SR images with sufficient information for pathological diagnosis. The low-cost digital method was first proposed for pathology, showing potential for clinical application. In future work, on a large-scale image dataset, the diagnostic difference between the generated and real images should be further evaluated.

Acknowledgments

Funding: This work was supported by the Emergency Management Science and Technology Project of Hunan Province (#2020YJ004, #2021-QYC-10050-26366 to G Yu).

Footnote

Reporting Checklist: The authors have completed the MDAR checklist. Available at <https://qims.amegroups.com/article/view/10.21037/qims-21-749/rc>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://qims.amegroups.com/article/view/10.21037/qims-21-749/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Institutional Review Board of Xiangya Hospital, Central South University (No. 2019030596), and individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the

original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Wittekind D. Traditional staining for routine diagnostic pathology including the role of tannic acid. 1. Value and limitations of the hematoxylin-eosin stain. *Biotech Histochem* 2003;78:261-70.
2. Arcega RS, Woo JS, Xu H. Performing and Cutting Frozen Sections. *Methods Mol Biol* 2019;1897:279-88.
3. Švajdler M, Švajdler P. Frozen section: history, indications, contraindications and quality assurance. *Cesk Patol Summer*;54:58-62.
4. Alturkistani HA, Tashkandi FM, Mohammedsaleh ZM. Histological Stains: A Literature Review and Case Study. *Glob J Health Sci* 2015;8:72-9.
5. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol* 2019;20:e253-61.
6. Pallua JD, Brunner A, Zelger B, Schirmer M, Haybaeck J. The future of pathology is digital. *Pathol Res Pract* 2020;216:153040.
7. Bueno G, Fernández-Carrobles MM, Deniz O, García-Rojo M. New Trends of Emerging Technologies in Digital Pathology. *Pathobiology* 2016;83:61-9.
8. Aeffner F, Adissu HA, Boyle MC, Cardiff RD, Hagendorn E, Hoenerhoff MJ, Klopffleisch R, Newbigging S, Schaudien D, Turner O, Wilson K. Digital Microscopy, Image Analysis, and Virtual Slide Repository. *ILAR J* 2018;59:66-79.
9. Tizhoosh HR, Pantanowitz L. Artificial Intelligence and Digital Pathology: Challenges and Opportunities. *J Pathol Inform* 2018;9:38.
10. Hanna MG, Reuter VE, Hameed MR, Tan LK, Chiang S, Sigel C, Hollmann T, Giri D, Sambay J, Moradel C, Rosado A, Otilano JR 3rd, England C, Corsale L, Stamelos E, Yagi Y, Schüffler PJ, Fuchs T, Klimstra DS, Sirintrapun SJ. Whole slide imaging equivalency and efficiency study: experience at a large academic center. *Mod Pathol* 2019;32:916-28.
11. Mukherjee L, Bui HD, Keikhosravi A, Loeffler A, Eliceiri K. Super-resolution recurrent convolutional neural networks for learning with multi-resolution whole slide images. *J Biomed Opt* 2019;24:1-15.
12. Neil DA, Demetris AJ. Digital pathology services in acute surgical situations. *Br J Surg* 2014;101:1185-6.
13. Liu S, Hua H. Extended depth-of-field microscopic imaging with a variable focus microscope objective. *Opt Express* 2011;19:353-62.
14. Xia P, Tahara T, Kakue T, Awatsuji Y, Nishio K, Ura S, Kubota T, Matoba O. Performance comparison of bilinear interpolation, bicubic interpolation, and B-spline interpolation in parallel phase-shifting digital holography. *Opt Rev* 2013;20:193-7.
15. Ying T, Jian Y, Liu X. Image Super-Resolution via Deep Recursive Residual Network. *IEEE Conference on Computer Vision & Pattern Recognition, IEEE Computer Society, 2017:2790-8.*
16. Wu Y, Yang F, Huang J, Liu Y. Super-resolution construction of intravascular ultrasound images using generative adversarial networks. *Nan Fang Yi Ke Da Xue Xue Bao* 2019;39:82-7.
17. Mahapatra D, Bozorgtabar B, Garnavi R. Image super-resolution using progressive generative adversarial networks for medical image analysis. *Comput Med Imaging Graph* 2019;71:30-9.
18. Hu Y, Gao X, Li J, Huang Y, Wang H. Single image super-resolution with multi-scale information cross-fusion network. *Signal Process* 2021;179:107831.
19. Chaudhari AS, Fang Z, Kogan F, Wood J, Stevens KJ, Gibbons EK, Lee JH, Gold GE, Hargreaves BA. Super-resolution musculoskeletal MRI using deep learning. *Magn Reson Med* 2018;80:2139-54.
20. Zhao C, Shao M, Carass A, Li H, Dewey BE, Ellingsen LM, Woo J, Guttman MA, Blitz AM, Stone M, Calabresi PA, Halperin H, Prince JL. Applications of a deep learning method for anti-aliasing and super-resolution in MRI. *Magn Reson Imaging* 2019;64:132-41.
21. Oyama A, Kumagai S, Arai N, Takata T, Saikawa Y, Shiraishi K, Kobayashi T, Kotoku J. Image quality improvement in cone-beam CT using the super-resolution technique. *J Radiat Res* 2018;59:501-10.
22. Christensen-Jeffries K, Couture O, Dayton PA, Eldar YC, Hynynen K, Kiessling F, O'Reilly M, Pinton GF, Schmitz G, Tang MX, Tanter M, van Sloun RJG. Super-resolution Ultrasound Imaging. *Ultrasound Med Biol* 2020;46:865-91.
23. Umehara K, Ota J, Ishida T. Application of Super-Resolution Convolutional Neural Network for Enhancing Image Resolution in Chest CT. *J Digit Imaging* 2018;31:441-50.
24. Ma J, Yu J, Liu S, Chen L, Li X, Feng J, Chen Z, Zeng S, Liu X, Cheng S. PathSRGAN: Multi-Supervised Super-Resolution for Cytopathological Images Using Generative Adversarial Network. *IEEE Trans Med Imaging*

- 2020;39:2920-30.
25. Anwar S, Khan S, Barnes N. A Deep Journey into Super-resolution: A survey. arXiv e-prints 2019; Preprint Available online: <https://arxiv.org/abs/1904.07523>
 26. Han D. Comparison of Commonly Used Image Interpolation Methods. Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering 2013.
 27. Karim SAA, Zulkifli NAB, Shafie AB, Sarfraz M, Ghaffar A, Nisar KS. Medical image zooming by using rational bicubic ball function. IGI Global 2020:146-61.
 28. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv e-prints 2014; Preprint Available online: <https://arxiv.org/abs/1409.1556>
 29. Inc G. Tensorflow. Version 1.15.0[software]. 2019; Available online: <https://pypi.org/project/tensorflow>
 30. Horé A, Ziou D. Image quality metrics: PSNR vs. SSIM. 2010 20th International Conference on Pattern Recognition, 2010:23-6.
 31. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004;13:600-12.
 32. Haris M, Shakhnarovich G, Ukita N. Deep Back-Projection Networks For Super-Resolution. IEEE Conference on Computer Vision and Pattern Recognition 2018:1664-73.
 33. Shi W, Caballero J, Huszár F, Totz J, Aitken AP, Bishop R, Rueckert D, Wang Z. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016:1874-183.
 34. Lim B, Son S, Kim H, Nah S, Lee KM. Enhanced Deep Residual Networks for Single Image Super-Resolution. IEEE Conference on Computer Vision and Pattern Recognition Workshops 2017:136-44.
 35. Kumar V, Cotran RS, Robbins SL. Robbins Basic Pathology. Saunders, 2003.

Cite this article as: Sun K, Gao Y, Xie T, Wang X, Yang Q, Chen L, Wang K, Yu G. A low-cost pathological image digitalization method based on 5 times magnification scanning. *Quant Imaging Med Surg* 2022;12(5):2813-2829. doi: 10.21037/qims-21-749

Table S1 Types of diseases in each system

ID	Human body systems	Types of disease	Numbers
0	Head and neck	1. Pleomorphic adenoma of salivary gland	3
		2. Parotid gland cellular schwannoma	3
		3. Papillary thyroid carcinoma	5
		4. Tongue squamous cell carcinoma	2
		5. Laryngeal squamous cell carcinoma	2
		6. Nodular thyroid cancer	5
			Total =20
1	Intracranial nervous system	1. Anaplastic oligodendroglioma	2
		2. Anaplastic astrocytoma	3
		3. Intracranial hair cell astrocytoma	3
		4. Intracranial meningioma	7
		5. Hippocampal sclerosis	5
			Total =20
2	Respiratory system	1. Large cell neuroendocrine carcinoma of the lung	6
		2. Lung adenocarcinoma	2
		3. Lymph node metastatic lung adenocarcinoma	5
		4. Lung adenocarcinoma in situ	4
		5. Keratinizing squamous cell carcinoma	1
		6. Anterior mediastinal thymoma	2
			Total =20
3	Digestive system	1. Omentum Malignant Mesothelioma	4
		2. Peritoneal mucous adenocarcinoma	2
		3. Small intestine stromal tumor	3
		4. Sigmoid colon adenocarcinoma	2
		5. Colon adenocarcinoma	4
		6. Rectal adenocarcinoma	5
			Total =20
4	Skin and soft tissue	1. Poorly differentiated adenocarcinoma of eyelid	3
		2. Highly differentiated myxoid liposarcoma of the buttocks	10
		3. Foot malignant melanoma	2
		4. Embryonic rhabdomyosarcoma	5
			Total =20

Table S1 (continued)

Table S1 (continued)

ID	Human body systems	Types of disease	Numbers
5	Female reproduction system	1.Uterine Leiomyoma	1
		2.High-grade serous carcinoma of the uterus	4
		3.Cervical endometrioid adenocarcinoma	2
		4.Cervical adenocarcinoma in situ	2
		5.Placenta adhesion and implantation	3
		6.Uterine pregnancy	1
		7.Cervical polyps	2
		8.Endometrial polyps	2
		9.Pelvic papillary serous carcinoma	3
		Total =20	
6	Urinary system	1.Adrenal sebaceous adenoma	4
		2.Retroperitoneal leiomyosarcoma	2
		3.Renal clear cell carcinoma	2
		4.Reninoma	3
		5.Kidney atrophy	4
		6.Renal medullary lipoma	5
		Total =20	
7	Lymphoid hematopoietic system	1.Reactive hyperplasia of peri-intestinal lymph nodes	5
		2.Reactive hyperplasia of inguinal lymph nodes	4
		3.Reactive hyperplasia of perigastric lymph nodes	5
		4.Ocular-MALT lymphoma	6
		Total =20	
8	Liver	1.Hepatocellular carcinoma	9
		2.Nodular hyperplasia of liver	11
		Total =20	
9	Breast	1.Invasive breast cancer	Total =20

Table S2 The distribution of PSNR and SSIM on each body system.

			0	1	2	3	4	5	6	7	8	9
Our Model	10×	SSIM	0.837±0.072	0.813±0.132	0.857±0.060	0.850±0.064	0.859±0.081	0.836±0.084	0.844±0.069	0.884±0.061	0.816±0.079	0.855±0.051
		PSNR	24.783±3.812	23.657±2.508	24.235±3.525	24.716±3.007	24.895±6.118	25.332±2.957	23.444±2.998	22.461±2.112	23.467±4.300	24.684±2.662
	20×	SSIM	0.680±0.121	0.619±0.198	0.694±0.116	0.694±0.103	0.705±0.162	0.686±0.110	0.662±0.127	0.731±0.101	0.624±0.165	0.703±0.093
		PSNR	23.011±4.061	21.491±2.975	22.184±3.853	22.996±3.288	23.211±6.777	23.623±3.267	21.351±3.471	20.477±2.243	21.336±5.317	23.045±2.988
	40×	SSIM	0.577±0.163	0.492±0.209	0.585±0.158	0.581±0.134	0.584±0.200	0.592±0.133	0.528±0.149	0.581±0.109	0.483±0.203	0.588±0.120
		PSNR	21.209±3.656	19.760±2.965	20.210±3.613	21.191±3.142	21.254±5.325	21.753±2.917	19.557±3.442	18.685±2.195	19.424±4.470	21.319±2.819
DBPN	10×	SSIM	0.832±0.065	0.816±0.118	0.848±0.050	0.845±0.054	0.855±0.079	0.837±0.049	0.837±0.059	0.876±0.053	0.810±0.074	0.841±0.040
		PSNR	24.064±3.712	23.310±2.598	23.515±3.563	23.961±3.078	24.300±6.573	24.752±3.031	22.718±3.094	21.842±2.060	22.787±5.077	23.767±2.655
	20×	SSIM	0.614±0.109	0.582±0.180	0.638±0.102	0.632±0.100	0.652±0.152	0.627±0.113	0.606±0.109	0.679±0.094	0.563±0.146	0.637±0.074
		PSNR	21.792±3.615	20.895±2.602	21.073±3.592	21.733±3.037	22.066±6.557	22.436±2.888	20.404±3.173	19.402±2.155	20.416±5.018	21.691±2.765
	40×	SSIM	0.551±0.167	0.469±0.203	0.545±0.159	0.551±0.132	0.563±0.239	0.568±0.128	0.490±0.144	0.526±0.092	0.451±0.227	0.550±0.113
		PSNR	20.992±3.847	20.027±2.753	19.995±3.944	20.926±3.233	21.223±6.867	21.608±3.047	19.468±3.507	18.201±2.453	19.610±5.428	20.900±2.878
ESPCN	10×	SSIM	0.842±0.069	0.814±0.133	0.849±0.060	0.856±0.063	0.862±0.084	0.848±0.062	0.840±0.064	0.866±0.054	0.812±0.086	0.856±0.046
		PSNR	24.073±3.464	23.227±2.532	23.362±3.542	23.973±2.934	24.095±5.561	24.666±2.903	22.771±3.117	21.490±2.378	22.879±4.132	23.814±2.491
	20×	SSIM	0.682±0.118	0.626±0.183	0.689±0.107	0.694±0.098	0.706±0.164	0.691±0.100	0.658±0.110	0.711±0.083	0.619±0.162	0.695±0.081
		PSNR	22.273±3.644	21.258±2.600	21.475±3.717	22.206±3.038	22.382±6.177	22.867±2.987	20.877±3.230	19.713±2.289	20.960±4.708	22.081±2.608
	40×	SSIM	0.575±0.159	0.492±0.208	0.573±0.154	0.580±0.131	0.589±0.224	0.591±0.129	0.521±0.146	0.568±0.098	0.476±0.221	0.581±0.115
		PSNR	20.987±3.559	19.973±2.524	20.103±3.565	20.932±2.966	21.151±6.135	21.521±2.831	19.568±3.180	18.603±2.159	19.610±4.668	20.891±2.644
EDSR	10×	SSIM	0.835±0.066	0.813±0.123	0.848±0.051	0.849±0.055	0.856±0.080	0.839±0.052	0.838±0.059	0.869±0.052	0.813±0.077	0.847±0.040
		PSNR	24.268±3.680	23.414±2.485	23.585±3.555	24.168±2.951	24.437±6.525	24.860±2.991	22.951±3.034	21.817±2.203	23.126±4.870	24.051±2.568
	20×	SSIM	0.659±0.118	0.611±0.172	0.670±0.102	0.670±0.090	0.687±0.163	0.670±0.090	0.640±0.105	0.698±0.079	0.600±0.158	0.667±0.073
		PSNR	22.108±3.798	21.173±2.640	21.345±3.746	22.038±3.125	22.389±6.954	22.770±3.104	20.723±3.229	19.656±2.192	20.809±5.409	21.944±2.713
	40×	SSIM	0.485±0.190	0.415±0.197	0.496±0.167	0.482±0.145	0.503±0.250	0.515±0.145	0.431±0.153	0.482±0.087	0.381±0.236	0.471±0.119
		PSNR	20.375±3.801	19.544±2.728	19.611±3.769	20.280±3.192	20.702±6.896	21.089±3.142	18.974±3.296	18.056±2.115	18.983±5.473	20.154±2.712
MDSR	10×	SSIM	0.835±0.066	0.813±0.123	0.848±0.051	0.849±0.055	0.856±0.080	0.839±0.053	0.838±0.060	0.869±0.053	0.812±0.077	0.847±0.040
		PSNR	24.258±3.682	23.406±2.485	23.578±3.549	24.160±2.950	24.428±6.532	24.850±2.995	22.943±3.030	21.816±2.187	23.117±4.876	24.045±2.568
	20×	SSIM	0.662±0.116	0.613±0.172	0.672±0.100	0.673±0.089	0.689±0.162	0.672±0.089	0.643±0.104	0.700±0.078	0.605±0.156	0.670±0.072
		PSNR	22.189±3.777	21.231±2.622	21.406±3.751	22.131±3.102	22.457±6.909	22.839±3.074	20.799±3.222	19.714±2.213	20.900±5.371	22.053±2.719
	40×	SSIM	0.511±0.163	0.441±0.194	0.515±0.152	0.513±0.130	0.530±0.223	0.536±0.127	0.459±0.140	0.508±0.086	0.414±0.222	0.512±0.111
		PSNR	20.703±3.631	19.825±2.598	19.842±3.654	20.673±3.087	20.945±6.081	21.361±2.927	19.303±3.275	18.301±2.173	19.313±4.809	20.690±2.763
BICUBIC	10×	SSIM	0.832±0.071	0.802±0.131	0.841±0.058	0.843±0.062	0.850±0.087	0.838±0.058	0.828±0.061	0.847±0.049	0.800±0.088	0.841±0.044
		PSNR	24.052±3.853	23.178±2.707	23.247±3.841	23.893±3.176	24.170±6.812	24.634±3.200	22.671±3.286	21.162±2.614	22.912±4.997	23.761±2.780
	20×	SSIM	0.659±0.130	0.595±0.187	0.660±0.119	0.667±0.105	0.677±0.184	0.670±0.102	0.625±0.113	0.658±0.076	0.588±0.177	0.662±0.085
		PSNR	22.075±3.941	21.043±2.791	21.109±4.051	21.946±3.294	22.233±7.204	22.679±3.254	20.593±3.489	19.053±2.649	20.812±5.465	21.851±2.888
	40×	SSIM	0.546±0.185	0.449±0.219	0.529±0.183	0.542±0.151	0.549±0.263	0.564±0.144	0.472±0.164	0.482±0.110	0.432±0.249	0.540±0.130
		PSNR	20.970±3.927	19.948±2.784	19.835±4.105	20.870±3.320	21.141±7.016	21.553±3.145	19.427±3.615	17.912±2.631	19.632±5.412	20.875±2.983

Table S2 (continued)

Table S2 (continued)

			0	1	2	3	4	5	6	7	8	9
BILINEAR	10×	SSIM	0.799±0.086	0.760±0.148	0.807±0.078	0.810±0.077	0.815±0.108	0.810±0.073	0.786±0.073	0.794±0.060	0.754±0.112	0.806±0.057
		PSNR	23.563±4.048	22.612±2.954	22.607±4.147	23.320±3.417	23.628±7.239	24.161±3.501	22.098±3.554	20.239±2.955	22.423±5.265	23.200±2.992
	20×	SSIM	0.631±0.147	0.559±0.197	0.628±0.139	0.637±0.120	0.645±0.207	0.647±0.116	0.586±0.125	0.606±0.086	0.549±0.198	0.632±0.098
		PSNR	21.783±4.054	20.720±2.917	20.711±4.250	21.602±3.441	21.900±7.469	22.405±3.434	20.255±3.657	18.468±2.870	20.541±5.623	21.519±3.022
	40×	SSIM	0.531±0.196	0.429±0.224	0.510±0.196	0.525±0.161	0.530±0.278	0.550±0.153	0.450±0.173	0.449±0.121	0.411±0.260	0.522±0.138
		PSNR	20.736±4.017	19.693±2.877	19.523±4.262	20.594±3.432	20.872±7.220	21.336±3.283	19.159±3.745	17.443±2.801	19.420±5.528	20.606±3.087
RND	10×	SSIM	0.830±0.066	0.809±0.121	0.841±0.051	0.843±0.055	0.851±0.081	0.835±0.051	0.833±0.058	0.861±0.051	0.808±0.076	0.841±0.039
		PSNR	24.110±3.571	23.293±2.447	23.358±3.519	23.986±2.897	24.163±5.841	24.682±2.948	22.814±3.047	21.563±2.260	23.027±4.664	23.864±2.499
	20×	SSIM	0.664±0.115	0.614±0.172	0.674±0.100	0.675±0.089	0.691±0.162	0.673±0.089	0.645±0.104	0.701±0.077	0.608±0.156	0.673±0.072
		PSNR	22.212±3.759	21.239±2.618	21.420±3.746	22.150±3.086	22.447±6.770	22.844±3.060	20.818±3.219	19.709±2.237	20.929±5.292	22.070±2.699

The bold indicates that the SSIM or PSNR value of the method on the generated and real images (10×, 20× or 40×) of one body system is the maximum among all methods. 0–10 represent the IDs of ten body systems, which are listed in *Table S1*.

Table S3 The visual score distribution of pathologists A, B, and C on each body system

	Head and neck	Intracranial nervous system	Respiratory system	Digestive system	Skin and soft tissue	Female reproduction system	Urinary system	Lymphoid hematopoietic system	Liver	Breast
Pathologist A	3.5±0.980	3.5±0.980	3.7±0.898	3.7±1.255	3.7±0.898	3.8±1.176	3.7±0.898	3.7±0.898	3.4±0.960	3.6±0.960
Pathologist B	3.7±1.255	3.6±0.960	3.7±1.255	3.6±0.960	3.8±1.176	3.7±0.898	3.7±1.255	3.9±1.055	3.7±0.898	3.6±1.300
Pathologist C	3.8±1.176	3.6±0.960	3.8±1.176	3.7±0.898	3.9±1.055	4.0±0.877	3.6±0.960	3.9±1.372	3.6±0.960	3.5±0.980