

Article

Not peer-reviewed version

RDAGAN: A Residual Dense Attention GAN for Microscopic Image Super-Resolution

[Sanya Liu](#), Xiao Weng, Xingen Gao, [Xiaoxin Xu](#), [Lin Zhou](#)*

Posted Date: 8 May 2024

doi: 10.20944/preprints202405.0461.v1

Keywords: single image super-resolution; microscopic image; generative adversarial network; image processing



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

RDAGAN: A Residual Dense Attention GAN for Microscopic Image Super-Resolution

Sanya Liu ^{1,†}, Xiao Weng ^{1,†}, Xingen Gao ², Xiaoxin Xu ³ and Lin Zhou ^{1,*}

¹ Xiamen Key Laboratory of Mobile Multimedia Communications, College of Information Science and Engineering, Huaqiao University, Xiamen 361021, China

² School of Opto-Electronic and Communication Engineering, Xiamen University of Technology, Xiamen 361024, China

³ Institute of Microelectronics Chinese Academy of Sciences, Beijing 100029, China

* Correspondence: linzhou@hqu.edu.cn

† These authors contributed equally to this work.

Abstract: With the development of deep learning, the Super-Resolution (SR) reconstruction of microscopic images has improved significantly. However, the scarcity of microscopic images for training, the underutilization of hierarchical features in original Low Resolution (LR) images, and the high-frequency noise unrelated with the image structure generated during reconstruction process are still challenges in the Single Image Super-Resolution (SISR) field. Faced with these issues, we first collected sufficient microscopic images through Motic, a company engaged in the design and production of optical and digital microscopes, to establish a dataset. Secondly, we proposed a Residual Dense Attention Generative Adversarial Network (RDAGAN). The network comprises a generator, an image discriminator, and a feature discriminator. The generator includes a Residual Dense Block (RDB) and a Convolutional Block Attention Module (CBAM), focusing on extracting the hierarchical features of the original LR image. Simultaneously, the added feature discriminator enables the network to generate high-frequency features pertinent to the image's structure. Finally, we conducted experimental analysis and compared our model with six classic models. Compared with the best model, our model improved PSNR and SSIM by about 1.5dB and 0.2, respectively.

Keywords: single image super-resolution; microscopic image; generative adversarial network; image processing

1. Introduction

Regardless of the magnification and numerical aperture of the objective lens used, the imaging throughput of current microscopes is typically only on the order of ten megapixels[1]. This leads to a compromise between high resolution and Field of View (FOV) when imaging. However, in biomedical research such as histopathology, hematology and neuroscience, there is a growing need for high-resolution imaging of large samples. In order to accurately resolve cellular-level life activities at the scale of the entire sample, a balance between global structure and microscale local details is required, along with quantitative analysis. Faced with this challenge, Super-Resolution (SR) imaging techniques were developed. Single Image Super-Resolution (SISR) reconstruction [2] aims to reconstruct a High-Resolution (HR) image from an input Low Resolution (LR) image. This technique is widely used in critical fields, including bright field micrographs[3], fluorescent imaging[4,5], remote sensing images[6,7], and surveillance videos[8]. Traditional SISR reconstruction algorithms can be divided into three categories: interpolation-based[9,10], reconstruction-based[11,12] and learning-based[13]. Interpolation-based algorithms have the advantages of simple principle and easy implementation. However, these methods are difficult to recover the detailed information of the image, and the reconstructed image is seriously distorted. Reconstruction-based algorithms first generate the constraints for SR reconstruction according to the imaging process of LR images, establish the corresponding mathematical model, and finally reconstruct HR images. However, when

the magnification is high, it is still difficult to recover enough high-frequency information. The learning-based reconstruction algorithm mainly establishes the mapping relationship between LR and HR hyperspectral images, and then, reconstructs according to the mapping relationship. In recent years, with the rapid development of deep learning, the excellent learning ability of convolutional neural networks also brings new opportunities for the development of SR imaging technology.

Specifically, Dong et al.[14] used deep learning in image SR reconstruction by introducing Convolutional Neural Networks (CNN). This groundbreaking work paved the way for subsequent innovations. Kim et al.[15] introduced the Very Deep Super Resolution (VDSR) model, which leverages a deeply constructed neural network combined with residual learning for image reconstruction. This approach effectively addresses the issue of image size reduction caused by successive convolutions. Lai et al.[16] proposed the Laplacian Pyramid Super-Resolution Network (LapSRN) model, which employs a Laplacian pyramid to progressively reconstruct the image and perform feature extraction through a residual structure. Ledig et al.[17] advanced the field by incorporating residual blocks within the structure of a Generative Adversarial Network (GAN) to create the Super-Resolution Generative Adversarial Network (SRGAN) model. The use of GANs significantly improves the visual perception quality of the generated images, making them more closely resemble real images. Lim et al.[18] developed the Enhanced Deep Super-Resolution (EDSR) model, which refines the residual structure found in the SRGAN model for improved performance. Zhang et al.[19] contributed by introducing the Residual Dense Block (RDB) structure and the Residual Dense Network (RDN) model, further advancing the capabilities of SR reconstruction.

In terms of microscopic image reconstruction, Zhang et al.[20] proposed Registration-Free GAN Microscopy (RFGANM) workflow by combining SRGAN network with optical microscope and degradation model to achieve deep learning SR in large FOV, and improve the resolution of wide-field microscopy and light-sheet fluorescence microscopy images. Wang et al.[21] used a GAN network to develop SR techniques for cross-modal fluorescence microimaging by simulating the mapping relationships between different imaging techniques, such as wide-field fluorescence imaging, confocal to STED, and TIRF to TIRF-SIM, through deep learning. Van Sloun[22] applied deep learning to ultrasound microscopy SR study and proposed Deep-ULM model, which is based on U-Net network to reduce the effect of diffraction and obtain HR images in real time. Li et al.[23] achieved high-quality reconstruction of ordinary wide-field fluorescence images to SIM SR imaging results using a Deep Fourier Channel Attention (DCFA) network, greatly contributes to the development of super-resolution in microscopic images.

It can be seen that although great progress has been made in the SR of microscopic images, most of them are focused on fluorescence images in a special way. We know that the features of fluorescence microscopy images are significantly different from those of cell images. Additionally, despite the above SR methods have achieved good results, there are some shortcomings in any of them. These include the absence of specialized datasets for microscopic images, which hampers model training; the inadequate use of hierarchical features in LR images by GAN-based SR models; insufficient utilization of information across convolutional layers due to varying receptive fields; a lack of prioritization in reconstructing feature map information, leading to suboptimal focus on critical details; and the reliance on pixel-level reconstruction errors as loss functions, which fails to capture high-frequency details and often results in overly smooth and potentially inaccurate images.

In order to response these challenges, this paper has carried out corresponding optimization work on cell microscopy images. The specific contributions are described as follows:

- (1) We produce a dataset of high and low resolution images of four cell types using microscope acquisition. You can find the dataset at [here](#).
- (2) In order to fully utilize the hierarchical features of the original image, we propose a Residual Dense Attention Generative Adversarial Network (RDAGAN), whose generator uses RDABs with increased attention mechanisms.

- (3) To reconstruct the high-frequency features associated with HR images, we add a feature discriminator to the original discriminator and optimize the loss function.
- (4) Our proposed optimized model is compared with six classic models. Compared to the best performing model among them, PSNR and SSIM have improved by 1.5 db and 0.2, respectively.

2. Proposed Methodology

Our model is based on the aforementioned RFGANM[20] that have been used for microscopic cell image SR, constructing a generator network architecture composed of four main sections: the feature extraction part, the residual dense attention part, the dense feature fusion part, and the reconstruction part. The input LR image first passes through the feature extraction section, where initial features are extracted. These features are then processed by the residual dense attention section. Here, the RDB extracts rich local features through densely connected convolutional layers, fully leveraging the hierarchical features of the layers. The local feature fusion within the RDB allows for an adaptive and more effective integration of previous and current local features, thus stabilizing the training of broader network. An attention mechanism is added after each RDB, first applying a channel attention module to obtain a weighted result, followed by a spatial attention module to refine the weighting. This enables the network to concentrate on information that is most beneficial for image reconstruction. The RDB and the attention module together constitute the Residual Dense Attention Block (RDAB). After extracting dense local features through a series of residual dense attention modules, dense feature fusion is proposed to mine multi-level features from a global perspective. Additionally, a feature discriminator[24] is introduced to differentiate the detailed features of SR and HR images, encouraging the reconstructed image to generate more high-frequency features instead of noise.

2.1. Microscopic Cell Image Dataset

We used a microscope to capture cell images using both 40× and 20× objective lenses simultaneously to establish our dataset. Due to factors such as lens aberration and autofocus issues, the collected images were subject to center cropping. HR images were cropped to a resolution of 1024×1024 pixels, while LR images were cropped to 512×512 pixels, reflecting the 2× objective relationship. In constructing the dataset, we faced challenges in precisely matching HR and LR images. To address this, we employed Bicubic Downsampling (BD) to generate corresponding LR images from their HR counterparts. These LR images were then used for both reconstruction purposes and performance comparison. Finally, we created 800 pairs of images for training data and 200 pairs of image val data for reconstruction for each of four cell images.

2.2. Residual Dense Blocks for Fusion CBAM

The structure of our generator network is shown in Figure 1, consisting of four parts: a shallow feature extraction network, a RDAB, a Dense Feature Fusion (DFF) and an upsampling network. The input and output of the generator network are represented by I_{LR} and respectively I_{SR} . A convolutional layer is used to extract shallow features F_0 of the LR image and perform global residual learning, while F_0 also serves as an input to the RDAB. F_0 denotes the convolution operation on the LR image. Assuming that there are D RDABs in the generator, the Dth RDAB can be represented as:

$$F_d = H_{RDAB,d}(F_{d-1}) = H_{RDAB,d}(H_{RDAB,d-1}(\cdots(H_{RDAB,1}(F_0))\cdots)) \quad (1)$$

where $H_{RDAB,d}$ denotes the operation of the dth RDAB and $H_{RDAB,d}$ is a composite function of operations such as convolution and Rectified Linear Unit(ReLU). Since F_d is generated by the dth RDAB fully utilizing each convolutional layer within the block, F_D can be considered as a local feature. The generator RDABs in this paper are 16 in total.

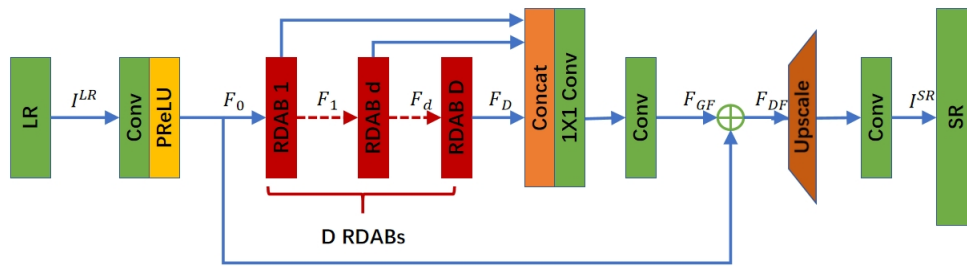


Figure 1. Residual Dense Attention Generative Adversarial Network (RDAGAN) framework

After extracting the hierarchical features with a set of D RDABs, further DFF is performed, including Global Feature Fusion (GFF) and Global Residual Learning (GRL). DFF fully utilizes the features of all previous layers, F_{DF} is the output feature mapping of the composite function H_{DFF} of the DFF module, which can be expressed as:

$$F_{DF} = H_{DFF}(F_0, F_1, \dots, F_D) \quad (2)$$

Local and global features are extracted and fused in the LR space to obtain F_{DF} , which is used as an input to the up-sampling network. The upsampling network used in this invention is the same as SRGAN. The entire network of the generator as a whole can be represented as:

$$I_{SR} = H_{Generator}(I_{LR}) \quad (3)$$

The structure of the RDAB is shown below in Figure 2. Our RDAB contains a dense connectivity layer, local feature fusion (LFF), Local Residual Learning (LRL), and CBAM, thus forming a Continuous Memory (CM) mechanism.

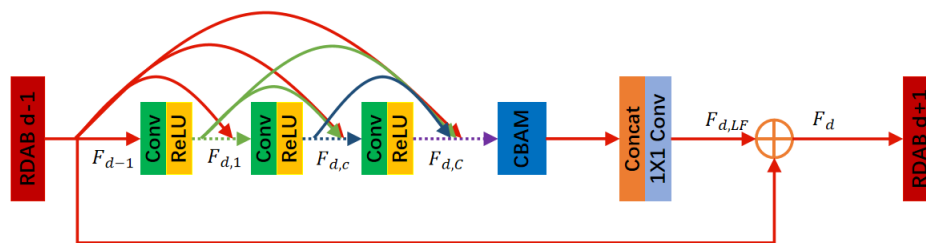


Figure 2. Residual Dense Attention Blocks (RDAB)

Denote the input and output of the d th RDAB by F_{d-1} and F_d , respectively. The outputs of the previous RDAB and each convolutional layer of that RDAB are directly connected to all subsequent layers, which not only preserves the feed-forward properties but also extracts the local dense features. $F_{d,c}$ denotes the output of the c th convolutional layer in that RDAB, and $F_{d,C}$ denotes the output of the last convolutional layer of the densely connected layers in that RDAB. The purpose of LRL is to further improve the information flow, and the final output of the d th RDAB is denoted as:

$$F_{d,LF} = H_{LFF}^d([F_{d-1}, F_d, \dots, F_{d,c}, \dots, F_{d,C}]) \quad (4)$$

$$F_d = F_{d-1} + F_{d,LF} \quad (5)$$

where $H_{LFF}^d(\cdot)$ denotes the fusion of all layer feature information in the d th RDAB. DFF is feature fusion and residual learning for each RDAB feature obtained to utilize hierarchical features in a global

manner. DFF includes GFF and GRL. The global features are extracted by fusing the features of all RDABs.

$$F_{GF} = H_{GFF}([F_1, \dots, F_D]) \quad (6)$$

$$F_{DF} = F_0 + F_{GF} \quad (7)$$

The RDAB introduces a continuous memory mechanism that allows previous RDABs to directly access each layer of the current RDAB. By using LFF, it is possible to train the deep network stably, while LRL can further optimize the information flow and gradient. In addition the RDAB uses GFF and GRL to further extract global features.

Convolutional Block Attention Module (CBAM)[25] is a simple and efficient forward convolutional neural attention module. Given a feature map, CBAM sequentially infers the attention map along two separate dimensions, channel and spatial, and then multiplies the attention map with the input feature map for adaptive feature refinement. CBAM is a lightweight and general-purpose module that can be seamlessly integrated into network architectures. The diagram of the attention module is shown in Figure 3.

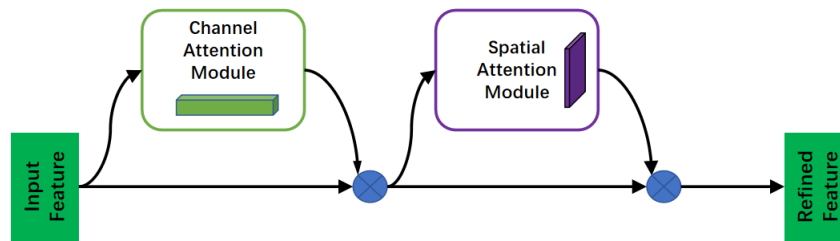


Figure 3. Convolutional Block Attention Module (CBAM)

2.3. Discriminator and Loss function

In order to make the reconstructed image generate real high-frequency information, we add a feature discriminator under the original image discriminator[21], which takes the image in the pixel domain as input, just like the previous method. On the other hand, the feature discriminator inputs images into a VGG network and extracts intermediate feature maps. The feature discriminator then attempts to distinguish between SR images and true HR images based on the extracted feature maps. Since the feature maps encode structural information, the feature discriminator distinguishes between SR images and true HR images based not only on high-frequency components but also on structural components. Ultimately, the generator is trained to synthesize real structural features instead of arbitrary high-frequency noise. The structure of our discriminator is shown in Figure 4.

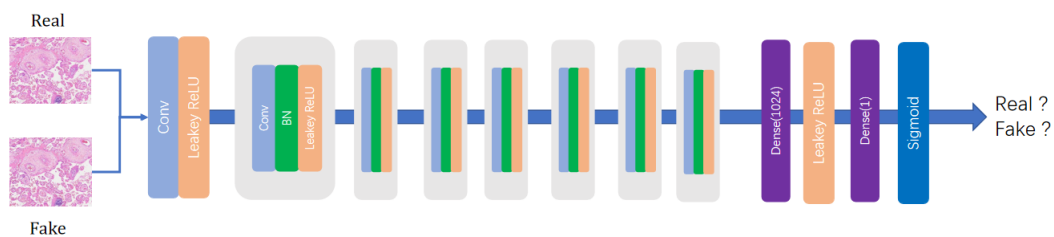


Figure 4. Architecture of our discriminator network

Image discriminator d^i and feature discriminator d^f . Image discriminator d^i distinguishes between real HR images and spurious SR images by examining the pixel values. Feature discriminator d^f distinguishes between real HR images and spurious SR images by examining their feature maps so

that the generator can be trained to synthesize more meaningful high-frequency details. The generator loss function is:

$$L_g = L_p + \lambda (L_a^i + L_a^f) \quad (8)$$

L_p is the perceptual loss[20]. L_a^i is the generator's image GAN loss for synthesizing high-frequency details in the pixel domain. L_a^f is the generator's feature GAN loss for synthesizing structural details in the feature domain. λ is the weights of the GAN loss terms. To train the discriminators d^i and d^f , we minimize the losses L_d^i and L_d^f , which correspond to L_a^i and L_a^f , respectively. The generator and the discriminator are trained by alternately minimizing L_g , L_d^i and L_d^f . The image GAN loss L_a^i and image discriminator loss L_d^i of the generator are defined as follows:

$$L_a^i = -\log(d^i(I^g)) \quad (9)$$

$$L_d^i = -\log(d^i(I^h)) - \log(1 - d^i(I^g)) \quad (10)$$

where $d^i(I)$ is the output of the image discriminator d^i . The feature GAN loss L_a^f and feature discriminator loss L_d^f of the generator are defined as follows:

$$L_a^f = -\log(d^f(\Phi^m(I^g))) \quad (11)$$

$$L_d^f = -\log(d^f(\Phi^m(I^h))) - \log(1 - d^f(\Phi^m(I^g))) \quad (12)$$

where $d^f(\Phi^m)$ is the output of the feature discriminator d^f . During training, the generator can be made to produce realistic structural high-frequency details instead of noise artifacts.

3. Experimental Results And Discussion

3.1. Experimental Settings

We used cell images obtained from a $40\times$ microscope as HR images and downsampled them using BD to obtain the corresponding LR images. Each of the four cells has 800 pairs of training datasets and 5 test datasets. The test datasets are each named Cell A, Cell B, Cell C, and Cell D. We use the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity index (SSIM) on the Y channel in YCbCr space as metrics. At the same time, we also calculated Learned Perceptual Image Patch Similarity (LPIPS), model parameters and single cell image reconstruction time.

The LR images are blur downsampled corresponding images with downscaling factors $\times 2$. The batch size is set to 16, and the input size is 192×192 patches randomly cropped from HR images. We used BD downsampling on HR images to obtain 96×96 LR images to be used as training. We add random horizontal flips and 90° , 180° , 270° rotations to the input. We set the weight λ in Equation (8) as 10^{-3} . For Φ^m in Equation (11), Equation (12) we use the Conv5 layer of the VGG-19. We train our model using ADAM optimizer with $\beta_1 = 0.9$, $\beta_2 = 0.999$. The learning rate is initialized to 2×10^{-4} and multiplied with 0.5 after 200-th epoch. Our model is implemented using the PyTorch framework and trained on an NVIDIA 3090 GPU.

3.2. Comparison With Classic Models

We conducted ablation experiments on our proposed model. We trained it using the collected microscope image dataset and evaluated it on the Cell A dataset. As can be seen from the results in Table 1, the CBAM, RDB and feature discriminator that we have added to the model have better PSNR value and SSIM value.

Table 1. The ablation experiments results ("✓" denotes the corresponding operation)

	CBAM	RDB	feature discriminator	PSNR/dB ↑	SSIM ↑
1				28.12	0.72
2	✓			28.75	0.75
3	✓	✓		32.83	0.85
4	✓	✓	✓	33.13	0.87

Then, we compare our model with six classic image SR models, including EDSR, Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN)[26], very deep Residual Channel Attention Networks (RCAN)[27], RDN, SRResnet[17](less a discriminator compared to SRGAN), RFGANM. And from Table 2, it can see that compared to other SR models, our model achieves the highest PSNR and SSIM, and good LPIPS, model parameters, and reconstruction times.

Table 2. Comparison of PSNR, SSIM, LPIPS, model parameters, and reconstruction times for Cell A, B, C, and D using our proposed model and six classic image SR models (the red indicates the best result)

Model	Dataset	PSNR/dB \uparrow	SSIM \uparrow	LPIPS \downarrow	Parameters/M	Time/s
RFGANM[20]	Cell A	28.12	0.72	0.2048	5.92	2.27
	Cell B	27.45	0.85	0.1108		
	Cell C	28.86	0.77	0.1645		
	Cell D	29.62	0.84	0.1513		
ESRGAN[26]	Cell A	28.74	0.75	0.1089	102.16	2.70
	Cell B	31.71	0.90	0.0292		
	Cell C	30.49	0.79	0.0775		
	Cell D	33.50	0.89	0.0414		
SRRResnet[17]	Cell A	30.98	0.82	0.1616	5.34	2.33
	Cell B	33.53	0.93	0.0385		
	Cell C	32.62	0.86	0.1356		
	Cell D	35.92	0.93	0.0572		
RCAN[27]	Cell A	31.06	0.82	0.1689	58.92	2.17
	Cell B	33.54	0.93	0.0412		
	Cell C	32.66	0.86	0.1466		
	Cell D	35.68	0.93	0.0590		
EDSR[18]	Cell A	31.63	0.84	0.1542	155.30	2.05
	Cell B	34.07	0.94	0.0362		
	Cell C	32.73	0.86	0.1414		
	Cell D	36.45	0.93	0.0606		
RDN[19]	Cell A	31.68	0.84	0.1531	50.14	3.23
	Cell B	34.13	0.94	0.0365		
	Cell C	32.81	0.86	0.1401		
	Cell D	36.89	0.94	0.0609		
RDAGAN	Cell A	33.13	0.87	0.1364	11.25	2.56
	Cell B	36.09	0.96	0.0259		
	Cell C	34.29	0.89	0.1272		
	Cell D	38.31	0.96	0.0550		

In addition, Figure 5 shows the reconstruction detail views of our proposed model and two classic models. It can be observed that the cell detail textures and edge parts reconstructed by our model are closer to the HR images. In summary, these results demonstrate that our network model has achieved excellent performance.

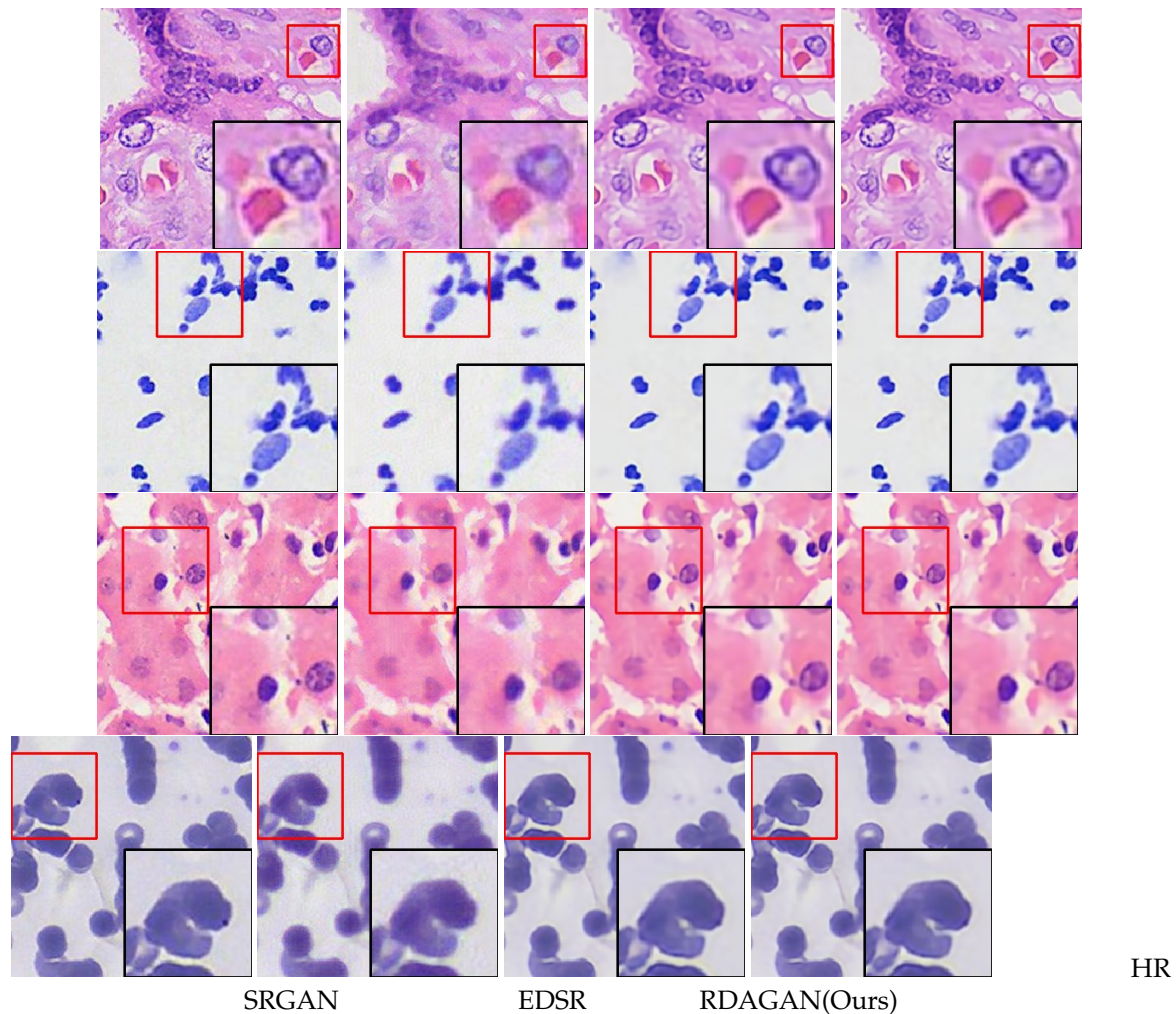


Figure 5. Comparison of reconstructed detail views for Cell A, Cell B, Cell C, and Cell D using our proposed model and two classical models

4. Conclusion

In this paper, we provide a microscopic images dataset. And the SRGAN model is improved by modifying the residual block of the generator to an RDB and combining the attention mechanism. Additionally, a feature discriminator is introduced to ensure that the reconstruction avoids generating artifacts and instead produces high-frequency features pertinent to the image. The loss function is also enhanced. Finally, the experimental results demonstrate that our model outperforms the classical models and is hopeful to meet the demand for high-resolution imaging of large volume samples.

Author Contributions: Conceptualization, S.L. and X.W.; methodology, S.L.; software, S.L.; validation, X.W. and S.L.; formal analysis, S.L.; investigation, S.L. and X.W.; resources, M.C.; data curation, X.W.; writing—original draft preparation, S.L. and X.W.; writing—review and editing, S.L., X.W., X.G. and X.X.; visualization, X.W.; supervision, S.L. and L.Z.; project administration, S.L.; funding acquisition, S.L. and L.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the Fundamental Research Funds for the Central Universities (Grant No. ZQN-1005) and the Scientific Research Funds of Huaqiao University (Grant No. 21BS118).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Many thanks Muwang Chen's contribution to the analysis method proposed and strong support in this article.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

FOV	Field of View
SR	Super-Resolution
LR	Low Resolution
HR	High-Resolution
SISR	Single Image Super-Resolution Reconstruction
RDAGAN	Residual Dense Attention Generative Adversarial Network
CNN	Convolutional Neural Networks
VDSR	Very Deep Super Resolution
LapSRN	Three letter acronym
GAN	Generative Adversarial Network
SRGAN	Super-Resolution Generative Adversarial Network
EDSR	Enhanced Deep Super Resolution
RDB	Residual Dense Block
RDN	Residual Dense Network
RF-GANM	Registration-Free GAN
DCFA	Deep Fourier Channel Attention
RDAB	Residual Dense Attention Block
BD	Bicubic Downsampling
DFF	Dense Feature Fusion
ReLU	Rectified Linear Unit
GFF	Global Feature Fusion
GRL	Global Residual Learning
LFF	Local Feature Fusion
LRL	Local Residual Learning
CBAM	Convolutional Block Attention Module
CM	Continuous Memory
PSNR	Peak Signal-to-Noise Ratio
SSIM	Structural SIMilarity index
LPIPS	Learned Perceptual Image Patch Similarity
ESRGAN	Enhanced Super Resolution Generative Adversarial Networks
RCAN	Residual Channel Attention Networks

References

1. J. Sanderson, "Multi-Photon Microscopy." *Current Protocols*, vol. 3, no. 1, 2023.[[CrossRef](#)]
2. R. Tsai and T. S. Huang, "Multiframe Image Restoration and Registration," *Advances in Computer Vision and Image Processing*, vol. 1, no. 2, pp.317-339, 1984.[[CrossRef](#)]
3. Y. Rivenson, H. C. Koydemir, H. Wang, Z. Wei, Z. Ren, H. Günaydin, Y. Zhang, Z. Göröcs, K. Liang, D. Tseng and A. Ozcan, "Deep Learning Enhanced Mobile-Phone Microscopy," *ACS Photonics*, vol. 5, no. 6, pp. 2354-2364, 2018. [[CrossRef](#)]
4. O. Wei, A. Andrey, L. Mickaël, H. Xian and Z. Christophe, "Deep Learning Massively Accelerates Super-Resolution Localization Microscopy," *Nature Biotechnology*, vol. 36, no. 5, pp. 460-468, 2018. [[CrossRef](#)]
5. E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-STORM: Super-Resolution Single-Molecule Microscopy by Deep Learning," *Optics*, vol. 5, no. 4, pp. 458-464, 2018.[[CrossRef](#)]
6. Y. Xiao, Q. Yuan, K. Jiang, J. He, Y. Wang, and L. Zhang, "From Degrade to Upgrade: Learning a Self-Supervised Degradation Guided Adaptive Network for Blind Remote Sensing Image Super-Resolution," *Information Fusion*, vol. 96, pp. 297–311, 2023.[[CrossRef](#)]

7. Y. Xiao, X. Su, Q. Yuan, D. Liu, H. Shen, and L. Zhang, "Satellite Video Super-Resolution via Multiscale Deformable Convolution Alignment and Temporal Grouping Projection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–19, 2021.[\[CrossRef\]](#)
8. Y. Xiao, Q. Yuan, K. Jiang, X. Jin, J. He, L. Zhang, and C. w. Lin, "Local-Global Temporal Difference Learning for Satellite Video Super-Resolution," *IEEE Transactions on Circuits and Systems for Video Technology*, 2023.[\[CrossRef\]](#)
9. T. Blu, P. Thévenaz, and M. Unser, "Linear Interpolation Revitalized," *IEEE Transactions on Image Processing*, vol. 13, no. 5, pp. 710-719, 2004.[\[CrossRef\]](#)
10. E. Meijering and M. Unser, "A Note on Cubic Convolution Interpolation," *IEEE Transactions on Image Processing*, vol. 12, no. 4, pp. 477-479, 2003.[\[CrossRef\]](#)
11. C. Fan, C. Wu, G. Li and J. Ma, "Projections onto Convex Sets Super-Resolution Reconstruction Based on Point Spread Function Estimation of Low-Resolution Remote Sensing Images," *Sensors*, vol. 17, no. 2, pp. 362, 2017.[\[CrossRef\]](#)
12. F. Li, X. He, Z. Wei, Z. Mu and M. Li, "Super-Resolution Image Reconstruction Based on an Improved Maximum a Posteriori Algorithm," *Journal of Beijing Institute of Technology*, vol. 27, no. 2, pp. 237-240, 2018.[\[CrossRef\]](#)
13. J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image Super-Resolution Via Sparse Representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861-2873, 2010.[\[CrossRef\]](#)
14. C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a Deep Convolutional Network for Image Super-Resolution," in *European Conference on Computer Vision (ECCV)*, pp. 184-199, 2014.[\[CrossRef\]](#)
15. J. Kim, J. K. Lee and K. M. Lee, "Accurate Image Super-Resolution Using Very Deep Convolutional Networks," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1646-1654, 2016.[\[CrossRef\]](#)
16. W. S. Lai, J. B. Huang, N. Ahuja, and M. H. Yang, "Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5835-5843, 2017.[\[CrossRef\]](#)
17. C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang and W. S, "Photo-Realistic Single Image Super Resolution Using a Generative Adversarial Network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4681–4690, 2017.[\[CrossRef\]](#)
18. B. Lim, S. Son, H. Kim, S. Nah and K. M. Lee, "Enhanced Deep Residual Networks for Single Image Super-Resolution," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 1132-1140, 2017.[\[CrossRef\]](#)
19. Y. Zhang, Y. Tian, Y. Kong, B. Zhong and Y. R. Fu, "Residual Dense Network for Image Super-Resolution," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2472-2481, 2018.[\[CrossRef\]](#)
20. H. Zhang, C. Fang, X. Xie, Y. Yang, W. Mei, D. Jin and P. Fei, "High-Throughput, High-Resolution Deep Learning Microscopy Based on Registration-Free Generative Adversarial Network," *Biomedical optics express*, vol. 10, no. 3, pp. 1044-1063, 2019.[\[CrossRef\]](#)
21. H. Wang, Y. Rivenson, Y. Jin, Z. Wei, R. Gao, H. Günaydin, L. A. Bentolila, C. Kural, and A. Ozcan, "Deep Learning Enables Cross-Modality Super-Resolution in Fluorescence Microscopy," *Nature Methods*, vol. 16, no.1, pp. 103-110, 2019.[\[CrossRef\]](#)
22. R. J. G. van Sloun, O. Solomon, M. Bruce, Z. Z. Khaing, H. Wijkstra, Y. C. Eldar, and M. Mischi, "Super-Resolution Ultrasound Localization Microscopy Through Deep Learning," *IEEE Transactions on Medical Imaging*, vol. 40, no. 3, pp. 829-839, 2021.[\[CrossRef\]](#)
23. C. Qiao, D. Li, Y. Guo, C. Liu, T. Jiang, Q. Dai, and D. Li, "Evaluation and Development of Deep Neural Networks for Image Super-Resolution in Optical Microscopy," *Nature Methods*, vol. 18, no. 2, pp. 194-202, 2021.[\[CrossRef\]](#)
24. S. J. Park, H. Son, S. Cho, K. S. Hong, and S. Lee, "SRFeat: Single Image Super-Resolution with Feature Discrimination," in *European Conference on Computer Vision (ECCV)*, pp. 439-455, 2018.[\[CrossRef\]](#)
25. S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," in *European Conference on Computer Vision*, pp. 3-19, 2018.[\[CrossRef\]](#)
26. X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao and C. C. Loy, "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks," in *European Conference on Computer Vision (ECCV) Workshops*, 2018. [\[CrossRef\]](#)

27. Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong and Y. R. Fu, "Image Super-Resolution Using Very Deep Residual Channel Attention Networks," in European Conference on Computer Vision (ECCV), pp. 286-301, 2018.[\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.