Multi-scale Deep Convolutional Neural Networks for Microscopic Image Super-resolution

Wazir Muhammad[1*], Nazia Ejaz[2], Ayaz Hussain[1], Jalal Shah[3], Sohrab Khan[3], and Inam Ul Ahad[4]

[1] Electrical Engineering Department, BUET, Khuzdar, Pakistan
[2] Biomedical Engineering Department, BUET, Khuzdar, Pakistan
[3] Computer System Engineering Department, BUET, Khuzdar, Pakistan
[4] School of Mechanical and Manufacturing Engineering Dublin City University, Dublin, Ireland
wazirlaghari@buetk.edu.pk*

Abstract. Deep convolutional neural networks (CNNs) have recently shown remarkable success in single image super-resolution (SISR), particularly in medical image super-resolution for microscopy. However, microscopy image reconstruction remains a challenging task through conventional approaches, which often require high hardware costs and yield unsatisfactory results. We propose a new multi-scale deep CNN architecture tailored to SISR for low-resolution (LR) microscopic images. To tackle the challenges of training deep CNNs, we utilize a residual learning approach, explicitly supervising the residuals using the disparity between high-resolution (HR) and LR images. The sum up of the recovered details to the LR image results in the reconstruction of the HR image. In addition, we employ gradient clipping to prevent gradient explosions that can occur with high learning rates. Furthermore, the choice of Depthwise separable convolution in our paper is to justified by its ability to reduce computational complexity and less memory usage while maintaining high accuracy. In contrast to current deep CNN-based SISR methods for natural images, where LR images are received by subsampling and blurring HR images, we evaluate our approach using lower objective lenses and thin smear blood samples. HR images captured with higher objective lenses are used as a benchmark to compare the performance. Extensive evaluations confirm that Multi-Scale Deep Convolutional Neural Networks for Microscopic Image Super-resolution (MDCM) outperform other methods. The proposed MDCM method addresses the critical need for accurate and fast reconstruction algorithms to improve temporal resolution in high-density super-resolution microscopy, particularly for live-cell imaging

Keywords: Super-resolution microscopy, Single image super-resolution (SISR), Biomedical imaging, Depthwise Separable Convolution

1 Introduction

Microscopy acts a critical part in biomedical examination and research by aiding the visualization and analysis of cellular structures and processes. Recent advancements in optics and computer vision have led to the development of various microscopy techniques that provide new insights into living cells [1]. The most commonly used transmitted light microscopy techniques are Phase contrast (PC) and Differential interference contrast (DIC) microscopy. PC microscopy approach is used to convert the phase information of the imaging field into intensity variations in the final

I. U. Ahad (ed.), *Proceedings of the 4th International Conference on Key Enabling Technologies (KEYTECH 2024)*, Atlantis Highlights in Engineering 35,
https://doi.org/10.2991/978-94-6463-602-4_6

reconstructed image. Similarly, DIC microscopy also visualizes phase gradients and enhances the contrast of unstained living cells [2]. A fluorescence microscope is an optical microscope that uses the phenomenon of fluorescence, rather than or in combination with scattering, reflection, and absorption, to investigate the characteristics of organic or inorganic materials. The approach of Fluorescence microscopy plays a cornerstone role in biological and cellular studies due to its ability to visualize specific cellular components with high specificity. However, conventional fluorescence microscopy is restricted by light diffraction, which restricts spatial resolution to a few hundred nanometers [3], because subcellular structures like mitochondria, microtubules, and proteins are frequently smaller than this threshold, viewing and studying them becomes difficult due to this constraint. Despite these challenges, various diffraction-limited super-resolution microscopy methods are currently developed, with resolutions as much as ten times greater compared to those of standard approaches. These methods encompass structured illumination microscopy (SIM), which uses an array of patterned light illuminations to increase resolution, stimulated emission depletion (STED), which uses the concept of stimulated emission to boost resolution, and non-linear SIM, which is a more advanced form of SIM that uses nonlinear effects to enhance resolution further. Super-resolution images can be produced using these experimental techniques, but they can be resource-intensive and frequently call for intricate hardware configurations.

Recently, in single-image super-resolution (SISR), deep learning and artificial intelligence (AI) have made substantial advances in image super-resolution (SR) problems [4-7]. Among the most well-known CNN-based models, Super Resolution Convolutional Neural Network (SRCNN) [8] is one of the cornerstone models in this field. It was one of the first to employ deep-learning methods to improve image resolution. Employing a simple 3-layer CNN, it develops a from beginning to end mapping between lower and high-quality images. For a better, more computationally efficient version of SRCNN that replaces the prior bicubic interpolation layer for upscaling with a deconvolution layer and uses smaller filters known as FSRCNN. By utilizing gradient clipping and residual learning, in paper [9], the authors developed the Very Deep Super-Resolution (VDSR) technique that employs twenty CNN layers to achieve cutting-edge results on benchmark datasets. In [10], the authors, for image super resolution tasks, proposed the Laplacian Pyramid Super-Resolution Network (LapSRN) algorithm which is a pyramid-based method. In LapSRN, they used a CNN layer with a Laplacian pyramid structure to enable the network to anticipate sub-band residuals at several stages, enlightening the excellence of the images. Although, these methods outperformed earlier methods in terms of performance, they still have certain drawbacks, but new CNN methods are trying to fix them. Initially, artificial datasets are used to train deep learning models. On these datasets, LR images are produced by subjecting HR images to a known degradation technique, such as bicubic downscaling. Nevertheless, the models have difficulty with the many degradations that real-world LR images frequently exhibit. To enhance generalization, methods such as domain adaptation and adversarial training are being investigated. Second, the majority of deep SR models, particularly those with extremely deep structures, can be slow and computationally expensive, which makes it challenging to implement on devices with limited resources. Approaches like as pruning, channel splitting, and Depthwise separable convolutions are being used in an attempt to create effective models. Even though deep CNN models are excellent at reducing pixel-by-pixel losses like mean square error (MSE), the reconstructed images might not be as clear and perceptually pleasing. Although perceptual quality is improved by using adversarial losses and pre-trained network-based perception losses, perceptual quality can still be enhanced. Our suggested model, called Multi-Scale Deep CNNs for Microscopic Image Super-resolution (MDCM), uses deep CNN to overcome the difficulties in image super-resolution.

Key contribution of our proposed model is as under:

1. Our proposed model uses the Depthwise separable convolution technique in place of standard convolution to drastically cut down on computational complexity and parameter count without sacrificing performance by the use of a 1x1 pointwise convolution after applying a single filter. Compared to a regular convolution, a Depthwise separable convolution is more efficient.
2. Leaky ReLU activation functions were used in our model after Depthwise separable convolution layers. Leaky ReLU creates a little slope for negative values with the aim of increasing gradient flow throughout training and assisting in avoiding the dying ReLU.
3. The proposed model upsamples the feature maps to the required output resolution by incorporating deconvolution layers, also referred to as transposed convolution. By upsampling the input features, deconvolution layers are able to effectively increase the spatial dimensionality of the feature maps.

In summary, Multi-Scale Deep CNNs for Microscopic Image Processing were proposed and employed a combination of Depthwise separable convolution, Leaky ReLU activation, deconvolution layers, and skip connections. The super-resolution model performs better on smear blood samples of microscopy images than other existing methods. This allows the reconstruction of high quality, aesthetically pleasing HR output images from low resolution microscopy inputs. Additionally, in the reconstructed HR images, the global residual skip connections aid in the preservation of significant fine details and structures, which is essential for precise analysis and diagnosis in microscopic applications.

2 Proposed Methodology

In this part, we present our suggested model, MDCM, to rebuild visually appealing superior HR output images from their comparable LR inputs, as shown in Figure 1. The model leverages a combination of convolutional layers, Depthwise separable convolution layers, shrinking layers, deconvolution layers, expanding layers, and global residual skip connections to effectively extract features and reconstruct the desired HR output images. Initially, model architecture used standard convolutional layer that operates on the novel input LR image. The main purpose of CNN layer for extracting low-level features, for example edges, textures, and basic patterns, from the original LR input. The output of this layer serves as the foundation for further feature extraction and reconstruction. Proposed model add five times Depthwise separable convolution layers to decrease the number of factors and mathematical complexity without sacrificing performance, because Depthwise separable convolution is a computationally efficient alternative to ordinary convolution. A Leaky ReLU activation function, which provides non-linearity and aids the model in learning the complicated mappings, comes after each Depthwise separable convolution layer. These layers are in charge of taking deeper, more intricate elements out of the input and resolving them so that higher-level details necessary for image reconstruction are captured. The model comprises a shrinking layer that comes after the Depthwise separable convolution layers. In order to reduce the number of feature channels, the shrinking layer is usually a convolutional layer with a kernel size of 1x1. This layer assists in lowering the memory and processing demands of the layers that come after it. A deconvolution layer, often referred to as transposed convolution, was used in our model to upsample the feature maps to the intended reconstructed output image. By upsampling the input features, deconvolution layers are able to effectively increase the spatial

dimensions of the feature maps. Reconstructing the output image's high-resolution details depends heavily on this layer. An extending layer, which is the last component of the model, is usually a convolutional layer with a 3x3 kernel size. The upsampled features are refined and made ready for the final output by this layer. In order to improve even more the reconstruction quality and training process stability, the model includes both a local and global residual skip connection. The input LR image is directly added to each layer's output via the local and global residual skip connection, which then adds the shrinking layer at the end. Moreover, skip connection facilitates training and enhances overall performance by teaching the model the residual mapping between the input and the intended HR output. The deep CNN model efficiently recovers both high-level and low-level features, upsamples the feature maps, and reconstructs the final HR output image by merging these elements. Important features are preserved, and a high-quality, aesthetically acceptable rebuilt image is produced due to the global residual skip connection.

2.1 Quality Evaluation Metrics

In the realm of medical image super-resolution, quality assessments are essential for determining the effectiveness of super-resolution algorithms, especially in the context of microscope images. These quality metrics enable researchers and professionals to evaluate how well their super-resolution models can produce high-quality microscopic images from low-resolution sources. The three most frequently utilized metrics for this evaluation are Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index Measure (SSIM). Each of these metrics offers distinct perspectives on the performance of super-resolution models.

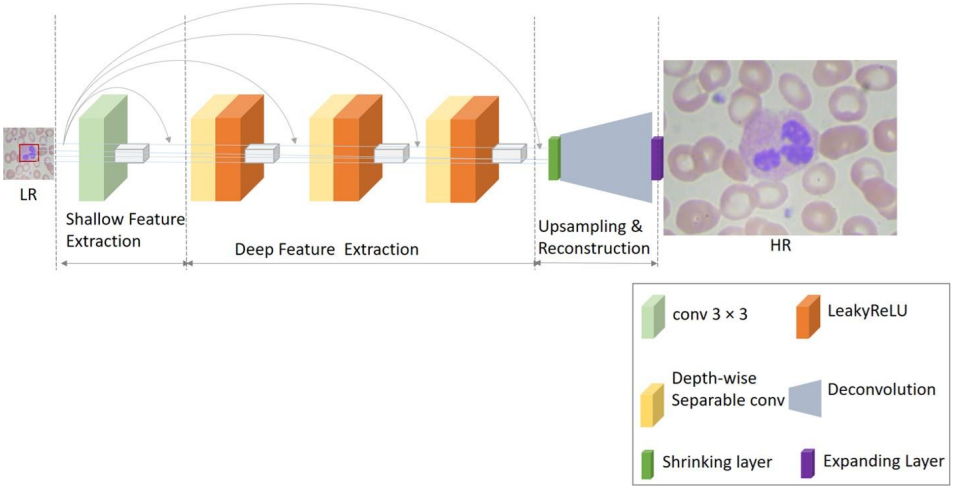

Fig. 1. Proposed Network architecture of Multi-Scale Deep Convolutional Neural Networks for Microscopic Image Super-resolution (MDCM).

Mean Squared Error (MSE). MSE is a statistical metric that measures the average squared difference between the anticipated pixel values (from the super-resolution model) and the actual pixel values. A lower value for the MSE suggests that the efficiency is higher since the anticipated output image is nearer to the original ground truth image.

$$\mathrm{MSE} = \frac{1}{mn}\sum_{i=1}^{N}(I_{\mathrm{true}}(i) - I_{\mathrm{pred}}(i))^2 \tag{1}$$

where I_{true} is the original ground-truth input image, I_{pred} is the predicted output image, and N is the total number of pixels.

Peak Signal-to-Noise Ratio (PSNR): PSNR is a statistic that calculates the ratio of the highest potential power of a signal (the image) to the strength of corrupting noise that influences the quality of its display. It is calculated from MSE and measured in decibels.

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}_I^2}{\mathrm{MSE}}\right) \tag{2}$$

$$\mathrm{PSNR} = 20 \cdot \log_{10}(\mathrm{MAX}_I) - 10 \cdot \log_{10}(\mathrm{MSE}) \tag{3}$$

$$\mathrm{PSNR} = 20 \cdot \log_{10}\left(\frac{\mathrm{MAX}_I}{\sqrt{\mathrm{MSE}}}\right) \tag{4}$$

Structural Similarity Index Measure (SSIM): SSIM is a perceptual statistic that measures the quality of image reduction due to the process, which includes compression. Unlike MSE, PSNR, and SSIM take into consideration modifications in structural information, brightness, and contrast.

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \tag{5}$$

where the average pixel values are denoted by μx and μy, and x and y are the two images under comparison. The covariance is σxy, and the variances are σ2 and σ2. Constants C1 and C2 are used to keep the division stable.

3 Experimental Results

Compare our proposed method results with those of other cutting-edge techniques, including Bicubic, SRCNN, VDSR, and LapSRN. We trained our proposed method for assessing parasitology slides of thin-smear blood samples using the MaMic image dataset [11]. Because this dataset has various properties, it is especially beneficial for clinical applications. High-quality microscopy images are crucial for researching biological structures. With the use of a virtual microscopy platform, great resolution and consistency were generated during the scanning process. Images were obtained at three distinct magnification objectives (10, 20, and 40) of dataset. This diversity makes it possible for us to evaluate the model's performance at various granularities, which is essential for precisely recognizing features in microscopic images. Moreover, JPEG format is used to save images, offering a 600 by 300-pixel resolution and 24

color bits per pixel. This format strikes a balance between file size and quality, making it appropriate for deep learning model training. To ensure that the model operates well whenever employed with previously unidentified data, the dataset is divided into training and testing sets. We can assess our proposed model's performance successfully by utilizing a dataset that is well-structured. Training was done using the Adam optimizer with a starting learning rate of 0.0001. Our proposed model MDCM is quantitatively compared with three benchmark publicly available methods including SRCNN, VDSR, LapSRN, and MDCM. The quantitative calculations (PSNR/SSIM) of our MDCM model greatly surpass those of the cutting-edge techniques, as shown in Table 1. The MDCM model's performance is assessed using enlargement factors of 2x, 3x, and 4x.

Figure 2 illustrates the comparison of computational costs for our MDCM model in relation to the number of parameters (K-times) versus PSNR / SSIM.

Table 1. Shows the quantitative assessment of Microscopic Smear Blood Image SR with our MDCM. The reported results in terms of PSNR/SSIM of factors 2, 3, and 4. The best value is represented by a red colour and bold quantitative numbers. The colour blue with underlined quantitative data represents the second-best value.

Method	Scale	PSNR	SSIM	Parameters
SRCNN	2X	49.53	0.9928	**57 K**
VDSR	2X	49.54	0.9929	665 K
LapSRN	2X	49.56	0.9930	812 K
Our (MDCM)	2X	**49.57**	**0.9931**	540 K
SRCNN	3X	46.52	0.9868	**57 K**
VDSR	3X	46.57	0.9869	665 K
LapSRN	3X	46.60	0.9871	812 K
Our (MDCM)	3X	**46.64**	**0.9872**	540 K
SRCNN	4X	45.04	0.9815	**57 K**
VDSR	4X	45.09	0.9816	665 K
LapSRN	4X	45.20	0.9819	812 K
Our (MDCM)	4X	**45.24**	**0.9821**	540 K

By incorporating Depthwise separable convolution layers, our MDCM model reduces its size compared to other deep CNN image super-resolution models. The performance of the proposed model is assessed on the MaMic image dataset with a scale factor of 2. Our proposed model has less parameters as compared to VDSR, and LapSRN model. Figure 3 depicts the perceptual quality of enlargement factor 3 on MaMic image image SR test datasets. The outcomes on challenging enlargement scale factor 3 observed that blurrier outcomes were produced by Bicubic, SRCNN, and VDSR. However, it is a difficult effort to improve an image for an enlargement factor of 3. Our MDCM model accurately retrieves exquisite texture detail while efficiently suppressing abnormalities and artifacts, because our approach follows the concept of Depthwise separable convolution layers.

4 Conclusion

This paper reports on the development of a new microscope image super-resolution model that makes efficient use of shrinking and expanding layers, Leaky ReLU activation, Depthwise separable convolution, and deconvolution layers. Our technique aims to improve the resolution of microscope images, especially those from the MaMic image collection. Incorporating Depthwise separable convolution into the model allows for efficient feature extraction without sacrificing the quality of the reconstructed images, all while preserving excellent performance and a significant reduction in model complexity. Depthwise separable convolution is

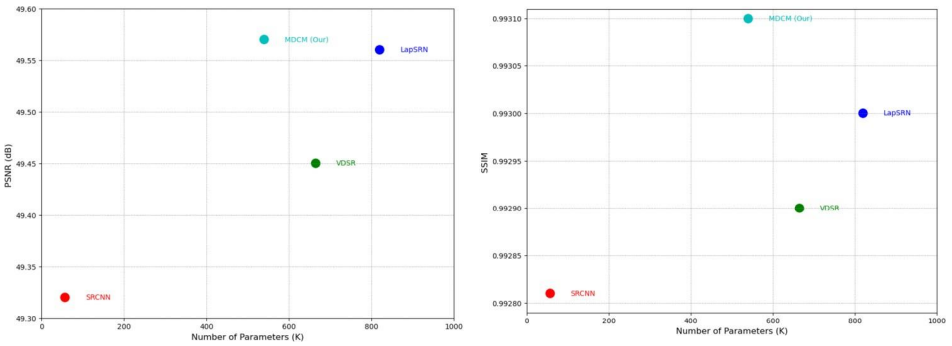

Fig. 2. The performance comparison in terms of model parameters versus PSNR / SSIM tested on MaMic image dataset with upscale factor 2.

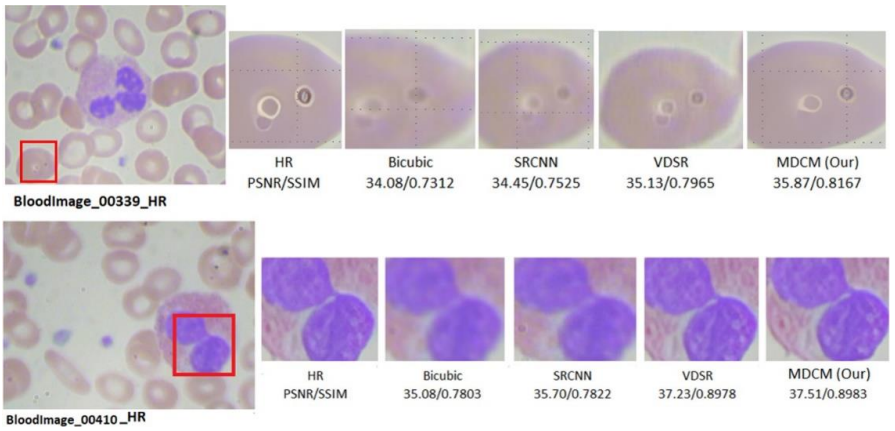

Fig. 3. Visual perceptual quality-wise improvement of Bloodimage 00339 and BloodImage 00410 obtained from MaMic image dataset enlargement factor 3x.

computationally more efficient as compared to normal convolution operation. By utilizing new variant of ReLU is a Leaky ReLU activation functions, because they easily dying ReLU issue is lessened and stable gradient flow is maintained throughout training. The network may successfully

manage the number of channels while preserving and effectively using the spatial information for the task at hand by including shrinking and expanding layers surrounding the deconvolution layer. In applications like image segmentation, where preserving both spatial and feature integrity is essential for precise predictions, this structure is especially helpful. The expanding and contracting layers help the model to catch and recreate intricate details in the images, and the deconvolution layers enable efficient upsampling. In contrast to current cutting edge techniques, the proposed MDCM proposed model outperformed them on a public dataset, obtaining better results regarding Peak Signal-to-Noise Ratio (PSNR), number of parameters, and Structural Similarity Index (SSIM). The outcomes highlight how well our design produces high-quality super-resolved images, which makes it a useful tool for biological imaging and digital pathology applications. All things considered, the proposed approach advances in the field of super-resolution microscopy images and lays the groundwork for further studies targeted at enhancing picture quality and interpretation in a range of biological applications. While Depthwise separable that significantly reduce both computational cost and the number of parameters compared to standard convolutions, they may also result in a modest reduction in accuracy when compared to regular convolutions. This is one of the constraints of our suggested model. Precision is crucial in medical imaging applications, where this trade-off can be very important. The diagnostic efficacy of the suggested model might be impacted if it is unable to adequately capture intricate characteristics in microscopic blood smear images. The capacity of the model to denote spatial relationships across many channels may be limited by the fact that Depthwise separable convolutions function independently on each channel. This constraint may make the model less useful in situations where contextual knowledge and interchannel correlations are essential, such identifying overlapping cells in blood smear images. Future research could examine hybrid designs that incorporate residual connections or attention processes along with other cutting-edge methods like Depthwise separable convolutions. This method might improve feature extraction performance while preserving Depthwise separable convolutions' efficiency advantages.

Acknowledgments. The authors acknowledge the financial support from Erasmus+ CBHE project BIOMED5.0, funded by the European Union (Project Number: 101129077). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

References

1. Melanthota SK, Gopal D, Chakrabart S, Kashyap AA, Radhakrishnan R, Mazumder N (2022) Deep learning-based image processing in optical microscopy. Biophys Rev 14 https://doi.org/10.1007/s12551-022-00949-3
2. V. Ntziachristos, Going deeper than microscopy: the optical imaging frontier in biology. Nat. Methods 7, 603–614 (2010)
3. Hafner, Ju¨rgen, Christopher Wolverton, and Gerbrand Ceder. "Toward computational materials design: the impact of density functional theory on materials research." MRS bulletin 31, no. 9 (2006): 659-668.
4. Muhammad, W., S. Aramvith, and T. Onoye, SENext: squeeze-and-excitationnext for single image super-resolution. IEEE Access, 2023. 11: p. 45989-46003.
5. Liu, S., et al., A Residual Dense Attention Generative Adversarial Network for Microscopic Image Super-Resolution. Sensors, 2024. 24(11): p. 3560.
6. Song, L., et al., Super-resolution reconstruction of structured illumination microscopy using deep-learning and sparse deconvolution. Optics and Lasers in Engineering, 2024.

174: p. 107968.
7. Muhammad, W., S. Aramvith, and T. Onoye, Multi-scale Xception based depthwise separable convolution for single image super-resolution. PLOS ONE, 2021. 16: p. e0249278.
8. Dong, C., C.C. Loy, and X. Tang. Accelerating the super-resolution convolutional neural network. in Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14. 2016. Springer.
9. Kim, J., J.K. Lee, and K.M. Lee. Accurate image super-resolution using very deep convolutional networks. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
10. Lai WS, Huang JB, Ahuja N, Yang MH. Deep laplacian pyramid networks for fast and accurate super-resolution. InProceedings of the IEEE conference on computer vision and pattern recognition 2017 (pp. 624-632).
11. WebMicroscope2019 http://fimm.webmicroscope.net/Research/Momic/mamic