



A Cascaded Refinement GAN for Phase Contrast Microscopy Image Super Resolution

Liang Han and Zhaozheng Yin^(✉)

Department of Computer Science, Missouri University of Science and Technology,
Rolla, USA
{lh248,yinz}@mst.edu

Abstract. Phase contrast microscopy is a widely-used non-invasive technique for monitoring live cells over time. High-throughput biological experiments expect a wide-view (i.e., a low microscope magnification) to monitor the entire cell population and a high magnification on individual cell's details, which is hard to achieve simultaneously. In this paper, we propose a cascaded refinement Generative Adversarial Network (GAN) for phase contrast microscopy image super-resolution. Our algorithm uses an optic-related data enhancement and super-resolves a phase contrast microscopy image in a coarse-to-fine fashion, with a new loss function consisting of a content loss and an adversarial loss. The proposed algorithm is both qualitatively and quantitatively evaluated on a dataset of 500 phase contrast microscopy images, showing its superior performance for super-resolving phase contrast microscopy images. The proposed algorithm provides a computational solution on achieving a high magnification on individual cell's details and a wide-view on cell populations at the same time, which will benefit the microscopy community.

1 Introduction

Phase Contrast Microscopy [19] allows researchers to acquire images on hundreds of live cells from different treatments over days or weeks without invasively staining them (Fig. 1(a)). The high-throughput experiments need a wide-view to monitor the entire cell population, so the magnification of phase contrast microscope is set low. But, the low magnification loses cell details and provides low-resolution (**LR**) images on individual cells (Fig. 1(b)). To obtain high-resolution (**HR**) images with cell details, we have to increase the magnification with a limited view on a small number of cells (Fig. 1(c)), which is not suitable for monitoring large-scale cell populations.

To simultaneously achieve a high magnification and wide-view, we propose a novel super-resolution approach to increase the image resolution on cell details while maintaining the wide-view on cell populations.

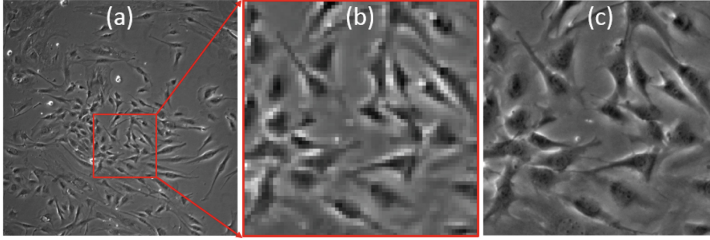


Fig. 1. (a) Monitoring cell populations with a low magnification; (b) A zoomed-in image shows the low resolution on individual cells; (c) A high magnification visualizes cell details but with a limited view.

2 Related Work and Our Proposal

The single image super-resolution problem is first tackled by some filtering approaches such as linear, bicubic [8] and Lanczos [4]. However, these filtering methods produce overly smoothed textures in the recovered HR images.

To preserve the edges in the recovered HR images, example-based approaches [17,18] aiming at learning a mapping between low- and high-resolution image patches are proposed. The example-based methods (or patch-based methods) exploit the self-similarity property and construct high-resolution image patches from the input image [5]. However, the example-based methods suffer from the heavy computational cost of patch search. Moreover, the found low- and high-resolution image patch pairs may not be sufficient to represent the large textural variations in the testing images.

In order to overcome the drawbacks of the example-based approaches, deep learning algorithms [3,9] are proposed to super-resolve an image. Instead of modeling the low- and high-resolution mapping in the patch space, these algorithms learn the nonlinear mapping in the image space. Dong *et al.* upscale an input LR image and train a three layer deep convolutional network to learn an end-to-end mapping between low- and high-resolution images [3]. Ledig *et al.* present a Generative Adversarial Network (GAN) for image super-resolution, which is capable of inferring photo-realistic natural images for $4\times$ upscaling factors [12]. Lai *et al.* propose a Laplacian Pyramid Super-Resolution Network to progressively reconstruct the sub-band residuals of the high-resolution images [11]. However, these deep learning algorithms mainly focus on the natural image super-resolution, and may not be suitable for super-resolving microscopy images as no optical properties of the microscopy imaging are taken into account.

In this paper, we propose a cascaded refinement GAN for phase contrast microscopy image super-resolution. The contributions of this study are mainly in three aspects. *First*, to our best knowledge, this is the first framework capable of super-resolving phase contrast microscopy images for $8\times$ upscaling factors. *Second*, we design a perceptual loss consisting of an adversarial loss and a content loss for our algorithm and achieve the best performance on phase contrast microscopy image super-resolution. *Third*, an optics-related data enhancement

is developed to improve the performance of our algorithm on phase contrast microscopy image super-resolution.

3 Preliminaries

3.1 Generative Adversarial Networks

There are one generator network G and one discriminator network D in the Generative Adversarial Networks (GANs) [6]. These two networks are trained alternatively to compete in a two-player min-max game. The generator G is optimized to simulate the true data distribution by synthesizing images that are challenging for the discriminator D to tell from the real ones. The discriminator D is optimized to try not to be fooled by the generator G by correctly distinguishing the real images from the synthetic images. These two networks play the min-max game with the following objective function

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))], \quad (1)$$

where x is a real image sampled from the real data distribution p_{data} , z is a noise vector drawn from distribution p_z (e.g., Gaussian or uniform distribution).

3.2 Optics-Related Data Enhancement

The properties of phase contrast microscopy images (e.g., halo artifact and low image contrast) motivated researchers to derive optics-related imaging models for microscopy image restoration [16], which include a series of Diffraction Pattern Filters (DPFs, Fig. 2(a1–a6)). Rather than solving the inverse problem of image restoration, we leverage the DPFs to enrich phase contrast microscopy images (Fig. 2(b0)) by convolving it with DPFs. As shown in Fig. 2(b1–b6), the convolution generates a set of images sensitive to different types of cell regions, which is an optics-related data enhancement to the original input.

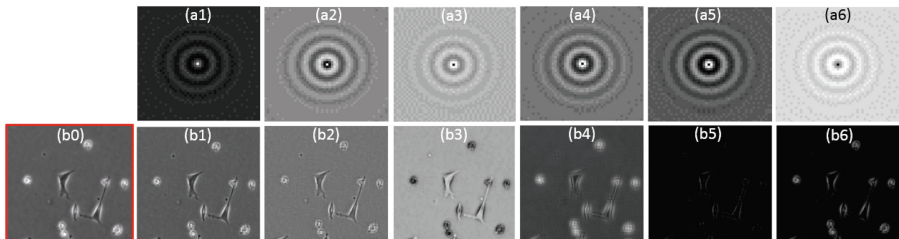


Fig. 2. DPFs and the enriched phase contrast microscopy images.

4 A Cascaded Refinement GAN for Super Resolution

In this section, we describe the network structure of the proposed cascaded refinement generative adversarial network (GAN) for phase contrast microscopy image super resolution, the loss function for network optimization, and the network implementation and training details.

4.1 Network Architecture

Generator Architecture: As shown in Fig. 3, the LR phase contrast image is first enhanced by convolving with some DPFs. The proposed generator takes the enhanced image stacks as input and then refines them with cascaded refinement modules. Each module operates at a certain resolution and the resolution is doubled between two consecutive modules.

All the modules $M_s (s \in \{1, 2, 3\})$ are structurally identical and consist of three layers: the input layer, the upsampling layer, and the convolutional layer (Fig. 3). The input layer of the first module is the enriched image stack of the LR image, while the input layers of the other modules are identical to the convolutional layers of the previous modules. The upsampling layer is obtained by bilinearly upsampling the input layer. As the upconvolution is prone to introduce characteristic artifacts to the output image [2, 14], we discard the upconvolution and use bilinear upsampling. The convolutional layer is obtained by implementing 3×3 convolutions, layer normalization [1], and Leaky ReLU nonlinearity [13] operations on the upsampling layer. A linear projection (1×1 convolution) is applied on the convolutional layer to project the feature tensor to the output image. Note that the output image at each upsampling level is not used as the input of the next module.

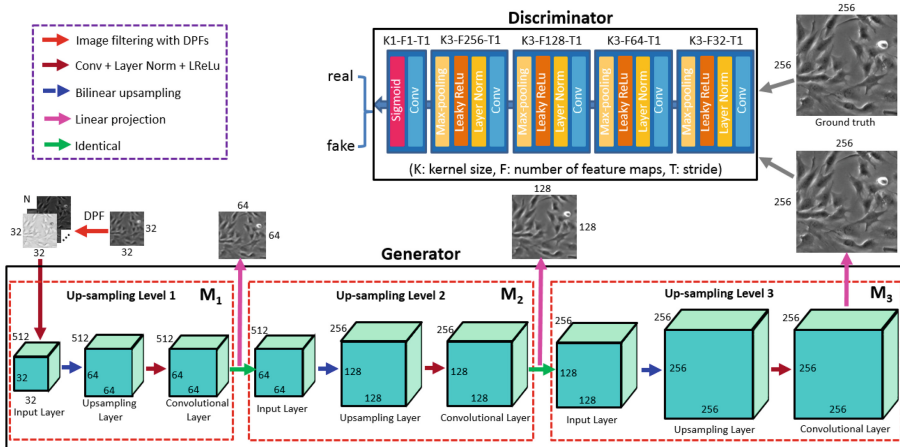


Fig. 3. Architecture of the Generator and Discriminator in our model.

Discriminator Architecture: There are four downsampling modules in the discriminator, and each downsampling module consists of 3×3 convolutions, layer normalization, Leaky ReLU nonlinearity, and 2×2 max-pooling operations. The discriminator tries to classify whether an input image is real or fake.

4.2 Loss Function

Let x be the input LR image and y be the corresponding HR counterpart. Our ultimate goal is to learn a generator G with parameters θ_G for generating a HR image $\hat{y} = G(x; \theta_G)$ which is as similar to the ground truth HR image y as possible. To learn the parameters (θ_G in the generator and θ_D in the discriminator), we formulate a perceptual loss function as the weighted sum of an adversarial loss and a content loss:

$$L(y, \hat{y}) = L_{adv}(y, \hat{y}) + \alpha L_{con}(y, \hat{y}). \quad (2)$$

Adversarial Loss: The adversarial loss in the GAN encourages the generator to generate images residing on the manifold of the ground truth images by trying to fool the discriminator. From Eq. 1 we can see the adversarial loss is formulated as $\mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z)))]$. In our case, the generator takes a LR image instead of a noise vector as the input, accordingly the adversarial loss $L_{adv}(y, \hat{y})$ over K training samples can be defined as:

$$L_{adv}(y, \hat{y}) = \frac{1}{K} \sum_{k=1}^K [\log(1 - D(G(x_k; \theta_G); \theta_D))], \quad (3)$$

Early in training, we can minimize $\log(-D(G(x_k; \theta_G); \theta_D))$ instead of $\log(1 - D(G(x_k; \theta_G); \theta_D))$ to mitigate the gradient saturation [6].

Content Loss: In addition to the adversarial loss in the GAN, we also add a content loss $L_{con}(y, \hat{y})$ to our perceptual loss function. Many previous state-of-the-art approaches rely on the widely used pixel-wise mean square error (MSE) loss to learn the parameters of the LR-HR mapping functions [3, 15], however, solutions of MSE optimization problems often lack the ability of learning high-frequency contents and result in overly smoothed textures [7, 12]. Instead, we propose a new content loss using the DPFs to emphasize the loss on different types of cell regions:

$$L_{con}(y, \hat{y}) = \frac{1}{K} \sum_{k=1}^K \sum_{s=1}^S \sum_{n=0}^{N-1} \|y_s * DPF_n - \hat{y}_s * DPF_n\|_1, \quad (4)$$

where y_s is the downsampled ground truth image at level s , \hat{y}_s is the generated upscaled LR image at level s (the output layer of the s^{th} module), S is the number of refinement modules in the generator, N is the number of DPFs, and $\|\cdot\|_1$ is the ℓ_1 distance.

4.3 Implementation and Training Details

The number of refinement modules in our generator is decided by the upscaling factor. In our experiments with the upscaling factor of $8\times$, there are 3 refinement modules, corresponding to a resolution increase from 32×32 to 256×256 in the generator. The Leaky ReLU in the generator and discriminator is with a negative slope of 0.2. Zero-padding is implemented before the convolution to keep the size of all feature maps unchanged in each module. The number of DPFs in our experiment is 12.

To train our networks, we alternate between one gradient descent step on the discriminator, and then one step on the generator. The minibatch stochastic gradient descent (SGD) is used in our experiment and the Adam solver [10] with momentum parameter 0.9 is applied to optimize the SGD. We use the batch size of 1. The learning rate is $1e-4$, and the algorithm is trained for 100 iterations. The weight α in the loss function is 1.

5 Experimental Results

Dataset: 11,500 high-resolution phase contrast microscopy images with different cell densities are captured at the resolution of $256 * 256$ pixels, and the low-resolution images are obtained by downsampling the high-resolution images. After getting the high-resolution and low-resolution image pairs, we randomly select 10,000 pairs of them as the training set, 1,000 as the validation set, and 500 as the testing set.

Evaluation Metrics: We evaluate the proposed super-resolution algorithm with a widely-used image quality metric: Peak Signal-to-Noise Ratio (PSNR).

Evaluation: We compare our algorithm with the baseline bicubic [8] and the current state-of-the-art LapSRN [11] algorithms. Figure 4 presents the visual comparison results on some randomly picked images. The bicubic filtering gives very blurry super-resolution results. LapSRN generates much sharper and clearer high resolution images than bicubic, however, the generated images are over-smoothed. It is mainly because the designed loss function of LapSRN is not suitable for microscopy image super-resolution. By taking the optical property of phase contrast microscopy imaging into consideration and designing a perceptual loss, our proposed algorithm generates HR images with clear cell details.

Given an input phase contrast image with a wide-view (e.g., Fig. 5(a)), we can divide the image into 32×32 patches, super-resolve each patch, and then combine the predicted patches to generate a HR image with megapixel resolutions (Fig. 5(b)).

Ablation Study: First, we investigate the effect of different loss functions. As shown in (Fig. 6(b1, b2)), only using the adversarial loss cannot generate HR images with clear overall contents though some cell details are exhibited. The generated images by only using the content loss (Fig. 6(c1, c2)) are not sharp enough and the cell structures are not presented well.

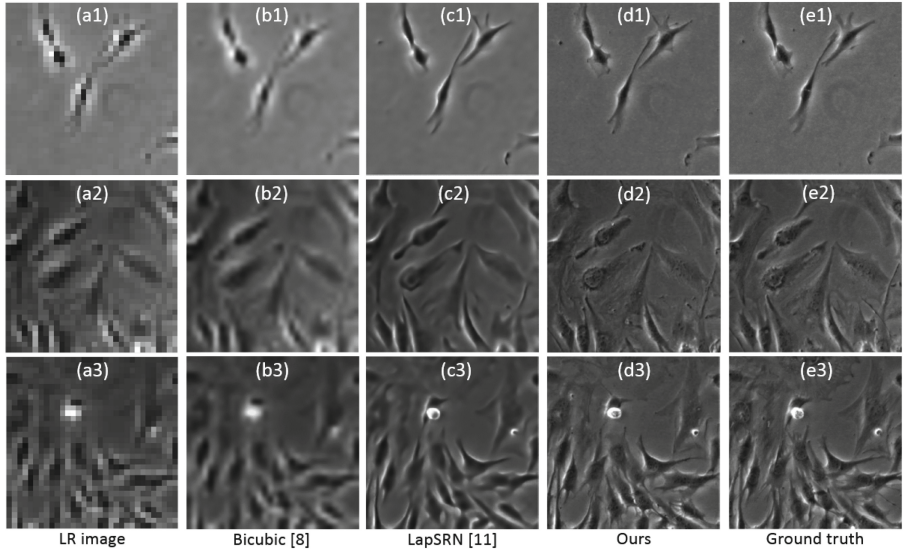


Fig. 4. Upsampling phase contrast images by different algorithms ($8\times$ upsampling).

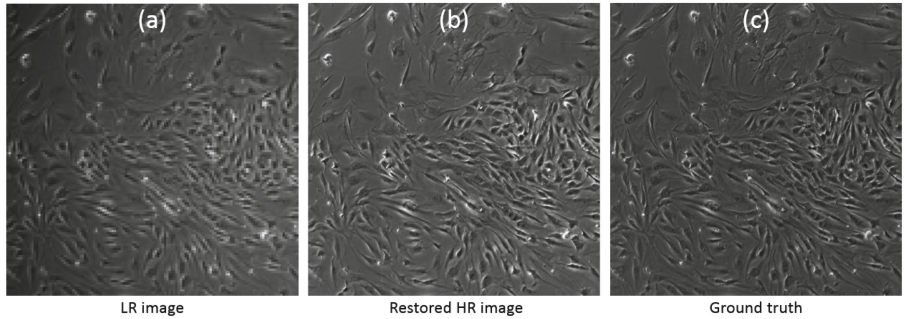


Fig. 5. Upsampling an input phase contrast image with a wide-view. Please zoom-in the online version to observe cell details.

Second, we investigate the effect of the optics-related data enhancement. If using the original image as the input to the generator in Fig. 3 (without enriching the input image by DPFs, the content loss is defined as $\frac{1}{K} \sum_{k=1}^K \sum_{s=1}^S \|y_s - \hat{y}_s\|$ accordingly), the generated images (Fig. 6(d1, d2)) provide sharp images but with some cell details missed, compared to the images generated using the enriched input (Fig. 6(e1, e2)). The enriched phase contrast images can present more feature information of cells, especially when the original input phase contrast image has low contrast and less textures.

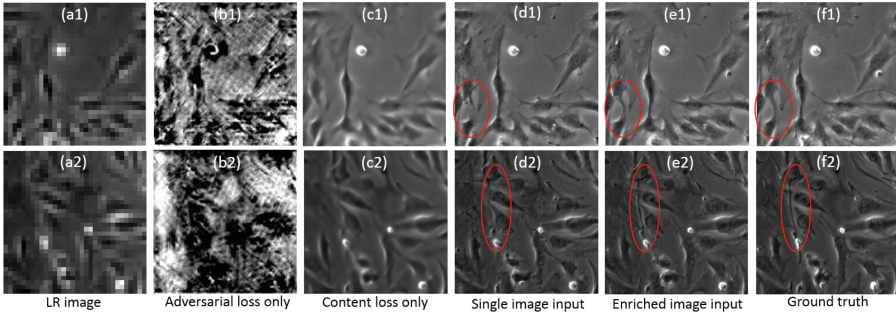


Fig. 6. The effect of different loss functions and the optics-related data enhancement.

The quantitative results of evaluation and ablation study are summarized in Table 1, which shows that our cascaded refinement GAN with optics-related data enhancement and perceptual loss (adversarial loss plus content loss) achieves the best performance.

Table 1. Quantitative evaluation.

	Bicubic [8]	LapSRN [11]	Ours	Content loss only	Adversarial loss only	Single image input
PSNR	20.6039	26.5055	27.8591	26.0928	11.6732	26.2914

6 Conclusion

In this paper, we investigate a super resolution algorithm to generate high-resolution phase contrast microscopy images from the low-resolution ones. Instead of upscaling the input image to the desired resolution directly, the proposed algorithm predicts the high resolution image in a coarse-to-fine fashion, i.e., increasing the spatial resolution of the input image progressively. A new loss function is designed for the proposed algorithm which consists of a content loss and an adversarial loss. The content loss forces the prediction of the high-resolution image to be similar to the real image in the feature domain enriched by optics-related data enhancement, while the adversarial loss encourages the prediction to be sharp and with more cell details.

The experiments demonstrate that our algorithm is very effective to recover high resolution phase contrast microscopy images from low resolution ones, and our algorithm outperforms the current state-of-the-art super-resolution algorithm. The research outcome provides a computational solution on achieving a high magnification on individual cells' details and a wide-view on cell populations at the same time, which will benefit the microscopy community.

Acknowledgement. This project was supported by NSF CAREER award IIS-1351049 and NSF EPSCoR grant IIA-1355406.

References

1. Ba, J.L., et al.: Layer normalization. [arXiv:1607.06450](https://arxiv.org/abs/1607.06450) (2016)
2. Chen, Q., Koltun, V.: Photographic image synthesis with cascaded refinement networks. In: ICCV (2017)
3. Dong, C., et al.: Image super-resolution using deep convolutional networks. TPAMI (2015)
4. Duchon, C.E.: Lanczos filtering in one and two dimensions. *J. Appl. Meteorol.* **18**, 1016–1022 (1979)
5. Freedman, G., Fattal, R.: Image and video upscaling from local self-examples. In: ACM TOG (2011)
6. Goodfellow, I., et al.: Generative adversarial nets. In: NIPS (2014)
7. Isola, P., et al.: Image-to-image translation with conditional adversarial networks. In: CVPR (2017)
8. Keys, R.: Cubic convolution interpolation for digital image processing. In: TASSP (1981)
9. Kim, J., Lee, J.K., Lee, K.M.: Accurate image super-resolution using very deep convolutional networks. In: CVPR (2016)
10. Kingma, D., Ba, J.: Adam: a method for stochastic optimization. In: ICLR (2015)
11. Lai, W.S., et al.: Deep laplacian pyramid networks for fast and accurate super-resolution. In: CVPR (2017)
12. Ledig, C., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: CVPR (2017)
13. Maas, A.L., Hannun, A.Y., Ng, A.Y.: Rectifier nonlinearities improve neural network acoustic models. In: ICML (2013)
14. Odena, A., Dumoulin, V., Olah, C.: Deconvolution and checkerboard artifacts. In: Distill (2016)
15. Shi, W., et al.: Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: CVPR (2016)
16. Su, H., et al.: Cell segmentation in phase contrast microscopy images via semi-supervised classification over optics-related features. In: Medical Image Analysis (2013)
17. Timofte, R., De Smet, V., Van Gool, L.: Anchored neighborhood regression for fast example-based super-resolution. In: ICCV (2013)
18. Yang, J., Lin, Z., Cohen, S.: Fast image super-resolution based on in-place example regression. In: CVPR (2013)
19. Zernike, F.: How I discovered phase contrast. *Science* **121**, 345–349 (1955)