

Blur-Robust Nuclei Segmentation for Immunofluorescence Images

*Devraj Mandal, *Abhishek Vahadane, Shreya Sharma, Shantanu Majumdar

Abstract— Automated nuclei segmentation from immunofluorescence (IF) microscopic image is a crucial first step in digital pathology. A lot of research has been devoted to develop novel nuclei segmentation algorithms to give high performance on good quality images. However, fewer methods were developed for poor-quality images like out-of-focus (blurry) data. In this work, we take a principled approach to study the performance of nuclei segmentation algorithms on out-of-focus images for different levels of blur. A deep learning encoder-decoder framework with a novel Y forked decoder is proposed here. The two fork ends are tied to segmentation and deblur output. The addition of a separate deblurring task in the training paradigm helps to regularize the network on blurry images. Our proposed method accurately predicts the instance nuclei segmentation on sharp as well as out-of-focus images. Additionally, predicted deblurred image provides interpretable insights to experts. Experimental analysis on the Human U2OS cells (out-of-focus) dataset shows that our algorithm is robust and outperforms the state-of-the-art methods.

I. INTRODUCTION & RELATED WORK

Automated nuclei segmentation in a microscopic image has been established as a key first step in digital pathology [1]. The challenge of nuclei segmentation in poor quality microscopic images has not been well-addressed in literature where more focus has been on differentiating touching cells [2][3]. Out-of-focus and motion blur are the two most common microscopic imaging artifacts that can severely degrade image quality. These can increase the chances of misinterpretation of pathological microscopic images, thereby resulting in misdiagnosis [4][5](Figure 1). Blur in microscopic images is prevalent due to the staining process of biopsy slides and image data acquisition. There are very few efforts in literature to directly predict nuclei segmentation from blurry microscopic images. This motivated us to develop a nuclei segmentation algorithm that gives robust performance on both sharp and out-of-focus images.

The study of out-of-focus images have been extensively carried out in [6] [7] [8] [9] [10]. Here, Quality Control (QC) was used to reject out-of-focus images which lead to either rejection of the image with an acceptable blur or tedious re-acquisition of the microscopic image.

A challenge of touching cell resolution in nuclei segmentation for sharp images has been addressed in the following works [12] [13] [14] [15] [16] [17] [18] [19] [20]. Traditional methods are typically unsupervised and based on shape and size constraints. Popular deep learning approaches like [3] [21] model nuclei boundary along with nuclei and background. The CIA-Net [2] address the problem by having a

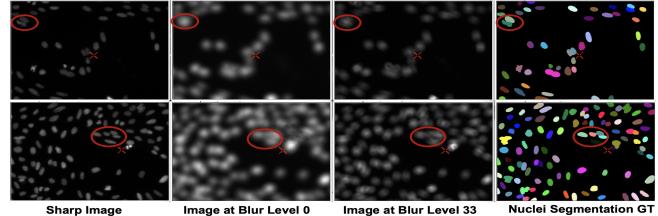


Fig. 1. Example from the Human U2OS cells (out-of-focus) dataset [11] showing the sharp version (1st col) and their out-of-focus images (2nd & 3rd col) along-with their expected ground-truth (last col). The touching nuclei become more difficult to distinguish in out-of-focus images (marked in red).

separate decoder for nuclei contour (boundary) segmentation. Hover-Net [22] proposed three decoders to predict nuclear pixel, horizontal and vertical map, and semantic nuclei segmentation. However, in these approaches, segmentation of blurry images is not addressed.

One of the trivial ways to address the blur is through image restoration techniques [23] [24] [25] [26] [27] [28] and then use the restored image for nuclei segmentation. These approaches incur loss of information in the restoration process which can further limit the nuclei segmentation performance. An interesting approach is to directly predict the nuclei segmentation from out-of-focus images. The DenoiSeg [29] algorithm is one such work which performs joint segmentation and restoration. DenoiSeg extends the idea of CNN3 [3] to segment nuclei, boundary, background separately and uses an additional channel (or the fourth class) in the segmentation model to predict the deblurred images. The network is jointly trained with two losses - cross-entropy loss for the first three outputs (segmentation task) and self-supervised loss for the fourth output (deblurring task). Our approach though motivated by DenoiSeg has many significant differences and was found to outperform DenoiSeg performance on the Human U2OS cells (out-of-focus) dataset [11] by a wide margin. In this work, we have made the following important contributions:

- A novel deep learning architecture with forked **Y** decoder. The fork end represent two outputs in the decoder for segmentation and deblurring objectives.
- Propose a combined loss function by addition of segmentation, deblurring, and regularization losses in supervised setting.
- A novel post-processing scheme for resolving the touching nuclei.
- Extensive ablation studies to understand the effect of proposed algorithm on real-world out-of-focus images as well as on more closely packed (touching) nuclei.
- A QC module in the post-processing scheme for further improvement in performance.

Authors from Rakuten Institute of Technology India, Rakuten Group, Inc.

*These authors (devraj.mandal; abhishek.vahadane@rakuten.com) have contributed equally

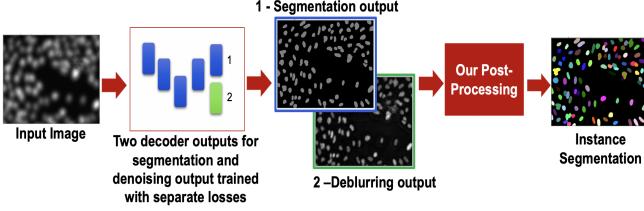


Fig. 2. Block Diagram of Blur Robust Nuclei Segmentation Model

II. PROPOSED MODEL

The proposed flow diagram of Blur-Robust Nuclei Segmentation (BNS) Model is shown in Figure 2. Given the input blur or sharp image I , the encoder-decoder network produces two outputs. The first output is three-channel segmentation I_{nuc} (nuclei, boundary, and background) and the second output is deblurred (sharp) image I_{deblur} . The total loss (Equation 1) used to train our encoder-decoder network is a combination of losses for the nuclei segmentation (Equation 2) and the deblurring (Equation 3) task. We hypothesize that the addition of deblurring loss as a separate task should help to regularize the network and give a better segmentation performance [30]. We use a combined mean squared error (mse) (L_{mse}), L-1 loss (L_{l_1}) for deblurring task and cross-entropy (L_{ce}) loss for nuclei segmentation task. We use different noise models to synthesize realistic blur images from their sharp counterparts and hence L_{deblur} (Equation 3) can be trained in a supervised capacity.

$$L_{total} = L_{nuc} + L_{deblur} \quad (1)$$

$$L_{nuc} = L_{ce}(I_{nuc}, I_{gt}) \quad (2)$$

$$L_{deblur} = L_{mse}(I_{deblur}, I_{sharp}) + L_{l_1}(I_{deblur}, I_{sharp}) \quad (3)$$

We can further post-process the nuclei segmentation output to resolve the touching nuclei whereas, the deblur output can be used by the pathologist for gaining interpretable insights in the results. The nuclei output I_{nuc} is first processed to resolve the touching cells by subtracting the boundary pixels from nuclei predictions. We next use marker-controlled watershed algorithm for further processing. The markers are obtained by computing local maxima of distance transform of I_{nuc} . We then remove the closely located markers by using a threshold. The threshold can be set by studying the minimum distance between the nuclei centers in the training data. We locate the unresolved touching cells by determining the nuclei predictions which are associated with multiple markers. Finally, we replace the original instance with the sub-segments (multiple nuclei) generated by running the marker-controlled watershed algorithm.



Fig. 3. Block Diagram of our Post-Processing Scheme

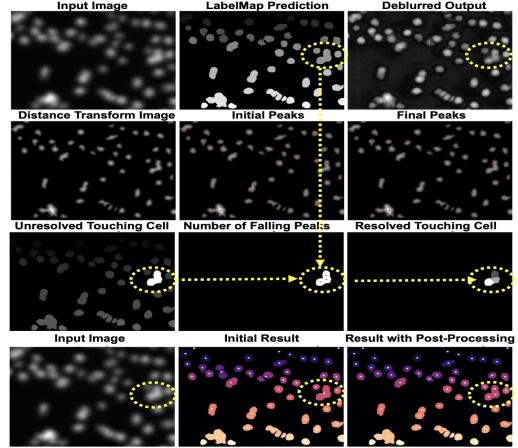


Fig. 4. Visual Illustration of our Post-Processing Scheme

The main steps of our post-processing scheme is shown in Figure 3 and visually illustrated in Figure 4. We show how the unresolved touching cells (marked in yellow) get resolved by taking into account the multiple markers generated from the peaks in the distance map. This greatly enhances our nuclei segmentation model performance.

III. EXPERIMENTAL RESULTS

A. Dataset and Network

In this work, we use the publicly available Human U2OS cells (out-of-focus) dataset [11] made publicly available by Broad Institute. The dataset consists of 32 image sets (z stacks) with $z = 16$ being considered to be the optimal focal plane that gives sharp (in-focus) images. Extra images were also captured at 15 focal plane above and 16 below the optimal plane. The out-of-focus (blur) and in-focus (sharp) image sets are considered to be ($z = 00$ to $z = 10$ and $z = 24$ to $z = 33$ plane) and ($z = 11$ to $z = 23$ plane) respectively [7]. Each image is 696×520 pixels. The instance nuclei in the images were manually annotated by an expert as considered as ground-truth for our experiments. The train and test set were from independent acquisition sites. We consider 10% of the training data to create the validation split. The objective in our work is to build a robust model to perform the segmentation task on both sharp as well as out-of-focus images. We separately measure the performance on the out-of-focus and sharp image sets using the F1-Score (between nuclei prediction and ground-truth) and provides the nuclei segmentation harmonic mean (NHM) to quantify the overall performance. NHM is defined as the harmonic mean between F1-scores for sharp and out-of-focus image sets. A higher NHM indicates that the model is robust to both sharp and out-of-focus images.

Our network is trained with image patch of 512×512 size. The network architecture and parameters such as number of convolution layers and filter were from the popular U-Net framework [31]. Hence, the encoder is same as in the [31]. However, we introduce two outputs in the decoder, one for the segmentation (output size $3 \times 512 \times 512$) and second for the deblur (output size $1 \times 512 \times 512$). These outputs have a separate preceding convolution layer of filter size 1×1

TABLE I

PERFORMANCE EVALUATION OF OUR PROPOSED MODEL.

Model	F1-Blur	F1-Sharp	NHM
U-Net (standard)	.598	.923	.726
U-Net (aug)	.815	.925	.867
BNS-1 (aug)	.830	.926	.876
BNS-2 (aug)	.832	.928	.877
BNS-3 (aug)	.834	.925	.878
BNS-4 (aug)	.845	.933	.887
BNS-4 + QC (aug)	.846	.947	.894

($CL_{1 \times 1}$). In the decoder, other convolution blocks prior to $CL_{1 \times 1}$ were common for both the outputs (hence, the Y fork decoder). This is the proposed architecture for BNS-1. We can have more variations of BNS-1 architectures. We change BNS-1 by separating one more preceding convolution block before $CL_{1 \times 1}$ to get BNS-2. Hence, BNS-2 have separate last convolution block (64 filters) and $CL_{1 \times 1}$ that connects to each respective output in the decoder. In the same way, BNS-3 have last two convolution blocks (64 and 128 filters) and $CL_{1 \times 1}$ separate to the each respective output in the decoder. BNS-4 have last three convolution block (64, 128, and 256 filters) and $CL_{1 \times 1}$ separate for each output. All the filters have size 3×3 if not mentioned explicitly.

B. Results

We consider that out-of-focus data is not available at the training stage since most of the images are acquired at sharp focus. We chose popular U-Net architecture as encoder-decoder as our baseline. The results have been provided in Table I. All models have the same post-processing as BNS for a fair comparison. We make the following important observations from the results. First, we observe that when the U-Net is trained without any augmentation examples (standard), the performance on the out-of-focus image set is significantly lower when compared with the sharp counterparts. It implies that out-of-focus or blur artifact reduces the performance. Next, we investigate the performance of U-Net model when trained with synthetic out-of-focus images generated by different noisy blur kernels (like Gaussian, Poisson, etc.) [7]. We observed that the use of Poisson blur and optical point spread function (psf) [7] gave the highest performance on the real out-of-focus test set which we reported here as U-Net (aug).

The performance of proposed model (BNS-4) uses the Poisson & optical psf [7] as augmentation strategy and gave a significant boost in performance over U-Net (aug) from $(0.815 \rightarrow 0.845, 0.925 \rightarrow 0.933 \text{ and } 0.867 \rightarrow 0.8887)$ on the out-of-focus, sharp and NHM respectively. We also show the performance of different variants of our BNS model in Table I and observed that all variants show better performance than U-Net (aug). Figure 5 shows a more detailed comparison between U-Net (aug) and BNS-4 (aug) to understand the segmentation performance on image sets of different blur levels (z -stack). We also report the relative gain in performance as $= \frac{\text{our} - \text{baseline}}{\text{baseline}} * 100$ (in %), where *our* and *baseline* are the NHM of the U-Net (aug) and BNS-4 (aug) respectively. We observe more relative gain in performance on the more blurry image sets (Figure 5). The

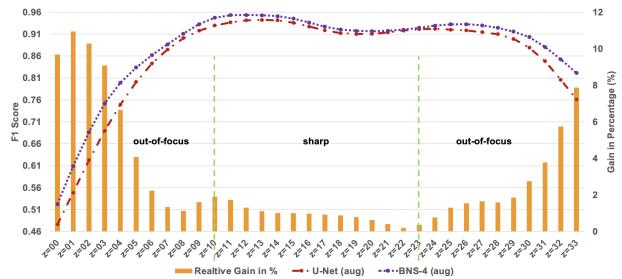


Fig. 5. Comparison of BNS-4 vs U-Net for different blur levels.

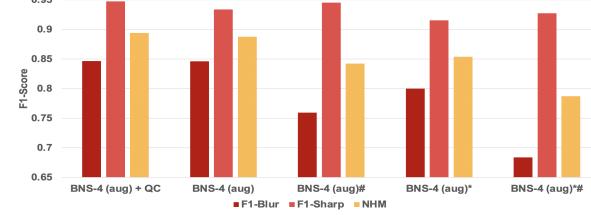


Fig. 6. Ablation studies of proposed BNS-4 model.

TABLE II
PERFORMANCE COMPARISON OF BNS-4 WITH DENOISEG [29].

Model	F1-Blur	F1-Sharp	NHM
DenoiSeg	.677	.869	.761
DenoiSeg*	.722	.878	.793
BNS-4 (aug)	.845	.933	.887

NHM performance of our model trained with *real out-of-focus images along-with their ground-truth* is 0.930.

Finally, we compare the performance of BNS against the state-of-the-art DenoiSeg algorithm [29] and report the results in Table II. We observe that BNS-4 significantly outperforms the DenoiSeg performance with a *relative gain of 14.2%* in NHM performance. Our model has two different paths for the two tasks which helps to regularize the network in a more elegant fashion than the 4-channel output. In addition, we also use a supervised deblur loss instead of self-supervision loss as done in DenoiSeg. We also incorporated our novel post-processing scheme in DenoiSeg (DenoiSeg*) and observed further improvement in the performance. This validates the importance of post-processing to resolve touching nuclei.

Finally, we perform ablation studies for our model and report the results in Figure 6. The # indicates results without post-processing scheme and * indicates that the evaluation has only been performed on difficult test images (with more than 50% touching nuclei). We observe that the performance without the post-processing scheme (#) degrades especially for the blur image sets. The post-processing scheme seems to degrade the performance slightly on the sharp image sets that can be recovered back by training a quality control (QC) module (Table I & Figure 6). The QC was trained to predict a given image as sharp or blur that can be used as a control parameter to determine whether to apply the post-processing scheme or not. We have trained a simple AlexNet model [32] with a two class classification objective (blur or sharp) as the QC module. Finally, we also observe that segmenting closely packed nuclei is harder as shown by the * results.

The proposed model enhances the interpretability of the

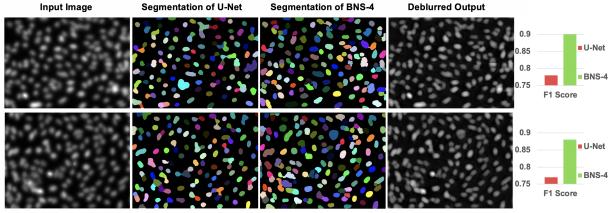


Fig. 7. Visual results showing the interpretability component of our model with the deblurred output clearly showing the different touching cells.

results. The predicted deblurred output of the network can enrich the information present in the input out-of-focus image. This surely helps the expert (pathologist) to interpret and validate our results. We visually illustrate in Figure 7 and show (from left to right), the original input image, the output of the U-Net (aug) and BNS-4 (aug) model along with the deblurred output. We also display the improvement in F1-Score for this example. We observe that the deblurred output is easy to interpret. Specifically, it improves the interpretation of closely packed nuclei which was virtually indistinguishable in the original input image.

IV. CONCLUSION

In this work, we have developed a Blur-Robust Nuclei Segmentation (BNS) model which gives significant performance improvement while segmenting real-world out-of-focus images against the current state-of-the-art. We observe that standard models trained without access to real out-of-focus images are unable to properly segment the nuclei in blurry images. This performance drop is greatly mitigated by our proposed model. In addition our model provides interpretable insights into the results which is of significant value to a pathologist. Extensive experiments with different variants our network architecture & an additional QC module shows the efficacy of our proposed method.

REFERENCES

- [1] Claire McQuin et al. Cellprofiler 3.0: Next-generation image processing for biology. *PLoS biology*, 16(7):e2005970, 2018.
- [2] Yanning Zhou, Omer Fahri Onder, Qi Dou, Efstratios Tsougenis, Hao Chen, and Pheng-Ann Heng. Cia-net: Robust nuclei instance segmentation with contour-aware information aggregation. In *IPMI*, pages 682–693. Springer, 2019.
- [3] Neeraj Kumar et al. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Trans. Med. Imag.*, 36(7):1550–1560, 2017.
- [4] Rafael Redondo et al. Autofocus evaluation for brightfield microscopy pathology. *J. Biomed. Optics*, 17(3):036008, 2012.
- [5] Kangkana Bora, Manish Chowdhury, Lipi B Mahanta, Malay Kumar Kundu, and Anup Kumar Das. Automated classification of pap smear images to detect cervical dysplasia. *Computer methods and programs in biomedicine*, 138:31–47, 2017.
- [6] Mark-Anthony Bray, Adam N Fraser, Thomas P Hasaka, and Anne E Carpenter. Workflow and metrics for image quality control in large-scale high-content screens. *J. Biomolecular Screening*, 17(2):266–274, 2012.
- [7] Samuel J Yang et al. Assessing microscope image focus quality with deep learning. *BMC Bioinf.*, 19(1):1–9, 2018.
- [8] Mark-Anthony Bray and Anne Carpenter. Advanced assay development guidelines for image-based high content screening and analysis. *Assay Guidance Manual [Internet]*, 2017.
- [9] Timo Kohlberger et al. Whole-slide image focus quality: Automatic assessment and impact on ai cancer detection. *J. Pathology Inform.*, 10, 2019.
- [10] Zhongling Wang, Mahdi S Hosseini, Adyn Miles, Konstantinos N Plataniotis, and Zhou Wang. Focuslitenn: High efficiency focus quality assessment for digital pathology. In *MICCAI*, pages 403–413. Springer, 2020.
- [11] Vebjorn Ljosa, Katherine L Sokolnicki, and Anne E Carpenter. Annotated high-throughput microscopy image sets for validation. *Nature methods*, 9(7):637–637, 2012.
- [12] Chenyi Ling, Michael Majurski, Michael Halter, Jeffrey Stinson, Anne Plant, and Joe Chalfoun. Analyzing u-net robustness for single cell nucleus segmentation from phase contrast images. In *CVPR Workshops*, pages 966–967, 2020.
- [13] Daniel Riccio, Nadia Brancati, Maria Frucci, and Diego Gragnaniello. A new unsupervised approach for segmenting and counting cells in high-throughput microscopy image sets. *IEEE J. Biomed. Health Inform.*, 23(1):437–448, 2018.
- [14] Amin Gharipour and Alan Wee-Chung Liew. Segmentation of cell nuclei in fluorescence microscopy images: An integrated framework using level set segmentation and touching-cell splitting. *Pattern Recognition*, 58:1–11, 2016.
- [15] Wanjun Zhang and Huiqi Li. Automated segmentation of overlapped nuclei using concave point detection and segment grouping. *Pattern Recognition*, 71:349–360, 2017.
- [16] Margarita Gamarra, Eduardo Zurek, Hugo Jair Escalante, Leidy Hurtado, and Homero San-Juan-Vergara. Split and merge watershed: A two-step method for cell segmentation in fluorescence microscopy images. *Biomed. Signal Process. Control*, 53:101575, 2019.
- [17] Qiang Zhang, Jinghan Wang, Zaihao Liu, and Dingwen Zhang. A structure-aware splitting framework for separating cell clumps in biomedical images. *Signal Process.*, 168:107331, 2020.
- [18] Costas Panagiotakis and Antonis A Argyros. Cell segmentation via region-based ellipse fitting. In *ICIP*, pages 2426–2430. IEEE, 2018.
- [19] Costas Panagiotakis and Antonis Argyros. Region-based fitting of overlapping ellipses and its application to cells segmentation. *Image Vision Comput.*, 93:103810, 2020.
- [20] Leonid Kostrykin, Christoph Schnörr, and Karl Rohr. Globally optimal segmentation of cell nuclei in fluorescence microscopy images using shape and intensity information. *Med. Image Anal.*, 58:101536, 2019.
- [21] Hirohisa Oda et al. Besnet: boundary-enhanced segmentation of cells in histopathological images. In *MICCAI*, pages 228–236. Springer, 2018.
- [22] Simon Graham et al. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Med. Image Anal.*, 58:101563, 2019.
- [23] Huangxuan Zhao et al. A new deep learning method for image deblurring in optical microscopic systems. *J. biophotonics*, 13(3):e201960147, 2020.
- [24] Sami Koho et al. Fourier ring correlation simplifies image restoration in fluorescence microscopy. *Nature Commun.*, 10(1):1–9, 2019.
- [25] Soonam Lee, Shuo Han, Paul Salama, Kenneth W Dunn, and Edward J Delp. Three dimensional blind image deconvolution for fluorescence microscopy using generative adversarial networks. In *ISBI*, pages 538–542. IEEE, 2019.
- [26] Da He, De Cai, Jiasheng Zhou, Jiajia Luo, and Sung-Liang Chen. Adaptive weighting depth-variant deconvolution of fluorescence microscopy images with convolutional neural network. *arXiv preprint arXiv:1907.03217*, 2019.
- [27] Da He, De Cai, Jiasheng Zhou, Jiajia Luo, and Sung-Liang Chen. Restoration of out-of-focus fluorescence microscopy images using learning-based depth-variant deconvolution. *IEEE Photon. J.*, 12(2):1–13, 2020.
- [28] Adrian Shajkofci and Michael Liebling. Spatially-variant cnn-based point spread function estimation for blind deconvolution and depth estimation in optical microscopy. *IEEE Trans. Image Process.*, 29:5848–5861, 2020.
- [29] Tim-Oliver Buchholz, Mangal Prakash, Deborah Schmidt, Alexander Krull, and Florian Jug. Denoiseg: joint denoising and segmentation. In *ECCV*, pages 324–337. Springer, 2020.
- [30] Sebastian Ruder. An overview of multi-task learning in deep neural networks. *arXiv preprint arXiv:1706.05098*, 2017.
- [31] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, pages 234–241. Springer, 2015.
- [32] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *NIPS*, 25:1097–1105, 2012.