DIGITAL STAINING OF MITOCHONDRIA IN LABEL-FREE LIVE-CELL MICROSCOPY

Ayush Somani[1], Arif Ahmed Sekh[2], Ida S. Opstad[2], Åsa Birna Birgisdottir[3],
Truls Myrmel[3], Balpreet Singh Ahluwalia[2], Krishna Agarwal[2],
Dilip K. Prasad[4], Alexander Horsch[4]

[1]Department of Mathematics and Computing, IIT (ISM) Dhanbad, India
[2]Department of Physics and Technology, UiT The Arctic University of Norway
[3]Department of Clinical Medicine, UiT The Arctic University of Norway
[4]Department of Computer Science, UiT The Arctic University of Norway
`alexander.horsch@uit.no`

Abstract. Examining specific sub-cellular structures while minimizing cell perturbation is important in the life sciences. Fluorescence labeling and imaging is widely used for introducing specificity despite its perturbative and photo-toxic nature. With the advancement of deep learning, digital staining routines for label-free analysis have emerged as a replacement for fluorescence imaging. Nonetheless, digital staining of sub-cellular structures such as mitochondria is sub-optimal. This is because the models designed for computer vision are directly applied instead of optimizing them for the nature of microscopy data. We propose a new loss function with multiple thresholding steps to promote more effective learning for microscopy data. Through this, we demonstrate a deep learning approach to translate the label-free brightfield images of living cells into equivalent fluorescence images of mitochondria with an average structural similarity of 0.77, thus surpassing the state-of-the-art of 0.7 with L1. Our results provide insightful examples of some unique opportunities generated by data-driven deep-learning enabled image translations.

1 Introduction

The study of function and structure of nanoscale sub-cellular organelles such as mitochondria is considered vital for understanding sub-cellular mechanisms of various diseases [1]. Its small size can fall below the diffraction limit [2] and poses a considerable challenge. However, mitochondria specific live-cell friendly fluorescent dyes [3] and high resolution fluorescence microscopes have enabled to image them in high contrast [4]. Nonetheless, there are trade-offs involved with the use of fluorescent dyes. Such dyes are chemical additions to the living cell system. They perturb the natural function and sometimes even damages the cells [5]. Furthermore, the phenomenon of photo-bleaching limits the photon-budget and long-term imaging. Finally, there are practical aspects such as imperfect labeling and expensive fluorescence microscopy instrumentation [6]. Therefore,

Fig. 1. Proposed method for label-free to digitally labeled image translation.
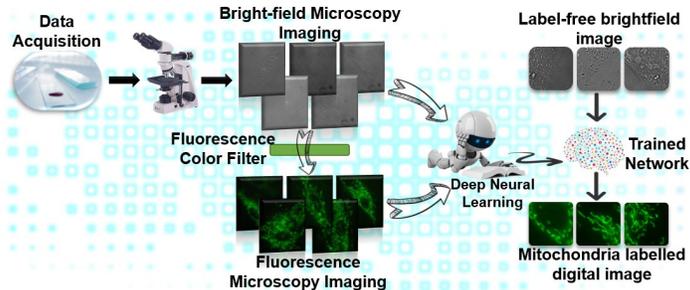
it is of high interest to explore label-free microscopy solutions such as brightfield differential interference and phase-contrast microscopy which use the inherent optical contrast of the sub-cellular structures in the image formation [7]. Unfortunately, the images obtained by these modalities encode the entire cell content, which makes the data difficult to interpret for studies concentrating on only a small portion of the sub-cellular content [8].

This study introduces a deep learning based approach for digital staining of mitochondria in label-free brightfield microscopy images of living cells. It involves the implementation of the conditional generative adversarial network (cGAN) [9] to transfer the unstained embryonic heart-derived cell-line to their corresponding mitochondria stained images, such as shown in Fig. 1. The highlight of the presented work is a novel loss function, custom-designed for learning the structures of interest while deprioritizing the learning for the remaining structures.

2   Materials and methods

Here, we present the data and deep learning approach used in this paper.

Data acquisition: The rat cardiomyoblast cell-line H9c2 (cells derived from embryonic heart tissue; Sigma Aldrich) was cultured in DMEM with 10% FBS on MatTek glass-bottom dishes. The cells were transiently transfected $24 - 48$ hours to express the mitochondrial fluorescence marker eGFP-OMP25-TM. During acquisition, the cells were kept at 37°C, 5% CO2, atmospheric oxygen, and a cell-culture medium (DMEM with 10% FBS). Time-lapse microscopy data were acquired using a DeltaVision OMX V4 Blaze imaging system (GE Healthcare Life Sciences, Marlborough, MA, USA) equipped with a 60X 1.42NA oil-immersion objective (Olympus) and sCMOS cameras. A total of five cells were imaged. Images were acquired by shifting the focal plane along the z-axis and repeating the acquisition of aligned pairs of brightfield and fluorescence images as separate channels using 2s as the temporal rate of revisiting the same plane. Each image is of size 1024×1024 pixels, where each pixel is of size 80 nm. For ensuring a good correlative imaging set-up, the images of the two modes were recorded using two different cameras with inbuilt pre-calibration for registration. Images of five cells were used for the approach. Each aligned image pair was randomly cropped into 25 smaller patches (256×256 pixels). The patches

Fig. 2. An overview of the pipeline to train a Deep Neural Network.

(a) Generator Network   (b) Discriminator Network
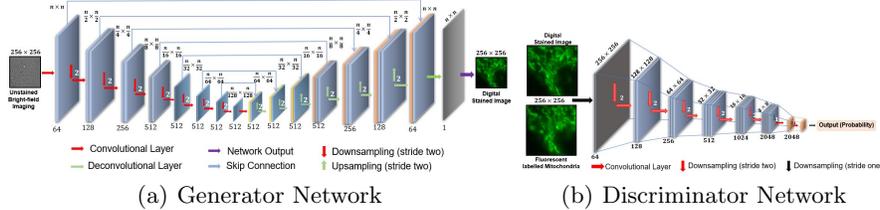
with mitochondrial region less than 20% were discarded, obtaining a total of 8480 correlated pairs of images, which were then proportionally split into 70-20-10% training, validation, and test sets. We had $5,936$ correlated image pairs for training and 848 correlated pairs of images for testing through this exercise.

cGAN architecture: The learning of non-linear mapping from label-free microscopy images to standard fluorescent stained image pair is accomplished using cGAN [9], which is an extension of the generative adversarial network (GAN). The model comprises of subsequent learning of a Generator network (G) and a Discriminator network (D), as shown in Fig. 2 using U-Net [10] as the backbone. G is an encoder-decoder network assigned to generate labeled images from the passed real brightfield image, inferring the probability distribution generating the data. The generator loss function computes the loss by comparing the generated image with the corresponding ground truth real fluorescence image. This generated image is then passed to the discriminator network with a comparatively more straightforward task of learning the rules to distinguish between the images generated by the generator network and the ground truth real of fluorescence images. The discriminator loss function seeks to maximize the probability of how realistic the generated image looks against the input target pair. This follows a zero-sum game of adversarial learning of the G-D network. When both function's losses attain a Nash equilibrium, we expect that the generator can generate close-to-reality digitally labeled images. Pix2pix [11] is used in this work, a variant of aligned image-to-image translation using cGAN. The primary motivation for choosing this model is the flexibility and adaptability to a wide variety of tasks without explicitly defining the relationship. The model learns the explicit density function describing the probability distribution.

Training details: The networks for all the combinations were trained from scratch for 200 epochs, and results were compared using Adam optimizer. Image rescaling was avoided to preserve the image spatial context. ReLu activation was employed and no dropouts were included.

3   Proposed loss function

The design or adoption of an optimal generator loss function is critical for the model's performance and often a challenging endeavor. MinMax and L1 norm based loss function are some conventional loss functions successfully used for low-noise medical images [12]. Sub-cellular fluorescence microscopy has a dominance

Fig. 3. Illustration of the proposed T-N based loss function.
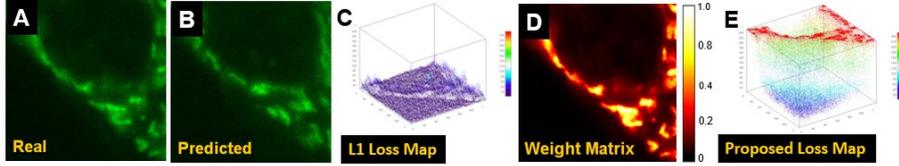
of noisy and empty background due to the nature of fluorescence. This results in the conventional loss functions prioritizing learning on the noisy background and sub-optimally learning the foreground pixels which are few in relative numbers. For example, Fig. 3(A) shows an actual fluorescence image, (B) is a digital candidate image generated by the generator, (C) the 3D rendering of pixel-by-pixel L1 loss shows that the sum of loss in background pixels can be higher compared to the foreground object due to large fraction of noisy background. If the generator minimizes the L1 loss, learning of the noisy background is prioritized. We propose to use a weighted loss function, in which the weight matrix shown in (D) when multiplied with the L1 loss (C) allows to make the net loss more sensitive to the foreground, as shown in (E). The details of the proposed weighted L1 loss appear in the following paragraph.

We propose using a custom multiple thresholding-normalized (TN) loss function [13] that effectively mitigates the problem mentioned above. The proposed loss aims to reduce the contribution (and not completely neglect it) of the noisy background to the loss estimation as compared to the foreground. Since the TN-based loss is not strictly zero in the background, learning of the low-intensity features such as out-of-focus light is still incorporated, only with lesser priority. We compute the weight matrix based on the three-sigma rule to a normal curve that can be derived from Chebyshev's inequality for a wide class of probability distributions. For a normal distribution with mean μ and standard deviation σ, the rule states that $\mu \pm \sigma$ contains $\sim 66.7\%$ of the image measurements, $\mu \pm 2\sigma$ covers 95% measurements and $\mu \pm 3\sigma$ covers 99.7% of the image measurements. The image is subsequently divided into three parts based on μ and σ and shifts the normalization to 0-20%, 20-70% and 70-100%, respectively. The eq. 1 below describes the computation of the weight matrix using the target image F

$$I_{weight} = \sum_{n=1}^{3} \sum_{i=m_{n-1}}^{m_n} \left(a_n \frac{F(x_i, y_i) - M_{n-1}}{M_n - M_{n-1}} + b_n \right) \tag{1}$$

where $m_0 = 0$, $M_0 = \min(F)$, $m_1 = M_1 = \mu + \sigma$, $m_2 = M_3 = \mu + 3\sigma$, $m_3 = 255$, $M_3 = \max(F)$. Further, heuristically chosen coefficients $a_1 = 0.2, b_1 = 0$, $a_2 = 0.5, b_2 = 0.2$, and $a_3 = 0.3, b_3 = 0.7$.

In order to choose the loss function for the discriminator network, we performed a comparative study of a variety of discriminator loss functions considering two options for the generator loss function: the conventional L1 loss and the TN weighted L1 loss (see Table 1). We found that mean square error loss and fo-

Table 1. Study of loss functions for the discriminator network in terms of SSIM.

Generator	Dice + BCE	Jaccard	Focal loss	Mean square error
L1	0.3292	0.3921	0.4058	0.4292
TN-weighted L1	0.5212	0.5877	0.6521	0.6874

cal loss provided a better convergence for both options of the generator loss. MSE loss providing the best result compared to other experimented losses was considered for the discriminator. This pixel-wise loss function emphasis the difference between the foreground and background statistical pixel data for learning.

4   Results

Structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR) were used for performance evaluation. We performed a comparative study for three set-ups, namely T-N(1), T-N(16), and L1, where T-N(x) represents the proposed loss function with batch size x. L1 represents the L1 based vanilla loss function used in state-of-the-art mitochondria labeling [14]. Quantitative results and representative examples are shown in Fig. 4. We consistently observed that the proposed method provides better SSIM than the state-of-the-art, with median SSIM of 0.77 (TN) being significantly better than the state-of-the-art (0.7 of L1). PSNR shows a similar trend, although not as prominently as SSIM. In terms of computation time, the digital staining takes 34 milliseconds per image of size 256×256 on average, using a standard desktop computer equipped with a single GPU. Even with a relatively moderate computer, the fast inference time suggests that the proposed digital staining can be easily integrated with an automated microscope.

SSIM Metric

	Q95	Q75	Q50	Q25
T-N (1)	0.81	0.79	0.77	0.71
T-N (16)	0.80	0.78	0.76	0.72
L1	0.77	0.72	0.70	0.67

PSNR Metric

	Q95	Q75	Q50	Q25
T-N	29.5	28.4	27.6	26.6
T-N (16)	29.1	28.3	27.5	26.4
L1	28.7	28.4	27.5	27.3

Fig. 4. Quartile comparision of different learning methods.

5 Discussion and Conclusion

This paper presents a method for artificially labelling mitochondria from a label-free brightfield microscopy image. The proposed TN loss function integrated with the adversarial network approach performs significantly better than the state-of-the-art. We envision this method to save researchers time, labor and costs, and enable better studies of cells without photo-toxicity, photo-bleaching, and perturbation of the natural cell composition. Since the method is data-driven, it does not need prior information regarding the experimental settings such as the camera, precise microscopy modality or the dyes used during experimentation. The model's precision can likely be further enhanced by training the network on a larger dataset with a higher number and variety of cells. We note that the presented digital staining approach can, in a straight forward manner, also be combined with other microscopy modalities by following a similar approach.

References

1. Samanta S, He Y, Sharma A, et al. Fluorescent probes for nanoscopic imaging of mitochondria. Chem. 2019;5(7):1697–1726.
2. Swayne TC, Gay AC, Pon LA. Methods Cell Biol.. vol. 80. Academic Press; 2007.
3. Chazotte B. Labeling Mitochondria with Fluorescent Dyes for Imaging. Cold Spring Harb Protoc. 2009;2009(6).
4. Kandel ME, Hu C, Kouzehgarani GN, et al. Epi-illumination gradient light interference microscopy for imaging opaque structures. Nat Commun. 2019;10(1):1–9.
5. Hoebe R, Van Oven C, Gadella TW, et al. Controlled light-exposure microscopy reduces photobleaching and phototoxicity in fluorescence live-cell imaging. Nat Biotechnol. 2007;25(2):249–253.
6. Christiansen EM, Yang SJ, Ando DM, et al. In silico labeling: predicting fluorescent labels in unlabeled images. Cell. 2018;173(3):792–803.
7. Zahedi A, On V, Phandthong R, et al. Deep analysis of mitochondria and cell health using machine learning. Sci Rep. 2018;8(1):1–15.
8. Vicar T, Balvan J, Jaros J, et al. Cell segmentation methods for label-free contrast microscopy: review and comprehensive comparison. BMC Bioinformatics. 2019;20(1):360.
9. Mirza M, Osindero S. Conditional Generative Adversarial Nets. arXiv preprint arXiv:14111784. 2014;.
10. Ronneberger O, Fischer P, Brox T; Springer. U-net: Convolutional networks for biomedical image segmentation. 2015; p. 234–241.
11. Isola P, Zhu JY, Zhou T, et al. Image-to-Image Translation with Conditional Adversarial Networks. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. 2017; p. 5967–5976.
12. Armanious K, Jiang C, Fischer M, et al. MedGAN: Medical image translation using GANs. Comput Med Imaging Graph. 2020;79:101684.
13. Kotte S, Kumar PR, Injeti SK. An efficient approach for optimal multilevel thresholding selection for gray scale images based on improved differential search algorithm. Ain Shams Med J. 2018;9(4):1043–1067.
14. Ounkomol C, Seshamani S, Maleckar MM, et al. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. Nat Methods. 2018;15(11):917–920.