Deep Learning Methods for Image Decomposition of Cervical Cells

Tayebeh Lotfi Mahyari
Department of Systems and Computer Engineering
Carleton University
Ottawa, Canada
tayebeh.lotfimahyari@carleton.ca

Richard M. Dansereau
Department of Systems and Computer Engineering
Carleton University
Ottawa, Canada
rdanse@sce.carleton.ca

***Abstract*—One way to solve under-determined image decomposition is to use statistical information about the type of data to be decomposed. This information can be obtained by a deep learning where convolutional neural networks (CNN) are a subset recently used widely in image processing. In this paper, we have designed a two-stage CNN that takes cytology images of overlapped cervical cells and attempts to separate the cell images. In the first stage, we designed a CNN to segment overlapping cells. In the second stage, we designed a CNN that uses this segmentation and the original image to separate the regions. We implemented a CNN similar to U-Net for image segmentation and implemented a new network for the image separation. To train and test the proposed networks, we simulated 50000 cervical cell cytology images by overlaying individual images of real cervical cells using the Beer-Lambert law. Of these 50000 images, we used 49000 images for training and evaluated the method with 1000 test images. Results on these synthetic images give more than 97% segmentation accuracy and gives decomposition SSIM scores of more than 0.99 and PSNR score of more than 30 dB. Despite these positive results, the permutation problem that commonly effects signal separation occasionally occurred resulting in some cell structure mis-separation (for example, one cell given two nucleoli and the other given none). In addition, when the segmentation was poor from the first stage, the resulting separation was poor.**

***Index Terms*—Machine learning; deep learning; image segmentation; image separation; translucent overlapped images.**

I. Introduction

Cervical cancer is one of the dangerous cancers in women around the world. Early detection of cancerous cells helps to improve the chance of treatment and some tests such as pap smear or Thinprep are important routine tests in screening to diagnose cervical cancer. Although cervical cell segmentation is a strong tool to detect abnormality of the nucleus and cytoplasm, it is difficult due to partial overlap of the cells. Also comparing the image values of the cervical cells give more information in a computer based diagnostic system but it needs cervical cell decomposition.

The goal in this work is to perform image separation of image structures and objects that overlap in the imaging plane but exhibit lucency/diaphaneity that could potentially be separated. Examples of imaging where object lucency commonly occurs includes cell cytology imaging, x-ray imaging, positron emission tomography imaging, hyperspectral imaging, photographing semi-reflective/semi-transparent materials, and document scanning. In this paper, we focus on the problem of separating overlapping cells in cervical cell cytology images.

Reviewing the literature, one form of image mixtures occurs when imaging through glass or some semi-reflective/semi-transparent medium that creates overlays of image information where the scene on the other side of the glass is mixed with a reflection on the glass. In [1], a joint diagonalization parameter estimation approach is taken for a frequency-domain separation. In [2], this type of reflection mixture is handled using a cyclic permutation approach while in [3], a generalized mixture ratio is estimated for the image separation. In [4], an iterative sparse blind separation algorithm is employed.

Other image separation techniques rely on multiple images or separate characteristics of the images. In [5], bleed-through in two-sided document scanning is considered where scans of either side of a document reveal shared elements with different cross-mixture levels. In [6] and [7], hyperspectral image information is used to help separate image elements based on the different spectral characteristics of those image elements. In [8], the two different but correlated images captured from dual energy x-ray imaging are used when performing image separation. In [9], information from x-ray imaging and visible-light imaging of a painting are used to reveal and separate hidden paintings on canvases that have been painted over.

Four different common approaches for image separation are principal component analysis (PCA), independent component analysis (ICA), singular value decomposition (SVD), and non-negative matrix factorization (NMF) [10]. However, none of PCA, ICA and SVD guarantee to have non-negative elements, and therefore the obtained source signals might not have physical meaning [10].

NMF is a strong tool that is being used in matrix decomposition where the decomposed elements are nonnegative [11]. Three common algorithms for image separation using NMF are the multiplicative update (MU) algorithm, alternating least squares (ALS) algorithm, and projected gradient (PG) algorithm [12], [13]. However, since the problem is not convex, there will be no promise to find a unique solution to the problem. One way to overcome this non-uniqueness issue is

to add more terms as constraints to the cost function in order to penalize unwanted solutions.

The main issue in these four categories of approaches is they usually need the number of observations to be equal to or greater than the number of sub images. However in some applications we need to separate a single image into two or more sub images and therefore the above methods will fail to give good results for these under-determined problems.

One solution to overcome the lack of information is to use statistical information about the data type that is going to be separated. Machine learning, a method that gives the ability to learn computers, formally focuses on the theory, performance, and properties of learning systems and algorithms [14]. However in all learning algorithms one of the basic and fundamental steps is extracting useful features from data.

Deep learning methods [15]–[17] have shown to find useful features with the cost of high learning time. Using the idea of the structure of human brain, deep learning methods use complex structures and are fed a lot of information. Given enough training, these structures automatically find important features and learn based on these extracted features. Deep learning has emerged as a new field of research in the machine learning since 2006 and has flourished in the past decade as a powerful tool in many learning applications such as signal processing.

One of the popular deep network structures for deep learning is the convolutional neural network (CNN). CNNs have been used widely in image processing fields such as computer vision, medical diagnosis, and robot remote controls.

One drawback for CNNs is the need of a large number of training data. However in our case the training dataset is limited. We focused on cervical cell cytology images and created synthetic images close to the real images by using Beer-Lambart law [18].

In this paper, we implement a two-stage CNN for image decomposition. We first segment partially overlapped translucent regions in the image and then use the segmentation information for the second part of the network to separate the regions. We implemented a network similar to UNet [19] for image segmentation and implemented a new network for image decomposition. We called our proposed two stage network UResNet because it has skip connections similar to UNet and it also has residual blocks.

Our proposed CNN and methods to create the synthetic datasets are explained in Sec. II and Sec. III. In Sec. IV, we evaluate the proposed method and conclusions are given in Sec. V.

II. Proposed CNN for Image Decomposition

Our proposed two-stage CNN is shown in Fig. 1 where we assume only two cells in the image resulting in two segmented cell images as output. Given the challenge of segmenting and separating cell images, an assumption of only two cells is made to simplify the problem in developing our initial proof-of-concept; extensions to generalize the cell separation in more complex images is left as future work.

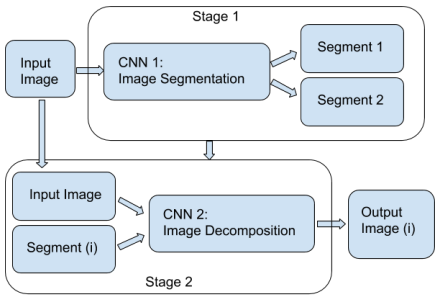

Fig. 1. Proposed two-stage CNN.

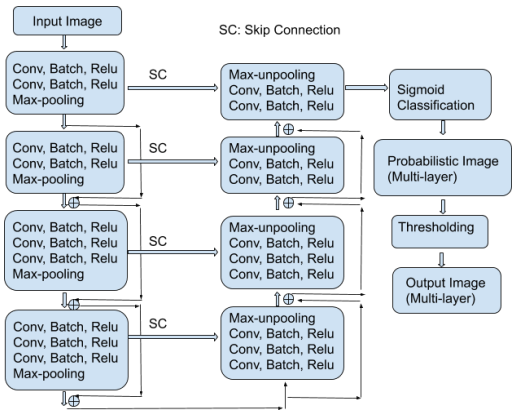

Fig. 2. Proposed residual network for image segmentation.

A. Two-Stage CNN for Image Decomposition

The first stage of the network, CNN1, uses our work in [20] applied to segment translucent partially-overlapped cervical cells in cytology images. The network structure in this first stage is shown in Fig. 2 where the cytology image with the two overlapping cells is fed in as input and the output is two images, each containing the segmentation of one of the cells.

Our proposed first stage for image segmentation consists of an encoder and a decoder. Both encoder and decoder have four blocks. First and second blocks at the encoder start with two sets of convolutional layers of size 3×3 followed by a batch normalization layer and a Relu layer. At the end of each block there is a max-pooling layer. Third and fourth blocks at the encoder consist of three sets of convolutional, batch normalization and Relu layers. The decoder structure is similar

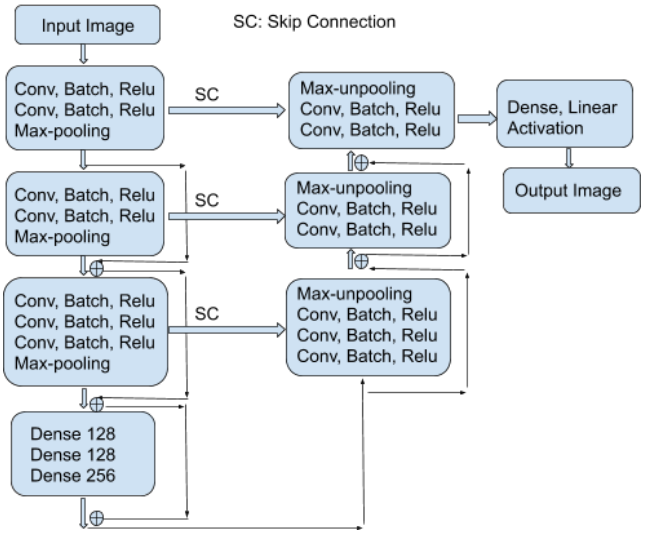

Fig. 3. Proposed residual network for image decomposition.

to encoder except that its blocks start with max-unpooling layer followed by convolutional layers, batch normalization layers and Relu layers. Similar to UNet [19] there is a skip connection between each block in the encoder to its peer block at the decoder. The convolutional layers for both encoder and decoder have $64, 128, 256$ and 512 number of filters for the four blocks respectively.

The second stage of the CNN, CNN2, performs the cell image separation. The CNN used is shown in Fig. 3 where the input is the original input image plus the image segmentation masks for each cell as determined from the first stage of the CNN. The output is two images with each image providing an individual cell that is segmented and separated. Similar to the first stage network, the image decomposition network consists of an encoder and a decoder each having four blocks. The first three blocks of encoder and decoder are similar to the first stage network, but the fourth layers in the second stage network consist of dense layers.

III. Data for Overlapped Cell Image Separation

Since training a CNN works best with a large amount of data, we created images from individual real cervical cells images that we synthetically manipulated, overlapped, and combined. We extracted 24 individual cervical cell images from the ISBI 2014 Overlapping Cervical Cytology Image Segmentation Challenge dataset [21]. Figure 4 shows a few samples from the dataset.

With these cell images, we applied random affine transforms to increase variation in our cell sample population. Then, we created synthetic images by overlapping and combining two cell images using the Beer-Lambert law [18].

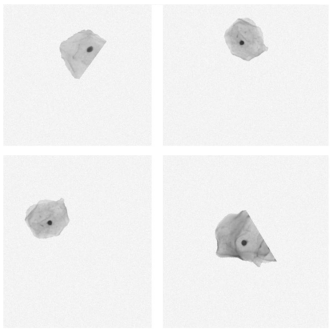

Fig. 4. A few samples of individual cervical cells from the ISBI 2014 Overlapping Cervical Cytology Image Segmentation Challenge dataset [21].

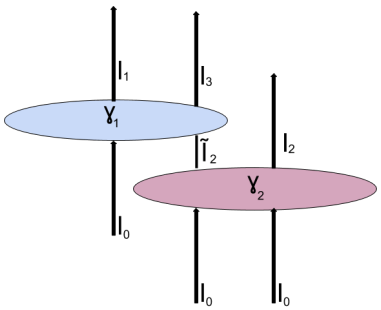

Fig. 5. Beer-Lambert law for two cells [18].

The Beer-Lambert law describes the attenuation of light traveling through materials [18]. Given a background light intensity of I_0, molar attenuation coefficient of objects as ϵ, molar concentration of the attenuating species in the material as γ, traveling distance of light through the object as L and the transmitted light from the object as I_T, the Beer-Lambert law gives the optical intensity profile as [18]

$$I_T = I_0 e^{-\epsilon\gamma L}. \tag{1}$$

Assume that we have two cells, C_1 and C_2, with the transmitted lights I_1 and I_2. Assume $\tilde{I}_2$ as the transmitted light through C_2 in the overlapped area and I_3 as the transmitted light through overlapped area of C_2 and then C_1 (see Fig. 5). By applying the Beer-Lambert law, we will have

$$\begin{aligned} I_1 &= I_0 e^{-\epsilon_1\gamma_1 L_1} \\ I_2 &= I_0 e^{-\epsilon_2\gamma_2 L_2} \\ \tilde{I}_2 &= I_0 e^{-\alpha_2\epsilon_2\gamma_2 L_2} \\ I_3 &= \tilde{I}_2 e^{-\alpha_1\epsilon_1\gamma_1 L_1} \end{aligned} \tag{2}$$

where L_1 and L_2 are the average thickness of C_1 and C_2 and α_1 and α_2 are the coefficients to control the difference of the thickness in the overlapping area [18]. Therefore, I_3 is obtained as [18]

$$I_3 = I_0^{1-\alpha_1-\alpha_2} I_1^{\alpha_1} I_2^{\alpha_2}. \tag{3}$$

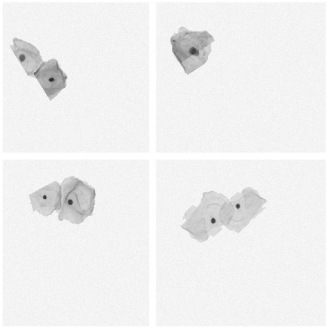

Fig. 6. A few synthetic mixtures created using Beer-Lambert law using (3).

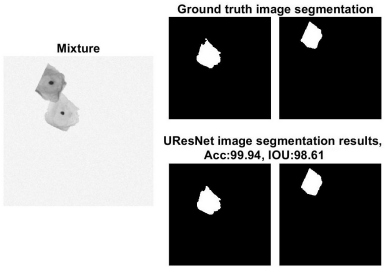

Fig. 7. Results for the segmentation CNN. Left: original mixture, top right: ground truth segmentation, bottom right: segmentation results.

Following the above approach, we generated 50000 synthetic images by overlapping and combining real cervical cell images using (3). In (3), we averaged over the background area for all the training images in the ISBI 2014 dataset [21] to find I_0 and used $\alpha_1 = 0.95$ and $\alpha_2 = 0.9$. We rotated and shifted the cells in ISBI 2014 dataset [21] and the ratio of overlap for the two cells in each image was randomly made. Fig. 6 shows a few synthetic mixtures we created using (3).

IV. Results

To evaluate our proposed network, we trained a CNN using Python Keras package with 50 epochs using 49000 learning data with image size 256×256. We used a learning rate of 0.001 and Adam optimizer [22] with first and second momentum parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$. Figures 7 and 8 show the segmentation and separation results for the proposed network, respectively.

We calculated the segmentation accuracy using

$$Acc = \frac{|\Sigma_G = \Sigma_R|}{N} \times 100\% \tag{4}$$

where $\Sigma_G = \Sigma_R$ is a set of nodes where the ground truth segmentation Σ_G is the same as the results Σ_R and N is the total number of pixels in the image. We also measured the intersection of union (IoU) using

$$IoU = \frac{\Sigma_G \cap \Sigma_R}{\Sigma_G \cup \Sigma_R} \times 100\% \tag{5}$$

where $\cap$ and $\cup$ represent intersection and union of ground truth segmentation and segmentation results, respectively.

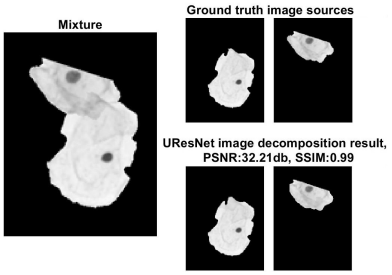

Fig. 8. Results for the separation CNN. Left: original mixture, top right: ground truth original sources, bottom right: decomposition results. The PSNR and SSIM values are averaged over the two decomposed images.

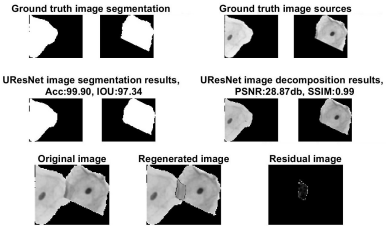

Fig. 9. Good image segmentation leads to good image decomposition and small residual image.

For image separation performance, we measured peak signal-to-noise ratio (PSNR) and structural similarity measure (SSIM) [23].

To the best of our knowledge there is no work done for decomposing partially overlapped cervical cells. But, for the segmentation part we compared our proposed segmentation method with two top segmentation methods for overlapping cervical cells [24]–[26]. Table I shows the accuracy of the segmentation results calculated by (4) and (5). The results are averaged over 1000 test images. It is seen that our method outperforms the two top methods on our dataset. Table II shows the PSNR and SSIM values for the decomposition results. The results are averaged over 1000 test images.

Although our proposed image decomposition method works for most of the cases in our dataset, it fails in some cases. Because our method has two stages, the results for the image decomposition stage depend on image segmentation results from the first stage. Therefore poor image segmentation results lead to poor image decomposition results. Figures 9 and 10 compare the results of image decomposition when the image segmentation is good and poor, respectively.

V. Conclusion and Future Works

In this paper, we used deep learning for overlapped cervical cell image separation. Although deep learning is widely used for other image applications, we believe that there is no work done yet for this type of image separation. We implemented a two-stage CNN for image decomposition. In the first stage, the network segments the partially overlapped translucent regions and this segmentation information is then fed into the second

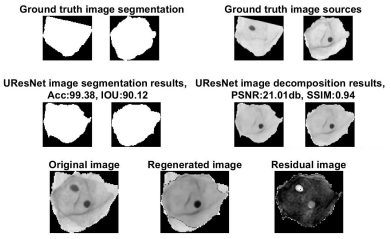

Fig. 10. Poor image decomposition and large residual image as a result of poor image segmentation.

TABLE I. Segmentation accuracy.

Description	Accuracy (%)		IOU (%)	
	mean	std	mean	std
Proposed Segmentation	99.87	0.11	96.95	2.35
Lu et al. [24]	99.13	1.26	84.16	18.54
Phoulady et al. [25], [26]	99.09	0.65	80.83	13.32

part of the network for image separation. Results for the synthetic images show segmentation accuracy of greater than 97% and decomposition SSIM score of more than 0.99.

However because the separation proposed model consists of two stages, poor results for the first stage lead to poor image decomposition results. In the future we intend to modify the network to address this problem. We will explore the situations where image decomposition results are poor and try to improve the results either by improving the image segmentation stage or by adding a post processing method to modify the results.

VI. Acknowledgment

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] E. Be'ery and A. Yeredor, "Blind separation of superimposed shifted images using parameterized joint diagonalization," IEEE Transactions on Image Processing, vol. 17, no. 3, pp. 340–353, 2008.

[2] K. Hara, K. Inoue, and K. Urahama, "Separation of layers from images containing multiple reflections and transparency using cyclic permutation," in 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Apr 2009, pp. 1157–1160.

[3] ——, "Generalized mixture ratio based blind image separation," IEEE Signal Processing Letters, vol. 20, no. 8, pp. 743–746, 2013.

[4] W. Souidene, A. Aissa-El-Bey, K. Abed-Meraim, and A. Beghdadi, "Blind image separation using sparse representation," in 2007 IEEE International Conference on Image Processing (ICIP), vol. 3, Sep 2007, pp. 125–128.

[5] B. Ophir and D. Malah, "Show-through cancellation in scanned images using blind source separation techniques," in 2007 IEEE International Conference on Image Processing (ICIP), vol. 3, Sep 2007, pp. 233–236.

[6] O. Tichý and V. Smidl, "Variational blind source separation toolbox and its application to hyperspectral image data," in European Signal Processing Conference (EUSIPCO), August–September 2015, pp. 1326–1330.

[7] S. Arberet, "Hyper-demix: Blind source separation of hyperspectral images using local ML estimates," in IEEE International Conference on Image Processing (ICIP), Sep 2010, pp. 1393–1396.

[8] Y. Chen, M. Maitre, and T. Fang, "Transparent layer separation for dual energy imaging," in IEEE International Conference on Image Processing (ICIP), Oct 2008, pp. 821–824.

TABLE II. Decomposition accuracy.

Description	PSNR (dB)		SSIM	
	mean	std	mean	std
Proposed Decomposition	28.33	1.88	.9855	0.0067

[9] N. Deligiannis, J. F. C. Mota, B. Cornelis, M. R. D. Rodrigues, and I. Daubechies, "X-ray image separation via coupled dictionary learning," in 2016 IEEE International Conference on Image Processing (ICIP), Sep 2016, pp. 3533–3537.

[10] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," IEEE Transactions on Geoscience and Remote Sensing, vol. 45, no. 3, pp. 765–777, 2007.

[11] Y. X. Wang and Y. J. Zhang, "Nonnegative matrix factorization: A comprehensive review," IEEE Trans. Knowl. Data Eng., vol. 25, no. 6, pp. 1336–1353, 2013.

[12] A. Janecek, S. S. Grotthoff, and W. N. Gansterer, "LIBNMF – a library for nonnegative matrix factorization," Computing and Informatics, vol. 30, pp. 205–224, 2011.

[13] C. J. Lin, "Projected gradient methods for nonnegative matrix factorization," Neural Comput., vol. 19, no. 10, pp. 2756–2779, 2007.

[14] J. Qiu, Q. Wu, G. Ding, Y. Xu, and S. Feng, "A survey of machine learning for big data processing," EURASIP Journal on Advances in Signal Processing, vol. 2016, no. 67, 2016.

[15] F. Jiang, A. Grigorev, S. Rho, Z. Tian, Y. Fu, W. Jifara, K. Adil, and S. Liu, "Medical image semantic segmentation based on deep learning," Neural Computing and Applications, vol. 29, March 2018.

[16] M. S. Badea, I. I. Felea, L. M. Florea, and C. Vertan, "The use of deep learning in image segmentation, classification and detection," CoRR, vol. abs/1605.09612, 2016. [Online]. Available: http://arxiv.org/abs/1605.09612

[17] J. Jotheeswaran and S. Tanwar, "Survey on deep learning for medical imaging," vol. 5, pp. 1608–20, 12 2018.

[18] J. Zhang, Z. Hu, G. Han, and X. He, "Segmentation of overlapping cells in cervical smears based on spatial relationship and overlapping translucency light transmission model," Pattern Recognition, vol. 60, 05 2016.

[19] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," CoRR, vol. abs/1505.04597, 2015. [Online]. Available: http://arxiv.org/abs/1505.04597

[20] T. L. Mahyari and R. M. Dansereau, "Deep learning methods for image segmentation containing translucent overlapped objects," in 2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Nov 2019, pp. 1–5.

[21] "Overlapping cervical cytology image segmentation challenge - ISBI 2014," https://cs.adelaide.edu.au/~carneiro/isbi14_challenge/.

[22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," International Conference for Learning Representations (ICLR), 2014.

[23] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612, 2004.

[24] Z. Lu, G. Carneiro, and A. Bradley, "An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells," IEEE Transactions on Image Processing, vol. 24, 01 2015.

[25] H. Ahmady Phoulady, D. B. Goldgof, L. O. Hall, and P. R. Mouton, "A new approach to detect and segment overlapping cells in multi-layer cervical cell volume images," in 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), April 2016, pp. 201–204.

[26] H. A. Phoulady, D. B. Goldgof, L. O. Hall, and P. R. Mouton, "A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images," Computerized Medical Imaging and Graphics, vol. 59, pp. 38–49, 2017.