Cell segmentation in colon tissue images using deep learning.

Zilli Matias Emanuel.
16/12/2016
Óbudai Egyetem University - Budapest, Hungary
Erasmus+ Programme

1 Introduction

Image segmentation is one of the most challenge tasks of computer vision. Its main purpose is to split the input image into parts, and then the identification of those parts. Respect to medical images, there are several methods to analyze histopathology samples, we can find many publications about new methods even if we only look for a specific subtopic, for example cell nuclei segmentation in colon tissue.

Computer aided image analysis of histopathology images could provide help for early detection and improved characterization of diseases such as breast cancer, pancreatic neuroendocrine tumor, and colon cancer.

Automated nuclei segmentation of cell images is a pre-requisite for various analyses including automatic morphological feature computation and classification. However, it remains to be a challenging problem due to, noises, shape, intracellular intensity heterogeneity, artifacts (e.g. blurred regions) introduced during image acquisition, potential poor contrast between the foreground and the background, also there exist significant variations on nuclei/cell size. Also, nuclei/cells are often clustered into clumps so that they might partially overlap with one another and finally, complex nature of histopathology stained with the standard hematoxylin and eosin images.

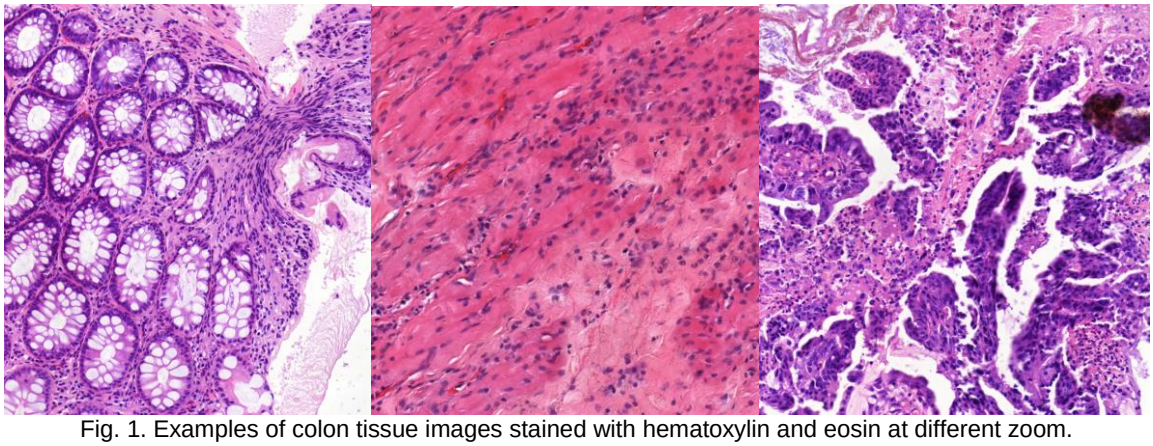


Fig. 1. Examples of colon tissue images stained with hematoxylin and eosin at different zoom.

Deep learning based techniques have been shown to produce promising results on histopathology images in various studies including mitosis detection, tissue classification, and immunohistochemical staining cell segmentation. It is ideally to learn in an implicit manner the diversity of patterns embedded within large datasets of images. Otherwise,

employing a feature engineering or "hand-crafted" approach may require more algorithmic load, and substantial effort to capture a similar gamma of characteristics. Many hand-crafted feature based approaches are not implicitly suitables to manipulate and distil large datasets in a efficiently manner. Deep learning approaches, on the other hand, run properly under these scenarios.

In this work, we propose a supervised deep learning-based model for accurate automatic cell nuclei segmentation. Given a tissue image, it begins with a deep convolutional neural network model to generate a probability map. Next, a threshold and morphological operations are applied to distinguish the background and the cells. One of the significant benefits of the proposed method is that it can be applicable to different staining histopathology images taken of different patients. Due to the feature learning characteristic of deep convolutional neural network and the high level shape prior modeling, the proposed method is general enough to work properly across different image scenario like healthy, adenoma, hyperplasia. Finally, we will validate the proposed algorithm on several histopathology images using a range of different tissue from various patients with differents diseases.

2 Project justification

Nowadays, computer aided image analysis, can benefit pathologists and patients and has attracted many attentions in research and clinical practice. In comparison with manual evaluation that is labor intensive and takes much time, computer aided image analysis can supply faster and reproducible analysis such that the science and clinician researchers can be released from boring and repeated routine to analyze images. Moreover, the complex nature of histopathology images presents significant challenges in manual image analysis, which might lead to inter pathologists variations; In addition, computer aided analysis can greatly reduce the inter pathologist variation providing an accurate diagnostic of diseases., it allows early detection of disease and personalized treatments that benefit the patients. Another advantage of automated methods is that they can provide measurements of important image features, which can be used in the clinical followup, and thus allow to do comparatives studies and provide a personalized medicine to the patient. A critical pre-requisite in computer aided diagnosis is cell nuclei segmentation, because of this, it is of interest develop a method to segment cell nuclei.

3 Proposed algorithm

3.1 Main deep learning approach

This section is divided into two parts. The first subsection describes the main deep learning method used to get the final segmentation result of tissue images. The second subsection details the methods used to construct a semi annotated ground truth to train the deep learning method.

Neural network architecture

To design a suitable neural network to accomplish a specific task, one needs to take various decisions such as number of layers, input patch size, types of layers, and layer attributes. We mitigate this dependency by instead opting a network inspired in the popular and successful VGG network[16]. The VGG approach has a design, where, as the network gets deeper and the input size shrinks, it increase the number of kernels so that essentially each layer is computed in the same amount of time. This affords us a better ability at predicting how much time the training and testing will take.

The network used accepts patches of 64 x 64, and contains interesting features. The first is the use of batch normalization layers before the activation layers, it allows us to use higher learning rates and be less careful about network initialization. It also eliminate the need for dropout.

The second interesting feature is this network does not have fully connected layers and replace them with convolutional layers. The benefit of this feature is that we can directly use the network to compute the output of a given image (after modifying to adjust for input image size) without having to do any network changes.

Patch generation

When the network is defined, the next step is generate image patches to construct the training and testing sets. This stage requires attention in order to ensure that the representation of a high diversity of characteristics presents in the image be included in the training set. Selecting non appropriate image patches for the specific task could have a dramatic consequence in the final result. Especially in the field of histopathology, where can be considerable variance present within a single class, such as nuclei. This is particularly pronounced in colon cancer nuclei, where nuclei areas can vary more than 100% between nucleis. Ensure that an enough set of patches is extracted from the images is one of the most key aspects in the final method effectivity.

The patch selection technique involves selecting patches from the positive and negative class. In our case the images do not come with a corresponding fully annotated ground truth segmentation map for nuclei (positive class) and non-nuclei (negative class). To get an approach to ground truth we use the unsupervised method described below, it is able of reach enough accuracy in generate a segmentation map to use in the training stage.

To augment the available training data with the objective of obtain enhance nuclei boundaries, we complement the approach, discussed above, with additional patches for the negative class generating an edge mask by morphological dilation of ground truth map. With the obtained dilated mask, we can select the negative patches, which are difficult to learn due to their similarity with the positive class. In addition, we generate rotations of 0 and 90 degrees to each patch, it is important to teach the network the desired robustness and invariance properties.

This patch selection approach gives a result which contain more separated nuclei and delineated boundaries.

Fig. 2 shows an example image with its associated semi annotated ground truth segmentation map.

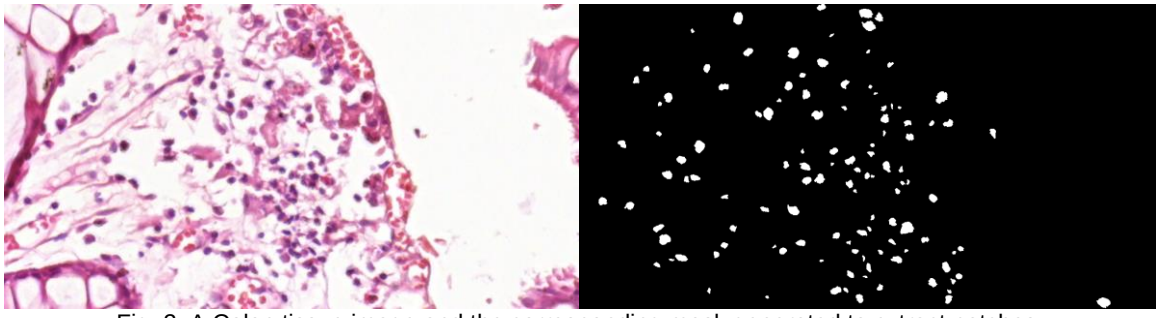


Fig. 2. A Colon tissue image and the corresponding mask generated to extract patches.

Training

The procedure to train a neural network is essentially the same and follows the well-known paradigm shown in [14]. The strategy utilizes an adaptive gradient descent approach, with a fixed batch size, a set of image patches are put to the network over a series of epochs, an error derivative is calculated, and it is backpropagated through the network updating the network weights. The learning rate start at 0.001 and is variable over time according to adaptive gradient method so that a local minimum can be reached.

In our case the input image patches for the train and test sets are taken of one big slide image and their corresponding label (i.e., nuclei or non-nuclei) are calculated according to the procedure explained in the patch generation section. In total we use 182000 patches of 64x64 pixels in both train and test sets. This data is used to train the network with Caffe framework [8] for 30 iterations in a Nvidia GTX770 GPU. To minimize the overhead and make maximum use of the GPU memory, we train with a large batch size of 128 images.

The obtained result in training was an accuracy of 92.93% in the testing set. The train was stopped at epoch number 30 in order to reduce the overfitting.
The set of resulting learned weights was stored to be used in the experiments stage.

Experiments

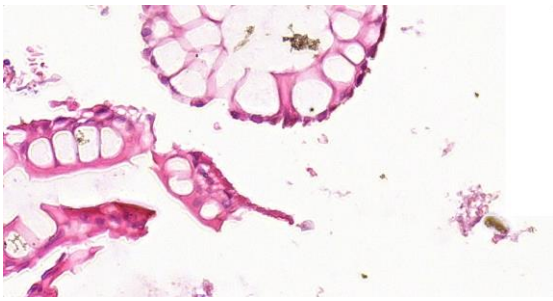
The dataset used in this experimentation is provided by the Obuda University, it is a set of 3 colon tissue microscopic images in big slide format (30208x71424 pixels) stained with Hematoxylin and eosin.

We demonstrate the application of the proposed method on three different image segmentation challenge tasks. The first task is the segmentation of an image taken of the same slide of training set. The second image used is from a big slide which belong a healthy patient. Finally, the last segmentation task uses an image from a big slide belonging a patient with cancer. The second and third image are from different big slide used in the training set, it is an interesting test because we can determine the capacity of the method of work with images never seen during the training stage and allow us to evaluate the generalization of the method. The images, the corresponding probability map and the final segmentation map are displayed in Fig. 3.

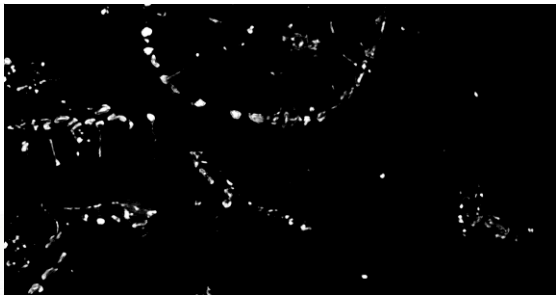
Also note that the nuclei detection rate is high. Additionally, we can see that in the recent review paper, [13] the performance is higher or comparable with several state of the art nuclear detection algorithms.

We can see also that the quality of segmentation map obtained in all cases is very good.

Note that the probability map obtained is follow of threshold at 0.5 and some morphological operations in order to obtain the segmentation map.

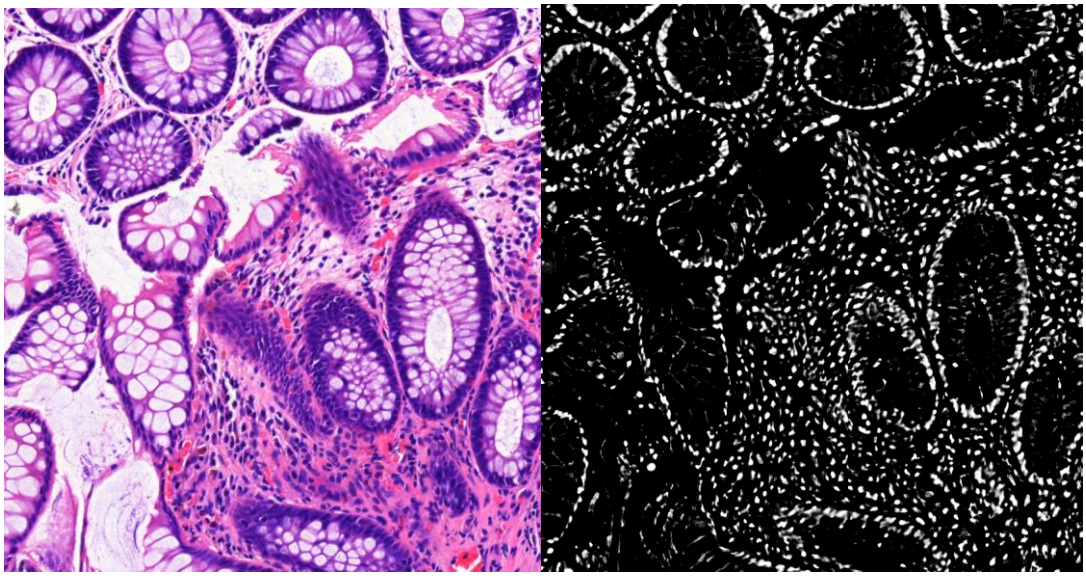


a

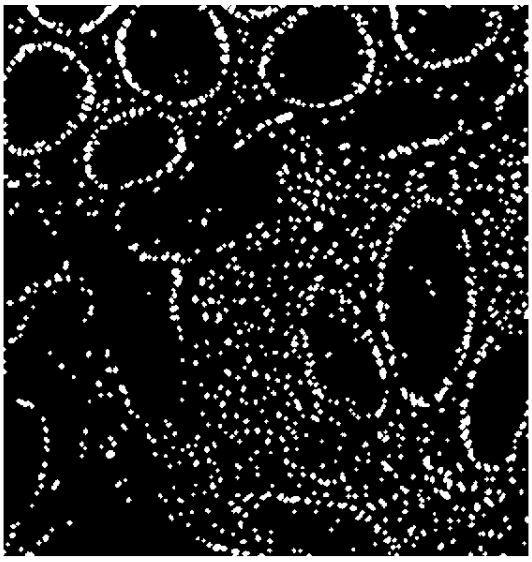


b

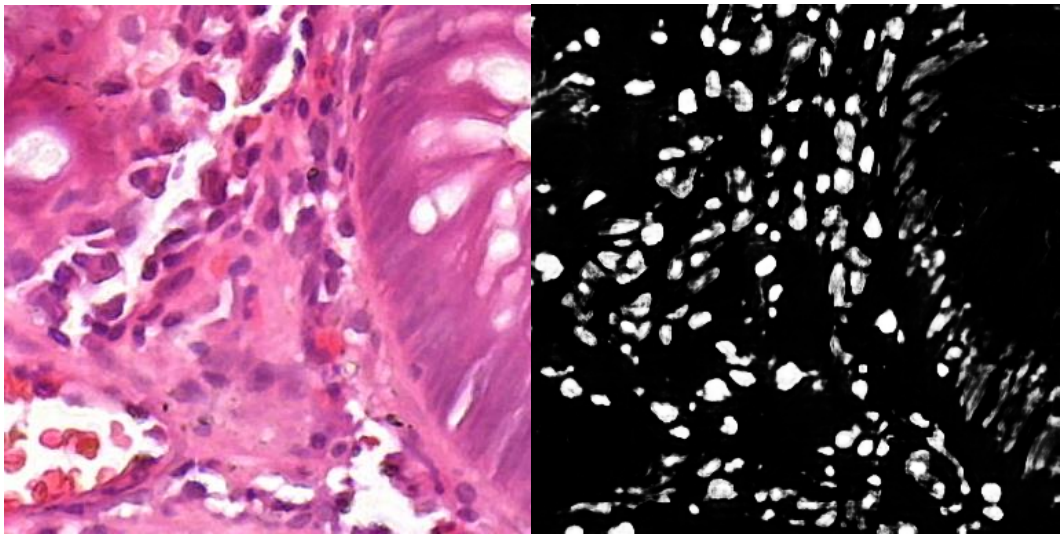
c



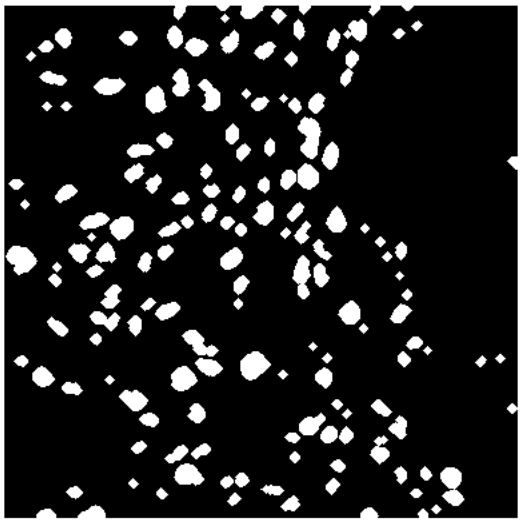
d e



f



g h



i

Fig. 3. Resulting image segmentated image using deep learning method.

In the above figure we can see the original image of the first (a), second (b) and third (c) task, the probability maps (b)(e)(h) and the final segmentation maps (c)(f)(i) obtained by the method.

3.2 Method to generate a semi annotated ground truth

The proposed algorithm consists in several techniques such as median filter, contrast stretching, color deconvolution[9] method and k-means clustering method[12] to accomplish the objective of segment the cells in a colon tissue image. Mostly the medical

images which are used for segmentation have low contrast and noise, hence median filter and contrast stretching are used to reduce the noise and improve the quality of the image respectively. After improving the quality of image, two methods are used to generate the binary mask that contains the cells. In first place we apply color deconvolution method getting a binary mask and in the second place we apply K-means method to get other binary mask. Both method give differents masks which contains differents cells, to get the final result we add the two last obtained masks.
The proposed algorithm is followed as below.

1. Load the image to be segmented.
2. Apply median filter.
3. Apply contrast stretching.
4. Run Color-deconvolution method and save the resulting mask.
5. Run K-means method and save the resulting mask.
6. Sum the above masks to get the final mask.

Color deconvolution method

1. Theory

Each pure stain will be characterized by a specific absorption factor c for the light in each of the three RGB channels. The detected intensities of light transmitted through a specimen and the amount A of stain with absorption factor c is described by Lambert-Beer's law

$$I_C = I_{0,C} \exp(-Ac_C)$$

with $I_{0,C}$ the intensity of light entering in the specimen, I_C the intensity of light detected by the camera after passing the specimen, and subscript C indicating the detection channel. As can be seen, this means that the transmission of light, and therefore the gray-values of each channel, depend on concentration of stain in a non-linear way, therefore the intensity values of the image can not directly be used for separation and measurement of each of the stains.

Then the optical density OD for each channel can be defined as

$$OD_C = -\log(I_C / I_{0,C}) = Ac_C$$

As can be seen, the OD for each channel is linear with the concentration of absorbing material, and can be used for separation of the contribution of multiple stains in a specimen.

Each pure stain will be characterized by a specific optical density for the light in each of the three RGB channels, which can be represented by a 3 by 1 OD vector describing the stain in the OD-converted RGB color space.

In the RGB images, the color system can be described as a matrix of the form

$$\begin{matrix} R & G & B \end{matrix}$$
$$\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \begin{matrix} hematoxilin \\ eosin \\ DAB \end{matrix}$$

with every row representing a specific stain, and every column representing the optical density as detected by the red, green and blue channel for each stain. Stain-specific values for the OD in each of the three channels can be easily determined by measuring relative absorbtion for red, green and blue on slides stained with a single stain.

The length of the vector will be proportional to the amount of stain, while the relative values of the vector describe the actual OD for the detection channels.
To perform the separation of the stains, we have to do an ortho-normal transformation of the RGB information, to get independent information about each stain's contribution, then for normalization, we divide each OD vector by its total length

$$\hat{p}_{ij} = \frac{p_{ij}}{\sqrt{\sum_k p_{ik}}}, k \in [1,3]$$

resulting in a normalized OD matrix M :

$$\begin{array}{c} R \quad G \quad B \\ \begin{bmatrix} \hat{p}_{11} & \hat{p}_{12} & \hat{p}_{13} \\ \hat{p}_{21} & \hat{p}_{22} & \hat{p}_{23} \\ \hat{p}_{31} & \hat{p}_{32} & \hat{p}_{33} \end{bmatrix} \begin{array}{l} hematoxilin \\ eosin \\ DAB \end{array} \end{array}$$

If c is the 3 by 1 vector for amounts of the three stains at a particular pixel, then the vector of OD levels detected at that pixel is $y = Mc$.

From the above it is clear that $c = M^{-1}y$. This means, that multiplication of the OD image with color-deconvolution matrix M^{-1}, results in orthogonal representation of the stains belonging on the the RGB image.

2. Implementation

An algorithm was developed to deconvolve the color information acquired with a red-green-blue (RGB) camera to calculate the contribution of each of the applied stains based on stain-specific RGB absorption.

This algorithm provides a robust and flexible method for objective immunohistochemical analysis of samples stained with up to three different stains using a laboratory microscope and a standard RGB camera.

First we calculate the OD image as follow from the RGB image I .

$$OD = -\log(I/256)$$

Then we determine the OD matrix for the combination of hematoxylin, eosin and DAB by estimation of stain absortion in each channel. To get those parameters, we select small ROIs areas which are all intensely stained with only one of the stains and then calculating the mean OD of each ROI. The resulting matrix for our dataset is the follow.

$$\begin{matrix} R & \text{G} & \text{B} & \\ \begin{bmatrix} 0.65 & 0.704 & 0.286 \\ 0.072 & 0.99 & 0.105 \\ 0.268 & 0.57 & 0.776 \end{bmatrix} & & & \begin{matrix} hematoxilin \\ eosin \\ DAB \end{matrix} \end{matrix}$$

The third step is calculate the OD normalized matrix M to get independent information about each stain's contribution.

$$\begin{matrix} R & \text{G} & \text{B} & \\ \begin{bmatrix} 0.65 & 0.721 & 0.286 \\ 0.704 & 0.99 & 0.57 \\ 0.268 & 0.10 & 0.776 \end{bmatrix} & & & \begin{matrix} hematoxilin \\ eosin \\ DAB \end{matrix} \end{matrix}$$

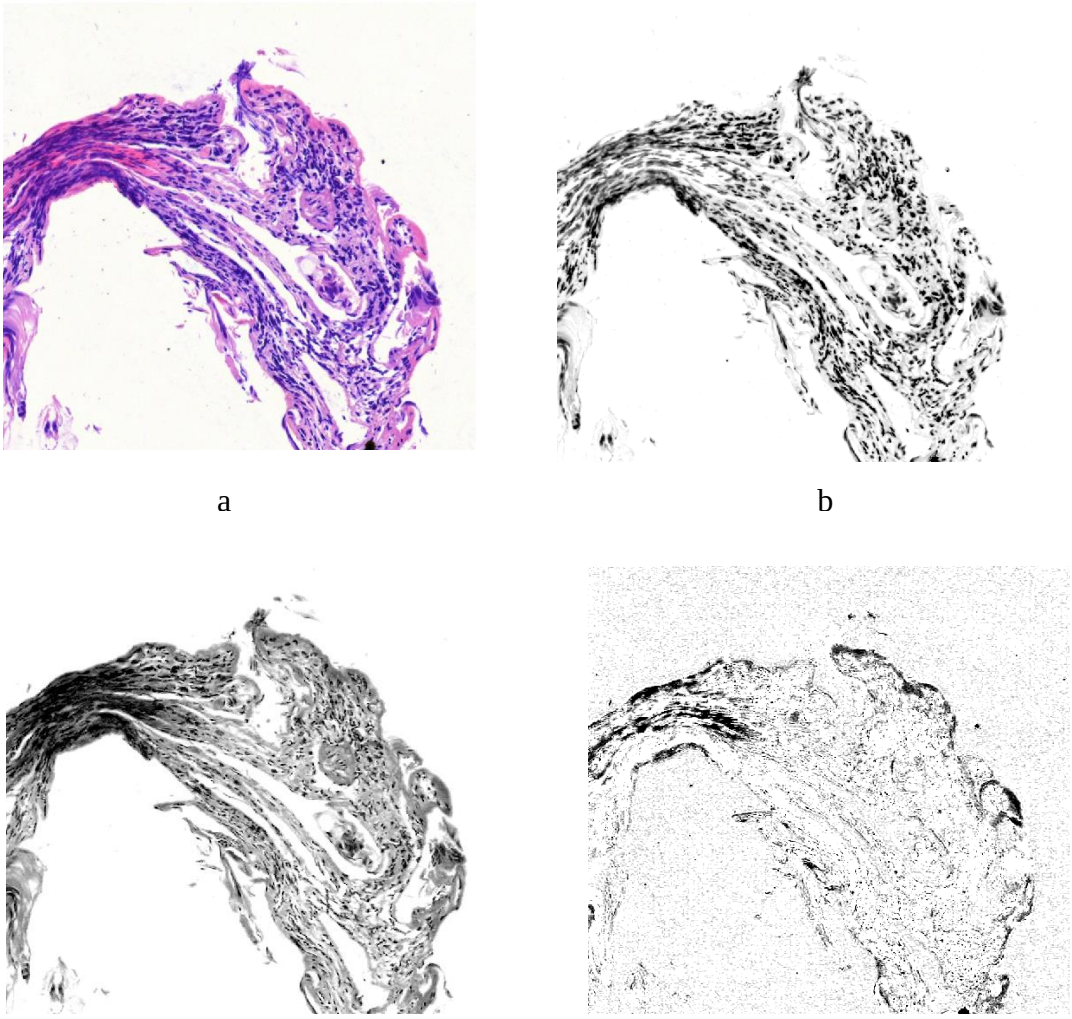
With the matrix calculated above we can calculate the color-deconvolution matrix M^{-1} to get

$$\begin{bmatrix} 1.88 & -0.07 & -0.59 \\ -1.01 & 1.13 & -0.48 \\ -0.55 & -0.12 & 1.57 \end{bmatrix}$$

In this matrix, the diagonal elements are greater than unity, while the non-diagonal elements are negative. The above equation implies that the corrected OD level values for each stain are formed by subtracting a portion of the green OD and the blue OD from the enhanced red OD to obtain the hematoxylin OD, subtracting a portion of the red OD and the blue OD from the enhanced green OD to obtain the eosin OD, and subtracting a portion of the red OD and the green OD from the enhanced blue OD to obtain the DAB OD.

To continue, we multiply the above matrix to each pixel in the RGB image, after that we have to do some contrast enhance operations to get a high dinamic range image. Our objective is get the cells, and suchs are in the hematoxylin channel, so we can discard the unnecessaries eosin and DAB channels.

A Result of experiment using the above method can be shown below.



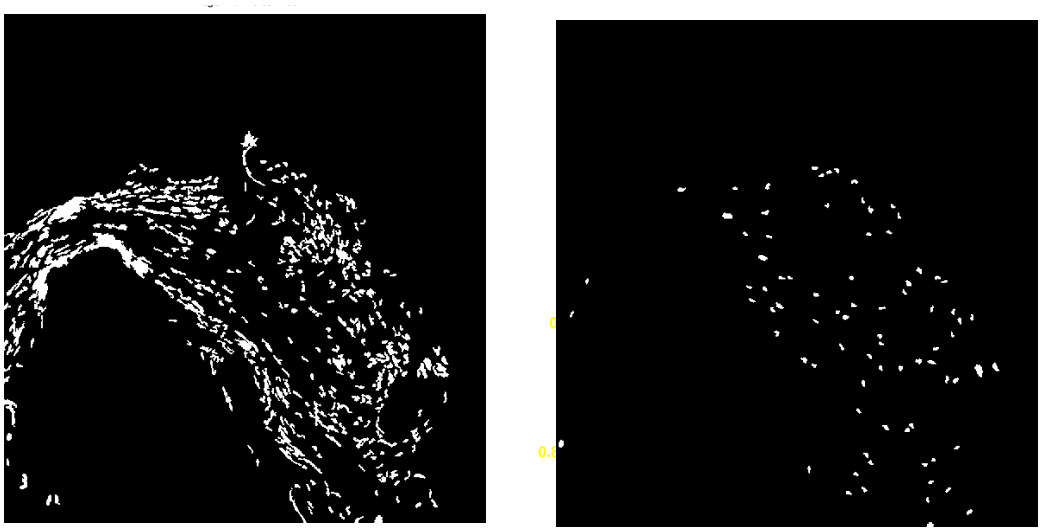
a b

c d

Fig. 4. Result of experiment using color deconvolution in a colon tissue image.

In the above figure we can see the original image (a), the hematoxylin channel (b), the eosin channel (c) and the DAB channel (d).

Until here, we have the Hematoxilin channel (shown in the image (b) above), this channel contains the cells and some undesired artifacts, then we apply a threshold to filter the most darkness objects, and some morphological operations to delete undesired artifacts; in first place erosion, later opening, and in the last step we delete the small remaining objects. To finalize we evaluate the roundness of each object in the image discarding the objects with low roundness.
A result of segment the above image can be shown below.



a b

Fig. 5. Resulting image segmentated image using color deconvolution method.

In the above figure we can see the morphological resulting image (a), and the final result after filter roundness objects.

K-means method

1. Theory

There are different techniques for image segmentation like threshold based, edge based, cluster based, neural network based. From the different technique one of the most efficient method is the clustering method. Also there are different types of clustering: K-means clustering, Fuzzy C-means clustering, mountain clustering method and subtractive clustering method. One of most used clustering algorithm is k-means clustering. It is simple and computationally faster than the hierarchical clustering. And it can also work for large number of variable. But it produces different cluster result for different number of number of cluster. So it is required to initialize the proper number of number of cluster k. Again, it is required to initialize the k number of centroid. Different value of initial centroid would result different cluster. So selection of proper initial centroid is also an important task.

K-means is a method to divide a set of data into a specific number of groups. In k-means clustering, it partitions a collection of data into a k number group of data. It classifies a given set of data into k number of disjoint cluster. K-means algorithm consists of two separate phases. In the first phase it calculates the k centroid and in the second phase it takes each point to the cluster which has nearest centroid from the respective data point. There are different methods to define the distance of the nearest centroid like Hamming, Euclidean, but one of the most used methods is Euclidean distance. Once the grouping is done it recalculate the new centroid of each cluster and based on that centroid, a new distance is calculated between each center and each data point and

assigns the points in the cluster which have minimum Euclidean distance. Each cluster in the partition is defined by its member objects and by its centroid. The centroid for each cluster is the point to which the sum of distances from all the objects in that cluster is minimized. So K-means is an iterative algorithm in which it minimizes the sum of distances from each object to its cluster centroid, over all clusters. Let us consider an image with resolution of $x \times y$ and the image has to be clusterized into k number of cluster. Let $p(x, y)$ be an input pixels to be cluster and c_k be the cluster centers. The algorithm for k-means clustering work as following:

1. Initialize number of cluster k and centre.
2. For each pixel of an image, calculate the Euclidean distance d, between the center and each pixel of an image using the relation given below

$$d = \left\| p(x, y) - c_k \right\|$$

3. Assign all the pixels to the nearest centre based on distance d.
4. After all pixels have been assigned, recalculate new position of the centre using the relation given below.

$$c_k = \frac{1}{k} \sum_{y \in c_k} \sum_{x \in c_k} p(x, y)$$

5. Repeat the process until it satisfies the tolerance or error value.
6. Reshape the cluster pixels into image.

Although k-means has the great advantage of being easy to implement, it has some drawbacks. The quality of the final clustering results is depending on the initial location and quantity of initial centroids. So if the initial centroid is randomly chosen, it will get different result for different initial centers. So the initial center should be carefully chosen so that we get our desire segmentation. And also computational complexity is another term which we need to consider while designing the K-means clustering. It relies on the number of data elements, number of clusters and number of iteration.

2. Implementation

An algorithm was developed to clusterize similar information in a image acquired with a red-green-blue (RGB) based in K-means method.
This algorithm provides a robust and flexible method for objective analysis of samples stained with hematoxylin and eosin obtained of laboratory microscope and a standard RGB camera.

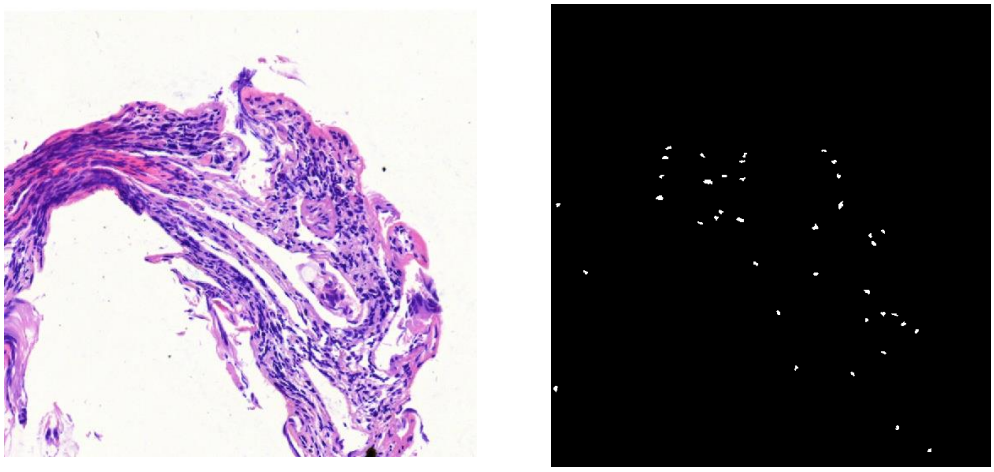
First, given a RGB image, we work with the R channel discarding the G and B channels, because the useful information about the cells is in this. Then we do K-means clustering with the heuristic method K-means++ to initialize the centroids. This initialization method demonstrates using a simulation study for several cluster orientations, that achieves faster convergence to a lower sum of within-cluster sums of squares point-to-cluster-centroid) distances than random centroid initialization. Therefore, the method improves the running time of K-means algorithm, and the quality of the final solution.

To continue we run K-means algorithm ten times with 7 clusters using new initial cluster centroid positions (based on K-means++ initialization) to get the solution with the lowest within-cluster sums of point-to-cluster-centroid distances. We do it because K-means++ cannot give the best initialization because it is heuristic.

The next step is get the cluster that contains the cells, to do it we calculate the mean of each cluster and we select the cluster that has the lowest mean, we use these criteria because the cells are the darkness objects in the image.

Until here we have a mask that contains the cells and some undesired artifacts, then we apply some morphological operations to delete undesired artifacts; in first place erosion, later opening, and in the last step we delete the small remaining objects. To finalize we evaluate the roundness of each object in the image discarding the objects with low roundness.

A segmentation result of the above method can be shown below.



a b

Fig. 6. Result of experiment using K-means method in a colon tissue image.

In the above figure we can see the original image (a) and the final result after apply the method (b).

4 Conclusion

We have shown how deep learning methods can be a valuable tool for the histopathology image analysis domain due to its innate ability to learn useful features directly from data. In addition, we have outlined a full approach containing the necessary methods to obtain a result using a supervised deep learning method from an unannotated image dataset.

Also, we have shown that a practical, and publicly available software framework[6] can perform on par, or better, than several state-of-the-art classification approaches for several digital histopathology analysis tasks.

The approach presented here are not intended to be a final ending point towards histological problems, but a surprisingly robust jumping off point for further research.

5 Acknowledgements

The author would like to thank the Erasmus+ Programme as well as the Obuda Egyetem for their support.

6 References

1. Ciresan, D.C., Gambardella, L.M., Giusti, A., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. In: NIPS. pp.2852{2860 (2012)
2. Veta M., Pluim JP., van Diest PJ., Viergever MA.: Breast cancer histopathology image analysis: A review. IEEE Trans Biomed Eng 2014;61:1400-11.
3. Janowczyk A., Madabhushi A.: Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases, J Pathol Inform (2016).
4. Olaf R., Philipp F., Brox T.: U-Net: Convolutional Networks for Biomedical Image Segmentation (2015) arXiv:1505.04597 [cs.CV]
5. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human level performance on imagenet classification (2015), arXiv:1502.01852 [cs.CV]
6. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding (2014), arXiv:1408.5093 [cs.CV]
7. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: NIPS. pp. 1106{1114 (2012)
8. Jia Y, Shelhamer E., Donahue J., Karayev S., Long J., Girshick R., et al. Caffe: Convolutional Architecture for Fast Feature Embedding. arXiv preprint arXiv: 1408.5093; 2014
9. Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution.
10. Szénási S., Vámossy Z., Kozlovszky M.:Evaluation and Comparison of Cell Nuclei Detection Algorithms.
11. Geert L, Clara I. Sánchez, Nadya T., Meyke H., Iris N., Iringo K., Christina H., Peter B., Bram van G., Jeroen van der L.: Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Scientific Reports 6, Article number: 26286 (2016).
12. Nameirakpam D, Khumanthem M, Yambem JC. Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm. Eleventh International Conference on Communication Networks, ICCN 2015.
13. Irshad H, Veillard A, Roux L, Racoceanu D. Methods for nuclei detection, segmentation, and classification in digital histopathology: A review-current status and future potential. IEEE Rev Biomed Eng 2014;7:97-114.
14. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436-44.
15. Xing F, Xie Y, Yang L. An Automatic Learning-Based Framework for Robust Nucleus Segmentation.
16. Simonyan K, Zisserma A.Very Deep Convolutional Networks for Large-Scale Image Recognition.(2014) arXiv:1409.1556 [cs.CV]