

Nucleus Classification in Colon Cancer H&E Images using Deep Learning

Ali Hamad¹, Filiz Bunyak¹ and Ilker Ersoy²

¹ Department of Computer Engineering and Computer Science, University of Missouri-Columbia, Columbia, MO 65211, USA.

² Department of Pathology and Anatomical Sciences, University of Missouri-Columbia, Columbia, MO 65211, USA.

Classification of nuclei, cells, and tissues is one of the most challenging tasks in biomedical image analysis. A large variety of computational approaches in the fields of image processing and machine learning including deep learning has been reported in the literature [1],[2],[3]. Performance of pattern recognition and classification tasks greatly depends on the features they use. Hand-crafted features may not always capture underlying structure of the data or generalize to other tasks. Deep learning uses multi-layer networks to reconstruct hierarchy of representations and enables automatic learning of complex features required for visual pattern recognition. We have recently developed an automated deep learning system for tissue classification in histopathology images [2]. Nuclei classification in the images of colon cancer H&E slides is also a challenging task due to large variation of cell morphologies. Hematoxylin & Eosin (H&E) is one of the most commonly employed stains in pathology. The digital images obtained from the H&E stained specimens can introduce additional challenges due to batch variations and illumination changes. To implement the nucleus classification task, we designed a Convolution Neural Network (CNN) [4]. The ability of CNN in learning high-dimensional and complex classes of objects makes it a very useful tool for a variety of classification tasks in biomedical imaging.

Our network was trained and evaluated by using a large colon cancer nuclei images dataset reported in [5] with 22,444 nuclei divided into four classes (Epithelial, Inflammatory, Fibroblast, and Miscellaneous) with the following number of samples (7,722, 6,971, 5,712, and 2,039), respectively. Our model consists of six layers (2 convolution layers, 2 subsampling layers, 1 fully connected layer, and output layer). The details of our deep learning network are shown in Figure 1. The original size of H&E images is 500x500x3 and the total number of colon cancer histology images is 100. A number of small patches with size 28x28x3 are extracted from images given the ground truth nucleus locations. Four-fold cross-validation is employed in our experiments. To make the classification fair among all nuclei types that have different number of patches and to evaluate the overall performance of classification, weighted average F1 score has been calculated. The weighted average F1 score for our network is 0.7749 which is higher than the previous published works as shown in Table 1. Our model converges to minimum error in 30 epochs while the network in [5] converges in 120 epochs. Our weighted precision and recall are 0.7739 and 0.7759, respectively. Figure 2 shows a sample image from the data set marked with ground truth locations and our classification results. Our future work includes applying deep learning strategies to 2D particle classification in cryo-EM images. Just as in cell, nucleus, and tissue classification, we anticipate that the deep-learning assisted 2D classification in cryo-EM images will better capture complex structural features of particles with less “human” bias compared to classical classification approaches. Deep learning based classification will be more reflective of the data and much less dependent upon input from the image analyst.

References:

- [1] F Bunyak *et al*, A segmentation-based multi-scale framework for the classification of epithelial and stromal tissues in h&e images. IEEE Int. Conf. Bioinformatics and Biomedicine (2015) p. 450–453.
- [2] Z Al-Milaji *et al*, Hybrid deep learning framework for the classification of epithelial and stromal tissues in h&e images. Proc. International Workshop on Deep Learning for Pattern Recognition (2016).
- [3] L Alzubaidi *et al*, Nucleus detection in h&e images with fully convolutional regression networks. Proc. First International Workshop on Deep Learning for Pattern Recognition (2016).
- [4] A Krizhevsky, I Sutskever and GE Hinton, Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems (2012) p. 1097–1105.
- [5] K Sirinukunwattana *et al*, Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans. Medical Imaging **35(5)** (2016), p. 1196–1206.
- [6] K Sirinukunwattana, DR Snead and NM Rajpoot, A novel texture descriptor for detection of glandular structures in colon histology images. SPIE Medical Imaging (2015) p. 94200S.
- [7] Y Yuan *et al*, Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. Science Translational Medicine **4(157)** (2012).

Method	Weighted F1 Score
Our CNN	0.775
Softmax CNN+SSPP [5]	0.748
Superpixel Descriptor [6]	0.687
CRImage [7]	0.488

Table 1. Comparative results for nucleus classification.

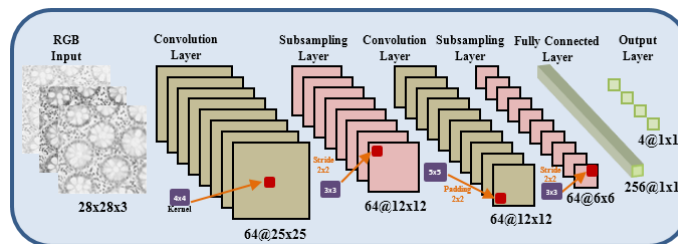


Figure 1. Architecture of our CNN model.

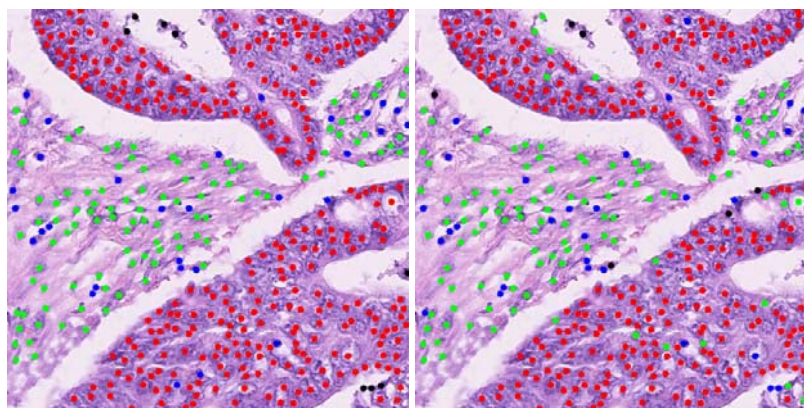


Figure 2. Ground truth and classification results on a sample image. Red, green, blue, and black dots represent epithelial, fibroblast, inflammatory, and other miscellaneous cells, respectively.