



Nuclei Classification Using Dual View CNNs with Multi-crop Module in Histology Images

Xiang Li, Wei Li^(✉), and Mengmeng Zhang

College of Information Science and Technology,
Beijing University of Chemical Technology, Beijing 100029, China
liwei089@ieee.org

Abstract. Histopathology image diagnostic technique is a quite common requirement; however, cell nuclei classification is still one of key challenge due to complex tissue structure and diversity of nuclear morphology. Cell nuclei categories are often defined by contextual information, including central nucleus and surrounding background. In this paper, we propose a Dual-View Convolutional Neural Networks (DV-CNNs) that captures contextual contents from different views. The DV-CNNs are composed of two independent pathways, one for global region and another for center local region. Noted that each pathway with “multi-crop module” can extract five different feature regions. Common networks do not fully utilize the local information, but the designed cropping module catches information for more complete features. In experiments, two pipelines are complementary to each other in score fusion. To verify the performance in proposed framework, it is evaluated on a colorectal adenocarcinoma image database with more than 20,000 nuclei. Compared with existing methods, our proposed DV-CNNs with multi-crop module demonstrate better performance.

Keywords: Histopathology image analysis
Convolutional neural network · Cell nuclei classification

1 Introduction

Histopathology is the most commonly used microscopic research for the diagnosis of cancer diseases. The cancer tissues are sampled from the body and then prepared for observing under the microscope. Stained with the standard hematoxylin and eosin (H&E) stain [1] can mark nuclei in histopathology images of cancer tissues, and pathologists need to identify the type of nuclei. The recognition of cell nuclei in histopathology images become one of the core challenge

This work was supported by the National Natural Science Foundation of China under Grant No. NSFC-61571033, Beijing Natural Science Foundation (4172043), Beijing Nova Program (Z171100001117050), and partly by the Fundamental Research Funds for the Central Universities (BUCTRC201615).

for qualitative and quantitative analysis at cell levels [5]. A single histopathology image may contain about thousands of nuclei, and pathologists are incapable of identifying all nuclei precisely. Traditional approach requires experienced pathologists to manually identify the cell, which is extremely laborious. Consequently, developing an automatic and reliable method [8] for classification tasks becomes an attractive research topic and automated images classification will allow pathologists to quickly obtain the specific information which can increase objectivity and less burden on observers.

Most of the existing automated attribute classification techniques for cells in histology images include two aspects, i.e., traditional machine learning methods and convolutional neural network (CNNs) [13]. Kumar et al. proposed a k-nearest neighbor based method for microscopic biopsy images, and the efficacy of other classifiers such as SVM, random forest, and fuzzy k-means was examined [12]. Recently, deep convolutional neural networks appear to be attracting considerable attention due to its excellent performance on visual recognition task. Different from traditional approaches, CNNs act more dynamically to provide multilevel hierarchies of features, have been extensively employed for histopathology image classification. Malon et al. [9] combined manually designed nuclear features with the learned features extracted by CNN which handled the variety of appearances of mitotic figures and decreased sensitivity to the manually crafted features and thresholds [16].

Different from usual scene classification, cell nuclei classification belongs to fine-grained image categorization which aims to classify sub-categories, such as different species of dogs. Therefore, variability in the appearance of the same type of nuclei is a critical factor that makes classification of individual nucleus equivalently difficult. Pathologists analyze the cell nuclei of the histopathological images, looking for detailed texture around nucleus and the shape of the nucleus [15, 17]. So in cell nuclei recognition, nuclei categories are often defined by local nuclei and global background around nuclei. But existing classification approaches are limited in considering different region of cell nuclei. In this paper, we present a Dual View convolutional neural networks by employing a dual network strategy to classify nuclei in routine H&E stained histopathology images of colon cancer. To be specific, CNNs at global way are able to capture cell background information, while CNNs at local way are capable of describing local details of center nucleus. Meanwhile, multi-crop module is added to the two subnets, which can catch diverse feature regions. DV-CNNs with multi-crop module integrate complementary diagnostic criteria of different hierarchy concepts, allowing them to explore the intrinsic connection between cellular background and nucleus.

We design different experiments to ensure a fair comparison. And the experimental results demonstrate that multiple feature regions around nuclei and different image views are important for cell nuclei classification in histopathology images and show superior performance to current methods on the HistoPhenotypes datasets.

2 Nuclei Classification Framework

Cell nuclei classification always couples global background with local part in a diagnosis. We need to consider this contextual information of histopathology images during building models. In order to address these challenges, we utilize the CNNs to propose an effective architecture that dual views of images are first conducted and then send to two branches for attaining multiple feature regions. In this section, we first describe the framework of DV-CNNs and then discuss multi-crop module.

2.1 Dual View Convolutional Neural Networks

To derive an efficient classifier in cell nuclei images, we need to overcome the challenge posed by variability in the appearance of the same type of nuclei [10]. Pathologists analyze the cell nuclei by looking for detailed texture around nucleus and the shape of the nucleus and then determine the possible nuclei type. Following the pathologist’s experiences [11], we propose DV-CNNs as illustrated in Fig. 1.

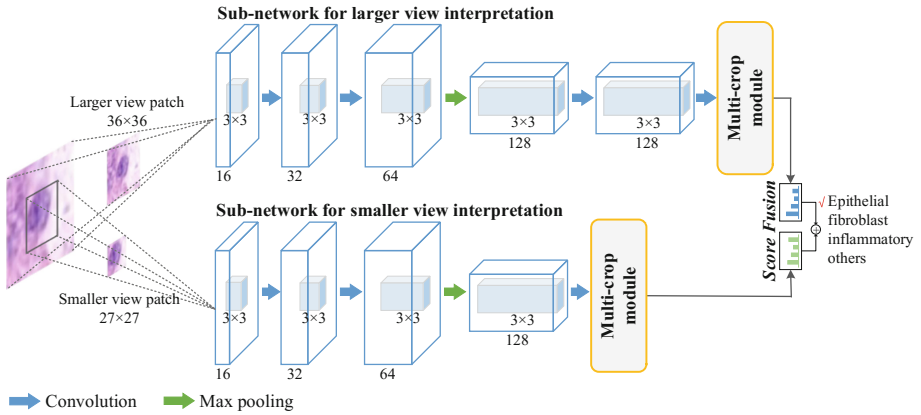


Fig. 1. The overall flowchart of the proposed Dual View CNNs.

The proposed model offers a dual view of images, with one pathway considering information from the larger area around the nucleus and another pathway focusing on local, nucleus level information. We observe contexts in the cell nuclei imagery, where nucleus is relatively small and it is easy to include larger region as input. If we focus only on larger region, the specific attributes of the nucleus will be ignored, and vice versa. The proposed approach is based on multiple cues extracted from a dual view of images such as the background and the center nuclei features, and then, they provide each other with complementary features to avoid missing important diagnostic information.

In this framework, two pathways process images in parallel. Then, two individual networks are concatenated to combine extracted features. As illustrated in Fig. 1, they are designed to describe cell nuclei at different views for contextual understanding. We feed larger view patches 36×36 and smaller view patches 27×27 to two sub-network to extract different features, respectively. For two sub-networks, each 3×3 convolution layer is followed with the Rectified Linear Unit activation layer and Batch Normalization layer [4]. A 2×2 max pooling operation with stride 2 for reducing half. Multi-cropping modules are added to two pathways, and next section will have detailed introduction about this module. Finally, the prediction results of our models are complementary to each other by fusing scores.

2.2 Multi-cropping Module

After feature extraction in both larger region and center local region, multi-crop module is discussed as shown in Fig. 2. There are two reasons by utilizing multi-crop module. The first reason is uncertainty of the central location of the nucleus. Input images with the same size but of different central location of the nucleus display different feature representations. Combinations of different location features help understanding image contents better. So compound features usually have better representation ability. The second advantage is robustness to variable appearance of the nucleus. Cropping patches from feature maps and fusing these feature patches are robust to diversity of nuclear morphology.

The multi-cropping module captures multiple local information from different regions in feature maps, while those previous works all rely on a fix feature size. Meanwhile, the multi-region crop is complementary to each other. The basic architecture as shown in Fig. 1 starts with convolutional layers and max pooling layers. After a series of convolutional layers, feature map P is obtained and send to multi-cropping module, then four different regions of feature are cropped around central nucleus, i.e., top-left, top-right, bottom-left, and bottom-right (denoted as TL, TR, BL, BR). Each region is about three quarters of the image P , where cropping patches are able to cover the majority area of a nucleus. Then, network is divided into five subnets for four region features (i.e., TL, TR, BL, BR) and the original feature map P . Each region is followed by one convolutional layer and a global average pooling layer. Batch Normalization layer is applied to the activations of convolutional layers, following by the ReLU layer for non-linearity. Those five networks produce five output as F_{TL} , F_{TR} , F_{BL} , F_{BR} , F_P , respectively. Finally, those five outputs are concatenated as,

$$Z = F_{TL} \oplus F_{TR} \oplus F_{BL} \oplus F_{BR} \oplus F_P, \quad (1)$$

where Z combines comprehensive information derived from different regions. Suppose the number of categories is N , then z is expanded to $N \times 1$ vector, named as \tilde{z} . The probability of each category is calculated as follows,

$$\sigma(\tilde{z})_j = \frac{e^{\tilde{z}_j}}{\sum_{k=1}^N e^{\tilde{z}_k}}, \quad (2)$$

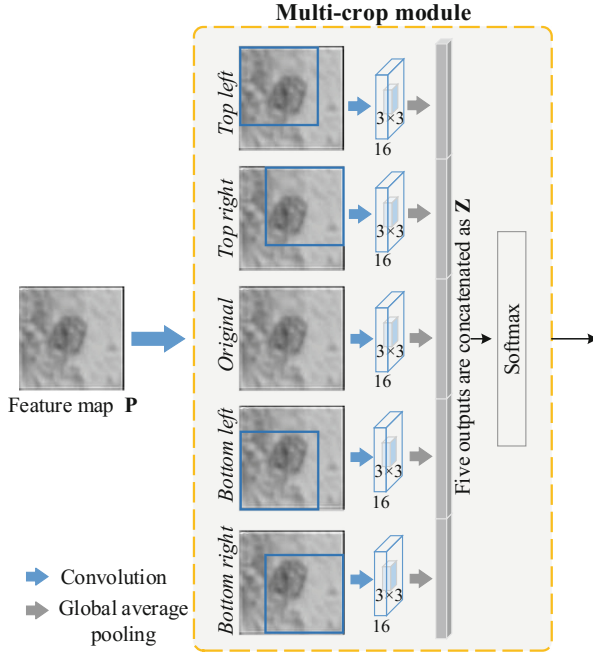


Fig. 2. Multi-cropping module in detail.

where \tilde{z}_j is the j -th value of $N \times 1$ vector. After a softmax layer, $\sigma(\tilde{z}_{global})_j$ and $\sigma(\tilde{z}_{local})_j$ are obtained. Finally, the prediction results of $\sigma(\tilde{z}_{global})_j$ and $\sigma(\tilde{z}_{local})_j$ are complementary to each other by fusing them. And the losses between outputs and label is computed through categorical cross entropy loss function. In the nuclei classification framework, the proposed DV-CNNs model is based on multiple cues extracted from dual views of images, plus multi-region features of the branch networks. As illustrated in Figs. 1 and 2, they are designed to describe cell nuclei at different regions for contextual understanding. At this final level, it is expected that the nuclear interaction with its background reaches the purpose to avoid missing important diagnostic information. It also leverages the idea of multi-region locations in features that effect the recognition of nuclei, making it suitable for building cell nuclei recognition networks.

3 Experiment Results and Analysis

The proposed CNNs are implemented in Keras, on Tensorflow backend. The training data are augmented by rotating and flipping all the images. In order to achieve effective fusion, two networks are trained to the best performance, respectively, and then dual CNNs are trained on this basis. We update parameters with Adam strategy [6]. All network models are trained for 80 epochs. A

batch size of 256 images is used. In the designed sub-network, the initial learning rate is 0.001, and is divided by 10 at 50 epochs. In the proposed Dual-View network, combination of two branches are fine-tuned by setting learning rate 0.00001.

3.1 The Experimental Data

HistoPhenotypes is a public database which involves 100 H&E stained histology images of colorectal adenocarcinomas. There are 22444 nuclei that classified into four class labels: 7722 epithelial, 5712 fibroblast, 6971 inflammatory and 2039 others. Figure 3 illustrates some examples of the nuclei in the dataset. Note that the dataset contains complex and wide variety shapes of nuclei where overlap is also present. As reported in [2, 14], 27×27 pixel nucleus-center patches are cropped to feed one of the networks. 36×36 pixel patches are extracted, including more contextual information to feed the another compensation branch.

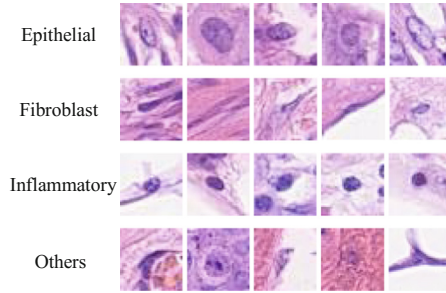


Fig. 3. Example patches of different types of nuclei found in the dataset. Each row corresponds to a cell nuclei class.

3.2 Classification Performance

Experiment One. First of all, experiments are conducted to compare with results in [14], where four classes are considered, i.e., epithelial, fibroblast, inflammatory and others. We employ 2-fold cross-validation the same as [14] for parameter tuning. And then, we calculate the F1 average score for all the classes, an area under the receiver operating characteristic curve for multiclass classification (multiclass AUC) [3] and overall accuracy.

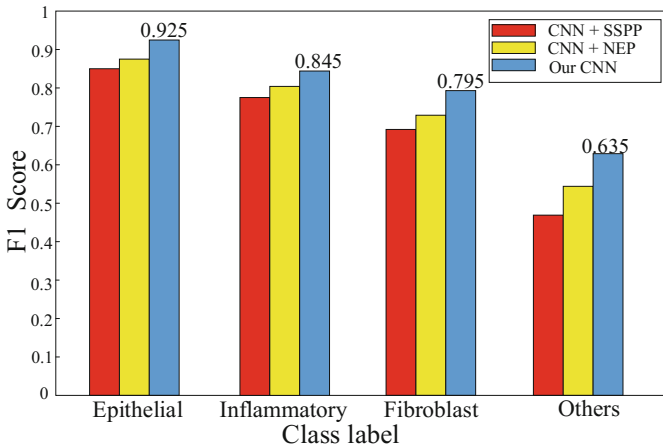
Classification performance of the proposed DV-CNNs with or without multi-crop module is provided as listed in Table 1. \times represents the sub-network without multi-crop module (denoted as MC.module) and vice versa. From comparison of “without multi-crop module” and “with multi-crop module”, it is observed that network with this module trained from different images patches (27×27 or 36×36) both yield better performance than “without multi-crop module” networks. Thus, multi-crop module actually contains more local details and richer

Table 1. Comparative results about DV-CNNs with or without multi-crop module for four-classes classification.

Input	+MC.module	Average F1 score	Multiclass AUC	Overall accuracy
36×36	×	0.827	0.939	82.63%
	√	0.835	0.945	83.75%
27×27	×	0.798	0.930	80.48%
	√	0.815	0.936	81.86%

Table 2. Comparative results about DV-CNNs and other classification approaches for four-classes classification.

Model		Average F1 score	Multiclass AUC
Softmax CNN+SSPP		0.748	0.893
Softmax CNN+NEP		0.784	0.917
Superpixel descriptor		0.687	0.853
CRImage		0.488	0.684
Dual-View CNNs	36×36 branch + MC.module	0.835	0.945
	27×27 branch + MC.module	0.815	0.936
	Score fusion	0.843	0.947

**Fig. 4.** Comparative results for nucleus four-classes classification stratified with respect to class label.

information of multi-region feature content. After verifying effectiveness of this extra module, an experiment is performed to investigate the proposed DV-CNNs.

Experimental results are listed in Table 2, where also includes the strategies of Standard Single-Patch Predictor (SSPP), Neighboring Ensemble Predictor (NEP), superpixel descriptor and CRImage proposed in [14]. From the results, it

is obvious that the existing methods are both lower than any single branch of the proposed CNNs. Furthermore, the larger input branch learned on whole images yields better results than the smaller input branch (i.e., 0.835 vs. 0.815 overall accuracy) by comparing the performance of CNNs at different views. This phenomenon indicates that the global branch has more discriminative information which may compensate for the lack of local branch. Finally, we fuse prediction scores of dual networks, and obtain the final performance with weighted F1 average score of 0.843 and multiclass AUC of 0.947. We obtain an improvement of 6% F1 scores over the CNN + NEP. Figure 4 further illustrates the class specific accuracy for the proposed method compared with existing methods. Generally, the proposed one produces higher accuracy for each class.

Experiment Two. Experiments are also designed to compare with results in [2], where only three classes are considered, i.e., epithelial, fibroblast and inflammatory. Others category consists of mixed cell nuclei; therefore, authors in [2] have excluded it from their study. The split of the training and testing set is the same; that is, there are 17004 training images and 3401 testing images.

Table 3. Comparative results about DV-CNNs with or without multi-crop module for three-classes classification.

Input	+MC.module	Overall accuracy
36×36	×	87.34%
	√	88.98%
27×27	×	86.34%
	√	88.06%

Table 4. Comparative results about DV-CNNs for three-classes classification.

Model		Overall accuracy
Fine-tuned VGG-16[2]		88.03%
Our CNN	36×36 branch + MC.module	88.98%
	27×27 branch + MC.module	88.06%
	Score fusion	90.40%

Classification performance of DV-CNNs with or without cropping module is provided as listed in Table 3. It is also proved that network with multi-crop module trained from different images patches (27×27 or 36×36) both yield better performance than “without module networks”. In [2], authors employed very deep architectures like GoogleNet, AlexNet [7] and VGG-16 trained with transfer learning, which applied a pre-trained model to fine-tune on this dataset.

It achieved the best result with the VGG-16 architecture with an overall accuracy of 88.03%. After fusing prediction scores of dual networks, we obtain the final performance with overall accuracy of 90.40% as listed in Table 4. From another point of view, DV-CNNs have much less parameter than the deep network VGG-16, but it can also reach the accuracy of the very deep network.

Through the above experiments, they demonstrate the efficiency of our methods in two aspects:

- Multi-crop module actually contains more local details and richer information of multi-region feature content. It is obvious that the multi-crop operation for training data has a direct influence for classification task; for example, evaluation metrics achieved with multi-crop module are higher than those without cropping module.
- As mentioned in Sect. 2, existing methods for histopathology images do not consider the global and local context. For the proposed DV-CNNs, we realize the best results compared with state-of-the-art work using four or three classes experimental data. These excellent results in Tables 1 and 2 demonstrate the effectiveness of the proposed method for cell nuclei recognition which allows us to hypothesize that joint global and local information is more beneficial to variability in the appearance of nuclei.

4 Conclusion

In this paper, we proposed an interesting Dual View CNNs with multi-crop module which are able to capture contextual information from different regions for nucleus classification in routine stained histology images of colorectal adenocarcinomas. The extracted features integrate complementary diagnostic criteria of different hierarchy concepts, allowing them to explore the intrinsic connection between cellular background and nucleus. We conducted experiments on a large dataset with more than 20000 annotated nuclei. The encouraging results compared with other approaches on cell nuclei classification demonstrated its effectiveness of the proposed method.

References

1. Adur, J., et al.: Colon adenocarcinoma diagnosis in human samples by multicontrast nonlinear optical microscopy of hematoxylin and eosin stained histological sections. *J. Cancer Therapy* **5**(13), 1259–1269 (2014)
2. Bayramoglu, N., Heikkilä, J.: Transfer learning for cell nuclei classification in histopathology images. In: Hua, G., Jégou, H. (eds.) *ECCV 2016 Part III. LNCS*, vol. 9915, pp. 532–539. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-49409-8_46
3. Hand, D.J., Till, R.J.: A simple generalisation of the area under the roc curve for multiple class classification problems. *Mach. Learn.* **45**(2), 171–186 (2001)
4. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. In: *International Conference on International Conference on Machine Learning*, pp. 448–456 (2015)

5. 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.* **7**(1–5), 97–114 (2014)
6. Kingma, D., Ba, J.: Adam: a method for stochastic optimization. In: *International Conference on Learning Representations* (2014)
7. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: *International Conference on Neural Information Processing Systems*, pp. 1097–1105 (2012)
8. Madabhushi, A., Lee, G.: Image analysis and machine learning in digital pathology: challenges and opportunities. *Med. Image Anal.* **33**, 170–175 (2016)
9. Malon, C.D., Cosatto, E.: Classification of mitotic figures with convolutional neural networks and seeded blob features. *J. Pathol. Inform.* **4**(1), 9 (2013)
10. Murthy, V., Hou, L., Samaras, D., Kurc, T.M., Saltz, J.H.: Center-focusing multi-task CNN with injected features for classification of Glioma nuclear images. In: *Applications of Computer Vision*, pp. 834–841. *IEEE* (2016)
11. Nguyen, K., Bredno, J., Knowles, D.A.: Using contextual information to classify nuclei in histology images. In: *IEEE International Symposium on Biomedical Imaging*, pp. 995–998 (2015)
12. Rajesh, K., Rajeev, S., Subodh, S.: Detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features. *J. Med. Eng.* **2015**, 457906 (2015)
13. Schmidhuber, J.: Deep learning in neural networks: an overview. *Neural Netw. Off. J. Int. Neural Netw. Soc.* **61**, 85 (2014)
14. Sirinukunwattana, K., Shan, E.A.R., Tsang, Y.W., Snead, D.R.J., Cree, I.A., Rajpoot, N.M.: Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Trans. Med. Imaging* **35**(5), 1196–1206 (2016)
15. Veta, M., Pluim, J.P.W., Diest, P.J.V., Viergever, M.A.: Breast cancer histopathology image analysis: a review. *IEEE Trans. Bio-Med. Eng.* **61**(5), 1400–1411 (2014)
16. Wang, H., Cruzroa, A., Gilmore, H., Feldman, M., Tomaszewski, J., Madabhushi, A.: Cascaded ensemble of convolutional neural networks and handcrafted features for mitosis detection. In: *SPIE Medical Imaging*, pp. 90410B–90410B-10 (2015)
17. Yu, Y., Lin, H., Meng, J., Wei, X., Guo, H., Zhao, Z.: Deep transfer learning for modality classification of medical images. *Information* **8**(3), 91 (2017)