World Scientific
www.worldscientific.com

Discrimination of cervical cancer cells via cognition-based features

Yue Liu*,†, Jiabo Ma*,†, Xu Li*,†, Xiuli Liu*,†, Gong Rao*,†,
Jing Tian*,†, Jingya Yu*,†, Shenghua Cheng*,†, Shaoqun Zeng*,†,
Li Chen‡,¶,** and Junbo Hu§,∥,**

*Britton Chance Center for Biomedical Photonics
Wuhan National Laboratory for Optoelectronics-Huazhong
University of Science and Technology
Wuhan, Hubei 430074, P. R. China

†MoE Key Laboratory for Biomedical Photonics
School of Engineering Sciences
Huazhong University of Science and Technology
Wuhan, Hubei 430074, P. R. China

‡Department of Clinical Laboratory, Tongji Hospital
Huazhong University of Science and Technology
Wuhan, Hubei 430030, P. R. China

§Department of Pathology, Hubei Maternal and Child Health Hospital
Wuhan, Hubei 430072, P. R. China
¶chenliisme@126.com
∥cqjbhu@163.com

Computer-assisted cervical screening is an effective method to save the doctors' workload and improve their work efficiency. Usually, the correct classification of cervical cells depends on the nuclear segmentation effect and the extraction of nuclear features. However, the precise nucleus segmentation remains a huge challenge, especially for densely distributed nucleus. Moreover, previous cellular classification methods are mostly based on morphological features of nucleus size or color. Those individual features can make accurate classification for severe lesions, but not for mild lesions. In this paper, we propose an accurate instance segmentation algorithm and propose cognition-based features to identify cervical cancer cells. Different from previous individual nucleus features, we also propose population features and cognition-based features according to

**Corresponding authors.

the Bethesda System (TBS) for reporting cervical cytology and the diagnostic experience of the cytologists. The results showed that the segmentation achieves better success in complex situations than that by traditional segmentation algorithms. Besides, the cell classification via cognition-based features also help us find out more about less severe lesions' nuclei than that based on conventional features of individual nucleus, meaning an improvement of classification accuracy for cervical screening.

Keywords: Cervical cancer; instance segmentation; nucleus classification; lesion cognition.

1. Introduction

Cervical cancer is the second most dangerous cancer for women in developing countries due to lack of extensive cervical cancer screening.[1] Developing a labor-saving and reliable method is a good choice to address this issue. Some FDA-certified computer-assisted devices, such as ThinPrep imaging system and FocalPoint guided screening system, have played a significant role in the diagnosis or treatment of diseases.[2–5]

Cervical cell classification is mainly based on the features, including the area and color of nucleus, the smoothness of nuclear membranes and nuclear-cytoplasmic ratio (N/C).[6] In other words, the cell nucleus is an important consideration for reporting cervical cytology. Therefore, a precise nucleus segmentation determines the performance of the final classification[7–11] in computer-assisted screening.

Currently, there are many segmentation algorithms[12–15] that perform well on isolated cells, but their effects on overlapping cells or cell clusters are not satisfactory. Since AlexNet[16] got a ground-breaking result in the ImageNet classification
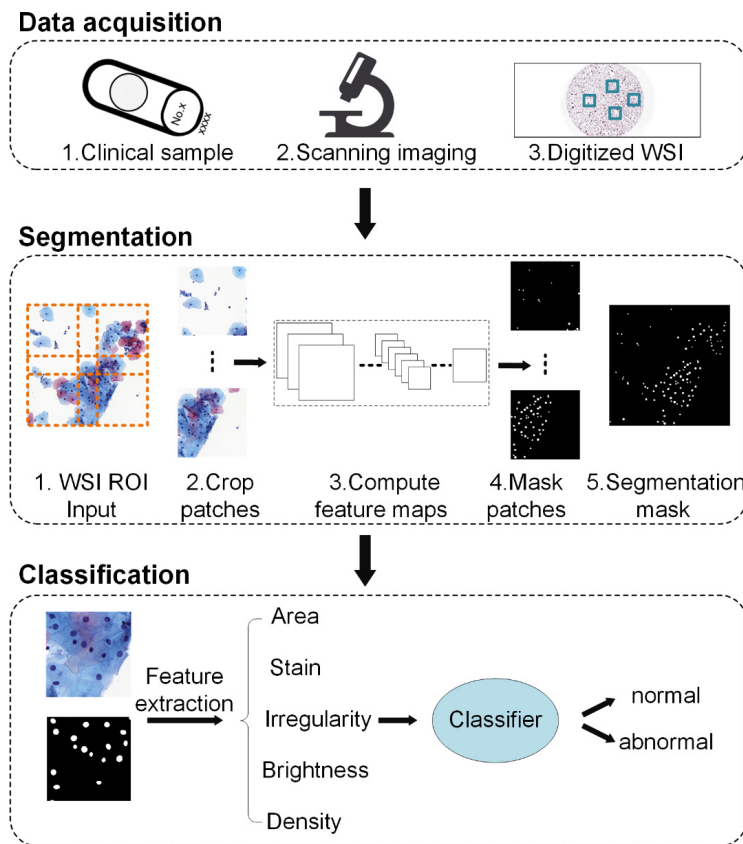

Fig. 1. The overview for the proposed work. Segmentation: we crop the data into patches which match up with the size of Mask-RCNN's input in an overlapping crop. Hence, there exists overlapping region between the neighboring mask patches output by the network.

competition in 2012, the deep learning technique has been widely used in many fields, such as natural language process, speech recognition, image recognition and so on. Some deep learning methods perform well in image segmentation, such as Fully Convolutional Networks (FCN),[17] U-net[18] and Deeplab.[19] Pan *et al.*[13] and Song *et al.*[14] adopted deep learning for nucleus segmentation in pathological images and made a great performance. But these semantic segmentation networks are not able to output nucleus instance in the case of densely distributed nuclei. In this paper, considering the effect in object detection and instance segmentation, we adopt the deep learning network—Mask Region Convolution Neural Network (R-CNN).[20]

For further nuclear feature, there are two main strategies for feature extraction, one is automatic extraction by Convolution Neural Network (CNN), and the other is handcrafted features based on the Bethesda System (TBS) rules. Zhang *et al.*[21] adopted CNN and generated high performance on both the Herlev Pap smear and the HEMLBC datasets. However, the features automatically extracted by CNN are extremely abstract, and it is difficult for pathologists to use these features to quantify the data. Plissiti *et al.*[8] and Bora *et al.*[11] have followed handcrafted features based on TBS rules. These handcrafted features do not perform well on overlapping cells due to the limitation by the performance of segmentation.

Figure 1 shows an overview of our work. Our contributions are summarized as follows. (1) A precise instance segmentation on the nuclei in cytology images is accomplished. (2) According to the pathologist's reading criteria, we not only extract conventional morphological features for single nucleus, but also add nuclear population feature. (3) The results show that our method demonstrates an accuracy of 87.8% and recall of 89.6% in nucleus classification, which has improved significantly compared with the features.[11]

2. Methods

2.1. *Segmentation*

In this study, firstly whole slide images (WSIs) were acquired. When making datasets, we collected 20,000 patches of 500 pixels ×500 pixels from whole slide images. After random rotation, translation and enhancement of pixel transformation, 197, 157 valid data were obtained as the final dataset. Test set was used for evaluating segmentation performance of the deep learning network.

During the training procedure of nuclear segmentation, Thin-Prep cytology test (TCT) data were used. After training, we applied the trained network to an input patch and got an output probability patch, in which each pixel had a probability between 0 and 1. As opposed to extracting patches, we stitched a global probability map by using all probability patches. Here, a threshold was set, and a pixel with probability greater than the threshold will be regarded as a part of a nucleus. After stitching the probability, we will get a segmentation output. When input patches belong to test set, the output and ground truth can be used to calculate intersection over union (IOU) to evaluate the performance of the network.

2.2. *Classification*

In TBS, for reporting cervical cytology,[6] High grade Squamous Intraepithelial (HSIL), Low grade Squamous Intraepithelial (LSIL), Atypical Squamous Cells of Undetermined Significance (ASC-US) and Atypical Squamous Cells-cannot exclude HSIL (ASC-H), are the main types of squamous cell lesions. At the same time, squamous cell lesions account for most of the clinical data. This paper mainly focused on the binary classification between squamous lesion cells and normal cells. As mentioned above, abnormal cells have many different features from normal cells according to Bethesda system. Although the color feature of nucleus differs due to different dye doses, the relative value for features between a random cell and a typical intermediate squamous epithelial cell usually keeps stable in different WSIs. Here, we filter typical intermediate squamous epithelial cells and calculate the mean of their every feature.

2.2.1. *Area*

Cervical cells cannot be completely consistent under various conditions of inflammation, physiological period and age. In cervical cytology, the size of the middle cell nuclei is used as a basic reference for judging other cell abnormalities. Therefore, in order to overcome the divergence of the cells on different slides, not the absolute nuclear area, but the relative

value of the area of the nucleus to the nuclear area of the middle surface cells on the same piece, that is, the relative area was used.

2.2.2. *Perimeter*

We use the total length of the outside edge of an area or a shape as perimeter.[22] As mentioned above, this relative perimeter is defined as the ratio of the perimeter of a nucleus to the mean of typical middle squamous epithelial cell's nuclear perimeter.

2.2.3. *Staining degree*

Usually, the color of abnormal cell's nuclei is deeper than normal cells. Considering that in hue saturation value (HSV) color model, saturation refers to the purity of color, the deeper the staining degree, the greater the saturation. So an original RGB patch was converted to an HSV patch, then the absolute value of staining degree of one single nuclei was described with the mean of the S-channel in an HSV patch. Referring to the above parameters, a relative value was used here, which is defined as the ratio of staining degree of a single nuclei to the mean of typical middle squamous epithelial cell's nuclear staining degree.

2.2.4. *Convex hull area ratio*

Irregularity of nuclear membranes is also an important feature of abnormal cells. Usually, a regular nuclear membrane is smooth and its convex hull is simply its boundary. However, mostly, an irregular nuclear membrane is a concave shape and its convex hull is a convex boundary that most tightly encloses it. So, this feature is defined as the ratio of convex hull area to original area to describe the irregularity of nuclear membranes.

2.2.5. *Brightness*

One kind of LSIL cells exhibit morphological features of koilocyte, which suggests an infection of the cell by human papillomavirus (HPV). It has an obvious cellular change: a clear area around the nucleus, known as a perinuclear halo. So the brightness of perinuclear region is close to the brightness of background, which means it is greater than nuclear brightness.

For some HSIL cells, there is less cytoplasm, almost no cytoplasm in some situations. In this case,

it will appear that perinuclear region of such nucleus is likely to remain in the background.

In the above two situations, the brightness of the perinuclear region may be significantly higher than the brightness of the nucleus. Similar to the staining degree, we convert an RGB patch to an HSV patch, and this feature is defined as the ratio of mean brightness of a perinuclear annulus to the mean brightness of the nuclei.

2.2.6. *Density*

In addition to the parameters considered by traditional methods, such as area, perimeter and so on, we find that the nucleus in densely distributed nuclear population also shows the characteristics of the lesion. Rodriguez *et al.*[23] developed a clustering algorithm based on density peaks, and Cheng *et al.*[24] applied density peak in localization of touching somas. Inspired by the above two methods, we defined this feature.

For a nucleus with center point P_i in the estimated region, the density is defined as

$$\rho_i = \frac{1}{Z} \sum_{j:||P_i-P_j||_2 \leq R} K(P_i - P_j), \quad (1)$$

$$\rho_i = \frac{1}{Z} \sum_{j:||P_i-P_j||_2 \leq R} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-||P_i - P_j||_2^2}{2\sigma^2}\right). \quad (2)$$

Here, $K(P_i, P_j)$ is a Gaussian kernel function with a kernel width σ, Z is a normalization constant, R is the window radius of the kernel function ($R = 2\sigma$), and $||.||_2$ is 2-norm. The kernel width σ is slightly more than half of the average value of typical middle squamous epithelial cell's radius.

2.2.7. *Population features*

In some case, cell has a similar feature to the cells surrounding it, so we add population feature to describe the population. The features is defined as

$$F_i = \frac{1}{Z} \sum_{j:||P_i-P_j||_2 \leq R} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-||P_i - P_j||_2^2}{2\sigma^2}\right) \cdot f_j. \quad (3)$$

Here, f_j is the single feature of a surrounding nucleus, F_i is the cluster feature of current nucleus. For the above features including area, perimeter and staining degree, the calculation is performed
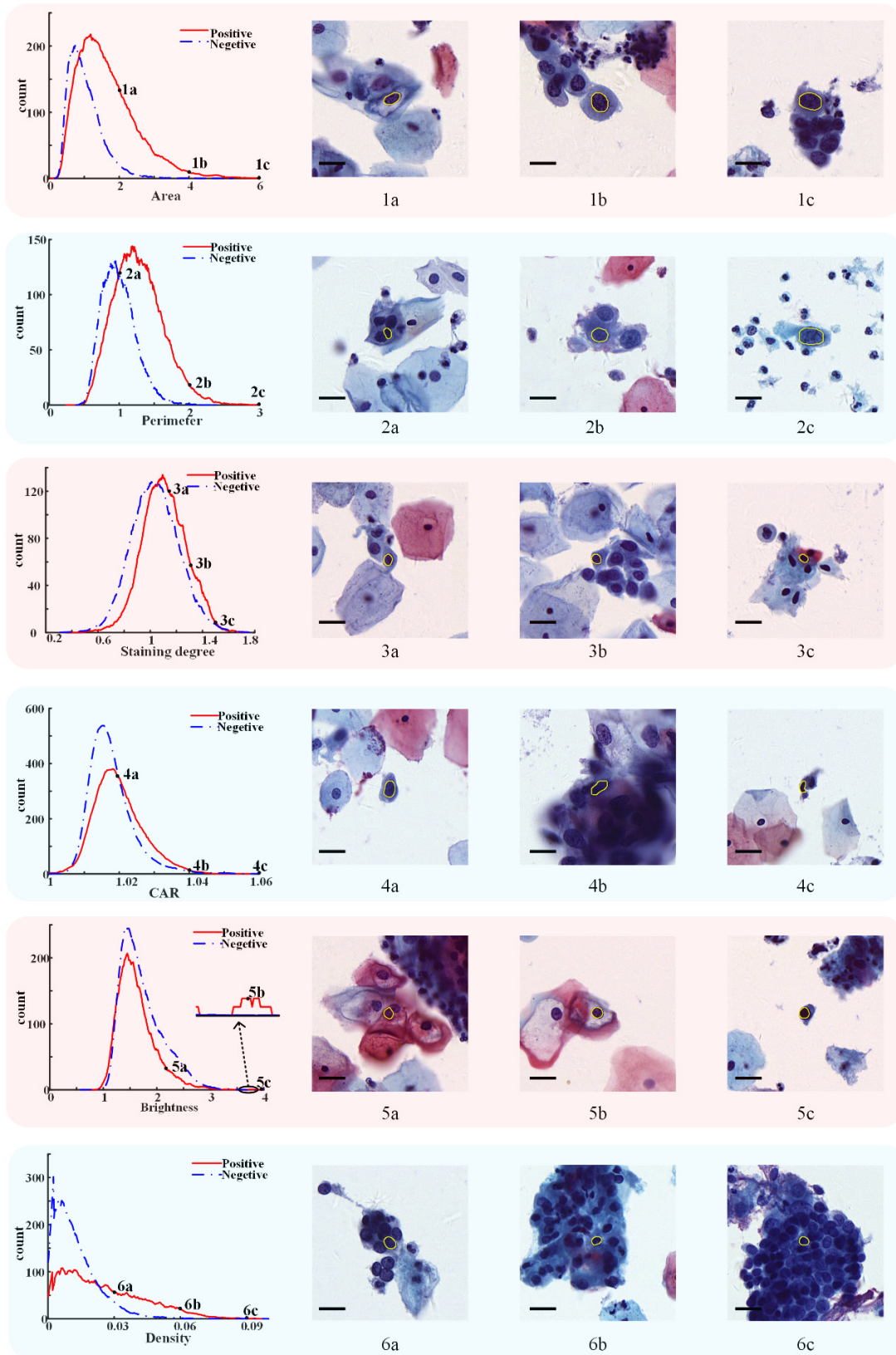
Fig. 2. Distribution of features and typical positive nuclei with different values of features. The nucleus circled by the yellow outline in each patch is a typical sample corresponding to the feature equal to a specific value, and these nuclei are from training set. Scale bar: 20 μm.

Table 1. Composition of datasets.

| Dataset | Segmentation (patch) | Classification (nucleus) | |
		Positive	Negative
Train	183878	19972	20340
Test	13279	4972	5120

using (3). The nuclei in a region of 500 pixels $\times$500 pixels (pixel size of $0.48\,\mu m \times 0.48\,\mu m$) around one nucleus is called the adjacent nucleus, and the distance between adjacent nuclei can be used to measure the density of cells.

After extracting features, the feature sets are normalized to between 0 and 1, and then the normalized data are employed in training and testing for a machine learning classifier. Random Forest (RF) was used in the research. Figure 2 shows distribution of features and typical positive nuclei with different values of features.

2.3. *Dataset*

In this research, 212 TCT slides from the Department of Clinical Laboratory, Tongji Hospital, Huazhong University of Science and Technology were used. These slides were digitized to WSIs by a slide scanning system (3DHISTECH Ltd.) with a $20\times$ magnification objective. Then contours of cell nucleus and its category in slides were precisely annotated with the Qupath, an open, powerful, flexible, extensible software platform for whole slide image analysis.[25] Here, the dataset contains a large number of squamous cells and a small number of glandular cells. Considering the actual incidence and data composition, we mainly consider the lesions of squamous cells in the classification. Some nontumor cytology changes are not considered. These contours of nucleus were employed in segmentation based on Mask-RCNN network, and the categories of nucleus were used for classification in traditional machine learning method. Annotations were performed by a lab technician and subsequently checked by experienced clinicians. The composition of datasets is shown in Table 1.

3. Results

3.1. *Evaluation of nucleus segmentation*

Recall rate and accuracy are indicators used to measure the effect of segmentation and classification.

When the test dataset is input into the Mask-RCNN network, the segmentation results for nuclei will be obtained.

In some cases, the network would make some mistakes such as inflammatory cells, etc. Compared with the ground truth, some nuclear segmentation results have a certain deviation. So we employ Intersection-over-Union (IOU) to define valid segmentation: for a nucleus, comparing the output by the network with the ground truth, if the IOU is above 0.5, it is regarded as valid, otherwise invalid. Finally, we count the number of nuclei in ground truth (N_{GT}), the number of nuclei output by the network (N_O) and the number of nuclei that are valid (N_V). And the recall and the accuracy are defined as

$$\text{recall}_s = N_V/N_{GT}, \qquad (4)$$

$$\text{accuracy}_s = N_V/N_O. \qquad (5)$$

Here, we compared the performance of classical methods and proposed methods on our own dataset, the statistical measures are listed in Table 2 and visual segmentation are illustrated in Fig. 3.

3.2. *Evaluation of nucleus classification*

The performance metrics for classification are accuracy, precision, recall and F-measure, which are given as:

$$\text{accuracy}_c = (TP + TN)/(TP + FP + FN + TN), \qquad (6)$$

$$\text{precision}_c = TP/(TP + FP), \qquad (7)$$

$$\text{recall}_c = TP/(TP + FN), \qquad (8)$$

$$F\text{-measure} = 2 * \text{precision} * \frac{\text{recall}}{\text{precision} + \text{recall}}. \qquad (9)$$

Here true positives (TP) represents the number of the correctly classified positive nucleus, and false positives (FP) represents the number of the wrongly

Table 2. Performances of different segmentation algorithms.

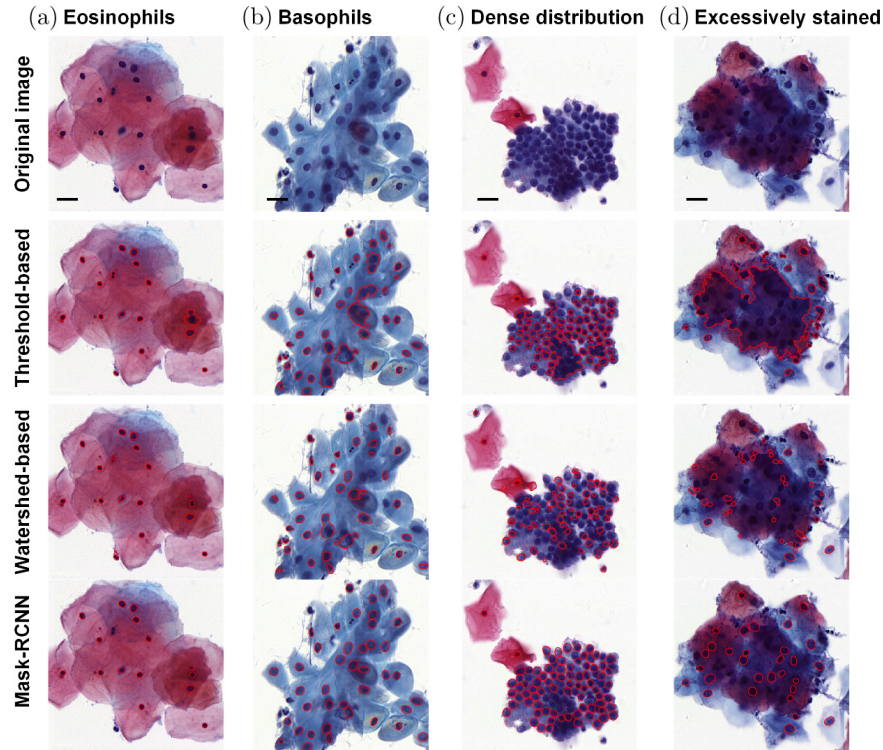| Methods	Accuracy	Recall
Threshold-based	0.8342	0.6923
Watershed-based	0.7964	0.6537
Proposed	0.9713	0.9720

Fig. 3.   Segmentation results of typical images obtained by different algorithms. Each column in the illustration showed the performance in the different situations. (a) Eosinophils, (b) basophils, (c) dense distribution, (d) excessively stained. In situation (a) and (b), conventional threshold-based and watershed-based algorithms still has a tolerable performance. However, in complex situation (c) and (d), they seem to be bad choices. Scale bar: 20 μm.

classified positive nucleus. For negative nucleus, the meanings of TN and FN are similar to the meanings of TP and FP.

Accuracy is the most common evaluation metric, the higher the accuracy, the better the classifier. However, when we have an asymmetric dataset, we need to refer other parameters to evaluate the performance of the classifier, such as precision and recall. The precision reflects the false positive rate, the higher the precision, the lower the false positive rate. The recall rate evaluates the false negative rate, the higher the recall, the lower the false negative rate. F-measure is a harmonic mean of recall and precision. Therefore, F-measure takes both false positive rate and false negative into account. In summary, the higher F-measure, the better performance of the classifier.

Here, the proposed features were compared with existing approach,[11] the statistical results of Ref. 11 are listed in Table 3 and the statistical measures of proposed are shown in Table 4. The morphological features adopted in Ref. 11 were area, perimeter, eccentricity, circularity and compactness.

From Tables 3 and 4, we can draw a conclusion that the performance of the classifiers trained by proposed features are better than that trained by features in Ref. 10. Compared with the previous methods, in addition to conventional features such as area and perimeter, we have introduced some new features from the following aspects.

Table 3.   Classification performance based on features in Ref. 10.

Classifiers	Accuracy	Precision	Recall	F-measure
SVM	0.7126	0.6930	0.7255	0.7086
RF	0.7168	0.7207	0.6738	0.6965
LR	0.6663	0.6550	0.6504	0.6527

Table 4.   Classification performance based on our features.

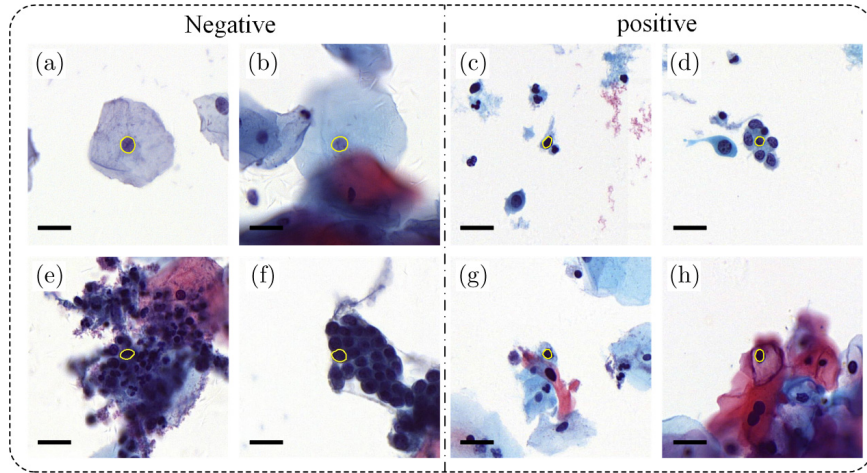| Classifiers	Accuracy	Precision	Recall	F-measure
SVM	0.8768	0.8610	0.8985	0.8794
RF	0.8984	0.8988	0.8979	0.8984
LR	0.8517	0.8780	0.8169	0.8463

Fig. 4. Typical nuclei that are not recalled. (a)–(d) nuclei which are not recalled by classifier based on Ref. 10. (e)–(h) nuclei which are not recalled by classifier based on proposed method. (a), (b), (e) and (f) represent negative nuclei, (c), (d), (g) and (h) represent positive nuclei. Scale bar: 20 μm.

(1) Irregularity of nuclear membranes: abnormal cells often exhibit irregular nuclear membranes instead of significant changes in eccentricity; (2) color information: brightness feature defined by us played a key role in discriminating koilocyte and abnormal cells with little cytoplasm; (3) density and population information improved the classification performance for nuclei which are densely distributed.

Typical nuclei which are not recalled by classifiers are illustrated in Fig. 4. Area has a great impact on the classifier trained by Ref. 11, negative nuclei with large area are mistaken, such as Figs. 4(a) and 4(b).

The distribution of abnormal scores of nuclei in test set are illustrated in Fig. 5. The abnormal scores that is returned by our classifier range from 0 to 1, and the closer the index is to 1, the more abnormal the slide. In the figure, the blue dots indicate negative slides and the red dots indicate positive
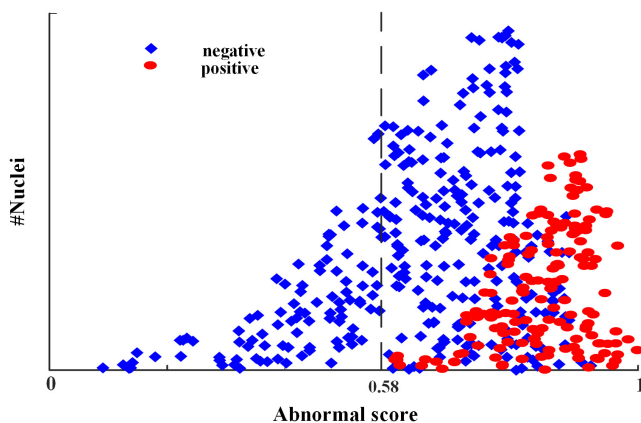
slides, and it can be seen from the distribution of the two types of slides that there is the potential to exclude partially negative slides.

4. Discussion and Conclusion

Traditional algorithm for the segmentation of nuclei perform well on discrete cells and does not perform well on overlapping cells. Segmentation of overlapping cells remains one of the most challenging problems in the image analysis. Semantic segmentation networks based on deep learning has a better performance than traditional methods.[14] The instance segmentation network overcomes a variety of complex situations except for rare cases, which lays a solid and reliable foundation for classification. The classifier trained by conventional handcrafted features, which mainly took conventional morphology features into account, are prone to misclassification of nuclei in Figs. 4(a)–4(d). Our method increased some features related to TBS rules and clusters information such as cell density.

Despite a perfect performance, our method still needs further optimization. (1) Need to improve the generalization of the method. It is also of great significance in the actual clinical application to improve the generalization ability of our method. Because the inconsistencies of various external factors including the method of sample preparation, dye dose and parameter of imaging instruments may result in difference of the raw data. Although these differences are not obvious in human vision, they may have great negative impact on the

Fig. 5. Distribution of abnormal score of nuclei in test dataset.

performance of algorithms. We are also working to convert raw data from different source into highly consistent data for analysis through Generative Adversarial Networks (GAN). (2) There still exists some misclassifications in situations like Figs. 4(e)–4(h). Besides, some densely distributed nuclei like endocervical canal cell in Fig. 4(f) also may be misclassified. The density feature defined by us describes a part of abnormal dense cell clusters, but the polarity of the arrangement is not considered here. As described in the TBS's rules, endocervical canal cell are arranged in a honeycomb shape and feather-like structure is a typical feature of AIS (endocervical adenocarcinoma *in situ*). So, next we will also define more useful and easily understood features for a better classification in the future.

Acknowledgment

References

1. A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, D. Forman, "Global cancer statistics," *CA Cancer J Clin.* **61**(2), 69–90 (2011).

2. E. J. Topol, "High-performance medicine: The convergence of human and artificial intelligence," *Nat. Med.* **25**(1), 44–56 (2019).

3. A. E. Dawson, "Can we change the way we screen?: The ThinPrep Imaging System," *Cancer* **102**(6), 340–344 (2004).

4. D. C. Wilbur, W. S. Black-Schaffer, R. D. Luff, K. P. Abraham, C. Kemper, J. T. Molina, W. D. Tench, "The Becton Dickinson FocalPoint GS Imaging System: Clinical trials demonstrate significantly improved sensitivity for the detection of important cervical lesions," *Am. J. Clin. Pathol.* **132**(5), 767–775 (2009).

5. Y. Jusman, S. C. Ng, K. Hasikin, R. Kurnia, N. A. A. Osman, K. H. Teoh, "A system for detection of cervical precancerous in field emission scanning electron microscope images using texture features," *J. Innov. Opt. Health. Sci.* **10**(2), 1650045 (2017).

6. Y. S. Erozan, "The Bethesda system for reporting Cervical cytology," *Acta Cytol.* **48**(5), 770 (2004).

7. Y. Marinakis, G. Dounias, J. Jantzen, "Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification," *Comput. Biol. Med.* **39**(1), 69–78 (2009).

8. M. E. Plissiti, C. Nikou, Cervical cell classification based exclusively on nucleus features, *Int. Conf. Image Anal. Recognit.* **7325**, 483–490 (2012).

9. Y. F. Chen et al., "Semi-automatic segmentation and classification of Pap smear cells," *IEEE J. Biomed. Health Inform.* **18**(1), 94–108 (2014).

10. E. Bengtsson, P. Malm, "Screening for cervical cancer using automated analysis of PAP-smears," *Comput. Math. Methods Med.* **2014**(2962), 842037 (2014).

11. K. Bora, M. Chowdhury, L. B. Mahanta, M. K. Kundu, A. K. Das, "Automated classification of Pap smear images to detect cervical dysplasia," *Comput. Methods Programs Biomed.* **138**, 31–47 (2017).

12. S. F. Yang-Mao, Y. K. Chan, Y. P. Chu, "Edge enhancement nucleus and cytoplast contour detector of cervical smear images," *IEEE Trans. Syst. Man Cybern. B. Cybern*, **38**(2), 353–366 (2008).

13. X. Pan et al., "Accurate segmentation of nuclei in pathological images via sparse reconstruction and deep convolutional networks," *Neurocomputing* **229**, 88–99 (2017).

14. Y. Song et al., "A deep learning based framework for accurate segmentation of cervical cytoplasm and nuclei," *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2014**, 2903–2906 (2014).

15. K. Li, Z. Lu, W. Liu, J. Yin, "Cytoplasm and nucleus segmentation in cervical smear images using Radiating GVF Snake," *Pattern Recognit.* **45**(5), 1255–1264 (2012).

16. K. Alex, S. Ilya, H. Geoff, "Imagenet classification with deep convolutional neural networks," *Proceedings of NIPS, IEEE, Neural Information Processing System Foundation* (2012), pp. 1097–1105.

17. E. Shelhamer, J. Long, T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(4), 640–651 (2017).

18. O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, *Int. Conf. Medical Image Computing & Computer-assisted Intervention* (2015).

19. L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.* **40**(4), 834–848 (2018).

20. K. He, G. Gkioxari, P. Dollár, R. Girshick, "Mask R-CNN," *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2018), pp. 1–1.

21. L. Zhang, L. Le, I. Nogues, R. M. Summers, S. Liu, J. Yao, "DeepPap: Deep convolutional networks for

cervical cell classification," *IEEE J. Biomed. Health Inform.* **21**(6), 1633–1643 (2017).

22. K. Benkrid and D. Crookes, "Design and FPGA implementation of a perimeter estimator," *Proceedings of the Irish Machine Vision and Image Processing Conference* (2000), pp. 51–57.

23. A. Rodriguez, A. Laio, "Machine learning. Clustering by fast search and find of density peaks," *Science* **344**(6191), 1492–1496 (2014).

24. S. Cheng, T. Quan, X. Liu, S. Zeng, "Large-scale localization of touching somas from 3D images using density-peak clustering," *BMC Bioinform.* **17**(1), 375 (2016).

25. P. Bankhead *et al.*, "QuPath: Open source software for digital pathology image analysis," *Sci. Rep.* **7**(1), 16878 (2017).