Hierarchy-NMS：Merging Candidate Bounding Boxes for Cerebrospinal Fluid Cell Image Segmentation

Xianwei Xu[1], Fangqi Li[1], Shilin Wang[1*], Zhenhai Wang[2*]

[1]School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China.

[2]Neurology Center, General Hospital of Ningxia Medical University, Yinchuan, China.
[*]Corresponding authors.
{xuxianwei,solour_lfq,wsl}@sjtu.edu.cn, wangzhenhai1968@163.com.

Abstract. In this article, Hierarchy-Non Maximum Suppression (H-NMS) is proposed to fix the defects of general NMS algorithms in cerebrospinal fluid cell image segmentation. To cope with the confusion caused by the hierarchy of candidate bounding boxes in this scenario, a tree structure is built on the top of all bounding boxes and only leaf nodes are saved as the final output. The experimental results showed that H-NMS outperforms some variants of NMS in the cerebrospinal fluid cell image processing in both image segmentation and cell counting.

1. Introduction

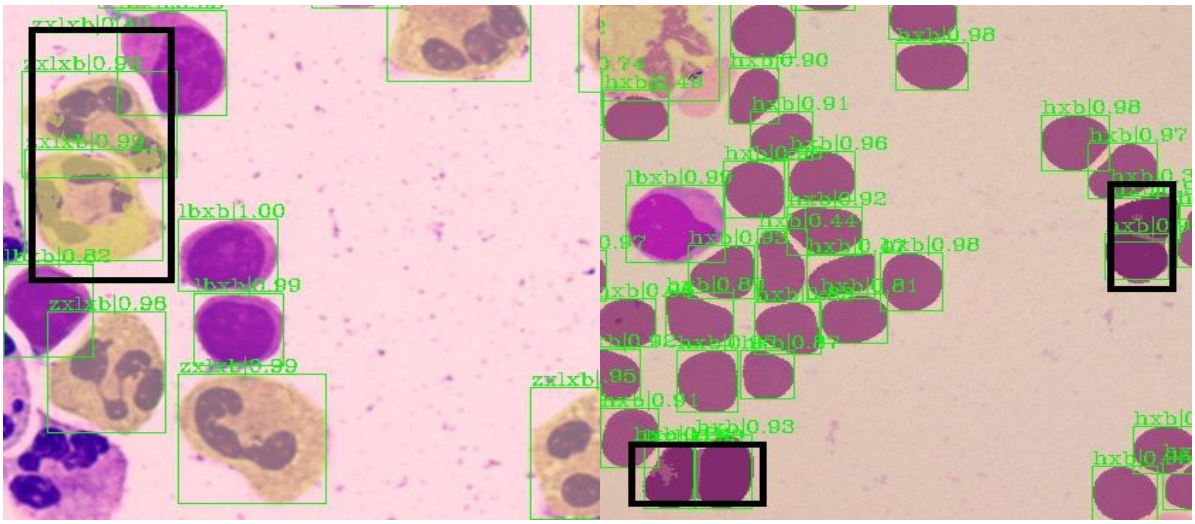
Cerebrospinal fluid (CSF) is a sort of transparent, viscous liquid between the arachnoid and pia mater of the meninges[1]. An adult' s body contains only about 140 ml of CSF. CSF plays an important role in protecting the brain from minor external damage, protecting nutrients in the central nervous system, and stabilizing chemical regulation of the central nervous system. The identification of the morphology and quantity of CSF cells is a crucial part of clinical CSF cytology examinations. From the perspective of cytotaxonomy, though little in number, CSF cells have diversified shapes. This fact makes CSF cells identification extremely difficult. In practice, CSF cytology is generally conducted by experienced specialists who firstly use a CSF cell slide centrifuge to collect CSF cells, stain them, and then observe them with an optical microscope. The entire process is done manually, which is inefficient and expensive. Therefore, it is urgent to develop an automatic system to segment the CSF images and numerate the different types of cells. With the popularization of deep learning, it has become a promising tool in image processing, e.g., segmentation in analysing the MRI images[2], medical image analysis[3], etc. Image segmentation is an indispensable pre-processing procedure for many image-based clinical applications, including the brain segmentation[4], the cardiac ventricle

segmentation[5], the abdominal organ segmentation[6], and cell segmentation in biological images[7]. The output of an image segmentation is usually a set of bounding boxes around objects of interest in the original image. In current applications, Non Maximum Suppression (NMS)[8] is often used conjugately with image segmentation to select the optimal set of bounding boxes. In NMS, candidate bounding boxes are sorted according to their confidences, a bounding box with an intersection over union (IoU)[9] ratio higher than some threshold with another candidate with a higher confidence would be deleted. However, the threshold in NMS is hard to determine. A low threshold would delete too many boxes that hinder the overall performance, while a threshold too high might cause false detection. Soft-NMS[10] has been proposed to cope with the defects of ordinary NMS. In Soft-NMS, the candidate box with a high IoU than another box gets a decline in confidence rather than is deleted directly. Soft-NMS neatly handled the problem of missing detection in NMS. Soft-NMS algorithm is a minute change to NMS algorithm without introducing additional parameters and is concise in implementation.

In the field of CSF cell image processing, a new problem emerges and new methods have to be developed. We propose a new candidate box merging algorithm called Hierarchy-NMS (H-NMS) in this scenario. The paper proceeds with the motivation behind our proposal, the proposal itself, extensive experimentation and the conclusion.

2. The Motivation and the Proposed Methods

For the segmentation algorithm, we adopted the state-of-the-art Cascade Mask-rcnn network model[11]. However, when using the traditional NMS to merge the bounding boxes, it was observed that many candidate bounding boxes encircled multiple cells and these boxes might not be eliminated during NMS. This is due to the fact that for CSF cell images, cells of a same type often form a cluster in image, as in Figure. 1, marking a crucial difference from general images. This phenomenon significantly increasing the number of remained boxes, hence the guess of the number of cells, which is mostly undesired for the task of cell counting. Meanwhile, it is empirically found that Soft-NMS could not help in alleviating this problem. Since NMS was designed for general images, it failed to address the specific problem in CSF cell images. Therefore, the Hierarchy-NMS is proposed.



(a) Candidate boxes for neutrophils. (b) Candidate boxes for red blood cells.

Figure. 1 Renderings of NMS algorithm in different types of cells. The black bounding boxes contain multiple cells and cannot be deleted by NMS.

To explicitly incorporate the inclusion relation between the candidate bounding boxes. We generalize the ordinary NMS by formulating the data structure behind from a list to a tree. The intuition after this formulation is: if one candidate bounding box is included in/include some target box with higher confidence (those boxes that are already fixed as candidate outputs), then it is located in the memory as the child/parent node of the node corresponds to the target bounding box. Formally, we still have to cope with contradiction and initialization. Our method is summarized in the following pseudo-program.

Algorithm 1 Hierarchy-NMS

1: **Input:** $B = \{b_1,..,b_N\}$ the list of candidate bounding boxes (of one type of cell in an image), $S = \{s_1,..,s_N\}$ the list of their corresponding confidence score, N_1 the IoU threshold and N_2 the IoL threshold of IoL, where IoL(box a, box b) is the ratio between the area of their intersection and that of b.
2: **Initialize:** Sort all the candidate boxes according to the confidence score.
3: **Initialize:** The root node r.
4: **Define:** Insert (candidate bounding box b, node n):
5: $C \leftarrow$the children of n
6: **for** c_i in C **do**
7: **If** IoU $(b, c_i) \geqq N_1$ **then exit**
8: **for** c_i in C **do**
9: **If** IoL$(c_i, b) \geqq N_2$ **then** Insert(b, c_i).
10: **for** c_i in C **do**
11: **If** IoL$(b, c_i) \geqq N_2$ **then** setting the parent of c_i as b, setting the parent of b as n.
11: Setting the parent of b as n, **exit**.
12:**for** b in B **do**
13: Insert(r, b)
13: **Output:** all candidate boxes represented by leaf nodes in the tree.

To see why H-NMS is a generalization of NMS and fixes the problem with CSF cell images, note the following two observations:

(1) If one bounding box is deleted according to NMS then so it be by H-NMS in the building state, *vice versa*. During the process of building up the H-NMS internal tree structure, each candidate bounding box is compared with those boxes that are located close to it (moreover, H-NMS prevents some comparisons between boxes that are far from each other, i.e., boxes correspond to the same depth but belong to different parents would never be compared). And the deletion rule is exactly the same as NMS, hence the sets of boxes being deleted are identical in both cases.

(2) For each node in the tree, its children nodes represent a series of candidate bounding boxes that are contained within the boxes corresponding to the parent node. This property is obviously implied by the operating process of the H-NMS.

To conclude, H-NMS is a generalization of NMS with a better structured internal data structure. Apart from realizing what NMS can do, it saves computation cost and better fits the task of bounding

box filter in CS cell images.

3. Experimental results and Discussions

The experiments were conducted on the CSF cell image dataset. Due to the lack of public dataset of CSF cell image research, we calibrated cells in CSF cell images manually and built our own dataset. At present, there are 659 images in the training set, 75 images in the validation set and 75 images in the test set. The dataset was converted into the format of COCO dataset[12], with the image resolution 912*608. And the results obtained by Hierarchy-NMS algorithm are shown in Figure. 2, it is observed that H-NMS successfully eliminate large boxes which contain multiple cells of the same class. For comparison, NMS and Soft-NMS were adopted as baseline methods in the post-processing stage of CSF cell image segmentation.

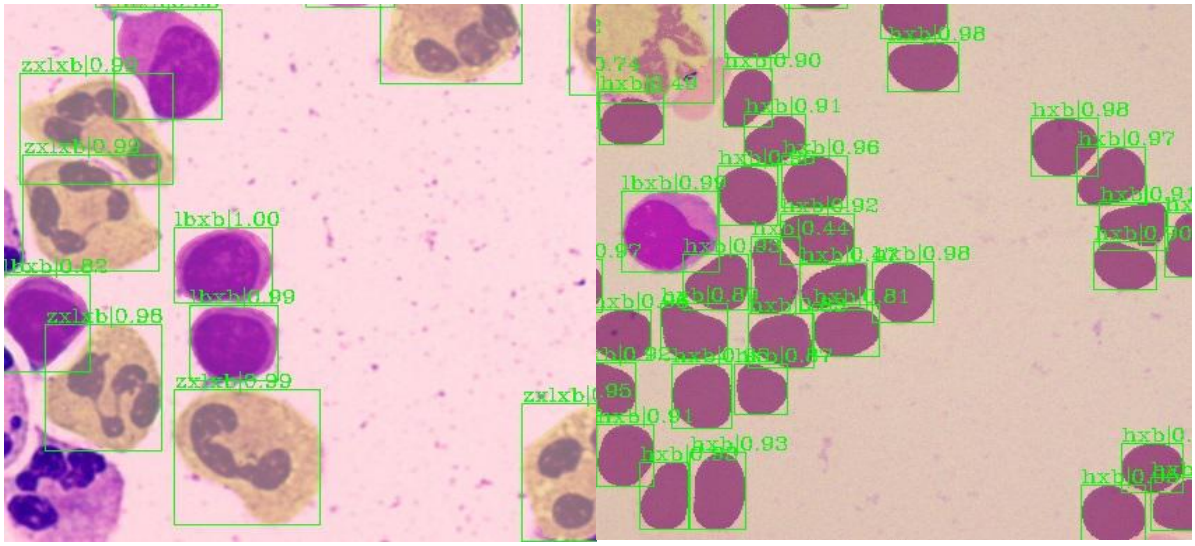


Figure. 2 Renderings of Hierarchy-NMS algorithm in different types of cells

3.1. Comparison of CSF cell image segmentation results with different algorithms

The benchmark COCO metrics including AP_{50}, AP_{75}, AP (averaged over IoU thresholds) and AP_S (all cells belong to S scale) were adopted as indices. Unless specified, AP is estimating utilizing mask IoU. In the post-processing stage of segmentation, we retrained the network model with NMS, Soft-NMS and H-NMS. We compare segmentation results of different algorithms in Table 1. It can be seen that the AP value of H-NMS algorithm is higher than that of NMS and Soft-NMS. And the image segmentation by H-NMS algorithm works best.

Table 1. Comparison of different AP values on segmentation with different algorithms

Algorithms	AP	AP_{50}	AP_{75}	AP_S
NMS	0.591	0.674	0.665	0.674
Soft-NMS	0.589	0.666	0.656	0.682
Hierarchy-NMS	0.611	0.683	0.677	0.712

3.2. Comparison of counting results in CSF cell image with different algorithms

To measure the result of counting cells, we proposed the average counting accuracy (ACA) index, and

adopted indexes of counting evaluation such as checkout accuracy (cAcc), average counting distance (ACD), mean category counting distance (mCCD), and mean category intersection of union (mCIoU) following Xiu-Shen Wei et al.[13]. For example, the ACA index is computed following: define M as the number of test images, PD_i as the predicted number of cells in the i-th image, and GT_i as the ground truth number of cells in the i-th image, then:

$$ACA = 1 - \sum_{i=1}^{M} \frac{|PD_i - GT_i|}{GT_i.M} \quad (1)$$

We list the counting results of different post-processing methods in Table 2. It can be seen that the performance of H-NMS is better than NMS and Soft-NMS.

Table 2. Comparison of counting results with different algorithms

Algorithms	ACA	cAcc	ACD	mCCD	mCIoU
NMS	36.93%	20.00%	4.19	3.03	30.30%
Soft-NMS	22.05%	16.00%	5.35	3.10	27.84%
Hierarchy-NMS	46.12%	29.33%	3.44	2.99	32.45%

3.3. Discussions

In the CSF cell image segmentation scenario, NMS fails to delete large boxes which contain multiple cells according to the IoU standard, which fact will reduce the counting performance of the system. Meanwhile, Soft-NMS is also not suitable for the CSF cell image scene. While H-NMS not only realizes the basic functions of NMS but also filters out confusing large boxes that contain multiple cells. It is conducive to the improvement of the image segmentation and counting results.

4. Conclusion

In general NMS algorithms, bounding boxes are linearly sorted according to their confidences. However, such linearity does not help in the CSF cell images. To cope with this problem, we explicitly integrate the inclusion relation between candidate bounding boxes into H-NMS and construct the set of bounding boxes into a tree, which successfully solves the problem with CSF cell images.

Although NMS algorithm and Soft-NMS algorithm have good performance in merging candidate bounding boxes in general images, they fail to succeed in CSF cell image segmentation, where H-NMS outperforms them in both segmentation performance and counting results. This fact is verified by our experiments on an actual medical dataset.

References

[1] Igual L, Soliva J C, Hernandezvela A, et al. A fully-automatic caudate nucleus segmentation of brain MRI: Application in volumetric analysis of pediatric attention-deficit/hyperactivity disorder[J]. Biomedical Engineering Online, 2011, 10(1): 105-105.

[2] Sundareswaran K S, Frakes D H, Fogel M A, et al. Optimum fuzzy filters for phase - contrast magnetic resonance imaging segmentation[J]. Journal of Magnetic Resonance Imaging, 2009, 29(1): 155-165.

[3] Sharma N, Aggarwal L M. Automated medical image segmentation techniques.[J]. Journal of Medical Physics, 2010, 35(1): 3-14.

[4] De Brebisson A, Montana G. Deep neural networks for anatomical brain segmentation[C]. computer vision and pattern recognition, 2015: 20-28.
[5] Avendi M R, Kheradvar A, Jafarkhani H, et al. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI[J]. Medical Image Analysis, 2016: 108-119.
[6] Roth H R, Oda H, Hayashi Y, et al. Hierarchical 3D fully convolutional networks for multi-organ segmentation.[J]. arXiv: Computer Vision and Pattern Recognition, 2017.
[7] Ronneberger O, Fischer P, Brox T, et al. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]. medical image computing and computer assisted intervention, 2015: 234-241.
[8] Hosang J, Benenson R, Schiele B, et al. Learning Non-maximum Suppression[C]. computer vision and pattern recognition, 2017: 6469-6477.
[9] Rezatofighi H, Tsoi N, Gwak J, et al. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression[C]. computer vision and pattern recognition, 2019: 658-666.
[10] Bodla N, Singh B, Chellappa R, et al. Soft-NMS — Improving Object Detection with One Line of Code[C]. international conference on computer vision, 2017: 5562-5570.
[11] Chen K, Pang J, Wang J, et al. Hybrid Task Cascade for Instance Segmentation[J]. arXiv: Computer Vision and Pattern Recognition, 2019.
[12] Lin T, Maire M, Belongie S, et al. Microsoft COCO: Common Objects in Context[C]. european conference on computer vision, 2014: 740-755.
[13] Wei X, Cui Q, Yang L, et al. RPC: A Large-Scale Retail Product Checkout Dataset.[J]. arXiv: Computer Vision and Pattern Recognition, 2019.