An Anti-biased TBSRTC-Category Aware Nuclei Segmentation Framework with a Multi-label Thyroid Cytology Benchmark

Junchao Zhu[1], Yiqing Shen[2], Haolin Zhang[3], and Jing Ke[4,5(✉)]

[1] School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China
junchaozhu@sjtu.edu.cn
[2] Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA
yshen92@jhu.edu
[3] School of Engineering, Shanghai Ocean University, Shanghai, China
m220851484@st.shou.edu.cn
[4] School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China
kejing@sjtu.edu.cn
[5] School of Computer Science and Engineering, University of New South Wales, Sydney, Australia

Abstract. The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) has been widely accepted as a reliable criterion for thyroid cytology diagnosis, where extensive diagnostic information can be deduced from the allocation and boundary of cell nuclei. However, two major challenges hinder accurate nuclei segmentation from thyroid cytology. Firstly, unbalanced distribution of nuclei morphology across different TBSRTC categories can lead to a biased model. Secondly, the insufficiency of densely annotated images results in a less generalized model. In contrast, image-wise TBSRTC labels, while containing lightweight information, can be deeply explored for segmentation guidance. To this end, we propose a TBSRTC-category aware nuclei segmentation framework (`TCSegNet`). To top up the small amount of pixel-wise annotations and eliminate the category preference, a larger amount of image-wise labels are taken in as the complementary supervision signal in `TCSegNet`. This integration of data can effectively guide the pixel-wise nuclei segmentation task with a latent global context. We also propose a semi-supervised extension of `TCSegNet` that leverages images with only TBSRTC-category labels. To evaluate the proposed framework and also for further cytology cell studies, we curated and elaborately annotated a multi-label thyroid cytology benchmark, collected clinically from 2019 to 2022, which will be made public upon acceptance. Our `TCSegNet` outperforms state-of-the-art segmentation approaches with an improvement of 2.0% Dice and 2.7% IoU; besides, the semi-supervised extension can further boost this margin. In conclusion, our study explores the weak annotations by constructing an image-

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14225, pp. 580–590, 2023.
https://doi.org/10.1007/978-3-031-43987-2_56

wise-label-guided nuclei segmentation framework, which has the potential medical importance to assist thyroid abnormality examination. Code is available at https://github.com/Junchao-Zhu/TCSegNet.

Keywords: Unbalanced Nuclei Segmentation · Semi-Supervised Learning · Thyroid Cytology · TBSRTC Diagnostic Category

1 Introduction

Thyroid cancer is the most common cancer of the endocrine system, accounting for 2.1% of all malignant cancers [1]. Clinically, pathologists rely on the six-category of "The Bethesda System for Reporting Thyroid Cytopathology" (TBSRTC) [2,3] to distinguish the cell morphology in the stained cytopathologic sections. The emergence of computational pathology allows automatic diagnosis of thyroid cancer, and nuclei segmentation becomes one of the most critical diagnostic tasks [4,5], as the shapes of nuclei, whether round, oval, or elongated, can provide valuable information for further analysis [6]. For example, small and scattered thyroid cells with a light hue and relatively low cell density are usually low-grade and indicative of early-stage cancer; whereas large and dark cells with extreme-dense agglomeration are usually middle- or late-grade [3]. Correspondingly, accurate location of cell boundaries is essential for both pathologists and computer-aided diagnosis (CAD) systems to assist decision [7].

However, nuclei segmentation in thyroid cytopathology is still challenged by the varying cellularity of images from different TBSRTC categories [3,8]. For example, benign cells (I & II) present high sparsity and are difficult to be distinguished from background tissues, thus may account for a relatively small proportion when equal images are involved in a training set [3]. By contrast, high-grade cells (V & VI) are densely packed and severely clustered, thus much more are presented in a training set. In this way, an unbalanced distribution across different categories resulted, correspondingly, the training leads to biased models with lower accuracy [9,10]. Such distinct morphological differences can be characterized by the TBSRTC category, which thus inspires us to utilize the handy image-wise grading labels to guide the nuclei segmentation model learning from unbalanced datasets. We also noticed that another challenge for accurate nuclei identification is the heavy reliance on large-scale high-quality annotations [11]. Moreover, amongst multiple annotation paradigms [12], pixel-level labeling is the most time-consuming and laborious, whereas the image-wise diagnostic labels, *i.e.* TBSRTC categories, are comparatively simpler. Despite the labeling intensity, prevalent nuclei segmentation methods, e.g., CIA-Net [13], CA$^{2.5}$-Net [14], and ClusterSeg [15], are limited to pixel-wise annotations, where the potential benefits of integrating accessible image-wise labels are unaware.

To narrow the gap discussed, we propose a novel TBSRTC-category-aware nuclei segmentation framework. Our contributions are three-fold. (1) We propose a cytopathology nuclei segmentation network named `TCSegNet`, to provide supplementary guidance to facilitate the learning of nuclei boundaries.

Innovatively, our approach can help reduce bias in the learning process of the segmentation model with the routine unbalanced training set. (2) We expand `TCSegNet` to `Semi-TCSegNet` to leverage image-wise labels in a semi-supervised learning manner, which significantly reduces the reliance on annotation-intensive pixel-wise labels. Additionally, an HSV-intensity noise is designed specifically for cytopathology images to boost the generalization ability. (3) We establish a dataset of thyroid cytopathology image patches of 224×224, where 4,965 image labels are provided following TBSRTC, and 1,473 of them are densely annotated [3] (to be on GitHub upon acceptance). To the best of our knowledge, it is the first publicized thyroid cytopathology dataset of both image-wise and pixel-wise labels. The annotated dataset well alleviates the insufficiency of an open cytopathology dataset for computer-assisted analysis (Fig. 1).

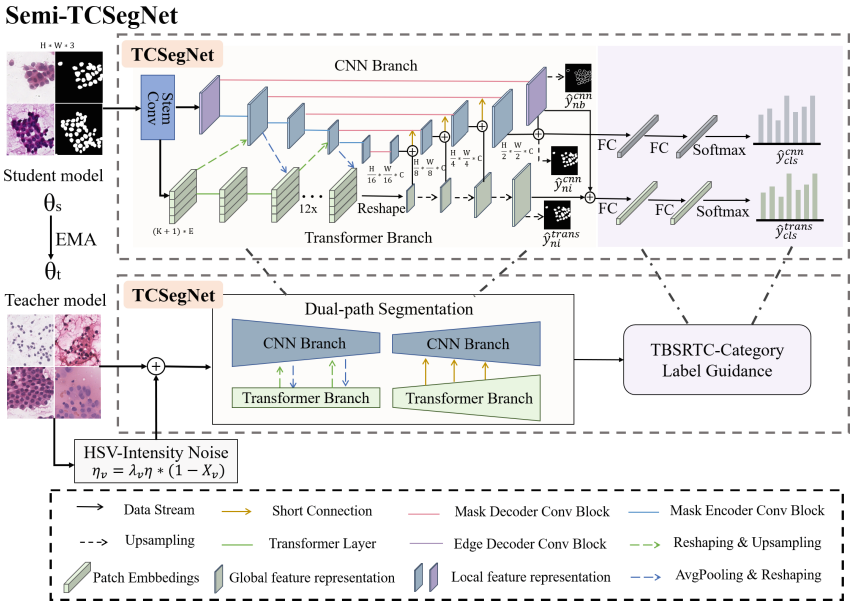

Fig. 1. An overview of the proposed `TCSegNet` and `Semi-TCSegNet`. `TCSegNet` utilizes annotation-lightweight image-wise TBSRTC-category labels to aid in the learning of unbalanced nuclei morphology in segmentation. `Semi-TCSegNet` reduces the heavy reliance on annotation-intensive pixel-wise labels through the use of a semi-supervised framework.

2 Methodology

Overview. We propose a novel TBSRTC-category Aware Segmentation Network (`TCSegNet`) to segment nuclei boundaries in cytopathology images, which

is guided by TBSRTC-category label to learn from unbalanced data. Our model uses a CNN and Transformer dual-path U-shape architecture, where the CNN captures the local features, and the Transformer extracts the global features for a more comprehensive representation of nuclei allocation [16]. Considering the spatial distributions of thyroid cells in cytopathology images, our design provides extended global information for more accurate segmentation. Our approach employs short connections to allow effective communication of local and global representations [17]. Formally, the overall segmentation loss $\mathcal{L}_{seg}$ to train our model is a combination of the binary cross-entropy loss (BCE), *i.e.*

$$\mathcal{L}_{seg} = \gamma_{ni} \cdot \texttt{BCE}(\hat{y}_{ni}^{cnn}, y_{ni}) + \gamma_{ni} \cdot \texttt{BCE}(\hat{y}_{ni}^{trans}, y_{ni}) + \gamma_{nb} \cdot \texttt{BCE}(\hat{y}_{nb}^{cnn}, y_{nb}), \quad (1)$$

where $\hat{y}$ is the prediction from TCSegNet and y is and pixel-wise annotation, respectively. Subscript ni and nb denote the nuclei area and boundary, respectively. Superscripts cnn and $trans$ write for the CNN branch and Transformer branch, respectively. We set the balancing coefficient γ_{ni} to 1 and γ_{nb} to 10. Additionally, to ensure the consistency between the two branches, we impose a dice consistency loss ($\mathcal{L}_{cons}$) between the nuclei instance predictions from the CNN branch and the Transformer branch, namely $\mathcal{L}_{cons} = \texttt{Dice}(\hat{y}_{ni}^{cnn}, \hat{y}_{ni}^{trans})$.

TBSRTC-Category Label Guidance Block. In TCSegNet, we introduce a TBSRTC-category label guidance block to address the learning issue from unbalanced routine datasets. This block consists of two learnable fully connected layers that process the feature extracted by the CNN and Transformer branches separately, which obtains image-wise TBSRTC-category prediction denoted as $\hat{y}_{cls}^{cnn}$ and $\hat{y}_{cls}^{trans}$. Correspondingly, to train this block, we use a cross-entropy loss function (CE) that provides an extra supervision signal to help the network learn from unbalanced datasets, defined as follows:

$$\mathcal{L}_{cls} = \texttt{CE}(\hat{y}_{cls}^{cnn}, y_{cls}) + \gamma_{cls} \cdot \texttt{CE}(\hat{y}_{cls}^{trans}, y_{cls}), \quad (2)$$

where y_{cls} is the image-wise TBSRTC-category label, and the balancing coefficient γ_{cls} is set to 3, as the global feature captured by the Transformer branch is tightly correlated with the image-level classification tag. Finally, the overall loss for TCSegNet becomes

$$\mathcal{L}_s = \mathcal{L}_{seg} + \mathcal{L}_{cls} + \mathcal{L}_{cons}. \quad (3)$$

Extension to Semi-supervised Learning. To leverage images that only have image-wise labels, we extend to a semi-supervised mean teacher [18] framework called Semi-TCSegNet. In this framework, both the student and teacher share the same full-supervised nuclei segmentation architecture of TCSegNet. The weights of the teacher θ_t are updated with the exponential moving average (EMA) of the weights of student θ_s, and smoothing coefficient $\alpha = 0.99$, following the previous work [19]. Formally, the weights of the teacher at e-th epoch are updated by

$$\theta_t^e = \alpha\theta_t^{e-1} + (1 - \alpha)\theta_s^e. \quad (4)$$

During the training stage, the teacher model assigns pixel-wise soft labels to the images with exclusive image-wise labels, thus expanding the scale of labeled data to the student model. The overall loss function for training the student is a combination of the supervised loss $\mathcal{L}s$ in Eq. (3) computed on fully-annotated data, and the semi-supervised loss $\mathcal{L}ss$ computed on data with only image-wise labels. It follows that the overall training objective in `Semi-TCSegNet` is to minimize the loss function $\mathcal{L} = \mathcal{L}_s + \lambda_{ss} \cdot \mathcal{L}_{ss}$, where $\mathcal{L}_{ss}$ measures the segmentation consistency between the teacher and student models via L_2-norm. The balancing coefficient λ_{ss} changes for every epoch e, namely

$$\lambda_{ss}^{e} = \exp(-5\,(1 - e/e_{\max})^2), \tag{5}$$

where $e_{\max}$ is the maximum epoch number.

HSV-Intensity Noise. The traditional method of integrating Gaussian noise in the mean teacher [18] may be problematic when working with cytopathology images that have an imbalanced color distribution. To address this issue, we generate a novel intensity-based noise, which can adaptively behave stronger in the dark nuclei areas and weaker in bright cytoplasm or background regions. We first sample η from a Gaussian distribution $\mathcal{N}\left(0, \sigma^2\right)$, where σ is the standard deviation computed from the pixel values of the V channel in HSV space. The Gaussian noise η serves as the basis for generating the intensity-based noise, which is obtained by $\eta_v = \lambda_v \cdot \eta \cdot (1 - X_v)$. Specifically, X_v is the pixel value of the image's V channel in HSV space, and hyper-parameter λ_v is set as 0.5 to control the amplitude of the intensity-based noise. Finally, the value of the obtained noise is clamped to $[-0.2, 0.2]$ before being added to the images.

3 Experiments

Image Dataset. We construct a clinical thyroid cytopathology dataset with images of both image-wise and pixel-wise labels as a benchmark (appear in GitHub upon acceptance) Some representative images are presented in Fig. 2, together with the profile of the dataset. The dataset comprises 4,965 H&E stained image patches and labels of TBSRTC, where a subset of 1,473 images was densely annotated for nuclei boundaries by three experienced cytopathologists and reached a total number of 31,064 elaborately annotated nuclei. Patient-level images were partitioned first for training and test images, and patch-level curation was performed. We divided the dataset with image-wise labels into 80% training samples and the remaining 20% testing samples. Our collection of thyroid cytopathology images was granted with an Ethics Approval document.

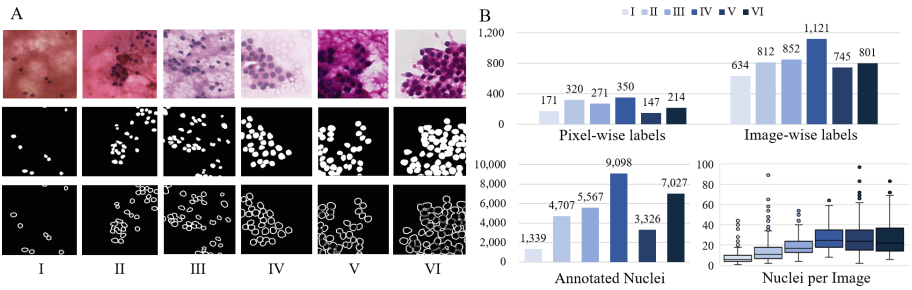

Fig. 2. The annotated thyroid cytopathology benchmark. A) Examples from I to VI of TBSRTC six diagnostic categories with pixel-wise nuclei mask and boundary annotations; B) The profile of the dataset.

Table 1. Quantitative comparisons in both fully-supervised and semi-supervised manners. The best performance is highlighted in **bold**, where we can observe that both `TCSegNet` and its semi-supervised extension outperform state-of-the-art.

Method		Dice	IoU
Fully-supervised	Mask R-CNN [20]	0.657	0.500
	Swin-Unet [21]	0.671	0.516
	SegNet [22]	0.676	0.587
	UNet++ [23]	0.784	0.691
	$Ca^{2.5}$-Net [14]	0.838	0.732
	CIA-Net [13]	0.854	0.775
	ClusterSeg [15]	0.857	0.761
	TCSegNet (Ours)	**0.877**	**0.788**
Semi-supervised	PseudoSeg [24]	0.734	0.612
	Cross Pseudo Seg [25]	0.737	0.618
	Cross Teaching [26]	0.795	0.704
	PS-ClusterSeg [15]	0.866	0.775
	MTMT-Net [27]	0.878	0.789
	Semi-TCSegNet (Ours)	**0.889**	**0.805**

Implementations. The proposed method and compared methods are implemented on a single NVIDIA GeForce RTX 3090 GPU card. We employ a conformer [16] with 12 Transformer layers and 5 CNN blocks as the encoder in `TCSegNet`. Both `TCSegNet` and `Semi-TCSegNet` use SGD optimizer with a momentum of 0.9 and a weight decay of 10^{-4}. The initial learning rate lr_0 is set to 5×10^{-3}, and the learning rate for e^{th} epoch is determined by the poly strategy [28], *i.e.*, $lr_e = lr_0 \times (1 - e/e_{max})^{0.9}$, where $e_{max} = 280$ is the total epoch number. We set the batch size for `TCSegNet` to 8, and for `Semi-TCSegNet` to 10, *i.e.* 8 fully-annotated images and 2 partially-annotated images per batch.

Compared Methods and Evaluation Metrics. We compared `TCSegNet` with the fully-supervised counterparts, including method specific for segmentation in general image [20,22], medical image [21,23], and nuclei [13–15]. We also compared `Semi-TCSegNet` with semi-supervised methods [15,24–27]. We used the officially released code published along with their papers for all the compared methods. Intersection over Union (IoU) and Dice score were applied as the evaluation metrics, where a higher value indicated a better semantic segmentation performance.

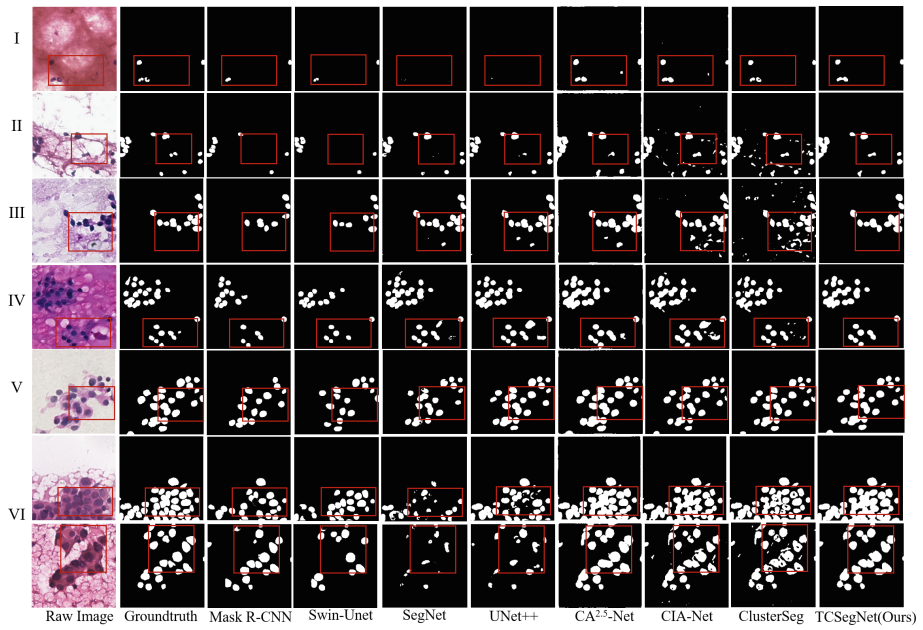

Fig. 3. Examples of segmented nuclei in thyroid cytopathology images by `TCSegNet` and prevent fully-supervised models methods. Observably, TCSegNet is more aware of scattered, small, or clustered nuclei.

Experimental Results. The results in Table 1 indicated that `TCSegNet` can achieve the highest performance by a Dice score of 87.7% and an IoU of 78.8%. The performance values in the challenging regions are highlighted with red boxes in Fig. 3, together with the line charts in Fig. 4 (A, B). Our approach is capable to address the current issue in the recognition and segmentation of small isolated cells graded in the I category, which is always ignored by the unbalanced pixel-wise cell morphology with other approaches. Also, it yields that the incorporation of TBSRTC-category can contribute to a partial alleviation of a biased model, resulting in more satisfying segmentation performance experimentally. Furthermore, the fact that the TBSRTC-category label is easy to obtain endows

the applicability of our model to various circumstances that nuclei in various sizes, shapes, and dyeing styles can be accurately recognized and segmented. Consequently, it can serve as a guarantee for the validity and accuracy of the subsequent analysis in real clinical practice. Moreover, with the semi-supervised learning, `Semi-TCSegNet` can further boost the performance to an 88.9% Dice score, and 80.5% IoU, by leveraging additional data with image-wise TBSRTC-category labels solely. The performance improvement of 1.2% Dice, 1.7% IoU, together with the general improvement is shown in the boxplot in Fig. 4 (C, D), as a demonstration of the advantage using full data resources with `Semi-TCSegNet`.

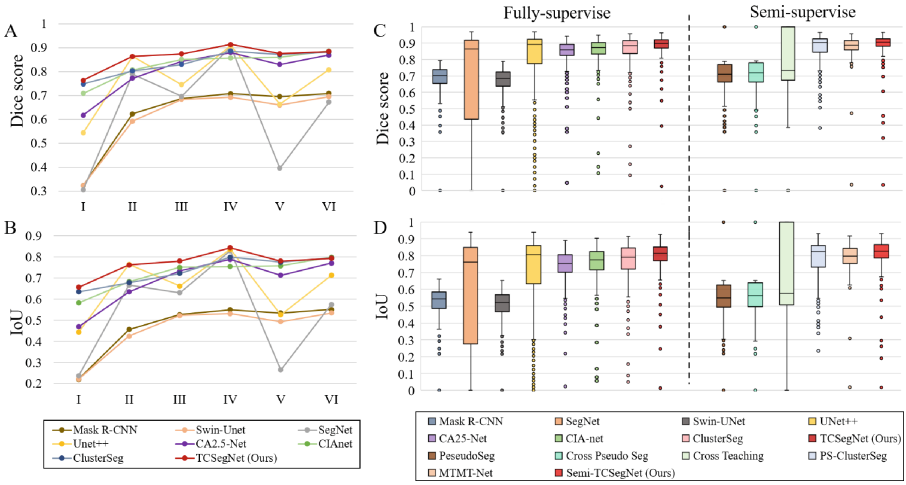

Fig. 4. Quantitative comparison presented by boxplot and line chart. A and B are the variation tendency of Dice score and IoU of six diagnostic categories in fully-supervised methods. C and D are the distribution of the Dice score and IoU of all mentioned models respectively. Our models presented a general improvement across the metrics.

Ablation Study. To evaluate the effectiveness of each functional block and demonstrate the functionality of semi-supervised learning, we illustrate the ablation study in Table 2. The results indicate that performance improvement is accumulated with increasing data size. Besides, training with a classification-learning block alone can increase the nuclei segmentation performance by 1.7% and 2.6% in the Dice score and IoU, respectively. Meanwhile, trained with specially designed HSV-Intensity noise can also increase the performance by 0.9% Dice and 1.4% IoU, showing its potential for generation ability improvement. Importantly, the benefits from the two blocks are orthonormal, where `Semi-TCSegNet` achieves the optimal performance with the utilization of both.

Table 2. Ablation study for our Semi-TCSegNet and functional blocks.

Classification Learning	HSV-Intensity Noise	Dice	IoU
		0.867	0.771
	✓	0.876	0.785
✓		0.884	0.797
✓	✓	**0.889**	**0.805**
W. image-wise data	+1k data	0.879	0.790
	+2k data	0.882	0.795

4 Conclusion

In this paper, we propose a TBSRTC-category aware nuclei segmentation framework `TCSegNet`, that leverages easy-to-obtain image-wise diagnostic category to facilitate nuclei segmentation. Importantly, it addresses the challenge of distinguishing nuclei across different cell scales in an unbalanced dataset. We also extend the framework to a semi-supervised learning fashion to overcome the issue of lacking annotated training samples. Moreover, we construct the first thyroid cytopathology dataset with both image-wise and pixel-wise labels, which we believe can it facilitate future research in this field. As the spatial distribution, shape, and area information from nuclear segmentation is supportive of diagnostic decisions, we will further leverage the segmentation result for malignancy analysis and also explore the potential of spatial information for unlabeled data exploration in the future.

Acknowledgement. This work was supported by National Natural Science Foundation of China (Grant No. 62102247) and Natural Science Foundation of Shanghai (No. 23ZR1430700).

References

1. Ferlay, J., et al.: Cancer incidence and mortality worldwide: sources, methods and major patterns in globocan 2012. Int. J. Cancer **136**(5), E359–E386 (2015)
2. Haugen, B.R., et al.: 2015 American thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the american thyroid association guidelines task force on thyroid nodules and differentiated thyroid cancer. Thyroid **26**(1), 1–133 (2016)
3. Cibas, E.S., Ali, S.Z.: The bethesda system for reporting thyroid cytopathology. Thyroid **19**(11), 1159–1165 (2009)
4. Kumar, N., et al.: A multi-organ nucleus segmentation challenge. IEEE Trans. Med. Imaging **39**(5), 1380–1391 (2019)
5. Veta, M., et al.: Prognostic value of automatically extracted nuclear morphometric features in whole slide images of male breast cancer. Mod. Pathol. **25**(12), 1559–1565 (2012)

6. Kakudo, K.: Thyroid FNA Cytology: Differential Diagnoses and Pitfalls. Springer, Heidelberg (2019). https://doi.org/10.1007/978-981-13-1897-9
7. Xing, F., Yang, L.: Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review. IEEE Rev. Biomed. Eng. **9**, 234–263 (2016)
8. Cibas, E.S., Ali, S.Z.: The 2017 bethesda system for reporting thyroid cytopathology. Thyroid **27**(11), 1341–1346 (2017)
9. Graham, S., et al.: Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. Med. Image Anal. **58**, 101563 (2019)
10. Gamper, J., Alemi Koohbanani, N., Benet, K., Khuram, A., Rajpoot, N.: PanNuke: an open pan-cancer histology dataset for nuclei instance segmentation and classification. In: Reyes-Aldasoro, C.C., Janowczyk, A., Veta, M., Bankhead, P., Sirinukunwattana, K. (eds.) ECDP 2019. LNCS, vol. 11435, pp. 11–19. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-23937-4_2
11. Greenwald, N.F., et al.: Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. Nat. Biotechnol. **40**(4), 555–565 (2022)
12. Everingham, M., Eslami, S.A., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes challenge: a retrospective. Int. J. Comput. Vision **111**, 98–136 (2015)
13. Zhou, Y., Onder, O.F., Dou, Q., Tsougenis, E., Chen, H., Heng, P.-A.: CIA-Net: robust nuclei instance segmentation with contour-aware information aggregation. In: Chung, A.C.S., Gee, J.C., Yushkevich, P.A., Bao, S. (eds.) IPMI 2019. LNCS, vol. 11492, pp. 682–693. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-20351-1_53
14. Huang, J., Shen, Y., Shen, D., Ke, J.: $CA^{2.5}$-net nuclei segmentation framework with a microscopy cell benchmark collection. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12908, pp. 445–454. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87237-3_43
15. Ke, J., et al.: Clusterseg: a crowd cluster pinpointed nucleus segmentation framework with cross-modality datasets. Med. Image Anal. **85**, 102758 (2023)
16. Peng, Z., et al.: Conformer: local features coupling global representations for visual recognition. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 367–376 (2021)
17. Hou, Q., Cheng, M.M., Hu, X., Borji, A., Tu, Z., Torr, P.H.: Deeply supervised salient object detection with short connections. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3203–3212 (2017)
18. Tarvainen, A., Valpola, H.: Mean teachers are better role models: weight-averaged consistency targets improve semi-supervised deep learning results. Adv. Neural Inf. Process. Syst. **30**, 1–10 (2017)
19. Laine, S., Aila, T.: Temporal ensembling for semi-supervised learning. arXiv preprint arXiv:1610.02242 (2016)
20. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask r-cnn. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2961–2969 (2017)
21. Cao, H., et al.: Swin-unet: unet-like pure transformer for medical image segmentation. In: Computer Vision-ECCV 2022 Workshops: Tel Aviv, Israel, 23–27 October 2022, Proceedings, Part III, pp. 205–218. Springer, Heidelberg (2023). https://doi.org/10.1007/978-3-031-25066-8_9
22. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **39**(12), 2481–2495 (2017)

23. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: UNet++: a nested u-net architecture for medical image segmentation. In: Stoyanov, D., et al. (eds.) DLMIA/ML-CDS -2018. LNCS, vol. 11045, pp. 3–11. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00889-5_1
24. Zou, Y., et al.: Pseudoseg: designing pseudo labels for semantic segmentation. arXiv preprint arXiv:2010.09713 (2020)
25. Chen, X., Yuan, Y., Zeng, G., Wang, J.: Semi-supervised semantic segmentation with cross pseudo supervision. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2613–2622 (2021)
26. Luo, X., Hu, M., Song, T., Wang, G., Zhang, S.: Semi-supervised medical image segmentation via cross teaching between cnn and transformer. In: International Conference on Medical Imaging with Deep Learning, pp. 820–833. PMLR (2022)
27. Chen, Z., Zhu, L., Wan, L., Wang, S., Feng, W., Heng, P.A.: A multi-task mean teacher for semi-supervised shadow detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5611–5620 (2020)
28. Liu, W., Rabinovich, A., Berg, A.C.: Parsenet: looking wider to see better. arXiv preprint arXiv:1506.04579 (2015)