

Edge-Aware Dual Branch Network for Nucleus Instance Segmentation

Junzhou Chen
Dept. of Computer Science
William & Mary
Williamsburg, USA
jchen57@wm.edu

Yanfu Zhang
Dept. of Computer Science
William & Mary
Williamsburg, USA
yzhang105@wm.edu

Sidi Lu
Dept. of Computer Science
William & Mary
Williamsburg, USA
sidi@wm.edu

Abstract—In mobile healthcare and remote diagnosis, nucleus segmentation is a critical step for pathological analysis, diagnosis, and classification, requiring real-time processing and high accuracy. However, variations in nucleus size, blurred contours, uneven staining, cell clustering, and overlapping cells hinder precise segmentation. Additionally, existing deep learning models often prioritize accuracy at the cost of increased complexity, making them unsuitable for resource-limited edge devices and real-world deployment. To address the aforementioned issues, we propose an edge-aware dual branch network for nucleus instance segmentation. The network simultaneously predicts target information and target contours. Within the network, we propose a context fusion block (CF-block) that effectively extracts and merges contextual information from the network. Additionally, we introduce a post-processing method that combines the target information and target contours to distinguish overlapping nuclei and generate an instance segmentation image. Extensive quantitative evaluations are conducted to assess the performance of our method. Experimental results demonstrate the superior performance of the proposed method compared to state-of-the-art approaches on the BNS, MoNuSeg, and CPM-17 datasets.

Index Terms—Nucleus segmentation, mobile healthcare, instance segmentation, medical imaging, dual-branch network

I. INTRODUCTION

Edge intelligence [1] plays a critical role in mobile healthcare [2], [3] by enabling real-time and accurate data processing near the patient, which is essential for applications such as remote surgery [4] and pathology image analysis [5]. For instance, in breast cancer diagnosis [6], edge intelligence can be used to analyze histological slides in real-time during a biopsy, detecting abnormal cell patterns instantly. This allows pathologists to view the results on-site, enabling rapid and accurate diagnoses without the delays of remote processing. As a result, the quality of care is enhanced through immediate feedback to surgeons and clinicians during the procedure.

One of the most widely used techniques in pathology is Hematoxylin and Eosin (H&E) staining [7], which provides the visual cues to differentiate nuclei from surrounding tissue. This has made nucleus image segmentation [8] an effective solution for the precise quantification and analysis of nuclei in cancer diagnosis, classification, and grading [9]–[13].

Advancing nucleus segmentation algorithms can significantly enhance accuracy in digital pathology, improving early disease diagnosis and treatment [14]. While deep learning

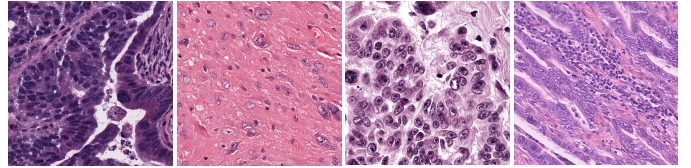


Fig. 1. Examples of nucleus images from different public datasets.

has substantially improved nucleus segmentation, challenges remain due to imaging conditions and the complex distribution of nuclei. As illustrated in Fig. 1, factors such as dense distributions, adhesive edges, varying sizes and shapes, blurred contours, uneven staining, and overlapping cell clusters often result in over- or under-segmentation, reducing accuracy. For example, in distance-based methods like StarDist [15] [16] [17], large nuclei’s centroids are too far from boundary pixels, impairing distance prediction, with supervision limited to individual distance values and lacking global shape constraints [18]. Similarly, traditional contour-based methods struggle with blurred boundaries [19] [20], further complicating segmentation.

In addition, many existing deep learning models for nucleus segmentation mainly focus heavily on maximizing accuracy, often leading to increased model complexity. This heightened complexity poses a challenge for deployment on resource-constrained edge devices, such as smart wearables and Internet of Things (IoT) devices, which have limited processing power, memory, and energy capacity, making it difficult to implement these models effectively in real-time applications [21], [22]. Therefore, developing a more efficient nucleus segmentation model that balances accuracy with resource efficiency, and is specifically optimized for edge devices, has become an urgent issue to address [21], [22].

To address the challenges in nucleus segmentation, we propose a novel neural network model called EADB-Net, which significantly reduces computational complexity while maintaining high segmentation accuracy. EADB-Net builds on the U2Net architecture [23] and integrates a hybrid attention mechanism [24]. By incorporating attention mechanisms into Residual U-blocks, EADB-Net effectively handles complex images. It introduces a context encoding layer to capture

contextual features and integrates hybrid attention modules at each layer to focus on relevant regions. Additionally, EADB-Net predicts both the foreground and contour of nuclei, combining these to achieve instance-level segmentation. Compared to traditional networks [25], [26], EADB-Net captures finer details, resulting in improved accuracy and robustness. Experimental evaluations on three publicly available datasets (BNS [27], MoNuSeg [28], and CPM-17 [29]) demonstrate that EADB-Net achieves state-of-the-art performance, making it an effective solution for nucleus segmentation tasks.

Specifically, the main innovations and contributions of this paper are as follows:

- We design a dual-task network for nucleus segmentation, combining a hybrid attention mechanism with nested residual U-blocks and leveraging enhanced contextual information for task prediction. This achieves lower computational complexity and higher accuracy, making it suitable for edge devices.
- We propose a novel Context Fusion Block (CF-Block) to effectively extract and integrate contextual information.
- By utilizing a dual-branch architecture, our network demonstrates strong edge-awareness capabilities. Combined with the introduced post-processing approach, it delivers effective and straightforward nucleus instance segmentation results.
- We evaluate EADB-Net on multiple datasets, including BNS, MoNuSeg, and CPM-17. Experimental results show that EADB-Net outperforms state-of-the-art models in terms of segmentation accuracy and robustness.

II. RELATED WORK

Currently, methods for nucleus image segmentation can be classified into traditional methods [30] and deep learning-based methods. Traditional segmentation methods include threshold-based segmentation [31], region-based segmentation [32], graph-based segmentation [33], superpixel-based segmentation [34], fuzzy clustering, and other methods [35]. These methods construct image segmentation models by manually selecting features, which results in good performance only on specific feature-rich datasets or samples. For example, C.H.Lin et al. [36] used a segmentation method based on a series of edge enhancement techniques, but it performed poorly on blurry nuclear contours. Yang Song et al. [37] employed a contrast-based adaptive version of the mean shift and SLIC algorithms, along with intensity-weighted adaptive thresholds, for segmenting nuclei in Papanicolaou (Pap) smear images. However, in many cases, the aforementioned traditional methods struggle to handle cervical cell images with irregular shapes and sizes. To address this issue, deep learning-based methods have gradually become mainstream. In recent years, research on image segmentation using convolutional neural networks (CNNs) [38] [39] has achieved remarkable results.

With the continuous improvement of computing performance, deep learning algorithms have demonstrated outstanding performance in image segmentation. In the field of nuclear

image segmentation, the most commonly used network architectures include FCN [40], Mask R-CNN [41], U-Net [42], and others. In particular, U-Net was initially applied to medical image segmentation tasks, utilizing skip connections to connect the intermediate downsample and upsample layers to extract contextual information. However, due to the direct fusion of low-level and high-level features through skip connections, there may be semantic gaps and difficulties in handling overlapping regions. To address this issue, subsequent researchers have proposed methods such as Attention U-Net [43], U2Net [44], CE-Net [44], CIA-Net [45], AL-Net [46], and others. These methods have demonstrated good performance on nuclear segmentation datasets. Among them, CE-Net [44] extends the application of U-Net in medical image segmentation by incorporating enhanced network structures such as DAC and RMP blocks. To obtain more advanced information while preserving spatial details, the Context Encoding Network (CE-Net) replaces the encoding module of U-Net with a pre-trained model, resulting in improved performance on 2D medical image segmentation tasks. The Attention U-Net [43] model is a U-Net model that incorporates attention mechanisms. This model has the ability to automatically learn the shape and size of the targets, significantly improving the sensitivity and accuracy of the model with minimal additional computational cost. Similarly, Xuebin et al. [44] designed a two-level nested U-shaped structure called U2Net for salient object detection (SOD). U2Net captures more contextual information from different scales, leading to better performance in SOD tasks.

In the context of nuclear segmentation and the need, for instance, delineation, DCAN [47] proposed a novel Deep Contour-Aware Network that simultaneously segments nuclei and their boundaries. The multi-task segmentation framework in DCAN has been widely used in nuclear segmentation methods. To capture multi-scale spatial information, a Spatial Perception Network (SpaNet) [48] was proposed. SpaNet's multi-scale dense units are equipped with a feature aggregation property that allows positional information to flow throughout the network. In the aforementioned methods, there is no feature fusion between the multi-task branches. An information aggregation module is introduced to fuse feature maps from different branches. Building upon this, a Boundary-assisted Region Proposal Network (BRP-Net) [49] is proposed. HoVer-Net [50] is another method proposed for simultaneous nucleus segmentation and classification. In this network, distance information between nucleus pixels and their centroids is introduced in both vertical and horizontal directions. These distance features are utilized to help the network learn information about nucleus shape and structure, thereby improving segmentation accuracy.

III. PROPOSED METHODOLOGY

In this section, we detail the proposed methodology and network architecture. Section III-A outlines the EADB-Net and its key components, including attention mechanisms. Section III-B introduces the Context Fusion Block for improved segmentation accuracy. Section III-C explains the supervi-

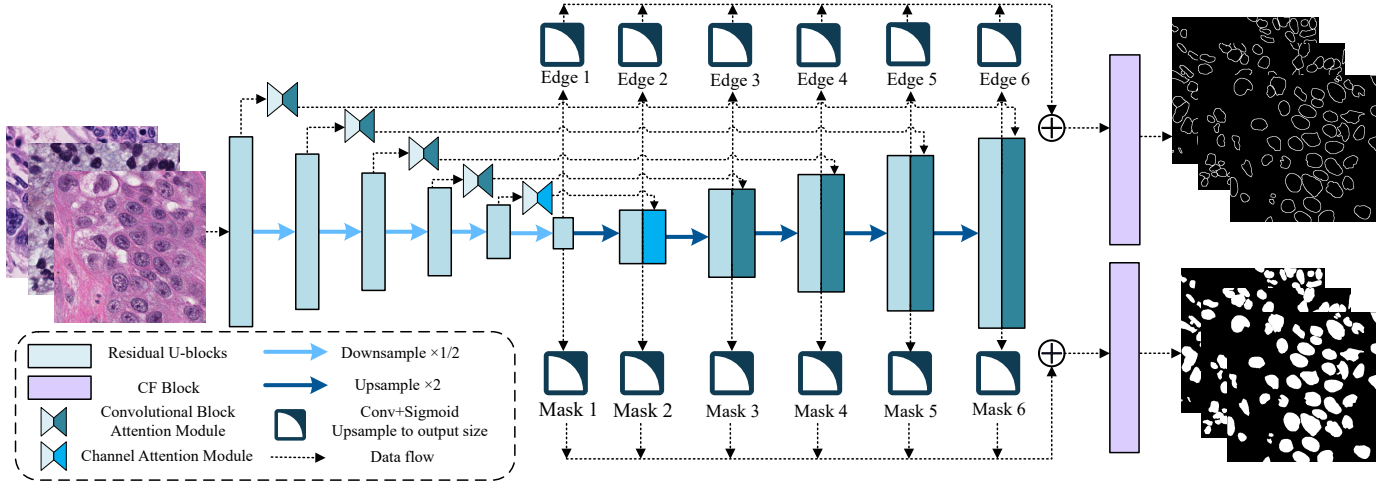


Fig. 2. Overall architecture of the network demonstrates the precise placement of modules within the network. The channel attention mechanism is applied at the deepest layer, while the remaining layers utilize a hybrid attention mechanism.

sion strategy using multi-branch loss functions. Lastly, Section III-D discusses the post-processing methods for refining nucleus instance segmentation.

A. Network Architecture

The overall structure of the proposed network is illustrated in Fig. 2. EADB-Net is an image segmentation model based on the U-Net architecture and consists of four major components: encoder, decoder, attention mechanism, and contextual mechanism.

The encoder phase comprises multiple downsampling layers utilized to subject the input image to successive convolutional and pooling operations, thereby extracting features at various hierarchical levels. Similarly, the decoder phase employs multiple upsampling layers to restore the resolution of the feature maps from the encoder phase and to produce prediction results through merging. Each downsampling and upsampling layer consists of the Residual U-block proposed by Qin *et al.* [23], and the nested U-structure facilitates the extraction of multiscale features within each phase and more effective aggregation of multilevel features across phases. As the network’s depth increases, the height of the Residual U-blocks is appropriately adjusted.

Each upsampling layer consists of a Residual U-block and a skip connection, where the Convolutional Block Attention Module (CBAM) [24] extracts features from the downsampling information and combines them with the upsampling information through the skip connection to fuse low-level and high-level feature information, as shown in Fig. 2. The information extracted at a coarser scale is used for CBAM to eliminate irrelevant and noisy influences in the skip connection. This is performed before the concatenation operation to merge only relevant activations. Meanwhile, for deeper layers of the network, using a channel attention mechanism alone often proves to be more effective, and replacing CBAM with a channel attention mechanism at the deepest layer has yielded better results.

The last component is the feature fusion module, which generates probability maps. Similar to U2Net [23], our EADB-Net constructs two output branches, where each branch generates six probability maps, $Mask_i$ $Edge_i$ $i \in [1, 6]$, from the last downsampling layer and all upsampling layers through a 3×3 convolutional layer and a sigmoid function. Subsequently, the feature maps from the side branches are upsampled to the size of the input image and fused using the CF Block mentioned earlier to generate the final probability maps, S_{mask} and S_{edge} .

EADB-Net aims to improve the network’s ability to learn both fine details and global features in high-resolution medical images while keeping the computational complexity relatively low. This network model performs exceptionally well in cell nucleus instance segmentation tasks and accurately identifies segmentations even with limited samples.

B. Context Fusion Block

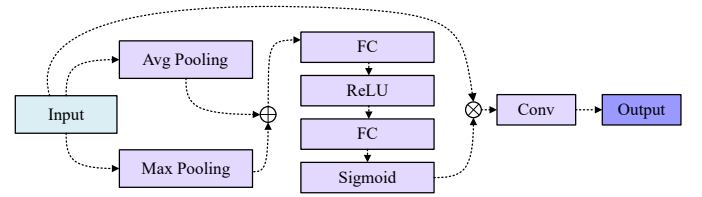


Fig. 3. Structure of the Context Fusion Block is shown, where the input is the concatenation of context features from the backbone network (6 layers). The information is fused within the block and then outputted.

In the context encoding layer of U2Net [23], only concatenation is performed for each side branch without considering the importance of each branch. Inspired by SE-Net [51] and CE-Net [52], we improved it by designing a CF-Block (context fusion) based on the SE-Block. This allows the model to capture different context information and improve the accuracy of semantic segmentation, as shown in Figure 1. The CF-Block helps the network focus on useful contextual information

in the image, thereby enhancing segmentation accuracy and efficiency.

The process, as depicted in Fig. 3, involves linearly interpolating different-sized contexts to match the output size and then concatenating them as input to the CF-Block. Next, the feature maps are globally average-pooled and max-pooled separately to generate a feature vector. The feature vector is then subjected to a non-linear transformation to generate context feature layer weights. Finally, the input features are multiplied by the weights and upsampled to the size of the input image using a 3×3 convolutional layer.

C. Supervision

The output of the network model consists of two branches: foreground and contour. Each branch is composed of 6 probability maps. The loss function, denoted as $Loss$, is defined as follows:

$$Loss = \omega_{edge}\tau_{edge} + \omega_{mask}\zeta_{mask} + \sum_{i=1}^M (\omega_{side}^{(i)}\tau_{side}^{(i)} + \omega_{side}^{(i)}\zeta_{side}^{(i)}) \quad (1)$$

Here, ζ_{side}^M and τ_{side}^M represent the losses for foreground and contour in the side branch, respectively. ζ_{mask} and τ_{mask} represent the losses for the final fused output feature maps S_{mask} and S_{edge} . ω represents the weights for each loss term. For each term ζ and τ , we calculate the loss as the sum of standard binary cross-entropy loss and Dice loss:

$$CE = - \sum_{(r,c)}^{(H,W)} [P_{G(r,c)} \log P_s(r,c) + (1 - P_{G(r,c)}) \log (1 - P_s(r,c))] \quad (2)$$

$$Dice = 1 - \sum_{(r,c)}^{(H,W)} \left(\frac{2 \times P_{S(r,c)} \times P_{G(r,c)}}{P_{S(r,c)}^2 + P_{G(r,c)}^2} \right) \quad (3)$$

$$\zeta = \tau = CE + Dice \quad (4)$$

H and W represent the height and width of the image, respectively. (r,c) represents a pixel in the image, $P_{G(r,c)}$ represents the ground truth label for that pixel, and $P_{s(r,c)}$ represents the predicted probability of the model that the pixel is a positive label.

D. Post Processing

For nucleus segmentation, it is often necessary to segment instances. To address the issue of significant overlap between nuclei, DCAN [47] proposed a contour-based instance segmentation method. By subtracting the contours from the foreground, all overlapping nuclei are separated, resulting in individual cell instances. In deep learning models, if the annotated contours are very thin, the model training can easily get stuck in local minima, leading to poor segmentation results. By thickening the annotated contours, the model can escape

local minima. However, this introduces a new problem. While subtracting the contours from the foreground allows obtaining instances of each cell, it also leads to the loss of foreground information.

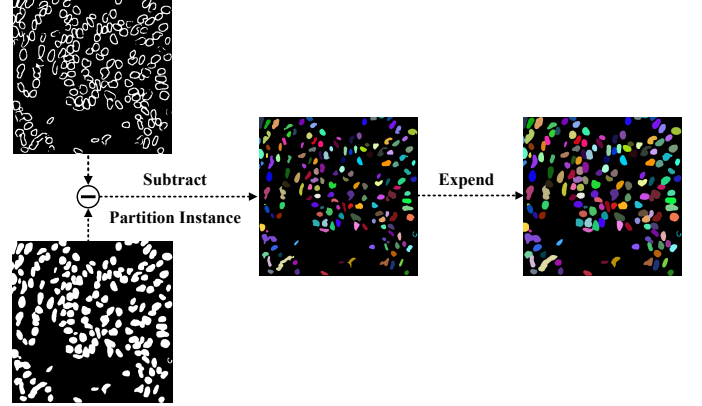


Fig. 4. The input to the EADB-Net is the predicted semantic information and contour information. The semantic information is subtracted by the contour information to obtain instances. Then, each instance is repaired or refined.

Algorithm 1 Instance Segmentation Processing

Input: mask, edge

```

1: kernel  $\leftarrow \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ 
2: mask_ins  $\leftarrow \text{mask} \odot (1 - \text{edge})$ 
3: num,objects,stats  $\leftarrow \text{connected\_components\_with\_stats}(\text{mask\_ins})$ 
4: while not is_mask_fully_covered(objects, mask) do
5:   for mask_id = 1 to num_objects do
6:     id_mask  $\leftarrow \text{objects}[\text{objects} == \text{mask\_id}]$ 
7:     id_mask  $\leftarrow \text{dilate\_mask}(\text{id\_mask}, \text{kernel})$ 
8:     objects[(id_mask != 0) & (mask != 0) & (objects == 0)]  $\leftarrow \text{mask\_id}$ 
9:   end for
10: end while
Output: objects

```

In this work, we propose a new method, as shown in Fig. 4. Here, we first subtract the contours from the foreground and perform instance labeling on each disconnected region. Then, we propagate the instance results outward until the foreground information is completely filled, generating the final segmentation mask. This method is simple and effective, as it can effectively utilize contour information to separate overlapping cells and restore cell instances using high-precision foreground information.

IV. EXPERIMENT SETUP AND RESULTS

In this section, we evaluate and test the performance and robustness of the proposed model through extensive experiments. We utilized three publicly available datasets, as described in Table I. Dice and AJI were employed as evaluation metrics.

For data augmentation, random rotation, random zooming, random shearing, random shifting, and horizontal flipping are used for training. The network runs on NVIDIA AGX Xavier, with a mini-batch size of 8, a learning rate of 0.001, a decay of 0.0005, and an epoch of 60.

TABLE I
DESCRIPTION OF DATASETS

| Datasets | Description | Annoations | Image Size |
|--------------|------------------------|----------------------------------|-------------|
| MoNuSeg [28] | 37 Training 14 Test | Seven Organs | 1000 × 1000 |
| CPM-17 [29] | 32 Training 32 Test | The Cancer Genome Atlas | 500 × 500 |
| BNS [27] | 36 Training 14 Test | Triple-Negative Breast Cancer | 512 × 512 |

A. Evaluation Metrics

In this experiment, we employed Dice, Panoptic Quality (PQ), and Aggregated Jaccard Index (AJI) as performance evaluation metrics. These metrics are defined as follows:

Dice: The Dice measures the similarity at the pixel level and is computed using the formula:

$$Dice = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (5)$$

TP (True Positive) represents the accurately segmented regions, FP (False Positive) represents the regions erroneously segmented as positive examples, and FN (False Negative) represents the regions erroneously segmented as negative examples.

Aggregated Jaccard Index (AJI) is utilized to assess the similarity between the instance segmentation results generated by an algorithm and the ground truth segmentations. Its formula is defined as follows:

$$AJI = \frac{\sum_{i=1}^N |G_i \cap P_M^i|}{\sum_{i=1}^N |G_i \cup P_M^i| + \sum_{F \in U} |P_F|} \quad (6)$$

In the formula for AJI, N represents the number of instances in the sample. G_i denotes the ground truth mask of the i -th instance, and P_M^i represents the predicted mask of the i -th instance. The term $\sum_{i=1}^N |G_i \cup P_M^i|$ denotes the total size of the union between the matched predicted and ground truth segmentations. The second term $\sum_{F \in U} |P_F|$ represents the total number of pixels in unmatched predicted segmentations. The numerator of the formula, $\sum_{i=1}^N |G_i \cap P_M^i|$, indicates the total number of pixels that are correctly matched.

B. Comparison with Other Methods

To further evaluate EADB-Net, we conducted comparative experiments on three publicly available datasets (MoNuSeg [28], BNS [27], and CPM-17 [29]) against state-of-the-art and classical models.

Quantitative Analysis: Table II illustrates the outstanding performance of the network across three datasets. In terms of evaluation metrics, the Dice index better reflects the effectiveness of semantic segmentation, while the AJI more

TABLE II
COMPARISON WITH OTHER METHODS ON MoNuSeg, CPM-17 AND BNS DATASET

| Model | MoNuSeg | | CPM-17 | | BNS | |
|----------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | Dice | AJI | Dice | AJI | Dice | AJI |
| U-Net [42] | 0.758 | 0.556 | 0.813 | 0.643 | 0.777 | 0.572 |
| Mask-RCNN [41] | 0.760 | 0.546 | 0.850 | 0.684 | 0.782 | 0.576 |
| DIST [53] | 0.786 | 0.560 | 0.826 | 0.616 | 0.779 | 0.582 |
| DCAN [47] | 0.793 | 0.525 | 0.828 | 0.561 | - | - |
| HoVer-Net [50] | 0.826 | 0.618 | 0.869 | 0.705 | 0.781 | 0.587 |
| AL-Net [46] | 0.823 | 0.649 | 0.873 | 0.707 | 0.790 | 0.624 |
| EADB-Net(Ours) | 0.838 | 0.656 | 0.895 | 0.721 | 0.812 | 0.651 |

accurately assesses the performance of instance segmentation. It is noteworthy that EADB-Net excels in both aspects. In the MoNuSeg [28] dataset, where the shapes of cell nuclei are more irregular, the morphology-based Hover-Net exhibits relatively poorer performance in instance segmentation. Conversely, DCAN focuses on segmentation through cell nucleus contours but struggles to address residual connections between cell nuclei. EADB-Net effectively strikes a balance between these two approaches, yielding significant results.

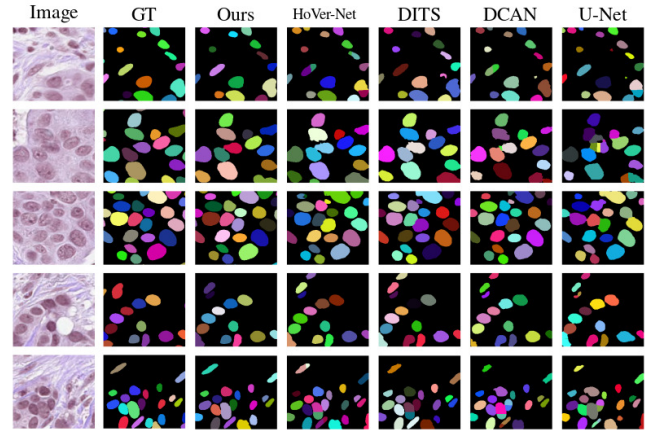


Fig. 5. Visual comparisons with SOTAs where GT indicates the ground truth.

Qualitative Analysis: The visual comparison results are presented in Fig. 5, showcasing the effectiveness of our proposed approach. Carefully selected challenging images with diverse sizes and textures were used for comparison. The performance of our model in accurately segmenting nuclei from diverse background scenes is evident. As depicted in the first row of Fig. 5, our method achieves boundary segmentation results that closely resemble the ground truth labels, effectively delineating nuclei. Moreover, the third row of Fig. 5 demonstrates the superior accuracy of our method in localizing nuclei and achieving precise contour delineation. This capability enables better differentiation of overlapping nuclei. These qualitative comparisons provide compelling evidence for the capabilities of our proposed approach in accurately segmenting nuclei, even in challenging scenarios involving variations in size, texture, and overlapping instances.

C. Ablation Study

To evaluate each module of EADB-Net, we conducted a series of ablation experiments on the MoNuSeg dataset. In these experiments, we chose U2Net as the baseline model and used watershed as the instance segmentation method for all experiments except the last one.

TABLE III
RESULTS OF EACH COMPONENT IN THE EADB-NET

| Model | Dice | AJI | PQ |
|---|--------------|--------------|--------------|
| Baseline U2Net(U-Net+RSU) | 0.790 | 0.533 | 0.570 |
| U2Net+CBAM | 0.820 | 0.590 | 0.583 |
| U2Net+CBAM+CF-Block | 0.827 | 0.598 | 0.589 |
| U2Net+CBAM+CF-Block+ Post Processing(ours) | 0.838 | 0.656 | 0.624 |

In the first experiment, CBAM integration in U2Net for nuclear segmentation resulted in significant improvements. The U-Net + RSU + CBAM model showed increases in Dice, PQ, and AJI scores. A comparison between CBAM applied at all levels and only at the deepest level demonstrated 0.6% improvements in Dice and PQ scores, favoring channel attention at the small spatial size of the deepest level.

In the second experiment, the CF-Block's effectiveness was confirmed. Allowing weighted fusion of contextual information, the U-Net + RSU + CBAM + CF-Block model exhibited improvements in the three evaluation metrics compared to the U-Net + RSU + CBAM model.

In the last experiment, post-processing in instance segmentation showed significant improvements in all three evaluation metrics. The auxiliary task of edge production contributed to enhanced pixel prediction for the nucleus by combining contour and semantic analysis, leading to superior performance.

V. OBSERVATIONS AND DISCUSSIONS

Pathology diagnosis plays a crucial role in medical diagnosis, especially in cancer diagnosis. To improve the efficiency and robustness of automated histopathology image analysis, we propose a hybrid attention dual-branch network based on Residual U-blocks. Extensive experiments are conducted on four challenging datasets to demonstrate the effectiveness of the network model and its individual modules. In the multi-task network, contour extraction and object segmentation are two complementary tasks. Previous works that divide instances based on contour extraction, such as DCAN and AL-Net, only consider nucleus instance division without focusing on the segmentation accuracy of each instance. Therefore, we propose to propagate around each nucleus instance until the target segmentation result is completely filled. Experimental results show that this method is simple yet effective.

From the MoNuSeg and BNS datasets, we observe that the model achieves high segmentation accuracy and can effectively segment overlapping nuclei. However, in the CPM-17 dataset, the instance segmentation accuracy is not satisfactory. Through careful investigation, we found that the shape prior of nuclei in the CPM-17 dataset deviates significantly from the general cancer dataset. The nuclei boundaries in this dataset are more

blurred, leading to unsatisfactory contour extraction results for overlapping regions. Therefore, for this dataset, we applied an erosion operation in the post-processing stage to improve the model's accuracy. After in-depth analysis, the following improvements can be made in future work:

- Address the issue of unclear contours and incomplete extraction of overlapping region contours. Consider incorporating prior knowledge of nucleus morphological features into the loss function to further optimize contour extraction accuracy or explore specific techniques for contour extraction in overlapping regions.
- EADB-Net demonstrates good performance in semantic segmentation and is not limited to cervical cell classification, indicating its potential application in other medical image classification tasks.
- In the post-processing stage, the instance spreading technique for each nucleus does not accurately handle overlapping boundaries. Further improvements can be made to refine the spreading strategy.

VI. CONCLUSION

This study addresses the complex problem of nucleus segmentation and proposes a hybrid attention dual-branch network based on Residual U-blocks. The network simultaneously predicts the semantic and contour information of nuclei and achieves high-precision instance segmentation through post-processing. Additionally, we discovered that the back-propagation from the auxiliary task further improves the accuracy of the main task. The key components of the network include Residual U-blocks, hybrid attention mechanisms, and context encoding layers. The network model achieves highly accurate semantic segmentation. Moreover, our post-processing method demonstrates strong generality and effectiveness by leveraging multi-task information (semantic and contour information) to delineate object instances. In the experiments, we validate the critical roles of different components in the segmentation task model. Extensive experimental results on four challenging histopathology image segmentation tasks demonstrate the superiority of our method, surpassing state-of-the-art approaches while being tailored for edge computing, thus advancing real-time mobile healthcare and remote diagnostic capabilities.

ACKNOWLEDGEMENT

This work is supported in part by the National Science Foundation (NSF) grant CNS-2348151 and Commonwealth Cyber Initiative grant HC-3Q24-048.

REFERENCES

- [1] X. Zhang, Y. Wang, S. Lu, L. Liu, W. Shi *et al.*, "OpenEI: An open framework for edge intelligence," in *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2019, pp. 1840–1851.
- [2] M. M. Baig, H. GholamHosseini, and M. J. Connolly, "Mobile healthcare applications: system design review, critical issues and challenges," *Australasian physical & engineering sciences in medicine*, vol. 38, pp. 23–38, 2015.

- [3] S. Lu, B. Wu, X. Cong, Y. Yao, and W. Shi, "SafeCampus: multimodal-based campus-wide pandemic forecasting," *IEEE Internet Computing*, vol. 26, no. 1, pp. 60–67, 2021.
- [4] C. Meng, T. Wang, W. Chou, S. Luan, Y. Zhang, and Z. Tian, "Remote surgery case: robot-assisted teleneurosurgery," in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*, vol. 1. IEEE, 2004, pp. 819–823.
- [5] A. Madabhushi, "Digital pathology image analysis: opportunities and challenges," *Imaging in medicine*, vol. 1, no. 1, p. 7, 2009.
- [6] A. Pati, M. Parhi, B. K. Pattanayak, D. Singh, V. Singh, S. Kadry, Y. Nam, and B.-G. Kang, "Breast cancer diagnosis based on IoT and deep transfer learning enabled by fog computing," *Diagnostics*, vol. 13, no. 13, p. 2191, 2023.
- [7] A. H. Fischer, K. A. Jacobson, J. Rose, and R. Zeller, "Hematoxylin and eosin staining of tissue and cell sections," *Cold spring harbor protocols*, vol. 2008, no. 5, pp. pdb-prot4986, 2008.
- [8] R. Hollandi, N. Moshkov, L. Paavolainen, E. Tasnadi, F. Piccinini, and P. Horvath, "Nucleus segmentation: towards automated solutions," *Trends in Cell Biology*, vol. 32, no. 4, pp. 295–310, 2022.
- [9] F. Clayton, "Pathologic correlates of survival in 378 lymph node-negative infiltrating ductal breast carcinomas. mitotic count is the best single predictor," *Ca Cancer J. Clin.*, vol. 68, no. 6, pp. 1309–1317, 1991.
- [10] C. W. Elston and I. O. Ellis, "Pathological prognostic factors in breast cancer. i. the value of histological grade in breast cancer: experience from a large study with long-term follow-up," *Histopathology*, vol. 19, no. 5, pp. 403–410, 1991.
- [11] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA Cancer J Clin.*, vol. 71, no. 3, pp. 209–249, 2021.
- [12] Q. Huang, W. Zhang, Y. Chen, J. Chen, and Z. Yang, "Review of cervical cell segmentation," *Multimedia Tools and Applications*, pp. 1–40, 2024.
- [13] L. Qian, Q. Huang, Y. Chen, and J. Chen, "A purified stacking ensemble framework for cytology classification," in *International Conference on Multimedia Modeling*. Springer, 2024, pp. 267–280.
- [14] B. E. Bejnordi, G. Litjens, N. Timofeeva, I. Otte-Höller, A. Homeyer, N. Karssemeijer, and J. A. Van Der Laak, "Stain specific standardization of whole-slide histopathological images," *IEEE Trans. Image Process.*, vol. 35, no. 2, pp. 404–415, 2015.
- [15] H. He, Z. Huang, Y. Ding, G. Song, L. Wang, Q. Ren, P. Wei, Z. Gao, and J. Chen, "Cdnnet: Centripetal direction network for nuclear instance segmentation," in *Proc. IEEE Int. Conf. Comput. Vision. (ICCV)*, 2021, pp. 4026–4035.
- [16] N. Alemi Koohbanani, M. Jahanifar, A. Gooya, and N. Rajpoot, "Nuclear instance segmentation using a proposal-free spatially aware deep learning framework," in *Med. Image Comput. Comput. Assist. Interv. (MICCAI)*. Springer, 2019, pp. 622–630.
- [17] U. Schmidt, M. Weigert, C. Broaddus, and G. Myers, "Cell detection with star-convex polygons," in *Med. Image Comput. Comput. Assist. Interv. (MICCAI)*. Springer, 2018, pp. 265–273.
- [18] S. Chen, C. Ding, M. Liu, J. Cheng, and D. Tao, "CPP-net: Context-aware polygon proposal network for nucleus segmentation," *IEEE Trans. Image Process.*, 2023.
- [19] S. Wienert, D. Heim, K. Saeger, A. Stenzinger, M. Beil, P. Hufnagl, M. Dietel, C. Denkert, and F. Klauschen, "Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach," *Sci. Rep.*, vol. 2, no. 1, p. 503, 2012.
- [20] C. Chen, W. Wang, J. A. Ozolek, and G. K. Rohde, "A flexible and robust approach for segmenting cell nuclei from 2d microscopy images using supervised learning and template matching," *Cytometry A*, vol. 83, no. 5, pp. 495–507, 2013.
- [21] L. Zhang, H. Kong, C. Ting Chin, S. Liu, X. Fan, T. Wang, and S. Chen, "Automation-assisted cervical cancer screening in manual liquid-based cytology with hematoxylin and eosin staining," *Cytometry A*, vol. 85, no. 3, pp. 214–230, 2014.
- [22] O. Sertel, J. Kong, H. Shimada, U. Catalyurek, J. H. Saltz, and M. Gurcan, "Computer-aided prognosis of neuroblastoma: classification of stromal development on whole-slide images," in *Med. Imaging 2008: Comput-Aided Diagnosis*, vol. 6915. SPIE, 2008, pp. 211–220.
- [23] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jager-sand, "U2-Net: Going deeper with nested u-structure for salient object detection," *Pattern Recognit.*, vol. 106, p. 107404, 2020.
- [24] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Lect. Notes Comput. Sci. (ECCV)*, 2018, pp. 3–19.
- [25] Q. D. Vu, S. Graham, T. Kurc, M. N. N. To, M. Shaban, T. Qaiser, N. A. Koohbanani, S. A. Khurram, J. Kalpathy-Cramer, T. Zhao *et al.*, "Methods for segmentation and classification of digital microscopy tissue images," *Front. Bioeng. Biotechnol.*, p. 53, 2019.
- [26] T. Ilyas, Z. I. Mannan, A. Khan, S. Azam, H. Kim, and F. De Boer, "TSFD-Net: Tissue specific feature distillation network for nuclei segmentation and classification," *Neural Netw.*, vol. 151, pp. 1–15, 2022.
- [27] P. Naylor, M. Laé, F. Reyral, and T. Walter, "Nuclei segmentation in histopathology images using deep neural networks," in *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn.* IEEE, 2017, pp. 933–936.
- [28] N. Kumar, R. Verma, D. Anand, Y. Zhou, O. F. Onder, E. Tsougenis, H. Chen, P.-A. Heng, J. Li, Z. Hu *et al.*, "A multi-organ nucleus segmentation challenge," *IEEE transactions on medical imaging*, vol. 39, no. 5, pp. 1380–1391, 2019.
- [29] Q. D. Vu, S. Graham, T. Kurc, M. N. N. To, M. Shaban, T. Qaiser, N. A. Koohbanani, S. A. Khurram, J. Kalpathy-Cramer, T. Zhao *et al.*, "Methods for segmentation and classification of digital microscopy tissue images," *Front. Bioeng. Biotechnol.*, p. 53, 2019.
- [30] F. Xing and L. Yang, "Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review," *IEEE reviews in biomedical engineering*, vol. 9, pp. 234–263, 2016.
- [31] J. C. Caicedo, J. Roth, A. Goodman, T. Becker, K. W. Karhohs, M. Broisin, C. Molnar, C. McQuin, S. Singh, F. J. Theis *et al.*, "Evaluation of deep learning strategies for nucleus segmentation in fluorescence images," *Cytometry A*, vol. 95, no. 9, pp. 952–965, 2019.
- [32] C. Wählby, I.-M. Sintorn, F. Erlandsson, G. Borgefors, and E. Bengtsson, "Combining intensity, edge and shape information for 2d and 3d segmentation of cell nuclei in tissue sections," *J. Microsc.*, vol. 215, no. 1, pp. 67–76, 2004.
- [33] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in nd images," in *Proc. IEEE Int. Conf. Comput. Vision. (ICCV)*, vol. 1. IEEE, 2001, pp. 105–112.
- [34] D. Stutz, A. Hermans, and B. Leibe, "Superpixels: An evaluation of the state-of-the-art," *Comput. Vision Image Understanding*, vol. 166, pp. 1–27, 2018.
- [35] X. Bai, C. Sun, and C. Sun, "Cell segmentation based on fopso combined with shape information improved intuitionistic fcm," *IEEE journal of biomedical and health informatics*, vol. 23, no. 1, pp. 449–459, 2018.
- [36] C.-H. Lin and C.-C. Chen, "Image segmentation based on edge detection and region growing for thinprep-cervical smear," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 24, no. 07, pp. 1061–1089, 2010.
- [37] Y. Song, W. Cai, D. D. Feng, and M. Chen, "Cell nuclei segmentation in fluorescence microscopy images using inter-and intra-region discriminative information," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*. IEEE, 2013, pp. 6087–6090.
- [38] F. Xing, Y. Xie, and L. Yang, "An automatic learning-based framework for robust nucleus segmentation," *IEEE transactions on medical imaging*, vol. 35, no. 2, pp. 550–566, 2015.
- [39] G. Lv, K. Wen, Z. Wu, X. Jin, H. An, and J. He, "Nuclei r-cnn: Improve mask r-cnn for nuclei segmentation," in *IEEE Int. Conf. Inf. Commun. Signal Process. (ICICSP)*. IEEE, 2019, pp. 357–362.
- [40] J. Dai, Y. Li, K. He, and J. Sun, "R-fcn: Object detection via region-based fully convolutional networks," *Adv. neural inf. proces. syst.*, vol. 29, 2016.
- [41] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proc. IEEE Int. Conf. Comput. Vision. (ICCV)*, 2017, pp. 2961–2969.
- [42] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Med. Image Comput. Comput. Assist. Interv. (MICCAI)*. Springer, 2015, pp. 234–241.
- [43] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz *et al.*, "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.
- [44] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, and J. Liu, "Ce-net: Context encoder network for 2d medical image segmentation," *IEEE transactions on medical imaging*, vol. 38, no. 10, pp. 2281–2292, 2019.
- [45] Y. Zhou, O. F. Onder, Q. Dou, E. Tsougenis, H. Chen, and P.-A. Heng, "Cia-net: Robust nuclei instance segmentation with contour-aware information aggregation," in *Lect. Notes Comput. Sci.* Springer, 2019, pp. 682–693.

- [46] J. Zhao, Y.-J. He, S.-Q. Zhao, J.-J. Huang, and W.-M. Zuo, "Al-net: Attention learning network based on multi-task learning for cervical nucleus segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 6, pp. 2693–2702, 2021.
- [47] H. Chen, X. Qi, L. Yu, Q. Dou, J. Qin, and P.-A. Heng, "DCAN: Deep contour-aware networks for object instance segmentation from histology images," *Med. Image Anal.*, vol. 36, pp. 135–146, 2017.
- [48] L. Sun, S. Cheng, Y. Zheng, Z. Wu, and J. Zhang, "SPANet: Successive pooling attention network for semantic segmentation of remote sensing images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 4045–4057, 2022.
- [49] S. Chen, C. Ding, and D. Tao, "Boundary-assisted region proposal networks for nucleus segmentation," in *Med. Image Comput. Comput. Assist. Interv. (MICCAI)*. Springer, 2020, pp. 279–288.
- [50] S. Graham, Q. D. Vu, S. E. A. Raza, A. Azam, Y. W. Tsang, J. T. Kwak, and N. Rajpoot, "Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images," *Med. Image Anal.*, vol. 58, p. 101563, 2019.
- [51] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7132–7141.
- [52] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, and J. Liu, "Ce-net: Context encoder network for 2d medical image segmentation," *IEEE transactions on medical imaging*, vol. 38, no. 10, pp. 2281–2292, 2019.
- [53] S. Lal, D. Das, K. Alabhya, A. Kanfode, A. Kumar, and J. Kini, "Nucleisegnet: robust deep learning architecture for the nuclei segmentation of liver cancer histopathology images," *Comput. Biol. Med.*, vol. 128, p. 104075, 2021.