Attention U-net for Cell Instance Segmentation

Qianze Liu
Faculty of Information Technology
Beijing University of Technology
Beijing,China
liuqianze0428@163.com

Tong Mo
Faculty of Information Technology
Beijing University of Technology
Beijing,China
tongmo02@icloud.com

Yiya Yao
Faculty of Information Technology
Beijing University of Technology
Beijing,China
1821893948@qq.com

Yufei Zhang*
Faculty of Information Technology
Beijing University of Technology
Beijing,China
* Corresponding author:denny_0601@126.com

***Abstract*—As a major cause of human death and disability, neurological diseases have profound impacts worldwide, yet there is still a lack of effective research to quantify such diseases. As a fundamental but challenging task in the field of medical imaging, segmenting neuronal cell types measured by microscopic observation images to help assess patient status plays a crucial role in the accurate quantification of the disease. To this end, we propose an instance segmentation network of Improved Attention U-net for automatic segmentation of different types of neuronal cells. Specifically, we introduce adaptive mechanism on the basis of U-net structure to enhance the robustness and discriminability of the model. This measure can capture the most important parts in the field of view to filter out redundant information in the target optimization process. In addition, considering the irregular and unique morphology of neuronal cells, we introduce deformable convolution kernels to adapt to the detection needs of various types of cells and lesions. Experimental results show that the algorithm has excellent performance in cell instance segmentation.**

Keywords-Instance Segmentation; Medical Imaging; U-net; Attention; Deformable Convolution

I. INTRODUCTION

Neurological diseases, including neurodegenerative diseases such as Alzheimer's disease, Huntington's disease, and brain tumors, are major causes of death and disability worldwide [1-3]. However, due to the complex and subtle responses of these fatal diseases to treatment, it is difficult to quantify them in a unified way by scientific means. A recognized method is to use optical microscopy to examine neuronal cells, which is convenient and non-invasive [4]. Unfortunately, identifying and segmenting individual neuronal cells in microscopic images will consume a lot of time and resources and be challenging.

The development of computer vision has created conditions for solving this problem [5-7]. Using machines to perform instance segmentation on these cells will save a lot of resources, facilitate the discovery of new effective drugs, and treat millions of people suffering from these diseases. However, the current solutions, although capable of accurate identification of small objects, have limited accuracy for neuronal cells: neuronal cells have very unique morphology associated with them. For example, the neuroblastoma cell line SH-SY5Y is difficult to segment with conventional masks in detection due to its concave and irregular unique cell body morphology [8]. Therefore, it is important to design a segmentation network that can adapt to the deformable and complex morphology of cells to achieve cell instance segmentation.

Inspired by this, we designed an instance segmentation network of Improved Attention U-net, called IAUnet. The design of IAU-net is based on the U-net structure, and the model is improved from two aspects: structure and feature. Structurally, we introduce Attention blocks and change the network layers to capture the most important parts in the field of view for cell instance segmentation, to filter out redundant information in the target optimization process. Feature-wise, considering the complex morphology and small size of neuronal cells, we use deformable convolution kernels to adapt to the detection needs of various types of cells and lesions. Our contributions are shown as follows:

- We explored the challenging task of cell instance segmentation in the field of medical imaging, providing a cell-based quantification method for neurological diseases.
- We proposed a novel instance segmentation network of Improved Attention U-net, which innovated from two aspects: structure and feature.
- We conducted experiments and visualization on Sartorius's benchmark dataset, achieving SOTA results, demonstrating the robustness and effectiveness of our method.

II. Related Work

A. Instance Segmentation

Instance segmentation is a challenging task in computer vision, aiming to locate and segment different categories of objects in various images [9]. It has many applications [10-12] in robotics, autonomous driving, surveillance, medical imaging, etc. In instance segmentation, each object instance is assigned a category label and a pixel-level mask. There are two main approaches [13]: top-down and bottom-up. Top-down methods first detect object regions with bounding boxes, then perform semantic segmentation within each detected region. Bottom-up methods first label the image at the pixel level, then use clustering or metric learning to group pixels into different object instances. However, instance segmentation also faces many difficulties, such as small objects, occlusion, geometric transformations, and image degradation [14]. The segmentation target of this study, neuronal cells, is difficult to segment with fixed pixel-level masks due to their unique irregular morphology and tiny volume.

B. Medical Imaging

Medical imaging [14] refers to the technology and process of obtaining internal tissue images of the human body or a part of it in a non-invasive way for medical or medical research purposes. It includes two relatively independent research directions: medical imaging systems and medical image processing. Medical imaging systems refer to systems that use various physical principles and instrument equipment [15-17], such as X-rays, ultrasound, magnetic resonance, positron emission tomography, etc., to obtain image information of the internal structure and function of the human body. Medical image processing refers to the process of using computer and mathematical methods to analyze, process, enhance, reconstruct, visualize, etc. medical images, in order to improve image quality, extract useful information, and assist diagnosis [18].

In recent years, with the development of computer technology and artificial intelligence, deep learning has made significant breakthroughs in the field of medical imaging, providing new ideas and methods for medical image analysis [19-21]. Deep learning is a machine learning technique based on multi-layer neural networks, which can automatically learn feature representation and abstract concepts from large amounts of data, thus achieving efficient pattern recognition and decision support [22,23]. Deep learning has shown excellent performance in multiple tasks of medical imaging, such as classification, detection, segmentation, registration, reconstruction, synthesis, etc. However, most of the related technologies are limited to static medical detection [24,25], and analysis of brain diseases, breast cancer, lung diseases with obvious disease characteristics. For the neurodegenerative diseases that this study focuses on, it is still a very challenging task.

III. Proposed Method

Improved Attention U-net (IAUnet) is an instance segmentation network for automatically segmenting different types of neuronal cells. Its main idea is to introduce adaptive mechanisms and deformable convolution kernels on the basis of the U-net structure, to dynamically segment cells of different shapes in microscopic images, and to improve the robustness and discriminability of the model.

Improved Attention U-net consists of an encoder and a decoder, where both the encoder and the decoder use residual networks as basic units. The encoder represents the downsampling module composed of residual networks, which is used to extract multiscale features of the input image. The decoder represents the reconstruction module composed of residual networks and upsampling modules, which is used to restore the resolution and details of the input image. The attention gates in the network are used to calculate the attention coefficients according to the feature maps of the corresponding level in the encoder, thereby suppressing the feature activation of irrelevant regions and highlighting the useful salient features. In the decoder, each upsampling module is connected to an attention gate, which can calculate the attention coefficients according to the feature maps of the corresponding level in the encoder, and multiply them with the input feature maps, thereby achieving adaptive focusing. In the last layer of the decoder, a deformable convolution kernel is used to adapt to different shapes and sizes of target structures. Finally, the network outputs the instance segmentation result predicted by Improved Attention U-net, that is, the probability of each pixel belonging to different types of neuronal cells or background regions.

IAUnet uses a generalized focal loss function based on the Tversky index as the objective function, which can solve the problem of data imbalance in medical image segmentation [26]. Compared with the commonly used Dice loss, it can achieve a better balance between precision and recall when training small targets (such as lesions). The objective function of Improved Attention U-net is as follows:

$$L_{FTL} = (1-\alpha)(1-\frac{\varphi_{pg}+\epsilon}{\varphi_{pg}+\alpha\varphi_{pg}^{'}+(1-\alpha)\varphi_{pg}^{''}+\epsilon})$$

$$s.t.\ \ \varphi_{pg} = \sum_{i=1}^{N} p_i\, g_i$$

$$\varphi_{pg}^{'} = \sum_{i=1}^{N} p_i\,(1-g_i)$$

$$\varphi_{pg}^{''} = \sum_{i=1}^{N} (1-p_i)\, g_i \tag{1}$$

where $p_i \in (0,1)$ represents the predicted probability of the i-th pixel, $g_i \in (0,1)$ represents the true label of the i-th pixel, ϵ represents α very small positive number, used to prevent the denominator from being zero. α represents the preference degree for positive samples (target regions), the larger α is, the more preference for positive samples, the more tendency to improve precision; the smaller α is, the more preference for negative samples, the more tendency to improve recall. γ represents the penalty degree for wrong predictions. The larger γ is, the more punishment for wrong predictions, the more tendency to solve the data imbalance problem.

The attention gate of Improved Attention U-net is an adaptive mechanism that can calculate the attention coefficients based on the input feature map and the corresponding level feature map in the encoder, thus suppressing the feature activation of irrelevant regions and highlighting the useful salient features [27,28]. The attention gate can achieve automatic focusing on the target structure in the input image, eliminating the need for using external localization modules, and improving the model sensitivity and accuracy. The formula of the attention gate is as follows:

$$A(x,g) = \left(1 + \sigma\left(W_x * x + W_g * g + b\right)\right) \odot x \qquad (2)$$

The deformable convolution kernel of Improved Attention U-net is a method that dynamically adjusts the size and shape of the convolution kernel according to the shape and size of the input feature map, thus better adapting to the changes of the target structure [29]. Considering the irregular and unique morphology of neuronal cells, using deformable convolution kernels can adapt to the detection needs of various types of cells and lesions. Deformable convolution kernels can be implemented by an offset field, adding a learnable offset to the original convolution kernel, thus giving the convolution kernel a certain degree of deformation ability. Deformable convolution kernels can be defined as:

$$y(p_0) = \sum_{p_n \in R} w(p_n) x(p_0 + p_n + \Delta p_n) \qquad (3)$$

where, $y(p_0)$ represents the value of the output feature map at position p_0; $x(p_0 + p_n + \Delta p_n)$ represents the value of the input feature map at position $p_0 + p_n + \Delta p_n$; $w(p_n)$ represents the weight of the original convolution kernel at position p_n, Δp_n represents the offset, which is predicted by an additional convolution layer during the training process.

IV. Experiment

In this section, we will introduce the experimental details of evaluating the performance of our proposed Improved Attention U-net, including three parts: data set, baseline algorithms, and experimental results.

A. Datasets

We selected the following three benchmark datasets for training and validation.

LIVECell [30]: This is a dataset specifically designed for cell instance segmentation, consisting of 5,239 manually annotated, expert-validated Incucyte HD phase contrast microscopy images. It has a total of 1,686,352 individual cell annotations, from eight different cell types (an average of 313 cells per image). LIVECell images are predefined into training (3188), validation (539) and test (1512) sets. Each split is also further subdivided into each of the eight cell types. The training set also has splits of different sizes (2, 4, 5, 25, 50%) to allow for dataset size experiments.

Sartorius-LIVECell [31]: This is a segmentation of the LIVECell dataset and also the benchmark dataset for the cell instance segmentation competition. Training annotations are provided in the form of run-length encoded masks and images are in PNG format. The number of images is small but the number of annotated objects is quite large. The hidden test set has about 240 images. Note: Although predictions are not allowed to overlap, training labels are fully provided (including overlapping parts). This is to ensure that complete data is provided for each object to the model. Eliminating overlaps in predictions is the task of the contestants.

where, $A(x,g)$ represents the feature map output by the attention gate; x represents the input feature map; g represents the feature map of the corresponding level in the encoder; σ represents the sigmoid function; $\odot$ represents the Hadamard product (element-wise multiplication); W_x, W_g, and b represent learnable parameters.

BBBC010 [32]: This is a dataset consisting of 200 microscopic images showing the morphological changes of human embryonic stem cells (hESCs) under different culture conditions. Each image has corresponding cell instance annotations, totaling 28,765 cells. This dataset is part of the Broad Bioimage Benchmark Collection (BBBC), which is a public collection of biomedical image datasets.

B. Baselines

We selected the following 6 baseline algorithms as the basis for the comparative experiment.

Mask R-CNN [33] is a two-stage instance segmentation algorithm based on Region Proposal Network and Fully Convolutional Network, which can predict a binary mask and a class label for each candidate region. The algorithm won the first place in the instance segmentation task of COCO 2017 challenge.

U-Net [34] is a fully convolutional network based on encoder-decoder structure, which can generate an output image with the same size as the input image. The algorithm is the most widely used method for medical image segmentation.

DeepLabv3+ [35] is a fully convolutional network based on atrous convolution and Atrous Spatial Pyramid Pooling module, which can enlarge the receptive field without reducing the resolution, and capture multi-scale context information. The algorithm achieves the best performance on semantic segmentation task on PASCAL VOC 2012 dataset.

StarDist [36] is an instance segmentation algorithm based on Star Distances Distance Transform (SDDT) and U-Net, which can represent each nucleus as a polygon, and predict the vectors from its vertices to its center point. The algorithm achieves excellent performance on MoNuSeg and BBBC010 datasets.

Cell-DETR [37] is a cell instance segmentation network based on Transformer and attention mechanism, which can directly predict the location and mask of cells from images. Cell-DETR also uses a Hungarian loss, which can effectively match the predicted and real cell instances.

HoVer-Net [38] is a nucleus instance segmentation network based on horizontal-vertical vector field and joint NMS, which can detect the nucleus center points from images, and calculate the nucleus boundaries according to the horizontal-vertical vector field.

C. Results

Robustness and performance. The proposed Improved Attention U-net framework is evaluated on three public benchmarks, including LIVECell, Sartorius-LIVECell and BBBC010, to compare its performance with state-of-the-art methods. First, we apply the model trained on the dataset directly to the test data, and select the average accuracy of five tests in three rounds of experiments as the result for recording, to observe the applicability of IAUnet and other baseline models on different datasets. The corresponding results are shown in Table 1. IAUnet model outperforms previous instance segmentation techniques, especially for Sartorius-LIVECell, the cell instance segmentation competition data that this study focuses on.

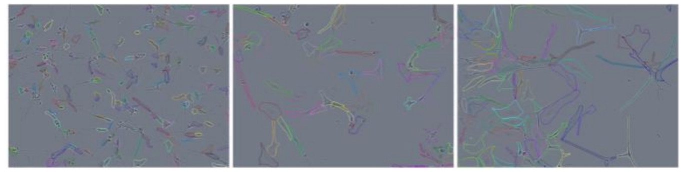


Figure 1. Visualization of instance segmentation results.

TABLE I. COMPARATIVE EXPERIMENTS (ACCURACY(%)) ON THREE BENCHMARK DATASETS

Method	LIVECell			Sartorius-LIVECell			BBBC010		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
Mask R-CNN	35.238	34.290	35.535	48.329	48.389	48.633	70.398	70.239	70.352
U-Net	42.893	42.034	42.298	52.902	52.893	55.289	70.219	70.532	69.651
DeepLabv3+	43.902	43.298	43.109	53.190	53.903	53.101	68.293	68.520	67.056
StarDist	36.230	36.083	36.230	48.013	49.290	49.239	65.525	65.652	65.982
Cell-DETR	39.902	39.118	40.028	49.983	49.455	49.190	68.941	68.954	67.251
HoVer-Net	40.289	40.289	41.930	50.190	50.239	50.981	73.465	73.236	74.448
IAUnet	46.832	46.239	45.923	59.239	57.903	58.388	79.239	78.328	78.657

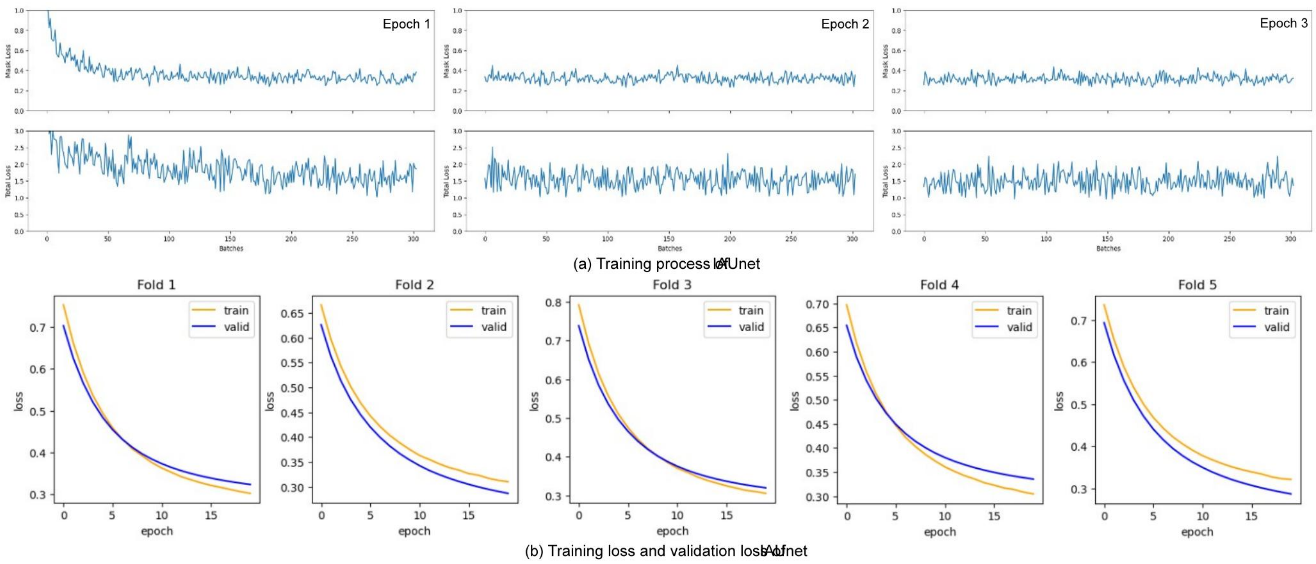
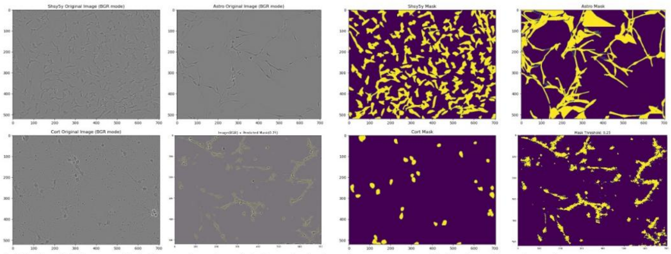


Figure 2. Loss and training process of our IAUnet.

In addition, in order to intuitively reflect the effectiveness and robustness of the model for cell instance segmentation tasks, we conducted experiments and visualization on microscope images containing multiple classes of cells, which is shown in Figure 1. It can be seen that the IAUnet model can accurately identify and segment cells of different types and shapes, and can handle some complex and challenging situations such as cell overlapping.

Training process. In order to analyze the training process and convergence of IAUnet more intuitively, we recorded and plotted the training inspection of the model. We take 300 batches as one training epoch, and Figure 2(a) shows the change of the loss function with the number of iterations (epoch) when we train the IAUnet model. It can be seen that the IAUnet model can converge to a lower loss value, and there is no overfitting phenomenon. At the same time, we plotted the training loss and validation loss of multiple folds of the model in Figure 2(b). This experiment proves the convergence of IAUnet learning.

Visualization. Our work aims to perform instance segmentation on irregular, unique, and complex neuronal cells. In order to intuitively understand the complexity and difficulty of the target in this task scenario, we visualized the neuronal cells in Figure 3. We used yellow to represent the mask information of the neuronal cells, and used the encoder and attention mechanism in U-net to calculate the importance of the cells. From the results, we can see that there is a huge difference between different types of cell lines, and it is difficult to calculate with a fixed-size patch, while our attention mechanism has well processed the microscope images.



(a) Origin BGR images (first 3 images) and predicted results (last one) (b) Mask and importance images

Figure 3. Importance mask and prediction results of neuron cells.

V. CONCLUSION

This paper focuses on neurological diseases and proposes a novel Improved Attention U-net framework for cell instance segmentation, which is a highly challenging problem in the medical image domain. Our method, based on the U-net framework, introduces Attention mechanism and deformable convolution kernel to solve the segmentation difficulty caused by the unique and complex morphological features of neuronal cells. The method has generality and modularity, and can be easily applied to image segmentation and achieve efficient and robust dynamic processing. Experimental results show that our work achieves SOTA performance on multiple benchmarks.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to several individuals who contributed to the completion of this research paper. First and foremost, I extend my deepest appreciation to my classmate Yufei Zhang, whose expertise, guidance, and continuous support have been invaluable throughout the entire research process. His insightful suggestions and constructive feedback greatly enhancing the quality of this work.

I am also grateful to friend Tong Mo for their valuable contributions to this study. His insightful discussions, and diligent efforts significantly enriching the research outcomes. His expertise and dedication were instrumental in overcoming challenges and achieving the desired results.

Furthermore, I would like to thank my friend Yiya Yao for her contributions to this paper. She assistance in data collection, analysis, and interpretation played a crucial role in the successful execution of this research. Their commitment and expertise greatly enhanced the rigor and credibility of the findings.

Additionally, I would like to acknowledge the support and cooperation received from our teacher and research participants, whose valuable insights and cooperation were indispensable to the successful completion of this study.

Lastly, I would like to express my heartfelt thanks to my beloved family for their unwavering support and unwavering confidence in me throughout these years. I am also sincerely grateful to my friends and fellow classmates who have offered their help and time, listening to me and assisting me in resolving difficulties during the challenging course of this thesis.

REFERENCES

[1] Scheltens, P., De Strooper, B., Kivipelto, M., Holstege, H., Chételat, G., Teunissen, C. E., ... & van der Flier, W. M. (2021). Alzheimer's disease. The Lancet, 397(10284), 1577-1590.

[2] Sathasivam, K., Hobbs, C., Mangiarini, L., Mahal, A., Turmaine, M., Doherty, P., ... & Bates, G. P. (1999). Transgenic models of Huntington'sdisease. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 354(1386), 963-969.

[3] DeAngelis, L. M. (2001). Brain tumors. New England journal of medicine, 344(2), 114-123.

[4] Gupta, C., Chandrashekar, P., Jin, T., He, C., Khullar, S., Chang, Q., & Wang, D. (2022). Bringing machine learning to research on intellectual and developmental disabilities: Taking inspiration from neurological diseases. Journal of Neurodevelopmental Disorders, 14(1), 28.

[5] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).

[6] Vuola, A. O., Akram, S. U., & Kannala, J. (2019, April). Mask-RCNN and U-net ensembled for nuclei segmentation. In 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019) (pp. 208-212). IEEE.

[7] Jingyao, W., Naigong, Y., & Firdaous, E. (2021, May). Gesture recognition matching based on dynamic skeleton. In 2021 33rd Chinese Control and Decision Conference (CCDC) (pp. 1680-1685). IEEE.

[8] Xicoy, H., Wieringa, B., & Martens, G. J. (2017). The SH-SY5Y cell line in Parkinson’s disease research: a systematic review. Molecular neurodegeneration, 12, 1-11.

[9] Hafiz, A. M., & Bhat, G. M. (2020). A survey on instance segmentation: state of the art. International journal of multimedia information retrieval, 9(3), 171-189.

[10] Xie, C., Xiang, Y., Mousavian, A., & Fox, D. (2021). Unseen object instance segmentation for robotic environments. IEEE Transactions on Robotics, 37(5), 1343-1359.

[11] De Brabandere, B., Neven, D., & Van Gool, L. (2017). Semantic instance segmentation for autonomous driving. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 7-9).

[12] Zhang, B., & Zhang, J. (2020). A traffic surveillance system for obtaining comprehensive information of the passing vehicles based on instance segmentation. IEEE Transactions on Intelligent Transportation Systems, 22(11), 7040-7055.

[13] Gu, W., Bai, S., & Kong, L. (2022). A review on 2D instance segmentation based on deep neural networks. Image and Vision Computing, 104401.

[14] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. IEEE transactions on pattern analysis and machine intelligence, 44(7), 3523-3542.

[15] Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine learning for medical imaging. Radiographics, 37(2), 505-515.

[16] Elangovan, A., & Jeyaseelan, T. (2016, February). Medical imaging modalities: a survey. In 2016 International Conference on emerging trends in engineering, technology and science (ICETETS) (pp. 1-4). ieee.

[17] Giger, M. L. (2018). Machine learning in medical imaging. Journal of the American College of Radiology, 15(3), 512-520.

[18] Olabarriaga, S. D., & Smeulders, A. W. (2001). Interaction in the segmentation of medical images: A survey. Medical image analysis, 5(2), 127-142.

[19] Wang, J., & Yu, N. (2022, July). SO-PERM: Pose Estimation and Robust Measurement for Small Objects. In 2022 International Joint Conference on Neural Networks (IJCNN) (pp. 1-7). IEEE.

[20] Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. Annual review of biomedical engineering, 19, 221-248.

[21] Yan, L. C., Yoshua, B., & Geoffrey, H. (2015). Deep learning. nature, 521(7553), 436-444.

[22] Dong, S., Wang, P., & Abbas, K. (2021). A survey on deep learning and its applications. Computer Science Review, 40, 100379.

[23] Wang, J., Ruan, X., & Huang, J. (2022). HDPP: High-Dimensional Dynamic Path Planning Based on Multi-Scale Positioning and Waypoint Refinement. Applied Sciences, 12(9), 4695.

[24] Wang, F., Casalino, L. P., & Khullar, D. (2019). Deep learning in medicine—promise, progress, and challenges. JAMA internal medicine, 179(3), 293-294.

[25] Piccialli, F., Di Somma, V., Giampaolo, F., Cuomo, S., & Fortino, G. (2021). A survey on deep learning in medicine: Why, how and when?. Information Fusion, 66, 111-137.

[26] Salehi, S. S. M., Erdogmus, D., & Gholipour, A. (2017). Tversky loss function for image segmentation using 3D fully convolutional deep networks. In Machine Learning in Medical Imaging: 8th International Workshop, MLMI 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 10, 2017, Proceedings 8 (pp. 379-387). Springer International Publishing.

[27] Wang, J., Mou, L., Ma, L., Huang, T., & Gao, W. (2023). AMSA: Adaptive Multimodal Learning for Sentiment Analysis. ACM Transactions on Multimedia Computing, Communications and Applications, 19(3s), 1-21.

[28] Galassi, A., Lippi, M., & Torroni, P. (2020). Attention in natural language processing. IEEE transactions on neural networks and learning systems, 32(10), 4291-4308.

[29] Dai, J., Qi, H., Xiong, Y., Li, Y., Zhang, G., Hu, H., & Wei, Y. (2017). Deformable convolutional networks. In Proceedings of the IEEE international conference on computer vision (pp. 764-773).

[30] Edlund, C., Jackson, T. R., Khalid, N., Bevan, N., Dale, T., Dengel, A., ... & Sjögren, R. (2021). LIVECell—A large-scale dataset for label-free live cell segmentation. Nature methods, 18(9), 1038-1045.

[31] Alzeeb, G., Arzur, D., Trichet, V., Talagas, M., Corcos, L., & Le Jossic-Corcos, C. (2022). Gastric cancer cell death analyzed by live cell imaging of spheroids. Scientific reports, 12(1), 1488.

[32] Chen, L., Strauch, M., Daub, M., Jiang, X., Jansen, M., Luigs, H. G., ... & Merhof, D. (2020, April). A CNN framework based on line annotations for detecting nematodes in microscopic images. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) (pp. 508-512). IEEE.

[33] Liu, J., & Li, P. (2018). A mask R-CNN model with improved region proposal network for medical ultrasound image. In Intelligent Computing Theories and Application: 14th International Conference, ICIC 2018, Wuhan, China, August 15-18, 2018, Proceedings, Part II 14 (pp. 26-33). Springer International Publishing.

[34] Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V. (2021). U-net and its variants for medical image segmentation: A review of theory and applications. Ieee Access, 9, 82031-82057.

[35] Azad, R., Asadi-Aghbolaghi, M., Fathy, M., & Escalera, S. (2020). Attention deeplabv3+: Multi-level context attention mechanism for skin lesion segmentation. In Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16 (pp. 251-266). Springer International Publishing.

[36] Stevens, M., Nanou, A., Terstappen, L. W., Driemel, C., Stoecklein, N. H., & Coumans, F. A. (2022). StarDist Image Segmentation Improves Circulating Tumor Cell Detection. Cancers, 14(12), 2916.

[37] Prangemeier, T., Reich, C., & Koeppl, H. (2020, December). Attention-based transformers for instance segmentation of cells in microstructures. In 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 700-707). IEEE.

[38] Graham, S., Vu, Q. D., Raza, S. E. A., Azam, A., Tsang, Y. W., Kwak, J. T., & Rajpoot, N. (2019). Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. Medical Image Analysis, 58, 101563.