AI Infused Underwater Analysis for Enhanced Coral Species

Abstract—**Among the most lively and varied ecosystems on our planet, coral reefs are yet more susceptible to the changes spurred by climate change, pollution, and human activities. Effective monitoring and a deep understanding of health changes in time are necessary elements for conserving these ecosystems. For the first time, we are presenting here a novel system integrating advanced deep learning techniques with computer vision to automatically detect, segment, and track coral reefs using images taken underwater. In this paper, we trained several models, namely YOLO, U-Net, DeepLabv3+, and FCN, on different sets of the selected coral reef images, demonstrating their feasibility for monitoring the health indicators of corals correctly. This approach is unique in combining the models optimized for different tasks; that is, while the model of YOLO v10 excels in detecting the coral object, the U-Net performs outstandingly in fine details and full-scale evaluation of coral structures. Furthermore, real-time monitoring integration would help scale and make coral conservation more efficient, thus enabling better decision-making in protecting these incredible marine habitats.**

Index Terms—**AI, coral reef, computer vision, monitoring, ROV, segmentation**

I. Introduction

Coral reefs constitute integral components of some of the most intricate and productive ecosystems on the planet, thereby offering critical habitats for a remarkably diverse array of marine organisms [1]. These delicate ecosystems are increasingly subjected to numerous threats, including climate change, pollution, and anthropogenic activities [2]. To comprehend these threats and formulate effective conservation strategies, it is essential to achieve precise monitoring and quantification of coral reef health [3]. Automated methodologies for the segmentation and tracking of underwater coral have the potential to transform the monitoring of coral reefs by facilitating a cost-efficient and scalable framework for data acquisition and analysis. These methodologies are intended to be employed in the assessment of long-term alterations in coral cover, species diversity, and other critical indicators of reef vitality, ultimately contributing to informed management decisions and conservation initiatives [4].

Recent innovations in computer vision and deep learning technologies have facilitated the creation of advanced algorithms for the segmentation and tracking of coral [5]. These methodologies capitalize on high-resolution underwater imagery and sophisticated neural network architectures to accurately identify and delineate individual coral colonies. By integrating these approaches with quantification techniques, researchers can derive significant insights regarding fluctuations in coral coverage, distributions of colony sizes, and species responses to environmental stressors [6]. This information may subsequently be employed to enhance decision-making processes and direct targeted interventions to protect and restore coral reefs.

1. Carried out comprehensive model training and evaluation on several algorithms in order to select the best-performing models for the task.
2. Designed and implemented an object detection system with tracking and area measurement in order to achieve better results.
3. Added more classes of corals for the purpose of making it more versatile and improving the quality of the model.
4. Improved the reliability of the model through development and validation cycles.

The organization of subsequent sections are as follows "Section II discusses the importance and necessity of technical initiatives towards coral reef monitoring. Section III highlights the proposed methodology for coral's inspection using the state of art technology. Section IV provides in depth analysis and comparison of obtained results with metric evaluations. Section V Concludes the proposed work with possible future research directions.

II. Related Works

Deep learning has significantly transformed image segmentation over the past few years, leading to impressive leaps within most applications, including underwater scenarios. Deep learning-based underwater image segmentation has attracted various reviews, with different surveys addressing this broad topic. This most notable work gives an overview of the progression of segmentation techniques, ranging from classic and up to state-of-the-art deep-learning methods - which became irreplaceable in underwater environments, where continually shifting conditions often make it impossible to accurately apply traditional methods, such as light refraction and water turbidity [7].
Underwater image segmentation, especially concerning the health status of coral reefs, has lately been picked to be of extreme interest on account of the very important role coral ecosystems play in marine biodiversity. A new method applies deep learning techniques to improve the accuracy

and speed at which monitoring is conducted on coral reefs by features that specifically present underwater imagery - including changing illumination conditions, water turbidity, and movement of marine organisms, which complicate the segmentation and tracking process [8].

The application of segmentation models in real-world uncontrolled underwater environments is highly challenging. The work points to the need for producing resilient models that adapt to varied and unpredictable underwater conditions, such as variations in water clarity, poor visibility, and mobile organisms. These factors limit the assessment, which calls for the development of models to monitor and analyze coral ecosystems reliably [9]. Further researches are made with deep learning and photogrammetry for coral segmentation, which has a great interest in the integration of advanced imaging techniques with deep learning to produce high-resolution models of coral reefs that are semantically segmented. It aims to be more detailed and accurate for environmental assessment. However, this method has limitations in scalability in dynamic underwater environments because it requires high-quality images and stable conditions in the capture process [10]. Deep learning has improved the detection and classification of fish in underwater video, hence the solution to the problem for real-time detection of fish species and their tracking within coral habitats for biodiversity monitoring and conservation purposes though the movement of fish and debris may interfere with the accuracy of some segmentation models and tracking algorithms [11].

The concern has chiefly been placed on the effects of global warming on coral reefs: coral bleaching. Models predicting when such a bleaching event will occur by utilizing machine learning have proven to be valuable tools in proactive reef conservation in the hands of scientists and conservationists. However, they rely very much on continuous accurate monitoring data, which in current times is hard to come by due to the uncontrolled nature of the underwater environment with the currently applied segmentation and tracking technologies [12]. Besides coral, deep-learning techniques have improved plankton classification. Classification of these marine organisms using convolutional neural networks offers a peep into the biodiversity of marine ecosystems. Due to their size and speed of movement, besides problems inherent to underwater imaging conditions, segmentation and tracking become difficult [13].

An efficient mode of coral surveying methodology based on a high-resolution 3D structure model acquired using a speedy sea scanner and U-Net segmentation presents a novel way of going about large-scale coral surveys. This method, through the integration of U-Net and 3D modeling, allows for effective and scalable monitoring of coral reefs, thereby facilitating a rather more comprehensive and frequent assessment of the health of the coral reefs. However, the dynamic nature of this underwater environment makes it challenging to create and maintain these 3D models, which are required by such scenarios with drastic changes in some states of the underwater environment and subsequently impacts segmentation accuracy and tracking consistency [14].

III. Methodology

Coral reef monitoring is increasingly becoming automated and in this regard, we have endeavored to use all the elements of object detection, segmentation, tracking, and looking at specific areas to characterize coral species underwater. Using deep learning models, the problems of identifying and counting coral species in well-designed underwater locations were well addressed. The flow of the research is discussed in Figure 1.

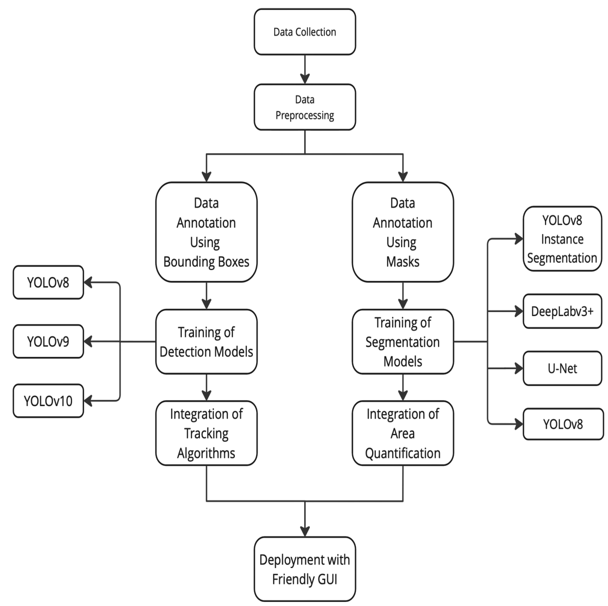

Fig. 1. Methodology of Coral Detection, Tracking and Area Quantification

1) *Collection and Preprocessing of Data*

The dataset that is utilized in this research has been prepared by combining the two publicly available datasets of coral reefs. These datasets include a range of underwater views taken at different times and places with various lighting and depth conditions, which ensures the performance stability of the model in different environments. After merging, the dataset was further augmented by horizontal flipping, rotation, scaling, and color jittering, in a bid to promote model generalization and correct class imbalance where necessary.

2) *Annotation Process*

In the annotation stage, the Roboflow tool was used for training data labeling. For instance, in object detection tasks, coral species were encircled with bounding boxes. In segmentation, such coral regions were outlined using polygons. The dataset was illustrative in that it had distinct 18 coral species, which was useful in ensuring accurate detection and segmentation. Object detection required bounding box annotations, and segmentation required mask annotations, all done with the utmost care using the Roboflow tool [15]. This was a very important manual annotation that was necessary to provide the correct ground truth labels, which would enrich the training of the models. The bounding boxes were mainly used for training YOLO models [16], while the segmentation masks were used with U-Net [17], DeepLabv3+ [18], and FCN models [19].

3) Model Training

The object detection models were trained using the architectures of YOLOv8, YOLOv9, and YOLOv10. These models were selected because of their speed and effectiveness in the ruthlessly complex environment of coral reefs. All models were trained on bounding box images by adjusting the hyperparameters for the best results on the underwater dataset. Training was done with a dynamic learning rate and augmentations imitating underwater distortions like light diffraction were used. For image segmentation, U-Net, DeepLabv3+, Fully Convolutional Networks (FCN), and YOLOv8 were used. These models were trained using polygonal annotations with an emphasis on the boundaries of corals. Each model was assessed based on intersection over union (IoU) measurement and pixel accuracy to ascertain if the coral boundaries were well segmented. The different architectures offered assessments of segmentation ability, with U-Net and DeepLabv3+ capability being rated superior based on how they extract features at different scales.

4) Quantification of the area

Once we finished the segmentation of the coral regions, we devised a novel approach to measure the space attributed to each species of coral. The pixel area for each separated coral figure was determined using the corresponding binary segmentation mask created by the segmentation models. To avoid miscalculation of the coverage area, pixel areas were converted to real-world measurements using reference lengths in the images, enabling precise area measurement. This offered an understanding of the growth patterns and distribution of different species of corals in various regions.

5) Tracking on Video

Video tracking was done using the ByteTrack [20] algorithm. Since ByteTrack is very good at associating frames of detected objects, an implementation had to be implemented even in these complicated underwater environments. Then we used this algorithm in our pipeline for tracking coral instances over time to understand temporal changes both in their coverage and movement. For improved video sequences of coral monitoring, we employed ByteTrack, an advanced approach for multi-object tracking. In this method, ByteTrack abided by the boxes created during object detection using YOLO series models, allowing monitoring of single coral species through multiple frames. The tracking algorithm coped with partial occlusions, disturbances such as underwater noise, and changes in illumination from one frame to another. Therefore, the tracking of the coral species was maintained over time. This method provided valuable information on the spatial and temporal order of movements and behavior of corals concerning their surroundings.

IV. Implementations and Results

This section presents in detail the performance outcomes of coral detection, segmentation, video tracking, and area quantification using various deep-learning models. The different deep models that were used for object detection include YOLO v8, YOLO v9, and YOLO v10, while for segmentation, the employed models are U-Net, DeepLabv3+, FCN, and YOLO v8. The metrics for evaluation are listed herein in the various tables: precision, recall, mAP-mean Average Precision, IoU-intersection over Union, and MIoU-Mean Intersection over Union.

A. Object Detection

Object detection is critical in defining and localizing corals in underwater images. The aim in this case is to demonstrate the capability of the models to detect different coral species with all their associated underwater factors, such as underwater obstruction, variation of light, other similar species, and so forth.

TABLE I
Performance Comparison of trained YOLO Models

Epoch	Model	Precision	Recall	mAP50	mAP50-95	AP
50	YOLOv8	95	93.9	94.4	72.8	80.5
100	YOLOv9	78.4	68.3	80.5	63.1	42
100	YOLOv8	85.3	87.9	90.4	72.3	54.2
100	YOLOv10	94.8	93.47	94.37	74.93	84.44

The results of the YOLO v10 model as seen in Figure 2 are impressively well-rounded over all the metrics of evaluation, compared to the rest of the variants tested. It shows high precision and recall, testifying that the model does quite proficiently in finding coral objects with high accuracy. mAP further supports the model's reliability in predicting accurate bounding boxes along with their respective class probabilities across a diverse range of coral species.

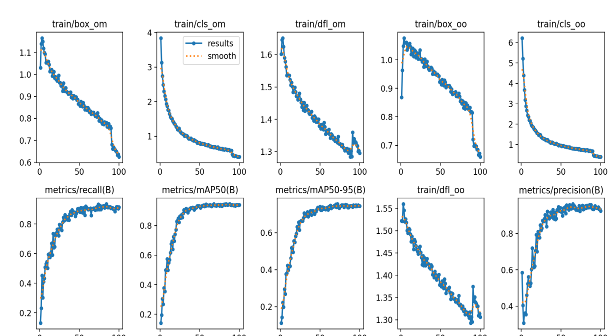

Fig. 2. Illustration of Yolov10 Model Performance

The better performance of YOLO v10, compared to YOLO v8 and YOLO v9 as seen in Table I, could be justified by its better architecture, considering improvements in feature extraction and multiscale detection capability. Lower precision and recall values of YOLO v9 indicate possible issues with the complexity of the coral pattern, especially in distinguishing between similar classes in difficult underwater conditions. Competitive is the YOLO v8, especially the 50-epoch variant, but it still lags behind the accuracy obtained by YOLO v10. This reflects further enhancements in the YOLO v10 model, hence responsible for improving its detection precision, and thus would be preferable when high accuracy at object location and classification is desired.

B. *Object Segmentation*

Segmentation is very important in eliminating ambiguities when it comes to the species of corals, thus making monitoring and Coral quantification processes clearer and easier. This section intends to test the capabilities of the different models in obtaining those details accurately for the estimation of the coral area.

TABLE II
COMPARISON OF MODELS BASED ON EPOCH AND IOU

MODEL	Epoch	IoU
U-Net	60 (Early Stopping)	85.6
DeepLabv3+	41 (Early Stopping)	83.48
FCN	100	80.83
YOLOv8	100	82.34

U-Net performed best among compared models for segmentation tasks with high precision and recall, along with a significant IoU score (Table II). This fact is indicative of the architecture of the U-Net model in capturing fine-grained details in coral images and hence providing superior segmentation performance to make it suitable for applications needing detailed boundary delineation of coral structure as shown in the Figure 3.

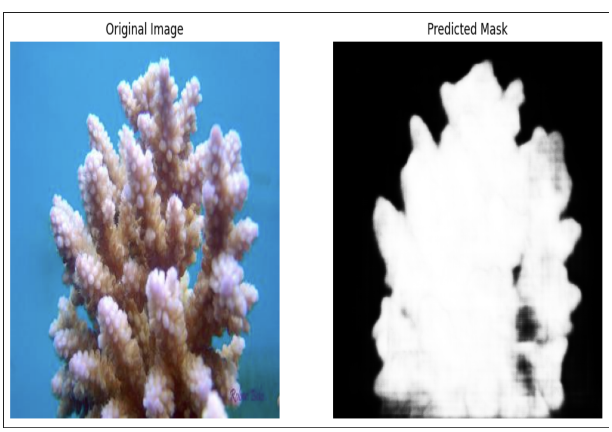

Fig. 3. Illustration of U-Net models Mask Prediction

DeepLabv3+ had reasonable success but was less accurate, especially at IoU, compared to U-Net. Performance for this model insinuates that while it can segment the coral regions, the detailing of the corals may not be as keen as that achieved by U-Net (Fig 3). FCN gave the least encouraging result out of all the segmentation models; it probably had simpler feature extraction mechanisms, hence its lower IoU score.

YOLO v8 was mainly developed for object detection tasks. However, segmentation tasks were also performed in the experiments. As can be seen in the results, despite being capable of performing segmentation, it is better adapted to detection, with lower mAP proven in segmentation tasks against U-Net and DeepLabv3+.

C. *Area Quantification*

Some coral surfaces must be measured over time, which contributes to the evaluation of general coral health and growth indiscriminately. In this sense, it is quite critical to have a clear understanding of coral and coral reef extent and be in a position to measure them for ecological studies, rather than evaluations.

Fig. 4. Illustration of Input images for Area Quantification

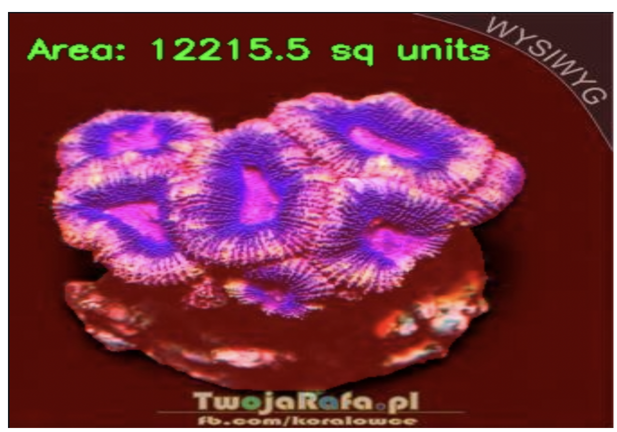

Fig. 5. Illustration of Quantified Area of the Input Coral

Incorporating the segmentation models into our custom pipeline made it possible to estimate the coral area with some reliability. U-Net accomplished coral-related tasks best, where they assessed coral area measured in pixels using the segmentation output for each coral area (Fig 4, 5). Subsequently, those pixel areas were transformed into real-life unit measurements, hence enabling effective measurement of coral cover and growth over a specified period. This level of accuracy is very important in ecological studies, for example, where coral cover is a significant factor for the health status of the ecosystem in general.

D. Video Tracking

Observing changes in corals through time allows researchers to understand the extent and patterns of growth and interactions of corals in a given period. The issue here is that it is difficult to keep track of the same corals in many frames due to the extreme nature of the aquatic environment and such factors as changing intensities of light and water currents.

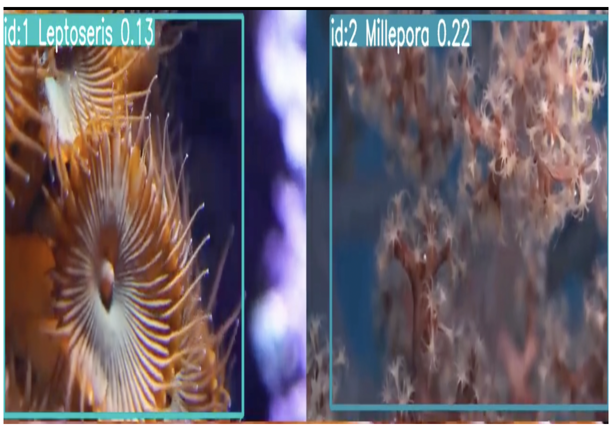

Fig. 6. Illustration of ByteTrack results

In all trials where the ByteTrack was applied, efficient results regarding the tracking of coral objects through the video images were achieved (Fig 6). The system coped well with the task of tracking coral species even in the presence of occlusion and other environmental dynamics. The ability to continuously know the presence of coral objects is essential for purposes of monitoring those objects over a long duration and also studying their behavior, hence ByteTrack is suitable for persistent tracking in underwater videos.

V. Conclusion

This study showcases the power of deep learning models in revolutionizing coral reef monitoring. The YOLO v10 model proved highly effective in detecting coral structures, while U-Net offered precise segmentation, which is critical for understanding reef health. By incorporating video tracking techniques like ByteTrack, we were able to dynamically monitor changes in coral coverage and movement over time, providing a more comprehensive view of reef conditions. The results suggest that these technological advancements can significantly improve how we monitor and conserve coral reefs, providing detailed and timely insights that help guide conservation strategies. Further, this proposed work will be useful as a base platform to deploy in the ROV for real-time validations handling the unique challenges posed by different underwater environments and to extend their use to monitoring other aspects of marine biodiversity.

References

[1] L. A. Courtney, E. Muller, K. S. Viglino, and C. S. Rogers, "Estimating 3-Dimensional Colony Surface Area of Field Corals," Journal of Experimental Marine Biology and Ecology, vol. 351, no. 1–2, pp. 234–242, June 2007. Available: https://doi.org/10.1016/j.jembe.2007.06.021.

[2] J. M. Roberts, A. J. Wheeler, and A. Freiwald, "Reefs of the Deep: The Biology and Geology of Cold-Water Coral Ecosystems," Science, vol. 312, pp. 543–547, April 2006. DOI: 10.1126/science.1119861.

[3] M. B. Lyons, R. M. Fisher, and K. R. N. Anthony, "New Global Area Estimates for Coral Reefs from High-Resolution Mapping," Cell Reports Sustainability, vol. 1, no. 2, Elsevier, p. n.p., 2024. Available: https://doi.org/10.1016/j.crsus.2024.100015.

[4] C. Roelfsema, S. Phinn, W. M. Udy, and G. S. Oxley, "Coral Reef Habitat Mapping: A Combination of Object-Based Image Analysis and Ecological Modelling," Remote Sensing of Environment, vol. 208, pp. 27–41, Feb. 2018. Available: https://doi.org/10.1016/j.rse.2018.02.005.

[5] A. Raphael, Z. Dubinsky, and D. Iluz, "Deep neural network recognition of shallow water corals in the Gulf of Eilat (Aqaba)," Sci Rep, vol. 10, p. 12959, July 2020. Available: https://doi.org/10.1038/s41598-020-69201-w.

[6] T. B. Razak, A. Munandar, I. Zainuddin, and W. A. Yahya, "Large Scale Study of Benthic Communities in Eastern Indonesia's Reef Systems," Regional Studies in Marine Science, vol. 44, p. 101731, March 2021. Available: https://doi.org/10.1016/j.rsma.2021.101731.

[7] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3523-3542, 2021. Available: https://ieeexplore.ieee.org/abstract/document/9356353

[8] H. Zhang, M. Li, X. Pan, X. Zhang, J. Zhong, and J. Qin, "Novel approaches to enhance coral reefs monitoring with underwater image segmentation," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 46, pp. 271-277, 2022. Available: https://isprs-archives.copernicus.org/articles/XLVI-3-W1-2022/271/2022/isprs-archives-XLVI-3-W1-2022-271-2022.pdf

[9] P. Drews-Jr, I. de Souza, I. P. Maurell, E. V. Protas, and S. S. C. Botelho, "Underwater image segmentation in the wild using deep learning," Journal of the Brazilian Computer Society, vol. 27, pp. 1-14, 2021. Available: https://link.springer.com/article/10.1186/s13173-021-00117-7

[10] H. Zhang, A. Grün, and M. Li, "Deep learning for semantic segmentation of coral images in underwater photogrammetry," ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 2, pp. 343-350, 2022. Available: https://www.research-collection.ethz.ch/handle/20.500.11850/557143

[11] Z. Sun, "A novel approach to coral fish detection and classification in underwater footage based on convolutional neural network," in *Journal of Physics: Conference Series*, vol. 1650, no. 3, p. 032012, Oct. 2020, IOP Publishing. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1650/3/032012/meta

[12] N. Boonnam, T. Udomchaipitak, S. Puttinaovarat, T. Chaichana, V. Boonjing, and J. Muangprathub, "Coral reef bleaching under climate change: Prediction modeling and machine learning," *Sustainability*, vol. 14, no. 10, p. 6161, 2022. Available: https://www.mdpi.com/2071-1050/14/10/6161

[13] A. Lumini, L. Nanni, and G. Maguolo, "Deep learning for plankton and coral classification," *Applied Computing and Informatics*, vol. 19, no. 3/4, pp. 265-283, 2020. Available: https://www.emerald.com/insight/content/doi/10.1016/j.aci.2019.11.004/full/html

[14] K. Mizuno, K. Terayama, S. Hagino, S. Tabeta, S. Sakamoto, T. Ogawa, K. Sugimoto, and H. Fukami, "An efficient coral survey method based on a large-scale 3-D structure model obtained by Speedy Sea Scanner and U-Net segmentation," *Scientific Reports*, vol. 10, no. 1, p. 12416, 2020. Available: https://www.nature.com/articles/s41598-020-69400-5

[15] B. Dwyer, J. Nelson, T. Hansen, et al., "Roboflow (Version 1.0) [Software]," 2024. Available: https://roboflow.com

[16] A. Wang, H. Chen, L. Liu, K. Chen, Z. Lin, J. Han, and G. Ding, "Yolov10: Real-time end-to-end object detection," arXiv preprint, vol. 2405.14458, 2024. Available: https://arxiv.org/abs/2405.14458

[17] R. Azad, E. Khodapanah Aghdam, A. Rauland, Y. Jia, A. H. Avval, A. Bozorgpour, S. Karimijafarbigloo, J. P. Cohen, E. Adeli, and D. Merhof, "Medical image segmentation review: The success of u-net," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024. Available: https://arxiv.org/abs/2211.14830

[18] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 801-818, 2018. Available: https://arxiv.org/abs/1802.02611

[19] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431-3440, 2015. Available: https://openaccess.thecvf.com/content_cvpr_2015/html/Long_Fully_Convolutional_Networks_2015_CVPR_paper.html

[20] Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu, and X. Wang, "Bytetrack: Multi-object tracking by associating every detection box," in *European Conference on Computer Vision*, pp. 1-21, Cham: Springer Nature Switzerland, 2022. Available: https://link.springer.com/chapter/10.1007/978-3-031-20047-2_1