

# Locating Bacterial Flagellar Motors

Yusuf Karamustafa, Emir Arda Eker, Faruk Rıza Öz, Doğukan Arslan  
karamustafay21@itu.edu.tr, ekere22@itu.edu.tr, ozf20@itu.edu.tr

Department of Artificial Intelligence and Data Engineering, Istanbul Technical University  
YZV416E Computer Vision - Spring 2025 Term Project Presentations



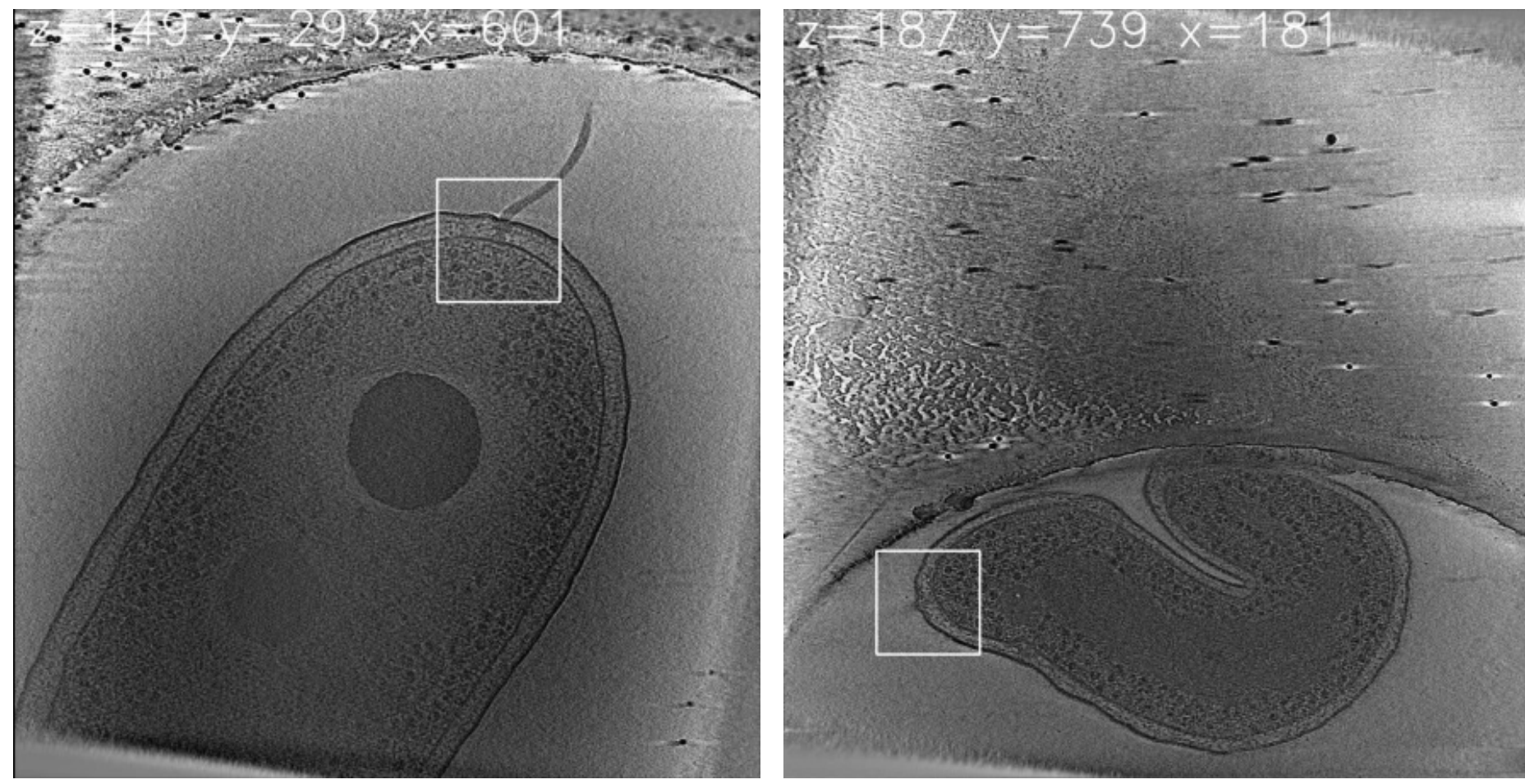
## 1. Abstract

The flagellar motor is a vital molecular machine driving bacterial motion. While cryo-electron tomography (cryo-ET) enables imaging in 3D, manually identifying these motors is slow and error-prone. In this project, we develop an automated approach to detect the presence and location of flagellar motors in tomograms.

## 2. Problem Definition

Identifying flagellar motors in 3D cellular images is a tedious and complex task due to the large number of slices and low image quality. The challenge is to replace manual effort with an automated system that can detect the presence and precise location of these motors.

This task is the focus of the *BYU Locating Bacterial Flagellar Motors* Kaggle competition, which provides 3D tomograms as stacks of 2D slices and requires predicting accurate 3D coordinates of motors. Our objective is to design a model that understands patterns across these slices to deliver reliable motor localization.



## 4. Motivation

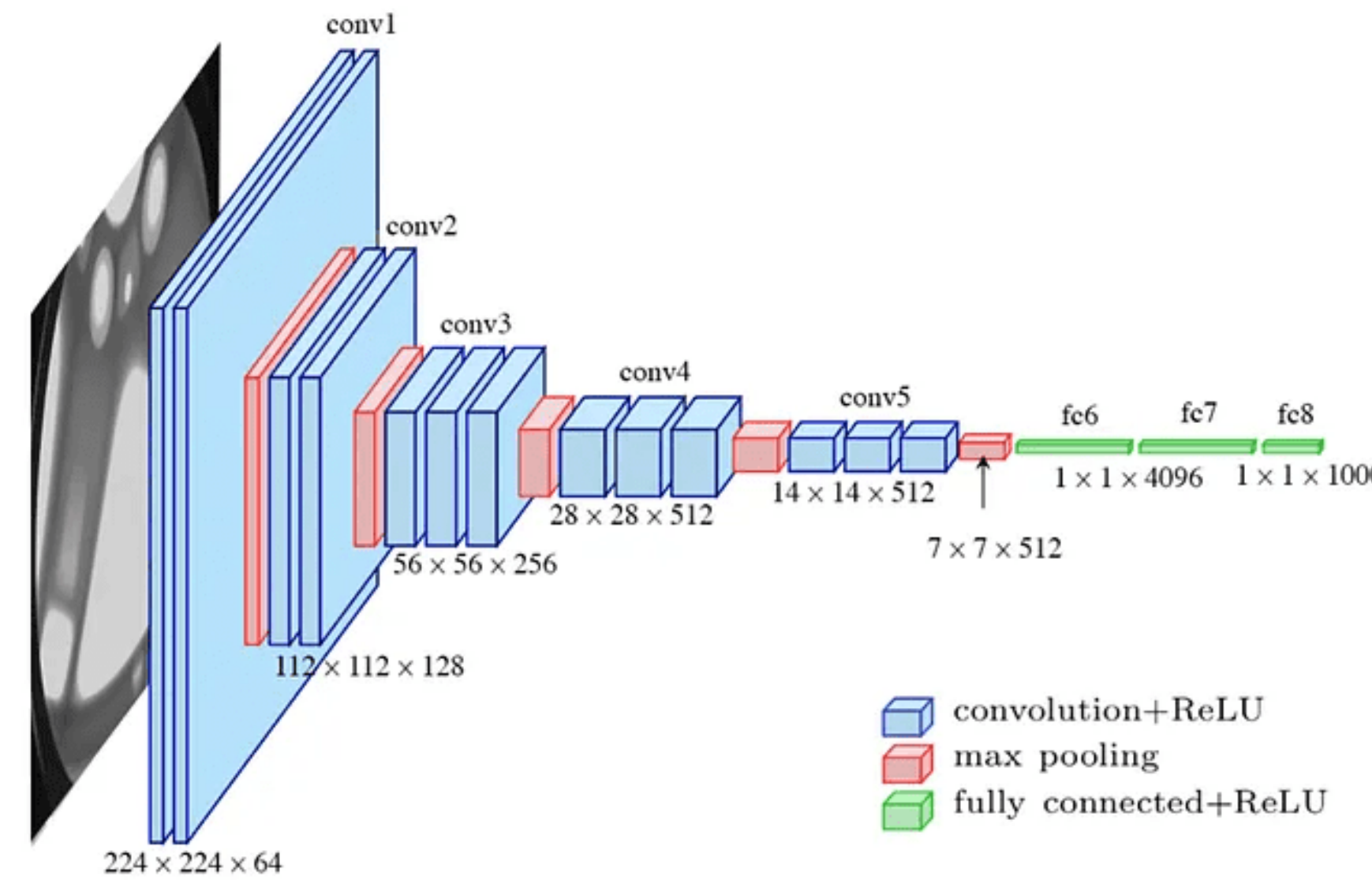
Detecting flagellar motors is critical for studying how bacteria move and interact with their environments. Despite its importance, this problem has seen little attention in the computer vision and deep learning community. The lack of existing solutions, combined with the biological significance of accurate motor detection, makes it both a challenging and valuable research task.

## 7. Conclusions

We presented an automated approach to detect bacterial flagellar motors in cryo-electron tomograms using deep learning. By combining YOLOv10 and MHAF-YOLO through a clustering-based ensembling strategy, we achieved a significant performance boost, outperforming individual models. Our method reduces the need for manual annotations and accelerates motor localization, making it valuable for microbiology and biomedical imaging research. We ranked **42nd out of over 1,000 teams** in the Kaggle competition, demonstrating the effectiveness of our approach. Future work could explore 3D-aware networks and temporal consistency across slices to further improve accuracy.

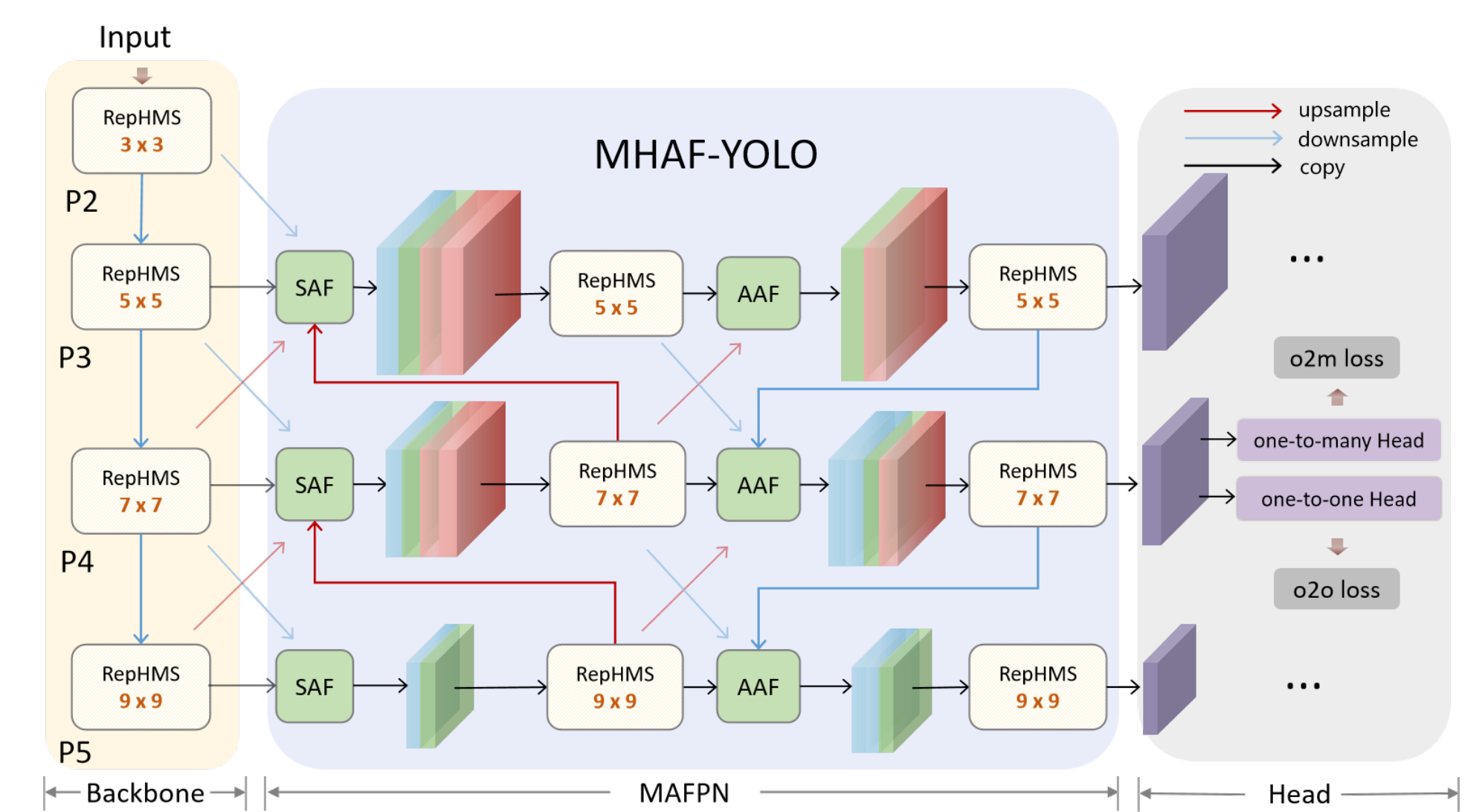
## 3. Proposed Method

Our pipeline gets inference from two different models and then combines these predictions using DBSCAN to obtain a final prediction. We used YOLOv10 and MHAF-YOLO as our models.



**YOLOv10:** This real-time object detector introduces four key components:

- *One-to-Many and One-to-One Detection Heads* for end-to-end detection without non-maximum suppression.
- A lightweight *C2f* module enabling efficient feature reuse and reduced computation in the backbone.
- *Dual Detection Heads* combining dense and sparse predictions to improve accuracy.
- A robust *Task-Aligned Label Assignment (TAL)* strategy aligning spatial and semantic cues for better supervision.

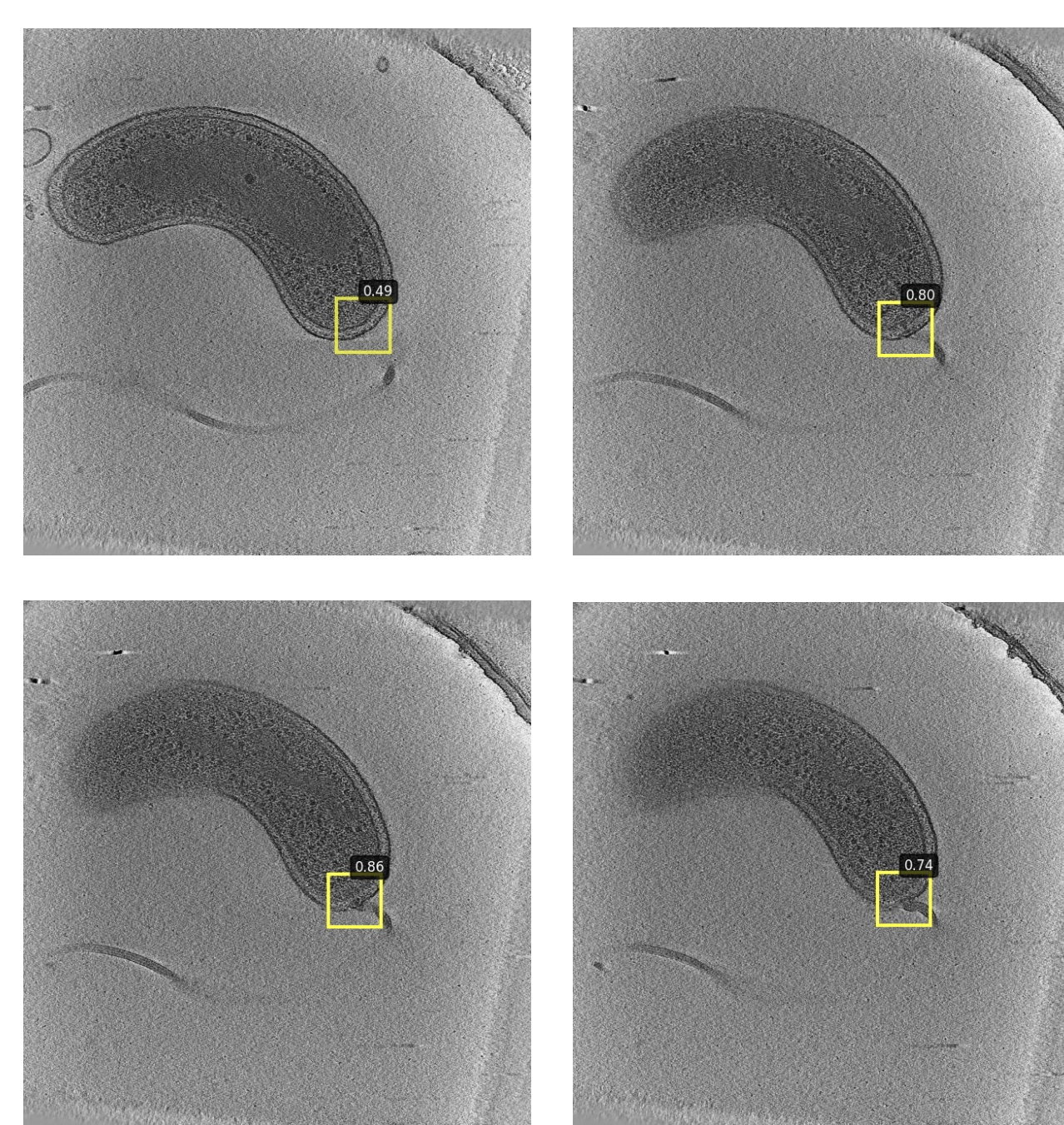


**MHAF-YOLO:** This model is built on top of YOLOv10 with four key innovations:

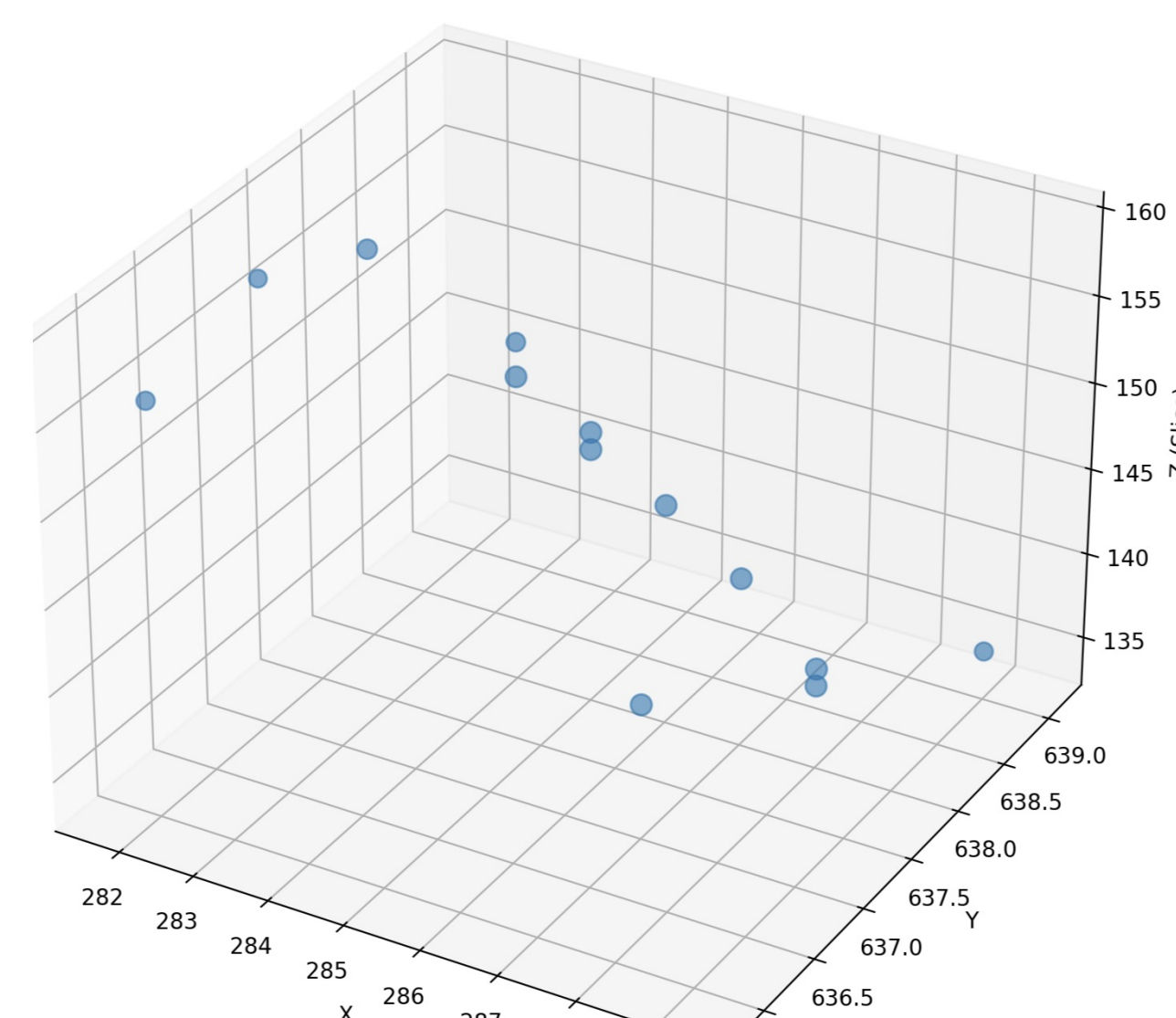
- A *Multi-Branch Auxiliary FPN (MAFPN)* that fuses shallow and deep features using SAF and AAF modules for better multi-scale representation.
- A *Superficial Assisted Fusion (SAF)* module that injects low-level spatial details into the neck, improving small object detection.
- An *Advanced Assisted Fusion (AAF)* module that enriches deep-layer gradients through dense multi-scale connections.
- A *RepHMS* block with re-parameterized large and small kernels, supported by the *GHFS mechanism*, which dynamically adjusts receptive fields for adaptive context extraction.

## 5. Qualitative Results

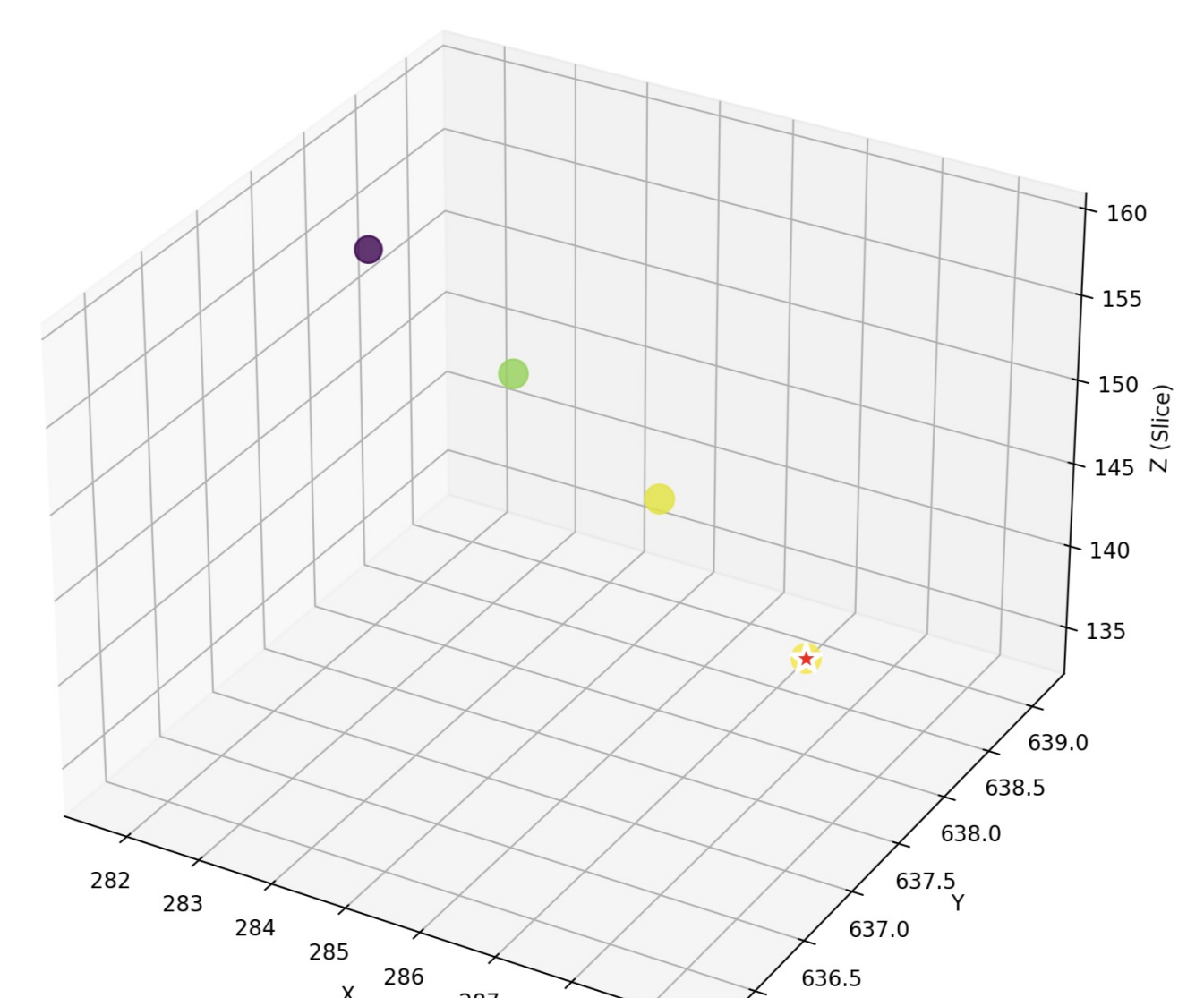
We first get the predictions from both of the models and then use DBSCAN to construct clusters, the cluster with the highest confidence is chosen as the final prediction.



Model Inference



Predictions in 3D



DBSCAN Clusters

## 6. Quantitative Results

A prediction is considered accurate if it is within a certain threshold:

$$\|y - \bar{y}\|_2 \leq \tau$$

The final score is calculated using the  $F_\beta$ -score with  $\beta = 2$ :

$$F_\beta = (1 + \beta^2) \cdot \frac{TP}{(1 + \beta^2)TP + \beta^2FN + FP}$$

**Results:** YOLOv10 scored an  $F_\beta$ -score of 0.776, while MHAF-YOLO achieved 0.802. The ensemble method reached a final score of **0.839**.

## 8. References

- [1] Wang, Ao, Chen, Hui, Liu, Lihao, Chen, Kai, Lin, Zijia, Han, Jungong, and Ding, Guiguang. *YOLOv10: Real-Time End-to-End Object Detection*. arXiv preprint arXiv:2405.14458, 2024.
- [2] Yang, Zhiqiang, Guan, Qiu, Yu, Zhongwen, Xu, Xinli, Long, Haixia, Lian, Sheng, Hu, Haigen, and Tang, Ying. *MHAF-YOLO: Multi-Branch Heterogeneous Auxiliary Fusion YOLO for Accurate Object Detection*. arXiv preprint arXiv:2502.04656, 2025.