A Comparison of Markerless Body Part Tracking Algorithms Toward a Virtual Reality Environment for Weakly Electric Fish

Alper Endeş[1], Amine Adıgüzel[1], Burak Karaçayır[1], İlyas Emre Vergili[1]
Supervised by Dr. İsmail Uyanık[1]
[1] Hacettepe University, Department of Electrical and Electronics Engineering

HACETTEPE UNIVERSITY

INTRODUCTION

Recent progress in deep learning algorithms enables the use of computer vision to study animal behavior in real time. Here, we investigated the behavioral responses of weakly electric fish via various convolutional neural network models. Our goal is to assess the performance of these algorithms toward building a real-time virtual reality environment for the weakly electric fish.

MOTIVATION

Eigenmannia virescens, a species of weakly electric fish, track the movement of a PVC refuge to remain hidden inside. To achieve this, *Eigenmannia* generates two counter propagating waves along its ribbon fin, which meet at a **nodal point**. Tracking this nodal point is especially challenging due to its dynamic behavior, which requires simultaneous analysis of consecutive frames.

METHODS

Design Requirements

- Markerless Pose Estimation
- Real-Time Object Detection
- Over 50 FPS Tracking Performance
- Accurate Tracking of Nodal Points
- Tracking Nodal Points on Video Motion

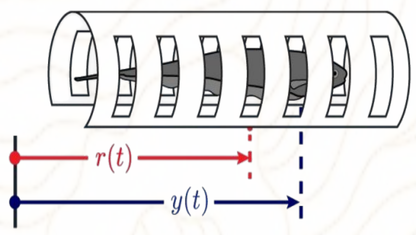
Promising Tracking Algorithms

Haar Cascades
- ✓ Simple Neural Network with three features
- ✓ Uses features brightness to detect the object

GOTURN
- ✓ Offline training, online tracking
- ✓ Predicts over previous and current frame

YOLOv3
- ✓ 106-layered CNN
- ✓ Uses Three Scale Method
- ✗ Requires GPU for speed

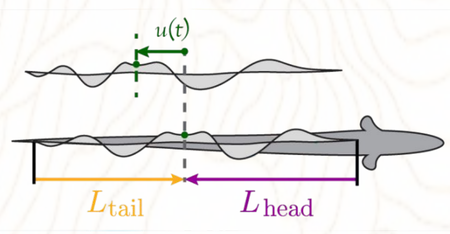
MOSSE
- ✓ Location of the max. correlation output indicates new position of the target
- ✓ Filter is computed in FFT

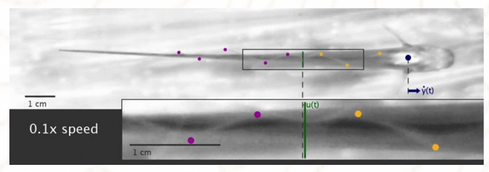


Tracking Nodal Point Methods

DeepLabCut
- ✓ Ability to Implement Deep CNNs
- ✓ Ability to Choose Different Models
- ✓ High Accuracy on Tracking Body Parts
- ✗ Takes Time to Train

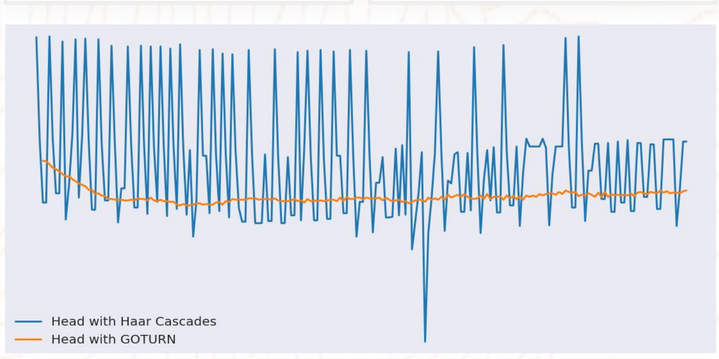
Custom CNN
- ✓ VGG16[2] Inspired Custom Model
- ✓ Well Documented PyTorch Library
- ✓ Customizable Layers and Hyperparameters
- ✗ Requires Hand Crafted Dataset

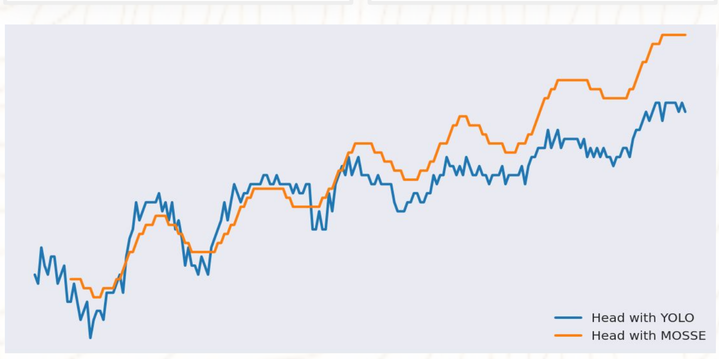


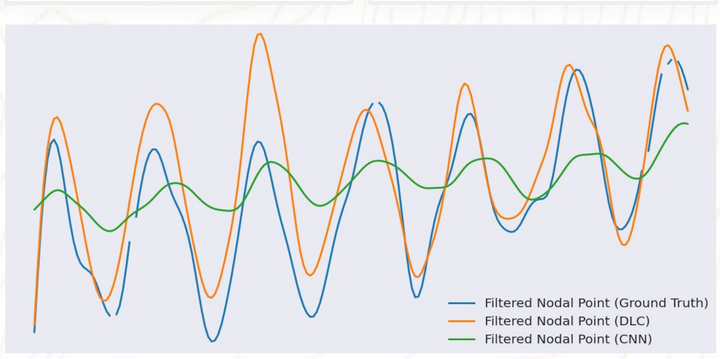


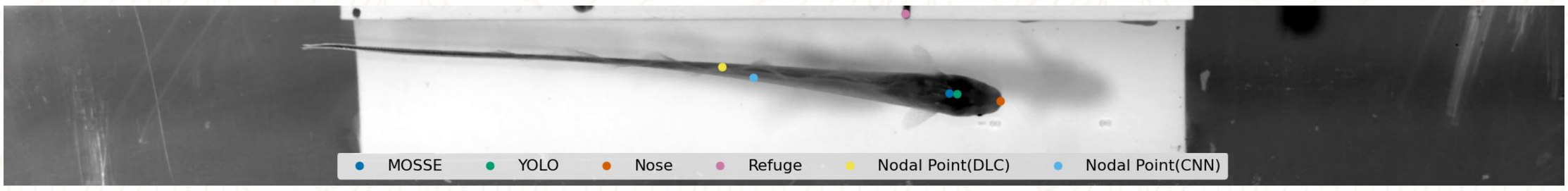
RESULTS AND DISCUSSION

Haar Cascades	GOTURN	YOLOv3	MOSSE	Custom CNN	DeepLabCut
⧗ Average 38 ms	⧗ Average 50 ms	⧗ 1300 ms with 1.70 GHz CPU	⧗ Average 10 ms	⧗ Average 84 ms	⧗ Average 138 ms
✗ Unreliable Detection Quality	✗ Bounding Box is not fixed size	✓ Satisfying quality	✓ Robust, stable, fast	✓ 0.009 MSE Loss	✓ Fits with Ground Truth
			✓ Real-Time	✗ Less Accurate	









ACKNOWLEDGEMENTS

This project was completed within the context of ELE401 - ELE402 Graduation Project courses in Hacettepe University, Faculty of Engineering, Department of Electrical and Electronics Engineering.

REFERENCES

1 Uyanik I, Sefati S, Stamper SA, Cho K, Ankaralı MM, Fortune ES and Cowan NJ (2020) Variability in locomotor dynamics reveals the critical role of feedback in task control. eLife 2020;9:e51219 DOI: 10.7554/eLife.51219
2 Simonyan, Karen & Zisserman, Andrew. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 1409.1556.