

FISH DETECTION AND TRACKING USING KINECT v2 SENSOR

ADITI BHATEJA



**AMAR NATH AND SHASHI KHOSLA SCHOOL OF
INFORMATION TECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY DELHI, INDIA**

JULY 2024

© Indian Institute of Technology Delhi (IITD), New Delhi, 2024

FISH DETECTION AND TRACKING USING KINECT v2 SENSOR

by

ADITI BHATEJA

**AMAR NATH AND SHASHI KHOSLA SCHOOL
OF INFORMATION TECHNOLOGY**

Submitted

in fulfilment of the requirements of the degree of Doctor of Philosophy

to the



INDIAN INSTITUTE OF TECHNOLOGY DELHI, INDIA

JULY 2024

dedicated to my family

CERTIFICATE

This is to certify that the thesis titled “**Fish Detection and Tracking using Kinect v2 sensor**” being submitted by **Ms. Aditi Bhateja** to the **Amar Nath and Shashi Khosla School of Information Technology, Indian Institute of Technology Delhi**, for the award of the degree of **Doctor of Philosophy**, is a record of bona-fide research work carried out by her under our guidance and supervision. In our opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted to any other university or institute for the award of any degree or diploma.

Prof. Brejesh Lall

Department of Electrical Engineering

Indian Institute of Technology Delhi

New Delhi – 110016

Prof. Prem Kumar Kalra

Department of Computer Science

and Engineering

Indian Institute of Technology Delhi

New Delhi – 110016

Date:

Place:

DECLARATION

I hereby certify that:

1. The work contained in the thesis is original and has been done by me under the guidance of my supervisors.
2. The work has not been submitted to any other institute for any other degree or diploma.
3. I have followed the guidelines provided by the institute in preparing the thesis.
4. I have conformed to ethical norms and guidelines while writing the thesis.
5. Whenever I have used materials (data, models, figures, text) from other sources, I have given due credit by citing them in the text of the thesis, and giving their details in the references, and taken permission from the copyright owners of the sources whenever necessary.

Aditi Bhateja

School of Information Technology

Indian Institute of Technology Delhi

Hauz Khas, New Delhi, Delhi-110016

ACKNOWLEDGEMENTS

I express hearty thanks and a deep sense of gratitude to my supervisors, Prof. Brejesh Lall and Prof. Prem K Kalra, who gave me an opportunity to undertake this research work. Without their constant guidance, encouragement and creative approach to problem solving, this work would not have reached completion.

I extend my heartfelt appreciation to the anonymous reviewers of my thesis for their insightful comments and valuable feedback during the review process. Their expertise significantly contributed to the improvement of the thesis. I would also like to thank my SRC members who have provided me quality inputs to improve my research work and unknown reviewers who accepted my papers for publication in reputed journals and conferences.

I am thankful to my father and mother, for their constant guidance and support throughout my research journey. Without their unwavering support, this task could not have been possible. I would give a big thanks to my husband, my guiding light, who has given constant encouragement, valuable time, and nudged me whenever required throughout my research journey.

I am thankful to all my other family members including my lovely daughter for giving her love, cheering me up whenever I felt down, and to support me through tough decisions I had to make.

Lastly, I am thankful to my colleagues and co-authors who engaged with me in doing research works and in brainstorming for new ideas and way forward whenever there seems to be a roadblock ahead.

ADITI BHATEJA

Abstract

Development of deep-sea observatories has led to increased demand of automatic fish detection and tracking. Detecting fishes, estimating fish number, and their tracking can help the marine biologists to understand underwater environment and nature of marine animals. Due to the bleak body deformation of fish, that results in their complex motion and multiple occlusions, robust fish tracking from video image sequences is a highly challenging task. This thesis explores the different techniques used for automatic fish detection and tracking followed by providing an application of the same for marine researchers in determining the critical velocity of fish. The thesis discusses the steps involved in the process, the challenges encountered, and different ways to proceed in the direction. The work entailed though focuses on fish detection and tracking but the same can be generalized as an application of object detection and tracking.

From the last few years, RGB-D cameras are widely used by researchers in various fields. Their reasonable cost and the ability to estimate distances at a high frame rate have made these sensors recommendable for applications in gaming accessories, robotics, computer vision, etc. In addition to colour, these sensors also provide depth information. Aspects like the stability, accuracy, and reliability of depth-sensing cameras like Kinect v2 must also be

considered before using the device for applications like that of 3D space modelling. In this thesis, an analysis of the error in the depth measurement as well as calculation of Depth Entropy given by Kinect v2 sensor in different mediums viz. air, glass and water has been done. We have validated our findings using the theories of optics. The findings from error analysis were used to make an error compensation model which can correct depth at each pixel of the image. The error analysis and error compensation model, proposed in our work, will help in improving the accuracy of present and future depth sensing devices.

Our thesis also presents Kinect v2 sensor based fish detection and tracking. The work builds on our knowledge of error analysis and error compensation to generate a 3-D depth map of aquarium.

Fish detection and tracking in a constrained environment (aquarium) aids in the research of marine biologists to understand and study the behaviour of fishes. The use of commodity hardware devices, like Microsoft Kinect, for detection and tracking is economical for the researchers and helps to automate the study of fishes round the clock without any requirement of manual intervention. In our work, we have calibrated depth and RGB images to detect and track multiple fishes present in each frame of a recorded video. A multi-tiered methodology involving classical clustering, classification, and image processing techniques is proposed to overcome the problems associated with

fish detection and tracking, and multiple fishes were tracked using centroid of the fish as central component.

Apart from utilizing the traditional machine learning (ML) algorithms for fish detection and tracking, we explored the domain of deep neural network for multi-fish detection and tracking. The data set provided by Kinect v2 was cleaned, aligned, and calibrated to make it suitable for object detection and tracking. We have utilised YOLOv5s model and proposed three different variants of the model namely YOLOv5-WAM, YOLOv5-RGBD, and YOLOv5-SpecialFeatures.

YOLOv5-WAM (With Attention Mechanism) used Transformer learning based attention mechanism to detect smaller objects like fishes. YOLOv5-RGBD incorporates 4 channel information i.e. 3 colour channels (R, G, B) and one Depth channel (D), to perform object detection on RGBD or multiple channel datasets. YOLOv5-SpecialFeatures utilised custom features derived from traditional feature detection techniques to experiment on amalgamation of custom features with deep learning model i.e. YOLOv5. YOLOv5-WAM and YOLOv5-RGBD can be adapted for use on devices having low compute availability, making them suitable for real time object detection and tracking using low-cost experimental setup. An object detection and tracking pipeline is also designed to detect and track underwater objects (fishes in our work), belonging to multiple classes in 3 dimensions.

Finally, the insights gained from fish detection were utilized to train a prediction model for predicting Critical Velocity (Ucrit) of fishes. Ucrit is one component which is calculated by marine researchers to identify the swimming capacity of fish. Traditional methods of computing Ucrit are human intensive, involve elaborate setup, and are time-taking. Hence, in this thesis, a model is trained by utilising the existing Ucrit data to predict Ucrit of fishes for which detection and tracking pipeline was developed previously.

सारांश

गहरे समुद्र में वेधशालाओं के विकास से स्वचालित रूप से मछली का पता लगाने और ट्रैकिंग की मांग में वृद्धि हुई है। मछलियों का पता लगाने, मछलियों की संख्या का अनुमान लगाने और उनकी ट्रैकिंग से समुद्री जीवविज्ञानियों को पानी के नीचे के पर्यावरण (जलमग्न पर्यावरण) और समुद्री जानवरों की प्रकृति को समझने में मदद मिल सकती है। मछलियों के शरीर की धूमिल विकृति के कारण, जिसके परिणामस्वरूप उनकी जटिल गति और कई रुकावटें होती हैं, वीडियो छवि अनुक्रमों से उपयुक्त मछली ट्रैकिंग एक अत्यधिक चुनौतीपूर्ण कार्य है। यह थीसिस स्वचालित तरीके से मछली का पता लगाने और ट्रैकिंग के लिए उपयोग की जाने वाली विभिन्न तकनीकों की पड़ताल करती है और इसके बाद मछली के महत्वपूर्ण वेग को निर्धारित करने में समुद्री शोधकर्ताओं के लिए इसका एक अनुप्रयोग प्रदान करती है। यह थीसिस प्रक्रिया में शामिल चरणों, सामने आने वाली चुनौतियों और दिशा में आगे बढ़ने के विभिन्न तरीकों पर चर्चा करती है। हालाँकि इसमें शामिल कार्य

मछली का पता लगाने और ट्रैकिंग पर केंद्रित है, लेकिन इसे वस्तु का पता लगाने और ट्रैकिंग के अनुप्रयोग के रूप में सामान्यीकृत किया जा सकता है।

पिछले कुछ वर्षों से, विभिन्न क्षेत्रों में शोधकर्ताओं द्वारा RGB-D कैमरों का व्यापक रूप से उपयोग किया जा रहा है। उनकी उचित लागत और उच्च फ्रेम दर पर दूरियों का अनुमान लगाने की क्षमता ने इन सेंसरों को गेमिंग एक्सेसरीज़, रोबोटिक्स, कंप्यूटर विज़न आदि में अनुप्रयोगों के लिए अनुशंसित बना दिया है। रंग के अलावा, ये सेंसर गहराई से जानकारी भी प्रदान करते हैं। 3D स्पेस मॉडलिंग (3D space modeling) जैसे अनुप्रयोगों के लिए डिवाइस का उपयोग करने से पहले Kinect v2 जैसे गहराई-संवेदन कैमरों की स्थिरता, सटीकता और विश्वसनीयता जैसे पहलुओं पर भी विचार किया जाना चाहिए। इस थीसिस में, Kinect v2 सेंसर द्वारा विभिन्न माध्यमों (हवा, कांच और पानी) में दिए गए गहराई मापन में त्रुटि के साथ-साथ गहराई एंट्रॉपी की गणना का विश्लेषण किया गया है। हमने प्रकाशिकी के सिद्धांतों का उपयोग करके अपने निष्कर्षों को मान्य किया है। त्रुटि विश्लेषण से प्राप्त निष्कर्षों का उपयोग एक त्रुटि क्षतिपूर्ति मॉडल बनाने के लिए किया गया था जो छवि के प्रत्येक

पिक्सेल पर गहराई को सही कर सकता है। हमारे काम में प्रस्तावित त्रुटि विश्लेषण और त्रुटि क्षतिपूरक मॉडल (एरर कॉम्पेनसेशन मॉडल) , वर्तमान और भविष्य की गहराई संवेदन उपकरणों (डेप्थ सेंसिंग डिवाइस) की सटीकता में सुधार करने में मदद करेगा।

हमारी थीसिस Kinect v2 सेंसर आधारित मछली का पता लगाने और ट्रैकिंग भी प्रस्तुत करती है। यह कार्य एक्वेरियम का 3-डी गहराई मानचित्र तैयार करने के लिए त्रुटि विश्लेषण और त्रुटि क्षतिपूर्ति के हमारे ज्ञान पर आधारित है।

एक सीमित वातावरण (मछलीघर) में मछली का पता लगाना और ट्रैकिंग करना समुद्री जीवविज्ञानियों के अनुसंधान में मछलियों के व्यवहार को समझने और अध्ययन करने में सहायता करता है। पता लगाने और ट्रैकिंग के लिए माइक्रोसॉफ्ट किनेक्ट जैसे कमोडिटी हार्डवेयर उपकरणों का उपयोग शोधकर्ताओं के लिए किफायती है और मैन्युअल हस्तक्षेप की आवश्यकता के बिना चौबीसों घंटे मछलियों के अध्ययन को स्वचालित करने में मदद करता है। हमारे काम में, हमने रिकॉर्ड किए गए वीडियो के प्रत्येक फ्रेम में मौजूद कई मछलियों का

पता लगाने और उन्हें ट्रैक करने के लिए गहराई और आरजीबी छवियों को कैलिब्रेट किया है। मछली का पता लगाने और ट्रैकिंग से जुड़ी समस्याओं को दूर करने के लिए शास्त्रीय क्लस्टरिंग, वर्गीकरण और छवि प्रसंस्करण तकनीकों को शामिल करने वाली एक बहु-स्तरीय पद्धति प्रस्तावित है, और केंद्रीय घटक के रूप में मछली के सेंट्रोइड का उपयोग करके कई मछलियों को ट्रैक किया गया था।

मछली का पता लगाने और ट्रैकिंग के लिए पारंपरिक मशीन लर्निंग (एमएल) एल्गोरिदम का उपयोग करने के अलावा, हमने बहु-मछली का पता लगाने और ट्रैकिंग के लिए गहरे तंत्रिका नेटवर्क के डोमेन का पता लगाया। Kinect v2 द्वारा प्रदान किए गए डेटा सेट को ऑब्जेक्ट डिटेक्शन और ट्रैकिंग के लिए उपयुक्त बनाने के लिए साफ़, संरेखित और कैलिब्रेट किया गया था। हमने YOLOv5s मॉडल का उपयोग किया है और मॉडल के तीन अलग-अलग वेरिएंट प्रस्तावित किए हैं, जैसे YOLOv5-WAM, YOLOv5-RGBD, और YOLOv5-SpecialFeatures।

YOLOv5-WAM (अटेंशन मैकेनिज्म के साथ) ने ट्रांसफार्मर लर्निंग आधारित ध्यान तंत्र का उपयोग मछलियों जैसी छोटी वस्तुओं का पता लगाने के लिए किया. YOLOv5-RGBD में RGBD या मल्टीपल चैनल डेटासेट पर ऑब्जेक्ट डिटेक्शन करने के लिए 4 चैनल जानकारी यानी 3 कलर चैनल (R, G, B) और एक डेप्थ चैनल (D) शामिल है। YOLOv5-SpecialFeatures ने गहन शिक्षण मॉडल यानी YOLOv5 के साथ कस्टम सुविधाओं के समामेलन पर प्रयोग करने के लिए पारंपरिक फीचर डिटेक्शन तकनीकों से प्राप्त कस्टम सुविधाओं का उपयोग किया। YOLOv5-WAM और YOLOv5-RGBD को कम गणना उपलब्धता वाले उपकरणों पर उपयोग के लिए अनुकूलित किया जा सकता है, जिससे वे कम लागत वाले प्रयोगात्मक सेटअप का उपयोग करके वास्तविक समय वस्तु का पता लगाने और ट्रैकिंग के लिए उपयुक्त हो जाते हैं। एक ऑब्जेक्ट डिटेक्शन और ट्रैकिंग पाइपलाइन को 3 आयामों में कई वर्गों से संबंधित पानी के नीचे की वस्तुओं (हमारे काम में मछलियों) का पता लगाने और ट्रैक करने के लिए भी डिज़ाइन किया गया है।

अंत में, मछली का पता लगाने से प्राप्त अंतर्दृष्टि का उपयोग मछलियों के क्रिटिकल वेलोसिटी (यूक्रिट) की भविष्यवाणी के लिए एक भविष्यवाणी मॉडल को प्रशिक्षित करने के लिए किया गया था। यूक्रिट एक घटक है जिसकी गणना समुद्री शोधकर्ताओं द्वारा मछली की तैराकी क्षमता की पहचान करने के लिए की जाती है। यूक्रिट की गणना के पारंपरिक तरीके मानव गहन हैं, इसमें विस्तृत सेटअप शामिल है और इसमें समय लगता है। इसलिए, इस थीसिस में, मछलियों के यूक्रिट की भविष्यवाणी करने के लिए मौजूदा यूक्रिट डेटा का उपयोग करके एक मॉडल को प्रशिक्षित किया जाता है जिसके लिए पता लगाने और ट्रैकिंग पाइपलाइन पहले विकसित की गई थी।

Table of Contents

Certificate	i
Declaration	ii
Acknowledgements.....	iii
Abstract	iv
List of Figures	xviii
List of Tables.....	xxi
Abbreviations	xxiii
1. Introduction	1
1.1 Scope and Objective	1
1.2 Major Contribution of the Thesis.....	10
1.2.1 Error Analysis of Kinect v2 RGBD Sensor	11
1.2.2 Multi Fish Detection and Tracking using Machine Learning Techniques	12
1.2.3 Multi Fish Detection and Tracking using YOLOv5	16
1.2.4 Predicting Ucrite: An Application of Fish Detection Pipeline	17
1.3 Layout of the Thesis	17
2. Literature Review	20
3. Validation Datasets.....	29
3.1 Custom Dataset.....	29
3.1.1 Kinect Calibration.....	34
3.2 Fish4Knowledge Dataset Dataset	36
3.3 COCO Dataset	38
3.4 Ground Truth	39
4. Error Analysis of Kinect v2RGB-D Sensor	41
4.1 Introduction.....	41
4.2 Methodology.....	47
4.3 Experimental Results	50
4.4 Angular Variation of the Error & its Validation through Optics	58

4.4.1 Medium 1: Air	59
4.4.2 Medium 2: Glass.....	61
4.4.2.1 Analysis of Depth with respect to angles	61
4.4.2.2 Validation through optics (v.1)	61
4.4.2.3 Vertical Field of View	64
4.4.3 Medium 3: Water	65
4.4.3.1 Analysis of Depth with respect to angles	65
4.4.3.2 Validation through optics (v.2)	65
4.4.3.3 Vertical field of view	68
4.5 Depth Entropy.....	69
4.6 Error Compensation Model	73
4.7 Inference	80
4.8 Conclusion	85
5. Multi Fish Detection and Tracking using RGBD Vector	87
5.1 Introduction.....	89
5.1.1 Background on Feature Detection	89
5.1.2 SURF	90
5.1.3 KAZE.....	92
5.1.4 SUZE (SURF+KAZE).....	96
5.1.5 Centroid Based Tracking	98
5.1.6 Fish Detection adTracking using SUZE	98
5.1.6.1 Results on aquarium dataset	101
5.1.6.2 Results on Fish4Knowledge database.....	105
5.1.7 3-D Detection and Tracking System.....	106
5.2 Methodology.....	110
5.2.1 Occlusion Handling	110
5.2.2 Fish Detection.....	112
5.2.2.1 K-means clustering	117
5.2.2.2 FuzzyC.....	120
5.2.2.3 DBSCAN	120
5.2.3 Fish Tracking.....	122
5.3 Experimental Results	124

5.3.1 System specification and Dataset preparation	124
5.3.2 Object detection results.....	124
5.3.3 Tracking Results	127
5.4 Conclusion	131
6. Multi Fish Detection and Tracking using YOLOv5	132
6.1 Introduction.....	132
6.2 YOLOv5 based Object Detection Techniques.....	140
6.2.1 Baseline Model YOLOv5	140
6.2.1.1 Backbone.....	142
6.2.1.2 Neck	144
6.2.1.3 Head	145
6.3 Dataset Preparation.....	146
6.3.1 Data gathering.....	146
6.3.2 Depth Correction	147
6.3.3 Frame Alignment.....	147
6.3.4 Data Calibration.....	148
6.3.5 Dataset Summary.....	149
6.4 Proposed Works.....	150
6.4.1 YOLOv5-WAM.....	151
6.4.2 YOLOv5-RGBD.....	155
6.4.3 YOLOv5-SpecialFeatures.....	156
6.4.4 Object Detection and Tracking Pipeline	158
6.5 Results and Discussion	159
6.5.1 Benchmark.....	160
6.5.2 Results of YOLOv5-WAM.....	163
6.5.3 Results of YOLOv5-RGBD.....	166
6.5.4 Results of YOLOv5-SpecialFeatures.....	167
6.6 Conclusion	168
7. Predicting Ucrit of Fishes: An Application of Fish Detection Pipeline.....	170
7.1 Introduction.....	170
7.2 Dataset Summary.....	173
7.2.1 Iberian Fish Dataset	173

7.2.2 Tropical Fish Dataset of Larval and Juvenile Fish	174
7.3 Methodology	175
7.4 Results and Discussion	180
8. Conclusion and Future Directions	186
8.1 Conclusion	186
8.2 Future Work	189
Bibliography	191
List of Publications Based on this Thesis	199

List of Figures

Figure 1: Diagrammatic representation of setup.....	30
Figure 2: Microsoft Kinect v2 sensor [1].....	31
Figure 3: Color (a) and Depth (b) frame captured by Kinect depth sensor	31
Figure 4: <i>Cyprinus rubrofasciatus</i>	32
Figure 5: <i>Pterophyllum scalare</i>	32
Figure 6: <i>Carassius auratus</i>	33
Figure 7: <i>Labeo chrysophekadion</i>	33
Figure 8: <i>Pangasianodon hypophthalmus</i>	33
Figure 9: <i>Chanos chanos</i>	33
Figure 10: Calculating Intrinsic Parameters	34
Figure 11: Fish4Knowledge dataset	36
Figure 12: Experimental Setup	48
Figure 13: Color and Depth Frames by Kinect v2.....	49
Figure 14: Plot of Error vs. Actual Depth value in a) air b) glass c) water	54
Figure 15: Field of view measurement	59
Figure 16: Experimental Setup of Kinect for depth measurements of objects in a glass box	63
Figure 17: Experimental Setup of Kinect for depth measurement of objects in a rectangular glass box filled with water	67
Figure 18: Accuracy distribution of Kinect v2 in water	69
Figure 19: Average Entropy Distribution	71

Figure 20: Average Entropy vs. Depth in a) air b) Glass c) Water.....	73
Figure 21: Error heatmap at depth 130 mm.....	78
Figure 22: Error Colormap of glass box filled with water.....	79
Figure 23: a) Original depth map of object in water box b) Corrected depth map c) Heatmap of error d) Frequency distribution of error	82
Figure 24: Error vs. Actual depth- Error estimation.....	83
Figure 25: Strongest keypoints obtained from a) SURF b) KAZE.....	92
Figure 26: Architecture of fish detection methodology.....	99
Figure 27: Evaluation Metrics	103
Figure 28: a) Sample fish and its pre-processing, b) sample frame for matching the features, c) detecting SUZE features in a frame, d) detected fish, e) trajectory of single fish and f) with multiple fishes.....	104
Figure 29: a) Sample image and SUZE features detected. b) SUZE features in a frame. c) Detected fishes. d) Centroid detection. e) Trajectories of fishes in fish4knowledge database	106
Figure 30: Detecting occluded fishes.....	112
Figure 31: Detection of objects using different models.....	1177
Figure 32: Fish detection using k -means clustering	119
Figure 33: Fish detection using FuzzyC	120
Figure 34: Fish detection using DBSCAN	121
Figure 35: Fish trajectories	130
Figure 36: Sharp turns of the fish trajectory	130
Figure 37: Architecture of YOLOv5	141
Figure 38: Important Blocks of YOLOv5.....	142

Figure 39: Timestamp plotting of RGB and Depth frames	148
Figure 40: Calibrated RGB-D frame	149
Figure 41: Dataset sample images	150
Figure 42: YOLOv5-WAM	152
Figure 43: YAML snippet for SpecialFeatures.....	158
Figure 44: Object Detection and Tracking Pipeline	159
Figure 45: Performance of YOLOv5 different models by Ultralytics	161
Figure 46: Detection Output different datasets.....	165
Figure 47: Fish trajectories on Custom Dataset.....	166
Figure 48: Flume setup for measuring Ucrit.....	172
Figure 49: Morphological features of fish	178
Figure 50: Actual Ucrit vs Predicted Ucrit	181

List of Tables

Table 4.1: Technical Specifications of Kinect v2	42
Table 4.2: Experimental Values of the Error and the Relative Error for measurements made in Air and Glass (for determination of Linear Accuracy)	55
Table 4.3: Experimental Values of the Error and the Relative Error for measurements made in Water (for determination of Linear Accuracy)	58
Table 4.4: Experimental Values of the Error and Relative Error for measurements (with air as a medium) for testing the horizontal field of view of the sensor.	60
Table 4.5: Experimental Values of the Error and Relative Error for measurements (with air as a medium) for testing the vertical field of view of the sensor	60
Table 4.6: Experimental Values of the Error and Relative Error for measurements (with empty glass as a medium)	62
Table 4.7: Experimental Values of the Error and Relative Error for measurements (with water filled in glass box, as a medium)	66
Table 4.8: Entropy variation in air.....	71
Table 4.9: Entropy variation in glass	72
Table 4.10: Entropy variation in water	72
Table 5.1: Precision and Recall for aquarium and Fish4Knowledge w.r.t SUZE features.....	102
Table 5.2: Classification Accuracy on Aquarium Dataset.....	103
Table 5.3: Classification Accuracy on Fish4Knowledge Dataset.....	105
Table 5.4: Accuracy comparison of clustering algorithms (in %).....	126

Table 5.5: Detection accuracy of RGB and RGBD model	126
Table 5.6: Classification accuracy using different models	127
Table 5.7: Technical parameters for trajectory	129
Table 6.1: YOLOv5 Results on Different Datasets	162
Table 6.2: YOLOv5-WAM Results on Different Datasets.....	163
Table 6.3: YOLOv5-RGBD Results on Custom Dataset	167
Table 6.4: YOLOv5-SpecialFeatures Results on Different Datasets.....	168
Table 7.1: Comparision of Ucrit Prediction models	183
Table 7.2: Ucrit Prediction on Custom dataset fishes.....	185

Abbreviations

ANN	Artificial Neural Network
Avg	Average
BoVW	Bag of Visual Words
CF	Correction Factor
COCO	Common Objects in Context
CSP	Cross Stage Partial
DBSCAN	Density Based Spatial Clustering
DECB	Depth Extended Code Book
FOV	Field of View
FPN	Feature Pyramid Network
GSS	Gaussian Scale Space
ICT	Information and Communication Technology
IoU	Intersection over Union
IR	Infrared
mAP	Mean Average Precision
MCMO	Multi-Class Multi-Object
MD	Measured Distance
MLP	Multilayer Perceptron
MOTA	Multiple Object Tracking Accuracy
MOTP	Multiple Object Tracking Precision
NA	Not Applicable

NPP	Nuclear Power Plant
PAN	Path Aggregation Network
RCNN	Region-Based Convolutional Neural Network
RE	Relative Error
SPP	Spatial Pyramid Pooling
SURF	Speeded Up Robust Features
SVM	Support Vector Machine
Taxa	Taxonomy
TOF	Time of Flight
Ucrit	Critical Swimming Speed
WAM	With Attention Mechanism
XGBoost	Extreme Gradient Boosting
YOLO	You Only Look Once