International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972

Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

INTERNATIONAL JOURNAL OF INNOVATION AND INDUSTRIAL REVOLUTION (IJIREV)
www.ijirev.com

GLOBAL ACADEMIC EXCELLENCE

ADVANCES IN FISH TRACKING TECHNOLOGIES FOR AQUACULTURE: OVERCOMING CHALLENGES AND SHAPING FUTURE RESEARCH

Samra Urooj Khan[1*], Kamarul Hawari Ghazali[1], Sundas Naqeeb Khan[2]

[1] Faculty of Electrical and Electronic Engineering, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia
Email: samrauroojkhan1615@gmail.com; kamarul@umpsa.edu.my
[2] Department of Graphics, Computer Vision and Digital Systems, Silesian University of Technology, Gliwice, Poland
Email: sundas.faisal@polsl.pl
[*] Corresponding Author

Article Info:

Article history:
Received date: 05.01.2025
Revised date: 19.01.2025
Accepted date: 10.02.2025
Published date: 03.03.2025

To cite this document:

Khan, S. U., Ghazali, K. H., & Khan, S. N. (2025). Advances in Fish Tracking Technologies for Aquaculture: Overcoming Challenges and Shaping Future Research. *International Journal of Innovation and Industrial Revolution, 7* (20), 29-61.

DOI: 10.35631/IJIREV.720003

Abstract:

Aquaculture is essential for developing countries to have food security and for fishermen's socioeconomic conditions to improve. Fish tracking is essential to intelligent fish farming since it helps with health assessments, behavior monitoring, and water quality maintenance. However, high individual resemblance, rapid movement, and occlusions from foam in tanks provide obstacles for multi-object fish tracking. This paper explores the level of digital technology in aquaculture today, emphasizing systems based on vision, acoustics, and biosensors. It draws attention to the benefits, drawbacks, and uses of various technologies while highlighting important areas that require more study. Development is still hampered by a lack of extensive fish datasets and standardized evaluation techniques. We outline future research possibilities and move into advanced deep learning techniques like tracking-by-detection and merging deep features with correlation filtering to address these issues. We also give an overview of pre-deep learning fish tracking systems. This review provides a comprehensive overview of the evolution of fish tracking technologies and outlines potential avenues for advancing research and technology in aquaculture.

Keywords:

Aquaculture; Fish Tracking; Intelligent Fish Farming; Vision-Based Systems; Deep Learning

International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

Introduction

Aquatic animals' stress responses to contaminants can be utilized to create biological early warning systems, and the behavioral alterations they undergo reflect their aquatic surroundings and ecosystems (Bae & Park, 2014; Mao et al., 2016; Z. Ren & Wang, 2010; Xia et al., 2018). The visual monitoring of changes in fish school behavior has gained importance in biological water quality monitoring in recent times due to significant advancements in computer vision and image processing technology (Bhargava, 2016; Kuklina et al., 2013; Ma et al., 2010; Papadakis et al., 2012; Zhao et al., 2019).

To identify items, including plants, fruits, cars, people, faces, animals, characters, and automobiles, computer vision employing 2D images has been employed extensively (Concepcion et al., 2020; De Luna et al., 2018; Fernandez et al., 2014; Quiros et al., 2015). Additionally, it is frequently utilized for multiple object tracking (MOT), which includes phenotyping plants, cars, animals, and people (Concepcion et al., 2020; Luo et al., 2021). One of the most important recent developments in fish behavioral biometrics monitoring, including anomaly detection, fish hunger, and responses to environmental circumstances, has also been recognized as MOT, which uses computer vision (Beyan et al., 2018; Xia et al., 2018; Yang et al., 2021).

Acquiring and statistically analyzing fish school motion data is the most informative technique to investigate schooling behavior and uncover underlying principles (Butail & Paley, 2012; Delcourt et al., 2009). Although manual collective motion analyses are laborious, time-consuming, and occasionally even unfeasible, video-tracking technology facilitates the swift and impartial measurement of collective motion. The ability to measure the trajectory of a large group of people has been made possible by the quick development of picture-capturing devices and video-tracking techniques (Delcourt et al., 2013; Ylieff & Poncin, 2003).

Scientists can learn more about the neurological and cognitive processes underlying such behaviors by examining collective behavior. The results of this research may also provide ideas for artificial systems. By accurately acquiring motion data of various organism groups without requiring laborious manual labor or pasting markers on the tracked objects, multi-object tracking using video cameras enables the discovery of new principles underlying these collective behaviors. Trajectory data is crucial for quantitatively analyzing these behaviors (Branson & Belongie, 2005; Mirat et al., 2013; Noldus et al., 2001; Qian et al., 2014).

A lot of study has been done on video surveillance systems in the last few years. The use of real-time monitoring systems requires the detection and retrieval of moving items as pre-processes (Javed & Shah, 2002; Zang & Klette, 2003). As a result, many excellent algorithms have been presented for tracking land objects. One method for capturing the full shape of tracked objects is background subtraction (Piccardi, 2004).

While most of them still relied on classical filters (Gordon, 2004) for correlation operations, fish tracking methods based on the Tracking by Detection (TBD) mechanism have applied deep learning-based methods on detectors, such as Faster convolutional neural network R-(CNN) (S. Ren et al., 2016) and you only look once (YOLO) (Bochkovskiy et al., 2020; Redmon, 2018), etc.

International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

To successfully handle camera motion, a novel tracking approach based on deformable multiple kernels (DMKs) is suggested. This algorithm combines the strengths of multiple kernel tracking and Deformable Part Models (DPM) detection. By incorporating a visual structure with kernel-based tracking, this technique offers an effective object-tracking solution that does not require training. Since monitoring objects underwater is more difficult than following people or cars on land, the algorithm works especially well for underwater cameras (Berclaz et al., 2011; Cai et al., 2013; Lee et al., 2014; Shitrit et al., 2013).

The researchers train the Siamese network to learn the Appearance Similarity (AS), and then use attention long short-term memory (LSTM) networks to record the motion similarity to monitor the location of the fish in subsequent frames. Furthermore, we use intersection-over-union (IoU) to decrease the search space surrounding the neighborhood region of the previous position where the spatial similarity score (SSS) of the subsequent bounding boxes needs to be matched. Finally, using the Hungarian formulation, joint optimization is used to determine the best tracking solution (Luong et al., 2024).

ZigBee and the generic packet radio service protocol were used by Luo et al. to build and install a real-time aquaculture monitoring system (Hongpin et al., 2015) that improved communication reliability between the sensor nodes and the central server. According to Zhang et al., fish farms that use Internet of Things (IoT) technology have much greater financial returns than those that do not (Zhang et al., 2013).

CNN (LeCun et al., 1989) exhibits an extremely low mistake rate (Parkhi et al., 2015) in its direct recognition of human faces. Additionally, a CNN can track a single item more accurately than conventional techniques in a complicated environment. The first tracker based on CNN tracking techniques documented in the literature was the deep learning tracker, which is built on a stacked denoising autoencoder network (N. Wang & Yeung, 2013).

Because deep learning performs exceptionally well in global feature extraction, it is also used in fish tracking. Fish were recognized and tracked by Wageeh et al. (Wageeh et al., 2021) using an optical flow algorithm in conjunction with YOLOV3 (Redmon, 2018) to determine the fish trajectory based on motion in each video frame. Fish detection using a segmented neural network Mask-RCNN (He et al., 2017) was developed by Arvind et al. (Arvind et al., 2019), and the detection outcomes were tracked using Generic Object Tracking Using Regression Networks (Held et al., 2016).

In aquaculture, the adoption of advanced fish tracking technologies is critical not only for optimizing production but also for addressing broader challenges related to food security. With the global population projected to reach 9.7 billion by 2050, the demand for protein sources, particularly fish, is expected to rise significantly (Action, 2020). As overfishing and environmental pressures continue to threaten wild fish stocks, aquaculture plays an increasingly vital role in meeting this demand (Naylor et al., 2021). Accurate monitoring of fish movement, behaviour, and health using tracking systems can improve feeding efficiency, reduce mortality rates, and help mitigate environmental impacts—ultimately enhancing sustainable production. For example, innovations in tracking technologies in Norway's salmon farming industry have led to a 20% increase in efficiency, directly contributing to food security (Bailey & Eggereide, 2020). Therefore, the wider application of these technologies is essential to ensuring global food security as the reliance on aquaculture intensifies.

International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

Table 1: Overview Of Various Fish Tracking Methods, Including Their Techniques, Algorithms, Performance Metrics And Limitations

Ref.	Method	Method Type	Application Area	Algorithm	Evaluation Metrics	Limitations
(Palconit et al., 2021)	Symbolic Regression vs. Gaussian Process Regression	Regression Techniques	Fish tracking	Symbolic Regression, Gaussian Process Regression	Accuracy: 74% to 100% vs. 81% to 91%	Longer computation time for Symbolic Regression
(Yuan et al., 2016)	Multiple-Fish Monitoring Algorithm (Otsu segmentation, Kalman filter)	Segmentation and Tracking	Multiple-Fish Monitoring	Otsu Adaptive Segmentation, Kalman Filter	Enhanced efficiency and reliability (general)	May not handle high-density fish schools as effectively
(Cheng et al., 2019)	3D Fish Tracking (extremum detection, ellipse fitting, Kalman filtering)	Tracking and Detection	Fish School Tracking	Extremum Detection, Ellipse Fitting, Kalman Filtering	Precision and reliability in fish school monitoring	Potential challenges with complex motion tracking
(S. H. Wang et al., 2016)	Zebrafish Tracking (Kalman filtering, rectangular chain model)	Tracking	Zebrafish Tracking	Kalman Filtering, Rectangular Chain Model	Effective occlusion handling and tracking accuracy	May struggle with very fast or erratic movements
(E. Fontaine et al., 2008)	Optical Tracking System	Optical Tracking	Fish Behavior Study	Optical Tracking Methods	Effective in studying fish behavior, posture prediction	May require calibration for different fish sizes
(Shiau et al., 2013)	Real-time Underwater Video System (bounding-surrounding boxes)	Video-Based Tracking	Aquaculture Fish Tracking	Bounding-Surrounding Boxes Method	High precision in fish tracking	Background water plants can still be challenging
(W. Li et al., 2022)	CMFTNet (Convolutional Multi-Fish Tracking Network)	Neural Network Tracking	Multi-Fish Tracking	Convolutional Neural Network, Anchor-Free Method	MOTA: 65.5%, IDF1: 27.4%	Moderate IDF1 score compared to some methods
(Chuang et al., 2016)	Flexible Multiple Kernels and Mean-Shift Tracking	Kernel and Mean-Shift	Underwater Camera Data Sets	Multiple Kernels, Mean-Shift	Efficient and cost-effective tracking	May not handle occlusions as well

International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

(Gupta et al., 2021)	Deep Fish Track Network (DFTNet)	Deep Learning Tracking	Fish Tracking	LSTM, Siamese Networks	60.9% reduction in identification switches	Can be complex to train and tune
(Pérez-Escudero et al., 2014)	IdTracker	Tracking and Re-Identification	General Animal Tracking	Unique Fingerprints Tracking	Effective across various species, handles occlusions and size matches	May require extensive training data
(C. Wang et al., 2023)	Global Association and Multi-View Data Fusion (3D zebrafish)	3D Tracking and Data Fusion	3D Fish Tracking	Global Association, Multi-View Fusion	MOTA: 67.9%, IDF1: 64.3%	May require extensive computational resources
(Bai et al., 2018)	Enhanced HOG for Zebrafish Tracking	Feature-Based Tracking	Zebrafish Behavior Analysis	Enhanced Histogram of Oriented Gradients (HOG)	Superior accuracy and efficiency, fewer samples needed	Requires high-quality input images
(Gao et al., 2019)	IoT-based Smart Aquatic Farming System	IoT and Data Management	Aquatic Farming	IoT-Based Forecasting, QR Codes	Low error rates in water quality prediction	Limited to water quality management and fish tracking
(S. Liu et al., 2024)	FishMOT	Object Tracking	General Fish Tracking	Object Identification, Intersection over Union (IoU)	Efficient accuracy and performance minimize computational complexity	May not perform as well with highly overlapping fish
(Y. Liu et al., 2024)	FishTrack (Pyramid Vision Transformer)	Vision Transformer-Based	Multi-Fish Tracking	Pyramid Vision Transformer, Spatiotemporal Fusion	IDF1: 82.5%, MOTA: 94.8%	Complex models may require high computational resources
(Shreesha et al., 2023)	Sillago Sihama-Vid Dataset	Dataset and Behavior Modeling	Fish Behavior and Tracking	Appearance, Location, Swim Direction Modeling	Improved MOTA, MOTP, IDSW, MT, and continuous behavior modeling	May need more diverse training data for generalization
(Xu & Cheng, 2017)	CNNs with Data Augmentation and Iterative Training	CNN and Augmentation	Zebrafish Behavioral Analysis	Convolutional Neural Networks (CNNs), Data Augmentation	Highly precise tracking, improved behavior analysis	Requires extensive training and computational resources
(Qian et al., 2016)	Multi-Fish Tracking (Head detection,	Head Detection and Tracking	Multi-Fish Orientation Tracking	Head Detection, Grayscale Features, Cost Function	High accuracy in position and orientation tracking	Limited by grayscale features for complex backgrounds

International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

	grayscale features)					
(Chuang et al., 2014)	Low-Contrast Stereo Films Tracking (Histogram back-projection, Viterbi approach)	Stereo and Back-Projection	Low-Contrast Fish Tracking	Histogram Back-Projection, Modified Viterbi Approach	Effective in low-contrast conditions, accurate fish length measurement	Performance may degrade in very noisy environments
(Saleh et al., 2022b)	Three-Phase Approach (Optical models, self-supervised learning)	Self-Supervised Learning	Fish Tracking and Segmentation	Optical Models, Self-Supervised Learning	Effective tracking and segmentation highlight improvement areas	Requires significant data for pseudo label refinement
(H. Wang et al., 2022)	Coupled Neural Network (Aberrant Porphyry Seabream Behavior)	Neural Network -Based	Fish Behavior Analysis	YOLOv5s, Siamese Region Proposal Network++ (SiamRPN++)	AP50: 99.4%, AP50: 76.7%	May not generalize well to other fish species or behaviors
(Barreiros et al., 2021)	YOLOv2 with Kalman Filter	Object Detection and Tracking	Fast-Moving Fish Tracking	YOLOv2, Kalman Filter	Effective tracking of fast-moving zebrafish groups	May have limitations with very fast movements

Fish tracking techniques are outlined in Table 1, with emphasis on the kind, application, and important algorithms of each method. Methods include neural networks such as CMFTNet and DFTNet for improved multi-fish tracking, symbolic regression and Gaussian process regression for accurate fish tracking. Optical tracking, video-based systems, and IoT-based solutions for behavioral analysis and real-time monitoring are some of the techniques. Important algorithms are used, including YOLOv2, enhanced HOG, and Kalman filtering. Each has advantages and disadvantages, including computing complexity or performance under difficult circumstances. All things considered, the table presents a variety of methods designed to provide precise tracking of fish and behavioral analysis in various settings.

Materials and Methods

The fish tracking methods discussed employ a range of techniques and materials tailored to different tracking challenges and environments. These methods include symbolic regression and Gaussian process regression, which focus on accuracy but differ in computation time, with symbolic regression generally requiring more processing power. Segmentation and tracking algorithms, combined with Kalman filtering, are used for efficient multi-fish monitoring, while 3D tracking methods utilize extremum detection and ellipse fitting to handle complex motion. Video-based systems, like those using bounding-surrounding boxes, and IoT-based smart systems manage real-time tracking and environmental monitoring. Advanced techniques offer high precision in challenging conditions, while datasets support behavior modeling. The diverse methods and materials, from optical models to neural networks, address various tracking needs, optimizing performance across different applications and scenarios. Fig. 1

International Journal of Innovation and Industrial Revolution
EISSN: 2637-0972
IJIREV
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003
illustrates the research overview according to the main key points of this research in detail. Fig. 2 illustrates the PRISMA methodology applied to the review of fish tracking studies, detailing the steps of literature search, screening, and categorization. It highlights the refinement process and analysis used to filter and assess relevant research.

3D Stereo Vision Tracking

To detect fish areas in the tank three fish captured at 4 frames per second using stereo cameras were converted to hue, saturation, value, and binarized. Fish centroids were computed, and these centroids were matched between photos using K-Nearest Neighbours to estimate depth. Models predicting centroid locations in the future were trained and validated using a 40-frame stereo video dataset (Palconit et al., 2021). Effective methods for estimating and tracking fish motions included linear regression, the Adaptive Neuro-Fuzzy Inference System (ANFIS) (Karaboga & Kaya, 2019) with clustering, GPR with Gaussian noise (P. Li & Chen, 2016), and symbolic regression using Multi-Granularity Genetic Programming (MGGP) (Palconit et al., 2020).

Zebrafish Tracking using CNN

CNN uses two fully connected layers, max-pooling, dropout to prevent overfitting, four convolutional layers, and 65x65 pixel images (Xu & Cheng, 2017). It was implemented in the matrix covariance network (Vedaldi & Lenc, 2015) and trained on GTX 980Ti for 30 epochs. To increase performance and enable precise classification and ID assignment for trajectory segments, the training set was enlarged. Complete trajectories were formed by consecutively connecting linked segments with matching IDs within 1,000 frames and 2,000 pixels after short segments with multiple IDs were filtered.

International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

Figure 1: Research Overview With All The Key Concepts Of Fish Tracking

Automated Planar Tracking

There are three phases to the tracking system (S. H. Wang et al., 2016):

Finding and Estimating Fish Pose: A technique that separates fish from the background based on pixel intensity and curvature at the head and tail is used to detect fish heads and estimate their pose. An average of 18,000 frames are used for background subtraction and binarization (Annadurai, 2007). To accurately distinguish the head and tail of fish, fish boundaries are determined and smoothed to calculate curvature. The perpendicular bisector between borders points determines the direction of the fish head, and the best- fitting rectangle around the fish is selected for pose estimation, omitting instances where coverage is less than 80% (E. I. Fontaine, 2008).

Fish Tracking: The head and body of the fish are tracked first by the system. Head movement is predicted by the Kalman filter (Hargrave, 1989), and the Kuhn-Munkres method with normalized cross-correlation and orientation metrics is used for data association. Random

International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

angles are used to fit body rectangles, and monitoring is stopped if coverage falls below 80%. Following that, split trajectories are relinked (Kuhn, 2004; Y. Liu et al., 2012).

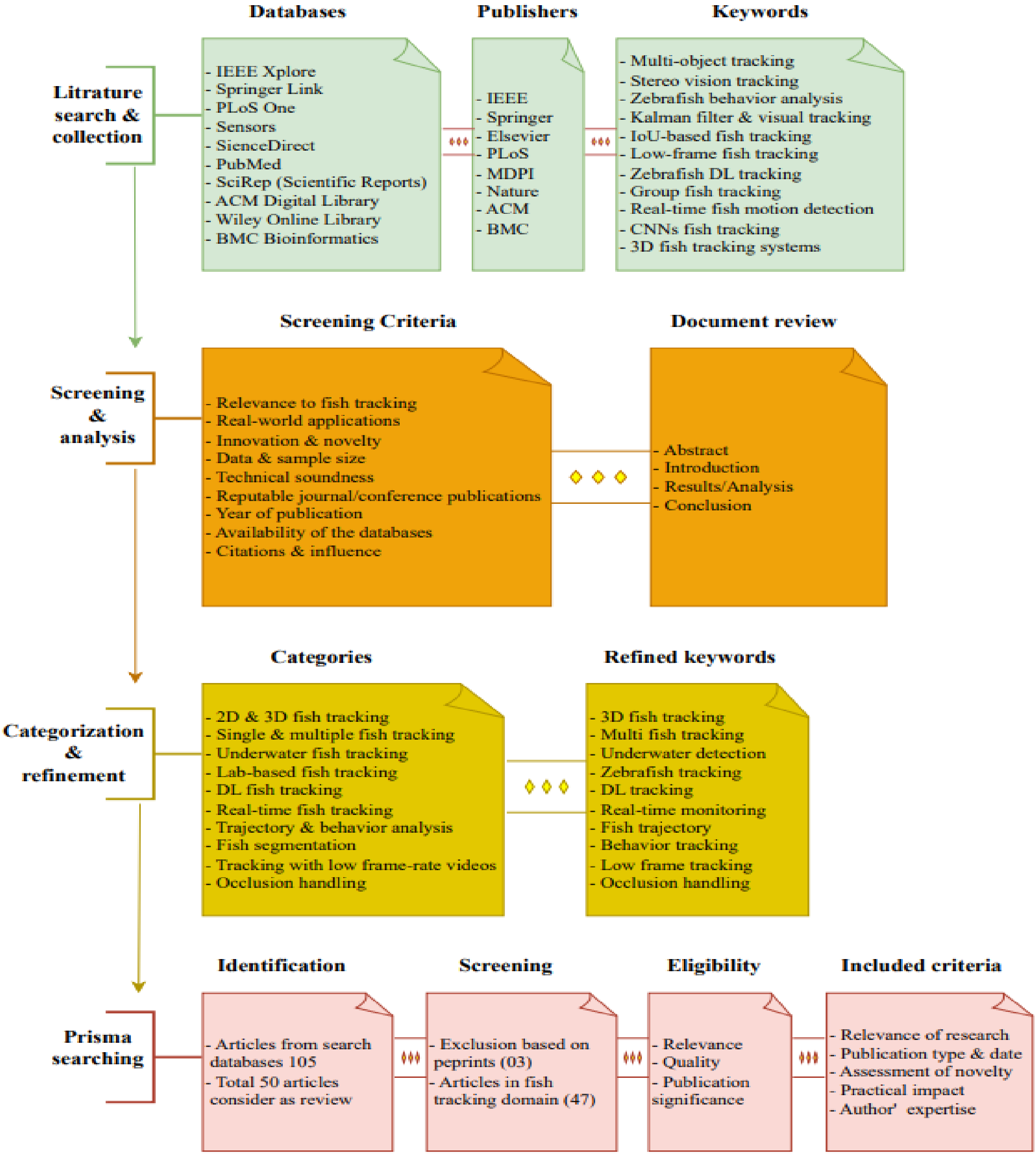

Figure 2: PRISMA Flowchart Showing Literature Search, Screening, Categorization And Analysis Process For Fish Tracking Studies

Tracklets Relinking: A minimum cost maximum flow (MCMF) technique is used to relink fragmented trajectories. This problem is represented with nodes indicating tracklets and capacities and costs determined by orientation and distance. The generalized Ford-Fulkerson method (Edmonds & Karp, 2003) is used to solve the MCMF issue, providing optimal relinking and removing unstable body rectangle information to minimize mistakes.

Using the Bounding-Surrounding Boxes Method for Fish Tracking

By distinguishing swimming fish from drifting water plants, the bounding-surrounding boxes (BSB) technique improves fish tracking and can be integrated with algorithms such as the

Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

Gaussian mixture model (GMM) (Stauffer & Grimson, 1999). To make dynamic adjustments and differentiate moving objects from the backdrop, the GMM uses a combination of Gaussian distributions to update the background model frame by frame. Morphological techniques are used to fine-tune foreground pixels after background subtraction and object identification is accomplished through segmentation. To match items between frames, the system makes use of color information and Pearson's correlation (Lee Rodgers & Nicewander, 1988).

Each foreground object is surrounded by a bounding box and a surrounding box that is T times larger and centered at the same location when using the BSB approach. Over time, objects that stay inside the surrounding box (such as water plants) are categorized as non-fish and are eliminated, whereas fish are detected and tracked when they go outside the box (Shiau et al., 2013).

CMFTNet

ResNet-101 with deformable convolutions is used by the CMFTNet multi-fish tracking network to manage the non-rigid nature of fish, enhancing geometric adaptation and identification accuracy. It addresses problems such as target missing and multi-detection, integrating CenterNet for fish detection with an emphasis on center points, bounding boxes, and size estimates. The three heads of the detection branch are box size, offset, and heatmap. The offset head refines positions with L1 loss, the box size head estimates dimensions with L1 loss, and the heatmap head employs Gaussian functions for center positions with logistic regression loss. The Re-ID branch uses a cross-entropy loss function to classify fish across frames. Through uncertainty loss, the network is trained with a combined loss function that balances detection and Re-ID losses (W. Li et al., 2022).

Fish Tracking Based on DMK

Through the combination of DPM detection and kernel-based tracking, the DMK tracking algorithm enhances object positioning (Chu et al., 2013). To improve tracking accuracy, it makes use of texture histograms, weighted color, and HOG characteristics. The process begins with n + 1 kernels, where a DPM component corresponds to the aspect ratio of the item, and the root kernel is assigned to the object's bounding box. HOG features, texture histograms, and color histograms are used to scale and position each component kernel (Comaniciu et al., 2003). To improve item placement from coarse to fine levels, the tracking uses rotation-invariant LBP texture histograms and iterates (Ojala et al., 2002).

Part kernels are aligned with the root kernel by adjusting them based on deformation costs and HOG similarity, following color and texture tracking. A weighted average of kernel centers determines the final bounding box, enhancing feature discrimination and resilience against occlusions. To manage abrupt scale variations, the algorithm additionally modifies the target scale in response to modifications in kernel bandwidth (Chuang et al., 2016).

YOLOv2 and Kalman Filter

The tracking system functions in three steps and was tested with fish in a glass tank that was sprayed with sand and illuminated by a full HD camera (Barreiros et al., 2021).

Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003
First Training: To automate detection, YOLOv2 is trained on manually labelled fish head areas. Detection and tracking: YOLOv2 address duplications by grid intersection, computes centroids and ellipses, and employs a Kalman filter for trajectory estimation. It also recognizes fish heads in new frames.

Trajectory reconstruction: By connecting fragmented trajectories with a cost function, fish behaviors such as agitation and speed may be measured.

YOLOv2, a fast and precise image division tool, uses bounding boxes with confidence scores, convolution, batch normalization, rectified linear units, and anchor boxes for detection (Z. Liu et al., 2018; Redmon & Farhadi, 2017). Centroids are calculated from greyscale images, fish heads are demarcated, and ellipses are altered to indicate the direction of movement. By re-connecting interrupted trajectories based on proximity and temporal order, the Kalman filter updates the fish's state vector based on uniform linear motion (Cheng et al., 2019; Yuan et al., 2016).

DFTNet
The tracking system models affinity measures for fish movements utilizing the TBD paradigm with ground-truth detections, emphasizing three important areas (Gupta et al., 2021):

Appearance Similarity: Using 66,000 training examples, a Siamese network analyses fish's appearance in successive frames to differentiate between species and produce matching scores even in the presence of varied backgrounds (Y. Li et al., 2017).

Motion Similarity: By concentrating on bounding boxes with notable motion changes, the attention LSTM forecasts fish movements. It uses overlap measurements to predict future locations after being trained on 62,094 trajectories (Graves & Schmidhuber, 2005).

Geographical Similarity: If the IoU metric surpasses 0.6, the bounding box overlaps across frames is assessed, and the same ID is assigned (Gupta et al., 2021).

Joint Optimization: The last track assignment uses the Hungarian Algorithm to optimize a hyperparameter λ that is based on appearance, motion, and spatial similarity scores. For precise tracking, a λ value of 0.2 successfully balances these variables (Gupta et al., 2021).

Improved YOLOV5 and SiamRPN++
The technique combines SiamRPN++ for tracking with an improved YOLOv5 network for aberrant fish detection. To increase detection accuracy, YOLOv5 uses focus, bottleneck cross-stage partial, spatial pyramid pooling, and PANet to partition images into grids for bounding boxes and confidence predictions. There are four variations of YOLOv5, the quickest and least accurate of which is YOLOv5. To improve multi-level feature utilization, SiamRPN++ uses a fully convolutional Siamese network with ResNet-50 for feature extraction and Layer-wise Aggregation. To lower parameters and increase precision, it additionally uses spatial awareness sampling and depth cross-correlation. For stable convergence, cross-entropy loss is used during training. To successfully handle multi-target settings, the approach combines detection and tracking results iteratively (B. Li et al., 2019; H. Wang et al., 2022).

Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

FishMOT

FishMOT uses YOLOv7 to detect fish and IoU values to manage missed detections and associate IDs. Trajectories are displayed using uniform colors and random IDs. To enable precise target association, the Kalman filter forecasts bounding box locations and high IoU values (~0.6) between neighbouring frames. A module that computes IoU and the Hungarian algorithm for optimal matching is used to further optimize this.

While the refined module takes care of detection omissions, the interaction module handles occlusions and morphological changes by merging IoU values and segmenting fish entities. To recover lost trajectories within a predetermined number of frames, linear interpolation is used in conjunction with provisional data within a buffer region (S. Liu et al., 2024).

Unsupervised Fish Trajectory Tracking

The framework uses a multi-step procedure to improve object tracking. A self-supervised model is first used to generate and refine pseudo labels, which are then utilized to train a segmentation network for higher accuracy compared to conventional techniques (Saleh et al., 2022a). Adaptive Gaussian thresholding is used in background subtraction to remove shadows and stationary elements from a median background image (Golilarz et al., 2019).

Recurrent all-pairs field transforms are used to compute optical flow, which corrects boundary imperfections and refines segmentation masks (Teed & Deng, 2020). With the help of conditional random fields (Krähenbühl & Koltun, 2011) and DeepLabv3 (Chen et al., 2017), additional refining is possible. To stabilize the pseudo-label quality, a historical moving average is used (Nguyen et al., 2019).

SoloV2 completes the segmentation process using matrix non-maximum suppression and dynamic instance segmentation for high-frame-rate processing. OpenCV is used to transform instance masks into revolving 2D bounding boxes (X. Wang et al., 2020). Motion estimation and data association are handled using the simple online and real-time tracking (SORT) (Bewley et al., 2016) framework, which updates target states based on bounding box positions and incorporates a Kalman filter and the Hungarian approach (Saleh et al., 2022b).

Tracking based on improved HOG

Through a series of phases, the tracking algorithm improves the precision of the zebrafish trajectory. Preprocessing is the first step, which involves elliptic fitting and background subtraction to stabilize and adjust the region of interest. Heuristics are used to generate initial tracklets, eliminating short tracklets to increase precision and manage crossings (Bai et al., 2018).

Enhanced HOG feature extraction classifies, binaries, and scales zebrafish regions of interest to further stabilize tracking. Two classifiers are used by the tracking module: one for tracklet extension and another for calculating the final trajectory. Tracklet matching is optimized by the Hungarian algorithm, and accuracy is ensured by manual error correction. A thorough explanation of the assessment metrics used to evaluate the outcomes is given in Table 2.

International Journal of Innovation and Industrial Revolution
EISSN: 2637-0972
IJIREV
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

Table 2: Description Of The Evaluation Metrics

Metric	Description
Precision	The sum of correctly tracked objects in all frames/total ground truth objects in all frames. Larger values are better.
Recall	The sum of correctly tracked objects in all frames/total tracked objects in all frames. Larger values are better.
FM, F1-measure	The harmonic means of precision and recall. Larger values are better.
Mostly Tracked Trajectories (MT)	Percentage of trajectories which are correctly tracked for more than 80% of their length. Larger values are better
Mostly Lost Trajectories (ML)	Percentage of trajectories that are correctly tracked less than 20% of their length. Smaller values are better.
Fragments (Frag)	Percentage of trajectories that are correctly tracked less than 80% but more than 20% of their length
ID Switch (IDS)	The frequency of identity switches after occlusion
Accuracy	The metric measures overall correctness by calculating the proportion of correctly identified instances out of the total.
Root Mean Square Error (RMSE)	RMSE measures the average magnitude of prediction errors by taking the square root of the average squared differences between predicted and observed values.
Miss Ratio	The Miss Ratio metric measures the proportion of actual positive instances that are not correctly identified by a detection or tracking system.
Error Ratio	The Error Ratio measures the proportion of incorrect predictions relative to the total number of predictions.
Correct Similarity Rate (CSR)	The CSR measures the proportion of correct predictions that match the ground truth.
Multiple Object Tracking Accuracy (MOTA)	The metric measures tracking system performance by combining false positives, false negatives, and identity switches to assess how well it maintains object identities over time.
Identification F1 Score (IDF1)	The metric evaluates the accuracy of object tracking by measuring the balance between precision and recall for correctly identified objects.
Identification Precision (IDP)	The IDP metric measures the proportion of true positive detections among all detected objects.
ID Recall (IDR)	The IDR metric measures the proportion of true positives among all actual objects that should have been detected.
Similarity Index (SI)	The SI measures the number of correctly detected objects, penalizing undetected objects, whether due to occlusion or detection error
Occlusion Ratio (OR)	The ratio of the total number of occlusions and the total number of targets
Occlusion Detection Ratio (ODR)	The ODR metric measures the proportion of occlusions detected by a tracking system relative to the total number of actual occlusions
Ground Truth (GT)	The number of trajectories in ground truth
Predicted Tracking (PT)	Partially tracked trajectories
Multiple Object Tracking Precision (MOTP)	It evaluates the intersection area over the union area of bounding boxes.
Average Precision (AP)	It refers to the area under the Precision recall (PR) curve
Frames Per Second on Graphics Processing Unit (GPU) (FPS_{GPU})	It is the number of frames per second transmitted by the algorithm under GPU conditions.

International Journal of Innovation and Industrial Revolution
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

Area Under the Curve (AUC)	The AUC metric measures classification performance by calculating the area under the Receiver Operating Characteristic curve.
Correct Tracking Ratio (CTR)	The metric measures the proportion of frames where objects are correctly tracked compared to the total number of frames
Average Interruption Times (AIT)	It measures the average number of times an object's tracking is interrupted or lost over a sequence of frames
Correct Identification Ratio (CIR)	It measures the proportion of correctly identified objects compared to the total number of objects that should have been identified
AR (Average Recall)	The metric measures the average recall of a tracking or detection system across different thresholds or conditions.

Results and Discussion

3D Stereo Vision Tracking

In fish tagging and tracking, regression-based algorithms performed better than ANFIS predictors; MGGP obtained the greatest F1 scores, while GPR produced the closest predictions and the lowest RMSE. For real-time applications, GPR is favored due to its accuracy and faster computing time for higher fish densities. Even with its superior precision, MGGP required 96% more computing time than GPR. Fish interactions affected tracking performance, including acceleration and speed. With a 100% F1 score for fish with high speed and low dispersion but a poor score for fish with high acceleration, MGGP demonstrated a strong association between fish speed and acceleration and F1 scores. Fish movements and tagging scores exhibited lesser connections in other models (Palconit et al., 2021). Fig. 3 shows the 3D stereo vision fish tracking with different prediction algorithms.

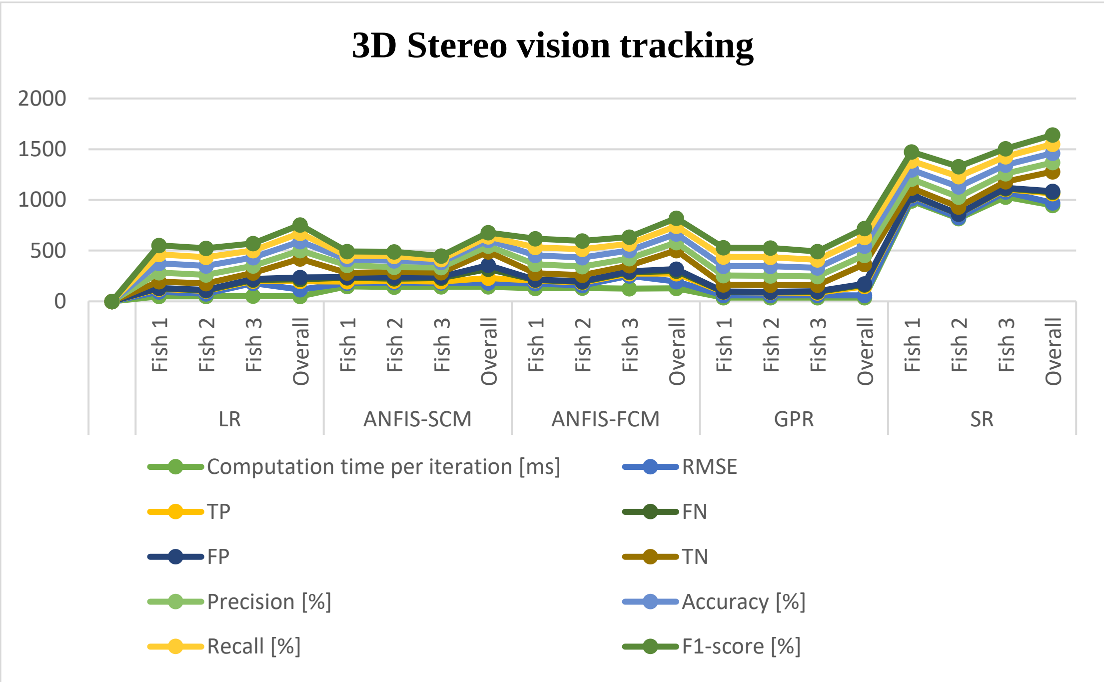

Figure 3: Results Of Fish Tagging And Tracking Using The Different Prediction Algorithms

(Palconit et al., 2021)

International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

Zebrafish Tracking Using CNN

Five video recordings were used to benchmark the tracking system (Xu & Cheng, 2017):
D1: 14 zebrafish, 2,048 x 2,040 resolution, 2,000 frames at 60 frames per second.
D2: 25, zebrafish, 2,048 x 2,040 resolution, 5,000 frames, 50 frames per second.
D3: 15,000 frames at 30 frames per second, with a resolution of 1,528 x 1,080.
D4: 11 zebrafish, 2,048 × 2,040 resolution, 4,000 frames, 50 frames per second (http://www.idtracker.es/).

D5: 25 zebrafish, divided into 10,000 frames on day 1 and 11,000 frames on day 3, with a resolution of 2,048 x 2,040.

In terms of monitoring and recognizing fish over several days, the CNN-based tracking system performed exceptionally well in all parameters. The stable head area is used for identification in this method, as opposed to 'idTracker,' (Sridhar et al., 2019), which uses fingerprinting. Fig. 4 illustrates the evaluation of two selected datasets (D1 and D5) on CNN.

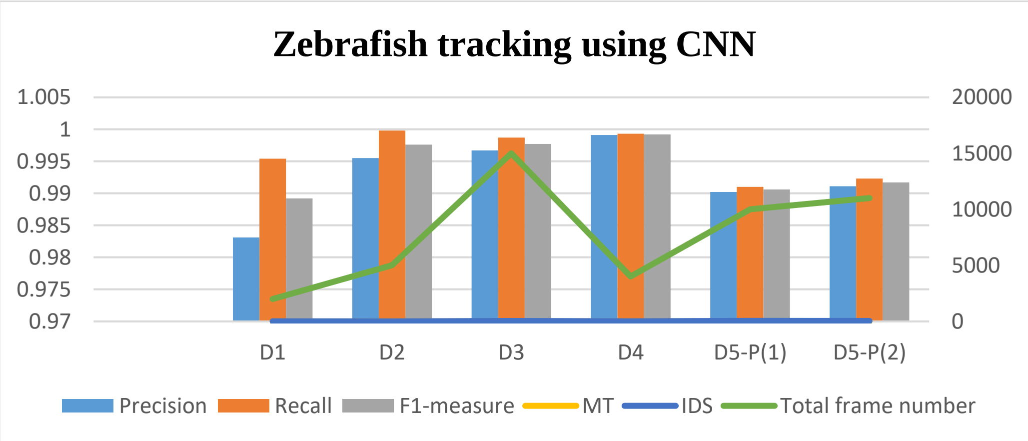

Figure 4: Evaluation Of Tracking Performance On D1 To D5

(Xu & Cheng, 2017)

Automated Planar Tracking

Results of tracking and assessments of performance are shown for two video datasets (D1 and D2) containing schools of 10 and 20 fish, respectively (S. H. Wang et al., 2016).

Performance of detection: Using 300 manually annotated frames from the DS1 and DS2 movies, the tracking system was assessed with an emphasis on nose point accuracy, body rectangle coverage, and proper body fitting. Higher fish density resulted in a considerable rise in occlusions, which raised the miss and error ratios, according to the performance metrics. Occlusions have a significant effect on detecting performance; they are responsible for almost all fitting failures. While body fitting accuracy is great when the system is not obstructed, occlusions during tracking must be addressed to increase overall detection reliability. Evaluation performance of fish body fitting detected in Fig. 5.

Performance of tracking: Two approaches were compared with the proposed tracking system (S. H. Wang et al., 2016): one that did not use body fitting, and the other that made use of the idTracker system (Pérez-Escudero et al., 2014), which performs well in extreme occlusion but may have identification problems when there is a large item density. CTR, running time, AIT,

International Journal of Innovation and Industrial Revolution
EISSN: 2637-0972
IJIREV
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

and CIR were among the evaluation measures that were used to make sure that tracking was done correctly and to reduce errors.

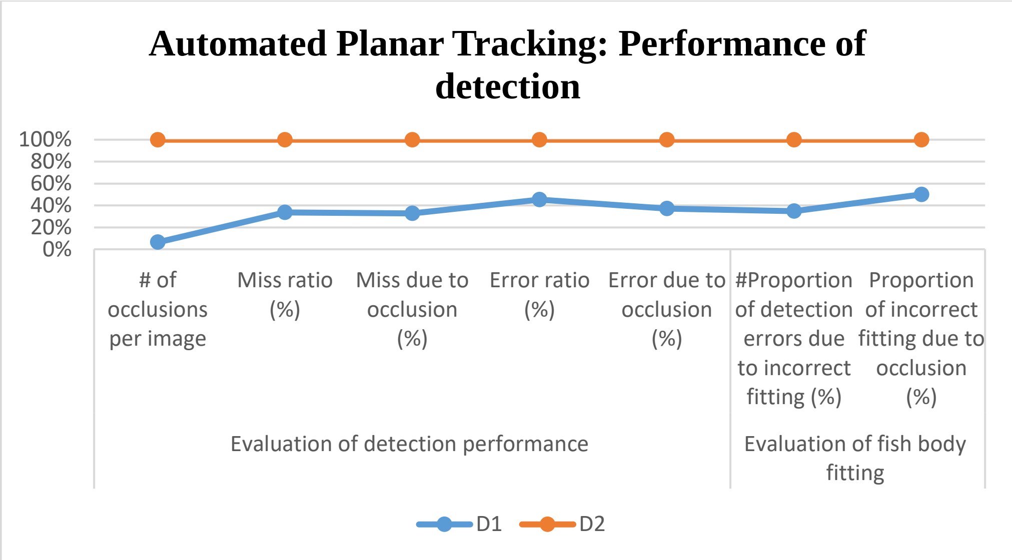

Figure 5: Evaluation Of Detection Performance And Fish Body Fitting

(S. H. Wang et al., 2016)

With over 95% of the running time going into fish body fitting, the suggested system requires substantially more time than idTracker due to its use of full-resolution images for precise body fitting. However, compared to idTracker, the suggested approach without body fitting is faster, suggesting that body fitting is a significant time-consuming aspect. Running time of the idTracker shows in Fig. 6.

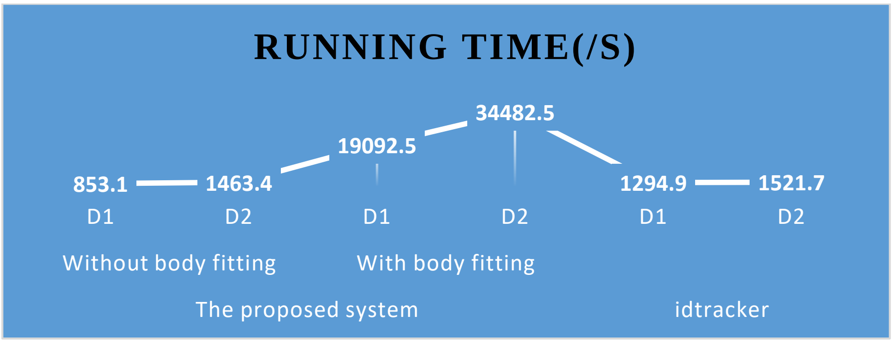

Figure 6: Running Time Of The System And idTracker

(S. H. Wang et al., 2016)

Because the system relies on head detection, which works even in occlusions, it performs better than idTracker, especially in situations with a high fish density. Due to unresolved head-tracking issues, the suggested system's CTR marginally lowers in the absence of body fitting. Tracklets relinking yield over 99% accuracy in body shape fitting and greatly improves tracking effectiveness. Tracklets relinking greatly enhances trajectory continuity, surpassing both idTracker and the system without body fitting, and lowering AIT by over 90%, particularly at denser fish populations. Compared to idTracker, the suggested approach exhibits greater trajectory disruption at higher occlusion frequencies while still maintaining superior

Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

continuity. Fig. 7 represents the CTR and AIT of the idTracker system. When fish are in low-density groups, idTracker performs exceptionally well at accurately identifying each fish, but identification errors cause problems when fish densities increase. Since body rectangles are not used in tracklets relinking, body fitting makes a minimal contribution to CIR.

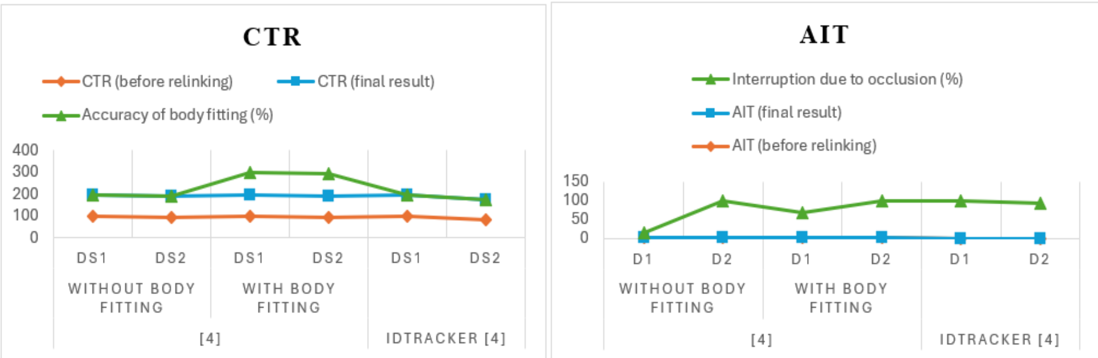

Figure 7: CTR and AIT of The System And idTracker

(S. H. Wang et al., 2016)

Detection performance is critical to tracking accuracy, particularly at larger fish numbers when occlusion is common. The suggested system's efficacy decreases with increasing fish density, even though it can fix some tracking-related detection errors. Because body fitting adds more complexity, the system without body fitting performs somewhat better in correcting detection mistakes than the one with body fitting. Fig. 8 represents the CIR measures with outcomes for the correct fish-tracking system.

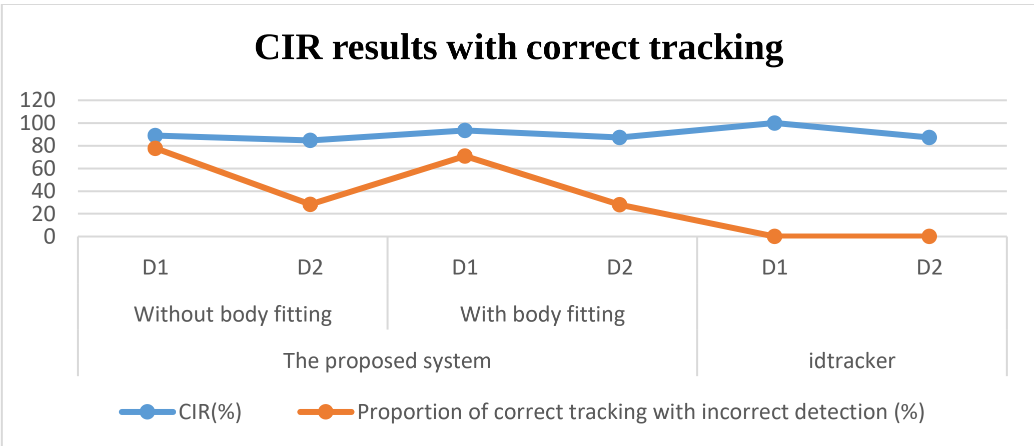

Figure 8: CIR of the System and idTracker While The Proportion Of Correct Tracking With Incorrect Detection

(S. H. Wang et al., 2016)

Using the Bounding-Surrounding Boxes Method for Fish Tracking

In complicated underwater settings, the suggested system (Shiau et al., 2013) processes 9–10 frames per second to accurately classify moving fish as foreground items and float water plants as background objects. Using surrounding boxes four times the size of the bounding box and observed times of one and two seconds, the CSR is computed. The precision with which foreground fish and background water plants may be distinguished is measured by CSR. Fig. 9 presents the empirical results, which suggest that the CSR can approach 90% along the running time of the CSR.

International Journal of Innovation and Industrial Revolution
EISSN: 2637-0972
IJIREV
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

FN: the tracking object's fish number.
PN: the number of tracking objects that are not fish.
FF: the quantity of fish that were appropriately identified as foreground items.
PG: the quantity of non-fish that were appropriately identified as background items.

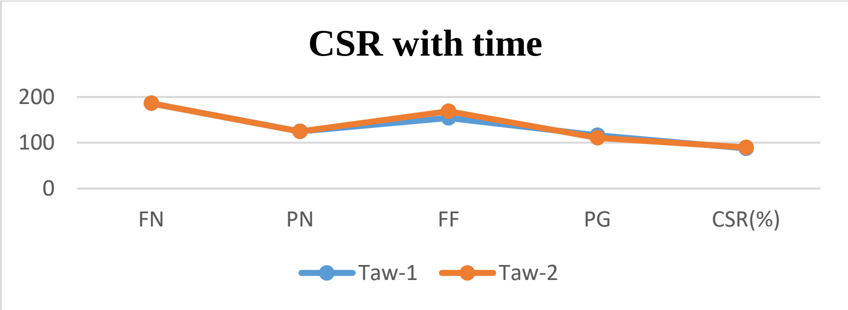

Figure 9: The CSR With Different Tracking Time

(Shiau et al., 2013)

CMFTNet

Using the OptMFT for training and testing under complicated settings, the CMFTNet model is contrasted with the traditional joint detection and embedding (JDE) and separate detection and embedding (SDE) paradigms. Benchmarks from the MOT Challenge and CLEAR Metrics are used to assess model performance. CMFTNet's higher accuracy and target retention in fish tracking is demonstrated through comparison with traditional tracking algorithms such as CenterTrack (Zhou et al., 2020) and FairMOT (Zhang et al., 2021). With a 38.6% assessment score, the IDF1 metric is used to measure performance. ID accuracy during frequent target swaps is given special attention, while IDP performance is 45.1% and IDR 33.6% on high ratios.

Using the OptMFT light dataset, testing traditional SDE models (SORT, DeepSORT, and Tracktor) with Faster RCNN backbones reveals that Faster RCNN-50-FPN outperforms Faster RCNN-101-FPN in terms of IDF1 scores. With the greatest IDF1 score of 38.6% and MOTA of 71.4%, CMFTNet beats JDE models and demonstrates better tracking accuracy and ID maintenance. On the other hand, FairMOT has great re-identification performance, while Center Track exhibits poor performance in keeping ID information despite high detection accuracy. On the OptMFT datasets, YOLO v5 and DeepSORT exhibit lower re-identification performance, with MOTA of 56.1% and IDF1 of 21.6% (W. Li et al., 2022). Fig. 10 shows the comparison of the existing systems with the measurements.

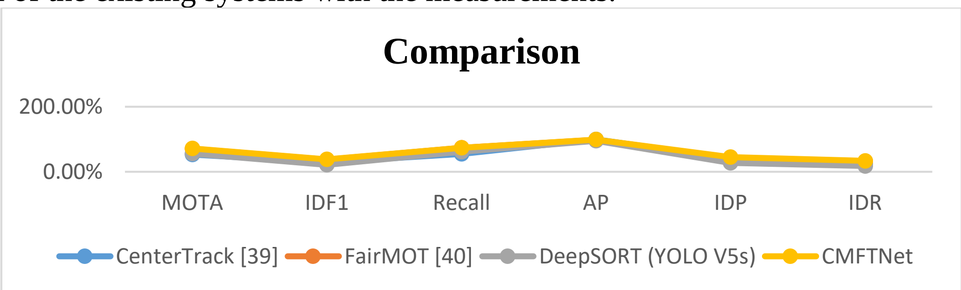

Figure 10: The Comparison with State-of-the-Art trackers and Applications In Challenging Scenarios

(W. Li et al., 2022)

International Journal of Innovation and Industrial Revolution
EISSN: 2637-0972
IJIREV
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

MFTNet handles occlusions and adhesion well, allowing it to follow spotted knifejaws effectively under a variety of situations. However, full occlusion may cause problems for long-term tracking. CMFTNet tracks fish consistently in a variety of conditions, even when faced with obstacles such as sudden turns and occlusions. Fig. 11 shows the results of the sequential masking with FL and PF.

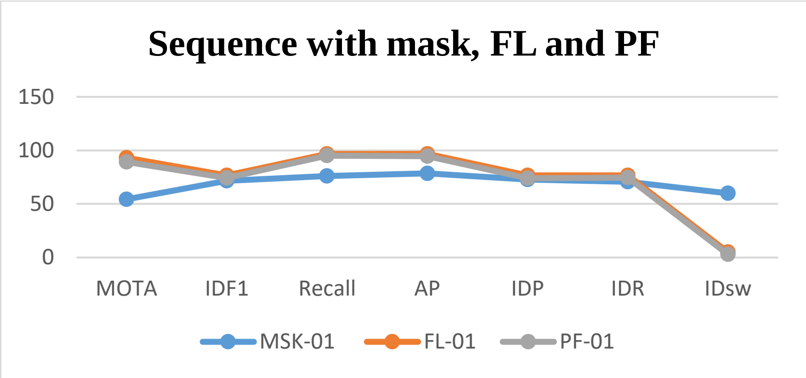

Figure 11: The Results of CMFTNet on the Mask Scoring R-CNN (MSK) Sequence along FL and PF

(W. Li et al., 2022)

Fish Tracking Based on DMK

The suggested DMK tracking (Chuang et al., 2016) technique makes use of DPM (Bochinski et al., 2018) for object identification and tracks numerous objects over video frames by utilizing color, texture, and HOG properties in conjunction with a multi-kernel technique. Trajectories are smoothed with a Kalman filter. Tested on extensive NOAA underwater video datasets, the technique uses moving cameras to track live fish in a variety of marine habitats while overcoming obstacles like changing viewpoints and considerable body deformation during swimming.

The DMK approach improves upon the previous motion similarity (MS) (Comaniciu et al., 2003) and combined multi-kernel (CMK) (Chu et al., 2013) methods in terms of accuracy and handling of occlusion by integrating texture information and deformation cost functions. These improvements enable precise and efficient tracking at reduced computing costs. The comparison of the DMK method with other fish-tracking methods is shown in Fig. 12.

International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

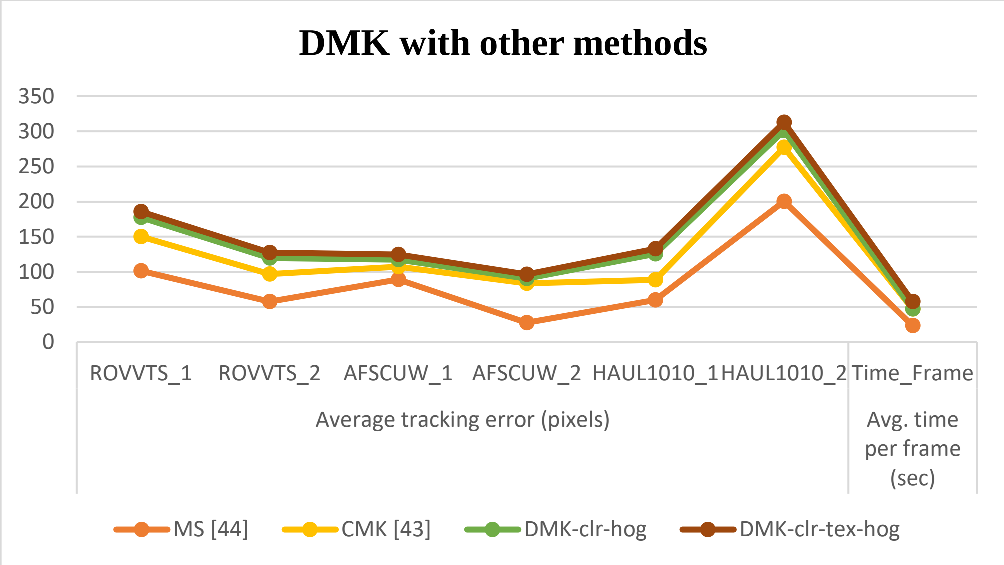

Figure 12: Comparison of DMK Method With Other Traditional Tracking Methods

(Chuang et al., 2016)

YOLOv2 and Kalman filter

The suggested tracking (Barreiros et al., 2021) technique obtained up to 0.99 accuracy and F-measure with YOLOv2-based detection when evaluated on eight video sequences with different fish counts and resolutions (Romero-Ferrero et al., 2019). Even in high-resolution datasets with up to 100 fish, it exhibited above 0.999 precision and handled both slow and fast motions with negligible detection loss. In videos with fewer fish, the system was able to obtain up to 100% CTR; however, in videos with more fish or quick motions, there was some tracking loss. All things considered, it did well in high-resolution datasets, keeping high CTR and CIR despite changing circumstances and occlusions. Fig. 13 represents the evaluation of the detection and fish-tracking method along with performance measures.

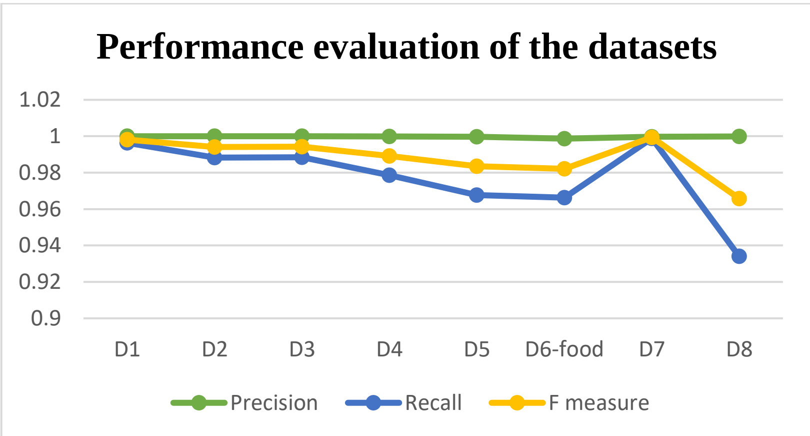

Figure 13: Performance Evaluation Of Proposed Detection And Tracking Method

(Barreiros et al., 2021)

DFTNet

Due to their comparable appearances, ablation experiments revealed that utilizing simply AS resulted in many ID shifts among fish. By using motion models—particularly Attention-

International Journal of Innovation and Industrial Revolution
EISSN: 2637-0972
IJIREV
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

LSTM—ID switches were lowered from 20,064 to 5,536. By combining AS, MS, and SSS with an ideal weight of $\lambda=0.2$, ID switches were further reduced by 95.16% (Gupta et al., 2021). Compared to visual IoU Tracker (Bochinski et al., 2018), IoU Tracker (Bochinski et al., 2017), DeepSORT (Wojke et al., 2017), Markov Decision Process (MDP) Tracker (Xiang et al., 2015), and ID switches were reduced by 37.9%, 39.1%, and 87.7%, respectively, compared to the DFTNet tracker, which combines appearance, motion, and IoU information. Fig. 14 is about the quantitative results of the traditional tracking methods.

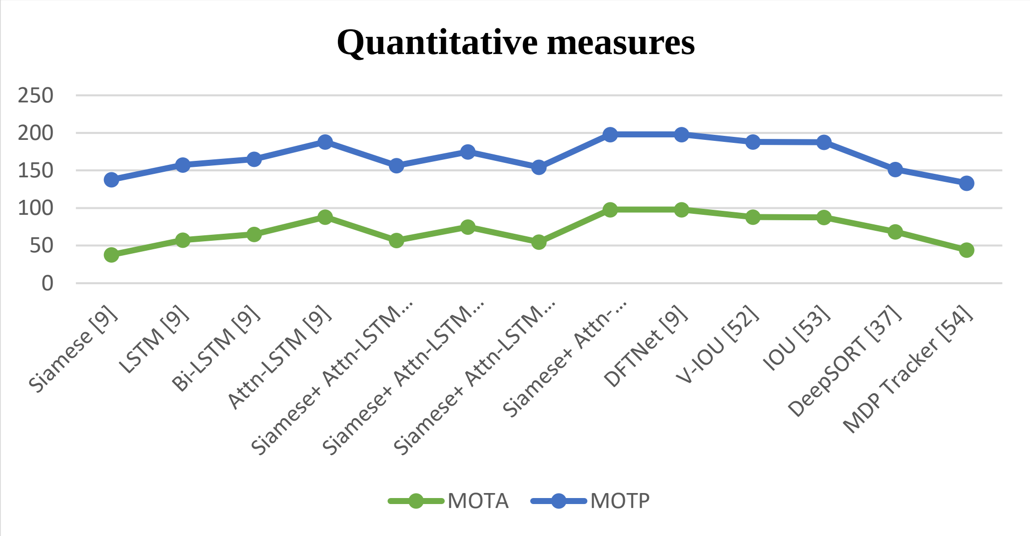

Figure 14: Quantitative Result Of Proposed And Traditional Methods

(Gupta et al., 2021)

Improved YOLOV5 and SiamRPN++

YOLOV5s-add outperforms Faster R-CNN, YOLOV3, YOLOV4, and YOLOV5s, boosting AP 50:95 from 70.6% to 79.4%. It was trained with 564 images and 100 iterations. Because of the multi-level feature fusion and mapping, it demonstrates enhanced detection, better management of missed targets, and fewer false identifications (H. Wang et al., 2022). Using stochastic gradient descent, SiamRPN++ was trained with 131 videos and 15,125 images. It outperforms kernelized correlation filters (KCF) (Henriques et al., 2014) in managing occlusions and deformations in aberrant fish tracking and excels in precision, success rate, and real-time tracking (at 27 frames per second). Fig. 15 displays precision, success, and AUC which demonstrates the superior accuracy and robustness of SiamRPN++ whereas looks for a comparison of the fish detection and tracking methods also.

International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

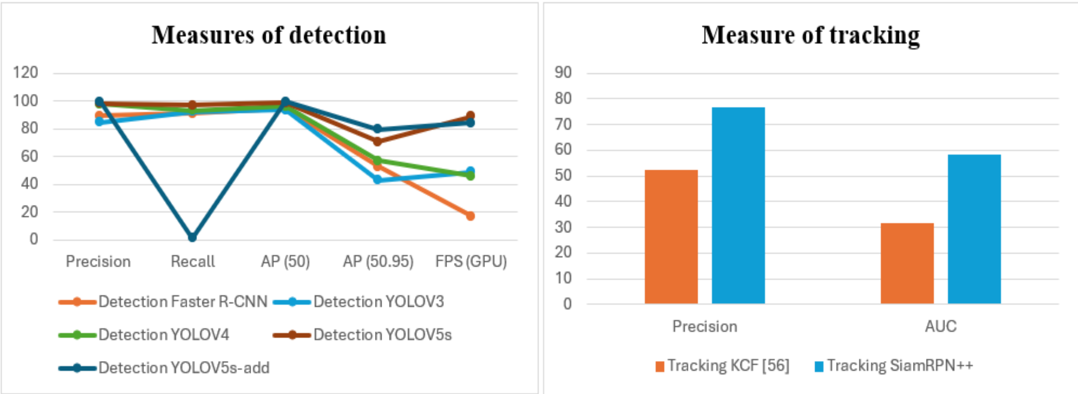

Figure 15: Comparison Of Detection And Tracking Algorithm

(H. Wang et al., 2022)

FishMOT

In four studies, FishMOT outperformed cutting-edge multi-object tracking algorithms, proving its superiority in a range of fish counts. It surpassed traditional MOT algorithms such as complete-bounding IoU (C_BIoU), which had a MOTA of 93.4% with 0 ID switches, to obtain the greatest MOTA. FishMOT was more than 6% over this. It achieved a MOTA of 99.06% on a 100-fish video in complicated circumstances with big fish schools, more than 75% higher than C_BIoU. With a MOTA of 99.92%, FishMOT fared better in difficult circumstances like uneven illumination than idtracker.ai and TRex. It also surpassed them in memory use, calculation time, and accuracy. Even in conditions that present challenges and with an increasing number of fish, its performance is resilient (S. Liu et al., 2024). Fig. 16 represents the existing MOT algorithms along with their accuracy measures comparative analysis.

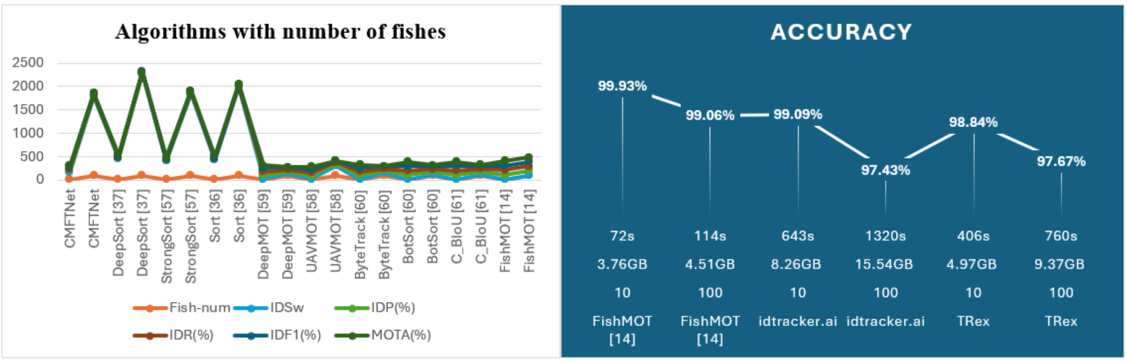

Figure 16: The Comparison Between State-Of-The-Art MOT Algorithms With Accuracy

(S. Liu et al., 2024)

Unsupervised Fish Trajectory Tracking

Concerning AP, AR, and IoU measures, the suggested unsupervised [68] approach for fish tracking and segmentation exhibits stable performance on the Seagrass, DeepFish, and YouTube-VOS datasets. It achieves accuracy close to fully supervised algorithms and exceeds earlier unsupervised systems, despite some difficulties in situations like Seagrass. Its efficacy in complicated circumstances with non-rigid objects and notable distortions is confirmed by qualitative data. Combining optical flow with background subtraction greatly increases segmentation accuracy, as demonstrated by an ablation study. The approach consistently beats the baseline, but accuracy only rises for a maximum of 100 epochs until overfitting causes a

International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

decline. A comparison of unsupervised and supervised fish detection and segmentations is shown in Fig. 17 on three different datasets. While Fig. 18 represents the unsupervised segmentation with its operations.

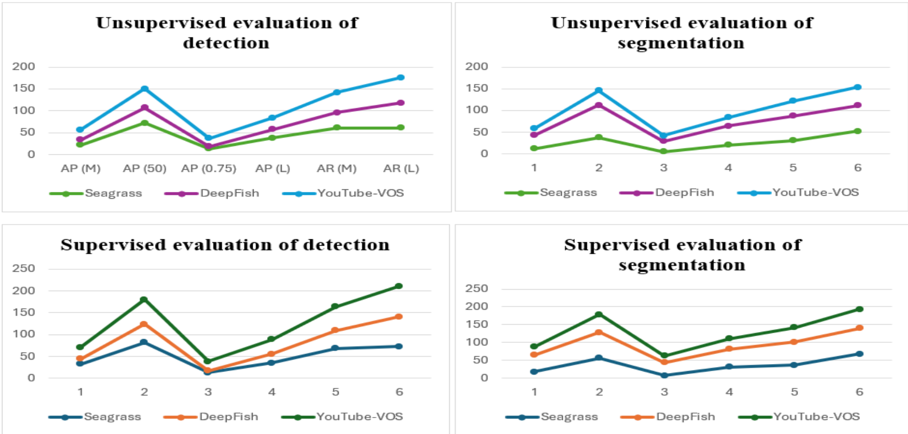

Figure 17: Comparison Of Unsupervised And Supervised Detection And Segmentation On Datasets

(S. Liu et al., 2024)

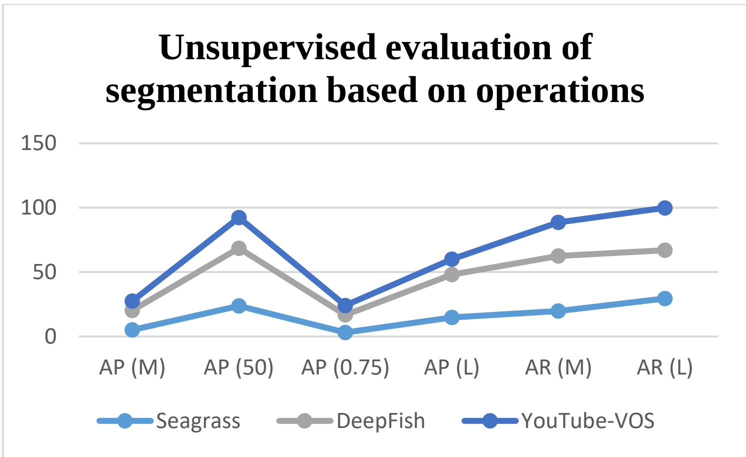

Figure 18: Comparison Of Unsupervised Segmentation Based On Optical Flow Without Background Subtraction

(S. Liu et al., 2024)

Tracking based on improved HOG

The enhanced HOG algorithm outperformed the prior technique, with an average classification accuracy of 93.19% as opposed to 55.33%. It achieved 100% recognition accuracy and 86.7% classification accuracy in 1500-frame videos in zebrafish shoaling investigations. After a month, accuracy for 18 out of 30 zebrafish steadied at 60% by the sixth week. With a tracking accuracy of 99.27%, which is 4.33% better than idTracker, the approach outperformed it. It also worked well with low-quality and sparsely sampled videos, retaining stability in the face of crossing frequencies and challenges with video clarity (Bai et al., 2018). Fig. 19 shows the identification of Zebrafish and shows the predictive probabilistic results regarding targets.

International Journal of Innovation and Industrial Revolution
EISSN: 2637-0972
IJIREV
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

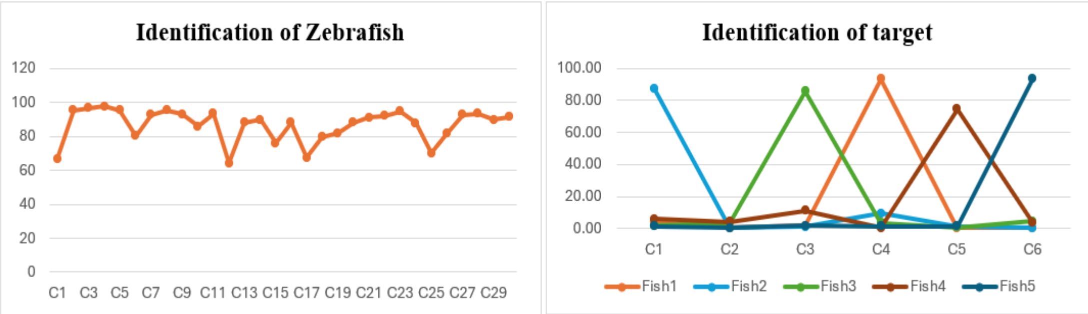

Figure 19: Evaluation Of Zebrafish And Predictive Probabilities Of Dosing Targets

(Bai et al., 2018)

Recommendations

3D Stereo Vision Tracking

Future studies should concentrate on fish motion dynamics, such as acceleration and speed, to investigate multiple fish tracking. Extending observation intervals to capture long-term behavioral trends and improving tracking systems are also critical for managing larger fish populations. These developments will result in tracking technologies that are more scalable and accurate in a variety of aquatic situations.

Automated Planar Tracking

Future developments should focus on improving state prediction and data association, particularly in low-frame-rate circumstances, to minimize misidentification and eliminate trajectory interruptions to overcome the constraints of the suggested tracking system (S. H. Wang et al., 2016). For precise head detection and body fitting, a high video resolution is essential. Furthermore, adding more detailed data to tracklet relinking than just the head rectangle could greatly increase the accuracy and efficiency of the system.

CMFTNet

Improving long-term tracking performance should be the top priority of further research, particularly for fish moving toward pond boundaries or light sources. It is imperative to improve the re-identification of fish targets, with a focus on developing deep learning-based correlation-matching techniques that are more effective. This is especially important for targets like fish fries that have few distinguishing features. Advances in these fields will result in tracking systems that are more accurate and robust for a range of demanding aquatic conditions.

Fish Tracking Based on DMK

Future improvements must concentrate on dynamically modifying kernel weights to raise accuracy in a variety of scenarios and more effectively manage occlusions and deformations. Further strengthening the system's resilience can be achieved by combining detection techniques with kernel-based tracking. The tracking algorithm's performance will need to be improved and its adaptation to different surroundings ensured, which will require expanding testing to include a larger range of scenarios and optimizing computing efficiency.

YOLOv2 and Kalman filter

To improve tracking accuracy significantly, especially in times of high occlusion, adding a specific fish recognition step is essential. By accurately recognizing each fish, this technique

International Journal of Innovation and Industrial Revolution
EISSN: 2637-0972
IJIREV
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003
lessens the spread of errors, improving tracking performance overall and enhancing robustness in complicated surroundings.

DFTNet

Subsequent endeavors will be focused on augmenting the present system through the integration of enhanced detector inputs and the optimization of association methodologies. Using more sophisticated detection techniques and effective data association procedures, these improvements seek to increase tracking accuracy and reliability, ultimately producing better performance in challenging tracking settings.

Improved YOLOV5 and SiamRPN++

Subsequent research ought to concentrate on a few crucial enhancements. First, the current single target tracking technique needs a lot of GPU power and is unsuitable for multi-target situations. It would be beneficial to develop a specialized multi-target tracking network that is adapted to fish behavior. Furthermore, for practical applications, the dataset must be expanded to encompass a larger spectrum of abnormal fish habits, as it now only includes one form of abnormal behavior. Finally, creating integrated software that takes these developments into account could help with counting and delivering real-time alerts for aberrant fish, which would be extremely beneficial to fish farmers.

FishMOT

Further initiatives will be made to expand FishMOT to encompass increasingly intricate fish-tracking situations, such as surroundings that are underwater and 3D. Advanced tracking algorithms and sensor technologies will need to be integrated to handle difficulties like complicated spatial relationships and fluctuating lighting. This also intends to investigate broader uses of FishMOT to better understand fish behaviour and enhance aquaculture management techniques. The aim will be to improve research outputs and operational efficiency in the aquaculture business by offering useful insights and tools through the integration of real-time monitoring and automatic alarm features.

Unsupervised Fish Trajectory Tracking

Prospective studies could improve the application of the method (Saleh et al., 2022b) and increase tracking accuracy in a variety of contexts by extending its scope to include other animal species that are frequently seen in watery environments. This model could also be modified for application in other domains, including tracking-by-detection in autonomous driving. These tracking techniques could improve object identification and vehicle navigation, resulting in safer and more effective autonomous driving technologies, when integrated with automotive systems. Investigating these options will increase the tracking framework's usefulness and spur innovation in a variety of fields.

Tracking Based On Improved HOG

Future research will focus on integrating multi-dimensional video analysis with cutting-edge deep learning approaches to improve feature extraction and tracking accuracy to address the problem of object overlapping. By improving the characteristics and spatial relationships of the items, this method will make use of deep learning to more effectively differentiate between overlapping objects. To enhance the identification and tracking of overlapping objects, it will also be investigated to incorporate motion parameters into deep learning models. This all-

International Journal of Innovation and Industrial Revolution IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003
encompassing approach seeks to greatly improve the system's performance in settings where item overlaps are frequent and challenging to control.

Economic Benefits
One of the key economic advantages of these technologies is improved feed optimization. Feed typically constitutes up to 50% of aquaculture operational costs. By using tracking systems to monitor fish feeding patterns and behavior, farmers can reduce overfeeding and minimize feed waste, directly improving profitability. Moreover, reducing mortality through early detection of diseases and stress can prevent major financial losses in the long term.

Another economic aspect is related to scalability and automation. Technologies that monitor fish in large aquaculture farms can reduce labor costs by automating tasks such as health monitoring and water quality control. This not only cuts operational expenses but also allows farms to expand without significantly increasing labor requirements.

Environmental Sustainability
On the environmental side, precise tracking technologies contribute to sustainability by improving resource management. For instance, better feed management reduces nutrient pollution, a major problem associated with overfeeding in fish farms that leads to water eutrophication. Furthermore, these technologies can help prevent overstocking, which often leads to habitat degradation and can stress both fish and the surrounding ecosystem.

Fish tracking also aids in assessing the environmental impact of aquaculture operations, ensuring that practices comply with environmental regulations. By closely monitoring water conditions, farms can proactively address potential issues, such as low oxygen levels or high concentrations of waste products, thereby minimizing negative effects on the surrounding marine or freshwater ecosystems.

Conclusion
In this review, GPR outperforms computational speed, whereas MGGP shows higher tracking accuracy for low-frame-rate stereo videos. Algorithms based on ANFIS are not as efficient in processing information nor as performant. CNNs categorize fish well through the analysis of head feature maps when they are augmented by data augmentation. The real-time underwater system effectively distinguishes fish from aquatic vegetation using sophisticated algorithms, and the suggested system monitors zebrafish in shallow water even in the presence of obstructions. When it comes to exceeding conventional techniques in multi-fish tracking, CMFTNet stands out for its high MOTA and IDF1.

Additionally, the publication presents several sophisticated techniques for fish tracking and detection. Stable tracking can be achieved by combining YOLOv2 with a Kalman filter, whereas DFTNet can handle intricate marine situations with fewer identification shifts. FishMOT provides reliable and effective multi-fish tracking, and improvements in YOLOV5s increase detection precision. Accurate tracking and segmentation are achieved by an unsupervised technique that combines optical flow with background subtraction. Furthermore, the enhanced HOG algorithm outperforms current techniques like idTracker in terms of tracking zebrafish accuracy and stability, allowing for in-depth behavioral research.

Acknowledgement
The authors would like to express their gratitude to the Ministry of Higher Education Malaysia and Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA) for their support through the Translational Grant RDU233601.

References
Action, S. I. (2020). World fisheries and aquaculture. *Food and Agriculture Organization, 2020*, 1–244.
Annadurai, S. (2007). *Fundamentals of digital image processing*. Pearson Education India.
Arvind, C. S., Prajwal, R., Bhat, P. N., Sreedevi, A., & Prabhudeva, K. N. (2019). Fish detection and tracking in pisciculture environment using deep instance segmentation. *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)*, 778–783.
Bae, M.-J., & Park, Y.-S. (2014). Biological early warning system based on the responses of aquatic organisms to disturbances: a review. *Science of the Total Environment*, *466*, 635–649.
Bai, Y.-X., Zhang, S.-H., Fan, Z., Liu, X.-Y., Zhao, X., Feng, X.-Z., & Sun, M.-Z. (2018). Automatic multiple zebrafish tracking based on improved HOG features. *Scientific Reports*, *8*(1), 10884.
Bailey, J. L., & Eggereide, S. S. (2020). Indicating sustainable salmon farming: The case of the new Norwegian aquaculture management scheme. *Marine Policy*, *117*, 103925.
Barreiros, M. de O., Dantas, D. de O., Silva, L. C. de O., Ribeiro, S., & Barros, A. K. (2021). Zebrafish tracking using YOLOv2 and Kalman filter. *Scientific Reports*, *11*(1), 3219.
Berclaz, J., Fleuret, F., Turetken, E., & Fua, P. (2011). Multiple object tracking using k-shortest paths optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *33*(9), 1806–1819.
Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and realtime tracking. *2016 IEEE International Conference on Image Processing (ICIP)*, 3464–3468.
Beyan, C., Katsageorgiou, V., & Fisher, R. B. (2018). Extracting statistically significant behaviour from fish tracking data with and without large dataset cleaning. *IET Computer Vision*, *12*(2), 162–170.
Bhargava, A. (2016). *A novel method for automated tracking and quantification of adult zebrafish behaviour during anxiety*.
Bochinski, E., Eiselein, V., & Sikora, T. (2017). High-speed tracking-by-detection without using image information. *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 1–6.
Bochinski, E., Senst, T., & Sikora, T. (2018). Extending IOU based multi-object tracking by visual information. *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 1–6.
Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *ArXiv Preprint ArXiv:2004.10934*.
Branson, K., & Belongie, S. (2005). Tracking multiple mouse contours (without too many samples). *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, *1*, 1039–1046.
Butail, S., & Paley, D. A. (2012). Three-dimensional reconstruction of the fast-start swimming kinematics of densely schooling fish. *Journal of the Royal Society Interface*, *9*(66), 77–88.

EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003
Cai, Z., Wen, L., Yang, J., Lei, Z., & Li, S. Z. (2013). Structured visual tracking with dynamic graph. *Computer Vision–ACCV 2012: 11th Asian Conference on Computer Vision, Daejeon, Korea, November 5-9, 2012, Revised Selected Papers, Part III 11*, 86–97.

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(4), 834–848.

Cheng, S., Zhao, K., & Zhang, D. (2019). Abnormal water quality monitoring based on visual sensing of three-dimensional motion behavior of fish. *Symmetry*, *11*(9), 1179.

Chu, C.-T., Hwang, J.-N., Pai, H.-I., & Lan, K.-M. (2013). Tracking human under occlusion based on adaptive multiple kernels with projected gradients. *IEEE Transactions on Multimedia*, *15*(7), 1602–1615.

Chuang, M.-C., Hwang, J.-N., Williams, K., & Towler, R. (2014). Tracking live fish from low-contrast and low-frame-rate stereo videos. *IEEE Transactions on Circuits and Systems for Video Technology*, *25*(1), 167–179.

Chuang, M.-C., Hwang, J.-N., Ye, J.-H., Huang, S.-C., & Williams, K. (2016). Underwater fish tracking for moving cameras based on deformable multiple kernels. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *47*(9), 2467–2477.

Comaniciu, D., Ramesh, V., & Meer, P. (2003). Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(5), 564–577.

Concepcion, R. S., Palconit, M. G. B., Dadios, E. P., Carpio, J. N., Bedruz, R. A. R., & Bandala, A. A. (2020). Arabidopsis Tracker: A Centroid-Based Vegetation Localization Model for Automatic Leaf Canopy Phenotyping in Multiple-Pot Cultivation System. *2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, 1–6.

De Luna, R. G., Dadios, E. P., & Bandala, A. A. (2018). Automated image capturing system for deep learning-based tomato plant leaf disease detection and recognition. *TENCON 2018-2018 IEEE Region 10 Conference*, 1414–1419.

Delcourt, J., Becco, C., Vandewalle, N., & Poncin, P. (2009). A video multitracking system for quantification of individual behavior in a large fish shoal: advantages and limits. *Behavior Research Methods*, *41*(1), 228–235.

Delcourt, J., Denoël, M., Ylieff, M., & Poncin, P. (2013). Video multitracking of fish behaviour: a synthesis and future perspectives. *Fish and Fisheries*, *14*(2), 186–204.

Edmonds, J., & Karp, R. M. (2003). Theoretical improvements in algorithmic efficiency for network flow problems. *Combinatorial Optimization—Eureka, You Shrink! Papers Dedicated to Jack Edmonds 5th International Workshop Aussois, France, March 5–9, 2001 Revised Papers*, 31–33.

Fernandez, M. C. D., Gob, K. J. E., Leonidas, A. R. M., Ravara, R. J. J., Bandala, A. A., & Dadios, E. P. (2014). Simultaneous face detection and recognition using Viola-Jones Algorithm and Artificial Neural Networks for identity verification. *2014 IEEE Region 10 Symposium*, 672–676.

Fontaine, E. I. (2008). *Automated visual tracking for behavioral analysis of biological model organisms*. California Institute of Technology.

Fontaine, E., Lentink, D., Kranenbarg, S., Müller, U. K., Van Leeuwen, J. L., Barr, A. H., & Burdick, J. W. (2008). Automated visual tracking for studying the ontogeny of zebrafish swimming. *Journal of Experimental Biology*, *211*(8), 1305–1316.

International Journal of Innovation and Industrial Revolution
IJIREV
EISSN: 2637-0972
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

Gao, G., Xiao, K., & Chen, M. (2019). An intelligent IoT-based control and traceability system to forecast and maintain water quality in freshwater fish farms. *Computers and Electronics in Agriculture*, *166*, 105013.

Golilarz, N. A., Demirel, H., & Gao, H. (2019). Adaptive generalized Gaussian distribution oriented thresholding function for image de-noising. *International Journal of Advanced Computer Science and Applications*, *10*(2).

Gordon, N. (2004). Beyond the kalman filter: Particle filters for tracking applications. *Fusion 2004: Seventh International Conference on Information Fusion*.

Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks*, *18*(5–6), 602–610.

Gupta, S., Mukherjee, P., Chaudhury, S., Lall, B., & Sanisetty, H. (2021). DFTNet: Deep fish tracker with attention mechanism in unconstrained marine environments. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1–13.

Hargrave, P. J. (1989). A tutorial introduction to Kalman filtering. *IEE Colloquium on Kalman Filters: Introduction, Applications and Future Developments*, 1.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. *Proceedings of the IEEE International Conference on Computer Vision*, 2961–2969.

Held, D., Thrun, S., & Savarese, S. (2016). Learning to track at 100 fps with deep regression networks. *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, 749–765.

Henriques, J. F., Caseiro, R., Martins, P., & Batista, J. (2014). High-speed tracking with kernelized correlation filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *37*(3), 583–596.

Hongpin, L., Guanglin, L., Weifeng, P., Jie, S., & Qiuwei, B. (2015). Real-time remote monitoring system for aquaculture water quality. *International Journal of Agricultural and Biological Engineering*, *8*(6), 136–143.

Javed, O., & Shah, M. (2002). Tracking and object classification for automated surveillance. *Computer Vision—ECCV 2002: 7th European Conference on Computer Vision Copenhagen, Denmark, May 28–31, 2002 Proceedings, Part IV 7*, 343–357.

Karaboga, D., & Kaya, E. (2019). Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey. *Artificial Intelligence Review*, *52*, 2263–2293.

Krähenbühl, P., & Koltun, V. (2011). Efficient inference in fully connected crfs with gaussian edge potentials. *Advances in Neural Information Processing Systems*, *24*.

Kuhn, H. W. (2004). The Hungarian method for the assignment problem. *Naval Research Logistics (NRL)*, *52*(1), 7–21.

Kuklina, I., Kouba, A., & Kozák, P. (2013). Real-time monitoring of water quality using fish and crayfish as bio-indicators: a review. *Environmental Monitoring and Assessment*, *185*, 5043–5053.

LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., & Jackel, L. (1989). Handwritten digit recognition with a back-propagation network. *Advances in Neural Information Processing Systems*, *2*.

Lee, K.-H., Hwang, J.-N., Okapal, G., & Pitton, J. (2014). Driving recorder based on-road pedestrian tracking using visual SLAM and constrained multiple-kernel. *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2629–2635.

Lee Rodgers, J., & Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, *42*(1), 59–66.

Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003
Li, B., Wu, W., Wang, Q., Zhang, F., Xing, J., & Yan, J. (2019). Siamrpn++: Evolution of siamese visual tracking with very deep networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4282–4291.

Li, P., & Chen, S. (2016). A review on Gaussian process latent variable models. *CAAI Transactions on Intelligence Technology*, *1*(4), 366–376.

Li, W., Li, F., & Li, Z. (2022). CMFTNet: Multiple fish tracking based on counterpoised JointNet. *Computers and Electronics in Agriculture*, *198*, 107018.

Li, Y., Tian, X., Shen, X., & Tao, D. (2017). Classification and Representation Joint Learning via Deep Networks. *IJCAI*, *2017*, 67.

Liu, S., Zheng, X., Han, L., Liu, X., Ren, J., Wang, F., Liu, Y., & Lin, Y. (2024). FishMOT: A simple and effective method for fish tracking based on IoU matching. *Applied Engineering in Agriculture*, *40*(5), 599–609.

Liu, Y., Li, B., Zhou, X., Li, D., & Duan, Q. (2024). FishTrack: Multi-object tracking method for fish using spatiotemporal information fusion. *Expert Systems with Applications*, *238*, 122194.

Liu, Y., Li, H., & Chen, Y. Q. (2012). Automatic tracking of a large number of moving targets in 3d. *Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part IV 12*, 730–742.

Liu, Z., Chen, Z., Li, Z., & Hu, W. (2018). An efficient pedestrian detection method based on YOLOv2. *Mathematical Problems in Engineering*, *2018*(1), 3518959.

Luo, W., Xing, J., Milan, A., Zhang, X., Liu, W., & Kim, T.-K. (2021). Multiple object tracking: A literature review. *Artificial Intelligence*, *293*, 103448.

Luong, C. T., Audira, G., Kurnia, K. A., Hung, C., & Hsiao, C. (2024). Fish 3D Locomotion app: a user-friendly computer application package for automatic data calculation and endpoint extraction for novel tank behavior in fish. *Journal of Fish Biology*, *105*(4), 1086–1108.

Ma, H., Tsai, T.-F., & Liu, C.-C. (2010). Real-time monitoring of water quality using temporal trajectory of live fish. *Expert Systems with Applications*, *37*(7), 5158–5171.

Mao, J., Xiao, G., Sheng, W., Qu, Z., & Liu, Y. (2016). Research on realizing the 3D occlusion tracking location method of fish's school target. *Neurocomputing*, *214*, 61–79.

Mirat, O., Sternberg, J. R., Severi, K. E., & Wyart, C. (2013). ZebraZoom: an automated program for high-throughput behavioral analysis and categorization. *Frontiers in Neural Circuits*, *7*, 107.

Naylor, R. L., Hardy, R. W., Buschmann, A. H., Bush, S. R., Cao, L., Klinger, D. H., Little, D. C., Lubchenco, J., Shumway, S. E., & Troell, M. (2021). A 20-year retrospective review of global aquaculture. *Nature*, *591*(7851), 551–563.

Nguyen, T., Dax, M., Mummadi, C. K., Ngo, N., Nguyen, T. H. P., Lou, Z., & Brox, T. (2019). Deepusps: Deep robust unsupervised saliency prediction via self-supervision. *Advances in Neural Information Processing Systems*, *32*.

Noldus, L. P. J. J., Spink, A. J., & Tegelenbosch, R. A. J. (2001). EthoVision: a versatile video tracking system for automation of behavioral experiments. *Behavior Research Methods, Instruments, & Computers*, *33*, 398–414.

Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(7), 971–987.

Palconit, M. G. B., Concepcion II, R. S., Alejandrino, J. D., Pareja, M. E., Almero, V. J. D., Bandala, A. A., Vicerra, R. R. P., Sybingco, E., Dadios, E. P., & Naguib, R. N. G. (2021). Three-dimensional stereo vision tracking of multiple free-swimming fish for

Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003
low frame rate video. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *25*(5), 639–646.

Palconit, M. G. B., Fuentes, R. B., Narvios, W. M. O., Rosales, M. A., Bandala, A. A., & Dadios, E. P. (2020). Adaptive compensator of magnetic levitation system using symbolic regression. *2020 IEEE REGION 10 CONFERENCE (TENCON)*, 1363–1367.

Papadakis, V. M., Papadakis, I. E., Lamprianidou, F., Glaropoulos, A., & Kentouri, M. (2012). A computer-vision system and methodology for the analysis of fish behavior. *Aquacultural Engineering*, *46*, 53–59.

Parkhi, O., Vedaldi, A., & Zisserman, A. (2015). Deep face recognition. *BMVC 2015-Proceedings of the British Machine Vision Conference 2015*.

Pérez-Escudero, A., Vicente-Page, J., Hinz, R. C., Arganda, S., & De Polavieja, G. G. (2014). idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nature Methods*, *11*(7), 743–748.

Piccardi, M. (2004). Background subtraction techniques: a review. *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, *4*, 3099–3104.

Qian, Z.-M., Cheng, X. E., & Chen, Y. Q. (2014). Automatically detect and track multiple fish swimming in shallow water with frequent occlusion. *PloS One*, *9*(9), e106506.

Qian, Z.-M., Wang, S. H., Cheng, X. E., & Chen, Y. Q. (2016). An effective and robust method for tracking multiple fish in video image based on fish head detection. *BMC Bioinformatics*, *17*, 1–11.

Quiros, A. R. F., Abad, A., Bedruz, R. A., Uy, A. C., & Dadios, E. P. (2015). A genetic algorithm and artificial neural network-based approach for the machine vision of plate segmentation and character recognition. *2015 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, 1–6.

Redmon, J. (2018). Yolov3: An incremental improvement. *ArXiv Preprint ArXiv:1804.02767*.

Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7263–7271.

Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(6), 1137–1149.

Ren, Z., & Wang, Z. (2010). Differences in the behavior characteristics between Daphnia magna and Japanese madaka in an on-line biomonitoring system. *Journal of Environmental Sciences*, *22*(5), 703–708.

Romero-Ferrero, F., Bergomi, M. G., Hinz, R. C., Heras, F. J. H., & De Polavieja, G. G. (2019). Idtracker. ai: tracking all individuals in small or large collectives of unmarked animals. *Nature Methods*, *16*(2), 179–182.

Saleh, A., Sheaves, M., Jerry, D., & Azghadi, M. R. (2022a). Transformer-based self-supervised fish segmentation in underwater videos. *ArXiv Preprint ArXiv:2206.05390*.

Saleh, A., Sheaves, M., Jerry, D., & Azghadi, M. R. (2022b). Unsupervised fish trajectory tracking and segmentation. *ArXiv Preprint ArXiv:2208.10662*.

Shiau, Y.-H., Chen, C.-C., & Lin, S.-I. (2013). Using bounding-surrounding boxes method for fish tracking in real world underwater observation. *International Journal of Advanced Robotic Systems*, *10*(7), 298.

Shitrit, H. Ben, Berclaz, J., Fleuret, F., & Fua, P. (2013). Multi-commodity network flow for tracking multiple people. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(8), 1614–1627.

Shreesha, S., Pai, M. M. M., Verma, U., & Pai, R. M. (2023). Fish Tracking and Continual Behavioral Pattern Clustering Using Novel Sillago Sihama Vid (SSVid). *IEEE Access*, *11*, 29400–29416.

Sridhar, V. H., Roche, D. G., & Gingins, S. (2019). Tracktor: image-based automated tracking of animal movement and behaviour. *Methods in Ecology and Evolution*, *10*(6), 815–820.

Stauffer, C., & Grimson, W. E. L. (1999). Adaptive background mixture models for real-time tracking. *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, *2*, 246–252.

Teed, Z., & Deng, J. (2020). Raft: Recurrent all-pairs field transforms for optical flow. *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part II 16*, 402–419.

Vedaldi, A., & Lenc, K. (2015). Matconvnet: Convolutional neural networks for matlab. *Proceedings of the 23rd ACM International Conference on Multimedia*, 689–692.

Wageeh, Y., Mohamed, H. E.-D., Fadl, A., Anas, O., ElMasry, N., Nabil, A., & Atia, A. (2021). YOLO fish detection with Euclidean tracking in fish farms. *Journal of Ambient Intelligence and Humanized Computing*, *12*, 5–12.

Wang, C., Wu, Z., Chen, Y., Zhang, W., Ke, W., & Xiong, Z. (2023). Improving 3-D Zebrafish Tracking With Multiview Data Fusion and Global Association. *IEEE Sensors Journal*, *23*(15), 17245–17259.

Wang, H., Zhang, S., Zhao, S., Wang, Q., Li, D., & Zhao, R. (2022). Real-time detection and tracking of fish abnormal behavior based on improved YOLOV5 and SiamRPN++. *Computers and Electronics in Agriculture*, *192*, 106512.

Wang, N., & Yeung, D.-Y. (2013). Learning a deep compact image representation for visual tracking. *Advances in Neural Information Processing Systems*, *26*.

Wang, S. H., Cheng, X. E., Qian, Z.-M., Liu, Y., & Chen, Y. Q. (2016). Automated planar tracking the waving bodies of multiple zebrafish swimming in shallow water. *PloS One*, *11*(4), e0154714.

Wang, X., Zhang, R., Kong, T., Li, L., & Shen, C. (2020). Solov2: Dynamic and fast instance segmentation. *Advances in Neural Information Processing Systems*, *33*, 17721–17732.

Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. *2017 IEEE International Conference on Image Processing (ICIP)*, 3645–3649.

Xia, C., Fu, L., Liu, Z., Liu, H., Chen, L., & Liu, Y. (2018). Aquatic toxic analysis by monitoring fish behavior using computer vision: A recent progress. *Journal of Toxicology*, *2018*(1), 2591924.

Xiang, Y., Alahi, A., & Savarese, S. (2015). Learning to track: Online multi-object tracking by decision making. *Proceedings of the IEEE International Conference on Computer Vision*, 4705–4713.

Xu, Z., & Cheng, X. E. (2017). Zebrafish tracking using convolutional neural networks. *Scientific Reports*, *7*(1), 42815.

Yang, L., Liu, Y., Yu, H., Fang, X., Song, L., Li, D., & Chen, Y. (2021). Computer vision models in intelligent aquaculture with emphasis on fish detection and behavior analysis: a review. *Archives of Computational Methods in Engineering*, *28*, 2785–2816.

Ylieff, M. Y., & Poncin, P. (2003). Quantifying spontaneous swimming activity in fish with a computerized color video tracking system, a laboratory device using last imaging techniques. *Fish Physiology & Biochemistry*, *28*.

International Journal of Innovation and Industrial Revolution
EISSN: 2637-0972
IJIREV
Volume 7 Issue 20 (March 2025) PP. 29-61
DOI 10.35631/IJIREV.720003

Yuan, X., Martínez, J.-F., Eckert, M., & López-Santidrián, L. (2016). An improved Otsu threshold segmentation method for underwater simultaneous localization and mapping-based navigation. *Sensors*, *16*(7), 1148.

Zang, Q., & Klette, R. (2003). Object classification and tracking in video surveillance. *International Conference on Computer Analysis of Images and Patterns*, 198–205.

Zhang, Y., Hua, J., & Wang, Y. Bin. (2013). Application effect of aquaculture IOT system. *Applied Mechanics and Materials*, *303*, 1395–1401.

Zhang, Y., Wang, C., Wang, X., Zeng, W., & Liu, W. (2021). Fairmot: On the fairness of detection and re-identification in multiple object tracking. *International Journal of Computer Vision*, *129*, 3069–3087.

Zhao, X., Yan, S., & Gao, Q. (2019). An algorithm for tracking multiple fish based on biological water quality monitoring. *IEEE Access*, *7*, 15018–15026.

Zhou, X., Koltun, V., & Krähenbühl, P. (2020). Tracking objects as points. *European Conference on Computer Vision*, 474–490.