



**EAST WEST UNIVERSITY**

Course Code: CSE475

Course Title: Machine Learning

Literature Review: Deep Learning for Fish Species Classification and Detection

Sec: 2

**Group - 09**

**SUBMITTED TO:**

Dr. Raihan Ul Islam

Associate Professor, Department of Computer Science and Engineering

East West University

**SUBMITTED BY:**

Sifat Noor Siam

ID: 2022-1-60-368

Md. Sakib Hasan

ID: 2022-1-60-098

Tasnim Saima Raita

ID: 2022-1-60-291

Sumona Sharmin Israt

ID: 2022-1-60-066

Date of Submission: 10 July 2025

## Literature Review: Deep Learning for Fish Species Classification and Detection

### 1. Importance and Challenges of Fish Classification

Fish classification plays a vital role in biodiversity monitoring, ecological research, and fisheries management. Traditional approaches—relying on manual identification or classical machine learning—struggle with scalability, accuracy, and robustness under real-world conditions, especially underwater. Key challenges include:

- Low inter-class and high intra-class variation
- Environmental noise (e.g., turbidity, lighting, occlusion)
- Limited and imbalanced datasets, especially in estuarine and turbid conditions

### 2. Traditional and Classical Machine Learning Approaches

Earlier methods employed:

- Manual inspection or taxonomic keys (labor-intensive)
- Image processing and classical ML models (e.g., SVM, KNN, Decision Trees)
- Handcrafted features like shape, texture, and color

These approaches are sensitive to image quality and lack generalization across datasets or environments.

### 3. Convolutional Neural Networks (CNNs) and Advances

CNNs marked a significant improvement by learning hierarchical features directly from raw data. Prominent architectures include:

- **ResNet, InceptionNet, MobileNet, EfficientNet, VGGNet, AlexNet**
- Applied in marine, freshwater, aquaculture, and coral reef settings

CNNs improved classification accuracy but faced limitations:

- Dependence on large datasets
- Difficulty in real-time or embedded deployment
- Poor generalization in turbid and complex underwater environments

### 4. Object Detection Models

Deep learning-based object detectors, particularly one-stage and two-stage architectures, became popular:

- **YOLO (v3, v4, v5, S2F-YOLO):** Fast, real-time detection with trade-offs in accuracy
- **Faster R-CNN, SSD:** More accurate but computationally intensive
- **Enhancements:** Anchor box tuning, backbone replacement (e.g., ShuffleNet V2), data augmentation (e.g., Mosaic, HSV)

Object detection frameworks have been tested in ecological monitoring, fish counting, and classification tasks with mixed results due to environmental noise and hardware limitations.

## 5. Vision Transformers (ViTs) and Transformer-based Models

Inspired by NLP, ViTs and DETR have been explored for their ability to model long-range dependencies via self-attention:

- **ViT, DeiT, Swin Transformer, DETR:** Outperform CNNs on large datasets
- Effective even in small-data settings with transfer learning and augmentation
- DETR shown to be more robust in cluttered and occluded underwater scenes

However, ViTs are data-hungry and have limited adoption in ecological applications, especially with imbalanced and small datasets.

## 6. Semi-Supervised and Weakly-Supervised Learning

To address the high cost of labeled datasets:

- **Semi-supervised learning (STAC, Soft Teacher, Unbiased Teacher):** Combines labeled and unlabeled data for training
- **Weakly-supervised learning:** Utilizes coarse labels or image-level tags
- **Contrastive learning and ensemble models (e.g., XGBoost)** have been used to improve accuracy with minimal supervision

These methods have shown promise in turbid and low-visibility underwater settings, where manual labeling is impractical.

## 7. Specialized Techniques and Contributions

Studies in the reviewed literature introduced innovations such as:

- **Custom preprocessing:** Color correction, noise filtering
- **Lightweight models:** For deployment on devices like Jetson Nano
- **Improved annotation quality:** To enhance training and model performance
- **New datasets:** Fish-Pak, iNat2021, estuarine and coral reef datasets, FDD (fjord-based)

Models like **S2F-YOLO** and **reduced AlexNet-based CNNs** have demonstrated strong performance in both speed and accuracy for real-time monitoring.

## 8. Summary of Research Gaps

Across studies, several key gaps remain:

- Limited datasets for estuarine and turbid environments
- Insufficient exploration of ViTs and DETR in underwater ecology
- Trade-offs between model accuracy, speed, and computational cost
- Generalization across diverse underwater habitats is still challenging
- Inadequate attention to annotation quality and dataset imbalance

## 9. Summary of Reviewed Paper

• SL	Dataset	Methodology	Limitations
[1]	14 datasets with 19 distinct subsets from varied environments (lab, onboard vessels, underwater in natural habitats)	Combined FGVC-PIM with Swin Transformer for hierarchical feature extraction and attention to discriminative regions.	Minor performance drop in four subsets (still >83%), indicating scope for improvement in harsh conditions.
[2]	Curated Estuarine Fish species dataset (EFD)	Vision Transformer (ViT) compared against VGG16/19, DenseNet121, ResNet50v2, Inception, and Xception.	Limited generalizability to non-estuarine fish and other environments was not fully assessed.
[3]	Temperate fish images and videos, fine-tuned from Fish4Knowledge.	YOLO for detection, SE-CNN for classification, transfer learning via ImageNet.	Lower post-training accuracy without augmentation; dependent on dataset expansion.
[4]	Custom fish dataset tested across YOLO variants.	S2F-YOLO (YOLOv5 + ShuffleNetV2 + focal loss), optimized for speed and precision.	Slight mAP drop (~2.24%) compared to YOLOv5x; model still needs tuning for different detection environments
[5]	Nine fish species, captured using camera in natural conditions.	FD_Net (YOLOv7-based) with MobileNetv3 and DenseNet-169 with ArcFace Loss and BNAM modules.	Struggles with low-quality images and challenging underwater conditions.
[6]	DIDSON sonar and optical video datasets from U.S. rivers.	YOLO adaptation + Norfair tracker + Kalman filter for classification and counting.	Only 8 species detected; system not yet generalized beyond original rivers.
[7]	FishInTurbidWater (custom, turbid conditions, weakly labeled).	Semi-supervised contrastive learning + weakly-supervised ensemble DNN (ImageNet-based transfer learning)	Potential lower accuracy under certain video noise levels; weak labeling may lead to occasional misclassifications.
[8]	Publicly available fish image datasets (e.g., Fish4Knowledge and custom datasets).	Utilized CNN architecture for feature extraction; experiments conducted with different hyperparameters and layers to enhance classification.	Classification accuracy may drop under varying underwater conditions, such as turbidity or occlusion.
[9]	High-quality datasets of 20 fish species with 2,000 images collected via underwater cameras.	Tested multiple deep learning architectures (VGG, ResNet, Inception) with data augmentation and transfer learning.	Performance affected by background noise, lighting variation, and insufficient samples for some species.
[10]	Labeled + unlabeled fish images from real-world underwater datasets (e.g., Fish4Knowledge).	A hybrid method combining convolutional neural networks with label propagation.	Model performance is sensitive to class imbalance and quality of unlabeled data.

## 10. Conclusion

The evolution from manual and classical ML approaches to CNNs, object detectors, and now transformer-based models marks significant progress in automated fish classification. However, success in real-world applications depends on addressing data scarcity, environmental variability, and deployment constraints. Future research should focus on:

- Enhancing transfer learning for small datasets
- Adopting semi-supervised frameworks for turbid environments
- Designing lightweight yet robust models for field use
- Building standardized, diverse, and annotated datasets

## References

- [1]. Veiga, R. J. M., & Rodrigues, J. M. F. (2024). Fine-Grained Fish Classification From Small to Large Datasets With Vision Transformers. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2024.3443654>
- [2]. Tejaswini, H., Manohara Pai, M. M., & Pai, R. M. (2024). Automatic Estuarine Fish Species Classification System Based on Deep Learning Techniques. *IEEE Access*, 12, 140412. <https://doi.org/10.1109/ACCESS.2024.3468438>
- [3]. Knausgård, K. M., Wiklund, A., Sjørdalen, T. K., Halvorsen, K. T., Kleiven, A. R., Jiao, L., & Goodwin, M. (2021). Temperate fish detection and classification: A deep learning based approach. *Applied Intelligence*, 52, 6988–7001. <https://doi.org/10.1007/s10489-020-02154-9>
- [4]. Wang, F., Zheng, J., Zeng, J., Zhong, X., & Li, Z. (2023). S2F-YOLO: An Optimized Object Detection Technique for Improving Fish Classification. *Journal of Internet Technology*, 24(6), 1211–1220. <https://doi.org/10.53106/160792642023112406004>
- [5]. Malik, H., Naeem, A., Hassan, S., Ali, F., Naqvi, R. A., & Yon, D. K. (2023). Multi-classification deep neural networks for identification of fish species using camera captured images. *PLOS ONE*, 18(4), e0284992. <https://doi.org/10.1371/journal.pone.0284992>
- [6]. Kandimalla, V., Richard, M., Smith, F., Quirion, J., Torgo, L., & Whidden, C. (2022). Automated detection, classification and counting of fish in fish passages with deep learning. *Frontiers in Marine Science*, 8, 823173. <https://doi.org/10.3389/fmars.2021.823173>
- [7]. Jahanbakht, M., Rahimi Azghadi, M., & Waltham, N. J. (2023). Semi-supervised and weakly-supervised deep neural networks and dataset for fish detection in turbid underwater videos. *Ecological Informatics*, 78, 102303. <https://doi.org/10.1016/j.ecoinf.2023.102303>
- [8]. Khan, S., Khan, M. A., Saba, T., Rehman, A., Mehmood, Z., & Tariq, U. (2020). Automatic fish species classification using deep convolutional neural networks. *Wireless Personal Communications*, 112, 207–222. <https://doi.org/10.1007/s11277-019-06634-1>
- [9]. Gao, T., Sun, Y., Wang, Y., Wang, H., & Wang, J. (2020). Classification of fish species using deep learning models. *Applied Artificial Intelligence*, 34(7), 513–527. <https://doi.org/10.1080/08839514.2020.1790301>
- [10]. Ali, F., Hassan, S. A., Naeem, A., & Raza, M. (2023). Semi-supervised learning for fish species recognition using CNNs and label propagation. *Ecological Informatics*, 71, 101784. <https://doi.org/10.1016/j.ecoinf.2022.101784>