2025-12

Machine-Vision-Based Extinct Fish Recognition: Bangladeshi Freshwater Case

Nil, Supprio Ghosh

IUB

Machine-Vision-Based Extinct Fish Recognition: Bangladeshi Freshwater Case

December 2025

Prepared by:

Supprio Ghosh Nil

ID: 2221077

Amit Bikram Roy

ID: 2222345

Department of Computer Science and Engineering
Independent University, Bangladesh

Supervised by

Md. Tarek Habib, PhD
Assistant Professor
Department of Computer Science and Engineering
Independent University, Bangladesh

Table of Contents

List of Figures

List of Tables

Attestation

We are aware of the fact that plagiarism is strictly prohibited and it conflicts with our university's rules and regulations. We guarantee the authenticity of our work. We have used some copyrighted material and models of others in our project which has been properly cited following the international standards proper guidelines.

Author Name:

Supprio Ghosh Nil

Signature:

....

Author Name:

Amit Bikram Roy

Signature:

....

Letter of Transmittal

Supprio Ghosh Nil
2221077
Independent University, Bangladesh
Department of Computer Science & Engineering
14 December 2025

To
Md Tarek Habib, PhD
Assistant Professor, Department of Computer Science and Engineering
Independent University, Bangladesh

Subject: Submission of Final Undergraduate Thesis

Dear Sir,
I am pleased to submit my undergraduate thesis titled “Machine-Vision-Based Extinct Fish Recognition: Bangladeshi Freshwater Case”, submitted in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science & Engineering.

This research, conducted under your supervision, presents a deep learning-based approach for the recognition and classification of six rare native freshwater fish species of Bangladesh. Using machine vision and convolutional neural networks (CNNs), the study aims to aid in biodiversity conservation by enabling accurate, automated identification of endangered fish species. Through the collection and analysis of 8,676 high-resolution images, the research compares multiple CNN architectures including ResNet-50, VGG-16, InceptionV3, and EfficientNet-B0 and introduces a custom CNN model that achieved the highest accuracy of 97.42%.

I would like to express my sincere gratitude for your continuous guidance, encouragement, and valuable insights throughout the research process. Your mentorship has been instrumental in the successful completion of this work.
Thank you for your time and consideration.

Sincerely,
Supprio Ghosh Nil
2221077
On behalf of Group

Evaluation Committee

Supervisor	
Name: Md. Tarek Habib, PhD	Signature: ______________
Internal Examiner 1	
Name: Mohammad Shidujaman, PhD	Signature: ______________
Internal Examiner 2	
Name: Md Zahangir Alam, PhD	Signature: ______________
External Examiner	
Name: Anisul Haq Bhuya	Signature: ______________

Acknowledgement

First of all, thanks to God for granting us the strength and patience to complete this research. We would like to express our sincere gratitude to our respected Assistant Professor Md. Tarek Habib, PhD, Independent University, Bangladesh. His dedicated guidance, insightful suggestions, remarkable patience, and the generous investment of his valuable time played a pivotal role in the successful completion of our project. His expertise and mentorship greatly enhanced the quality and depth of our work, and we are truly appreciative of his unwavering support throughout our academic journey.

We would also like to extend our heartfelt thanks to Saadia Binte Alam, PhD, Head of the Department of Computer Science & Engineering, for her continuous encouragement and supportive academic environment, which greatly facilitated our progress.

Our sincere appreciation also goes to Assistant Professor Md. Sadekur Rahman of Daffodil International University, whose motivation and supportive words encouraged us to stay focused and dedicated throughout this project.

Additionally, our sincere appreciation goes to the Fab Lab for their extraordinary help and consistent support. The collaborative and resourceful environment provided by Fab Lab significantly contributed to the development and execution of our project. We are grateful for the opportunities and resources extended to us, which undoubtedly enriched our learning journey.

We would like to express our gratitude to our senior brother, Md. Ataur Rahman, for his valuable guidance, motivation, and support throughout the project. His advice and encouragement helped us navigate challenges with confidence.

My acknowledgements also extend to my mother, younger brother, and younger sister, who provided continuous support, inspiration, and strength throughout this period and helped me in collecting data. Their unconditional love and encouragement kept me motivated at every step.

Finally, we acknowledge and appreciate the contributions of all individuals who supported us, both directly and indirectly, during this endeavor.

Abstract

The biodiversity of fish in Bangladesh is at great risk because of overfishing, poaching, destruction of habitats, pollution, and increased market demands. The safeguarding of such species requires the growth of rare native fish to be recognized immediately and precisely to ensure conservation projects can intervene in time. We concentrated on 6 rare species in this study, which are Bain, Deshi Magur, Kajoli, Khoilsha, Koi, and Tangra. A total of 8,676 high-resolution images were obtained in various districts. The pictures include real-life situations such as different backgrounds, light, and single and multiple fish per picture. Five deep learning models in Convolutional Neural Network (CNN), ResNet-50, VGG-16, Inception V3, and EfficientNet-B0 were tested and trained on the basis of the precision, the accuracy, the F_1-score and the recall. The most promising one was the Custom CNN, which had an overall accuracy of 97 percent and had good and stable performance with all the species in all conditions. The method offers an efficient and effective method to track fish populations in real-time and assist in the conservation of biodiversity. This approach offers a robust and practical solution for tracking fish populations and supporting biodiversity conservation efforts.

1 Introduction

The freshwater aquatic environments of Bangladesh harbor an impressive number of indigenous fish species that are having essential roles in ecology, culture, food, and livelihoods of millions of people, though most of them are currently on the verge of extinction because of overfishing and habitat destruction, as well as increasing market pressure. The biggest difficulty is that the people cannot be aware of what species are endangered, and as a result, they end up harvesting those species and society is being depleted every second, which is unintentionally in rural markets and local fishermen. In this regard, the current deep learning and machine-vision technologies have been able to provide powerful solutions by facilitating quick, precise, and convenient identification of the species using widely-available hardware, including smartphones. Continuing on the opportunities, the current chapter presents a deep learning-based recognition system that has been constructed to recognize six native Bangladeshi fish species that are starting to become rarer: Tara Baim, Deshi Magur, Kajoli, Khoilsha, Deshi Koi, and Tengra. Five state-of-the-art models, namely ResNet-50, VGG-16, InceptionV3, EfficientNet-B0, and a self-made CNN, were trained and tested using 8,676 real-life images that are gathered in different districts and under different conditions of lighting, background, and orientation. The best of them was the Custom CNN, which had the highest accuracy and the presence of high-stable levels of precision, recall, and F_1-scores, and is very adaptable to morphological changes and can be used in the situation of the real-time implementation. This chapter aims to discuss how this type of machine-vision system could contribute to conservation, specifically what benefits this type of system has in reducing misidentification, providing public awareness, helping researchers and providing scalable tools to patrol endangered species. Combining technology innovation with ecological need, the work can bring out the potential of AI-based fish recognition to play a significant role in preserving freshwater in the Bangladesh wilderness and promotes the future development of technology-based solutions, mobile or community-based, to make conservation more attainable and effective.

1.1 Background of the Project

Bangladesh is also called the land of rivers, which has one of the most complex systems of freshwater sources in the world, as it has rivers, floodplains, canals, beels, and haors. It is an enormous aquatic ecosystem that provides an unusually high rate of biodiversity particularly among freshwater fish species. The national records on biodiversity indicate that Bangladesh is home to over 260 native freshwater fish species, with several of them being endemic to its riverine systems. These species are not merely important ecologically, but they also form the main source of livelihood, food production as well as cultural legacy of millions of individuals. Fish forms almost 60 percent of animal protein consumed in the average Bangladeshi diet and the fisheries industry makes a substantial contribution to the GDP, rural employment as well as export earnings of the country. Therefore, sustainable management of the aquatic biodiversity has great significance on the environment and socio-economic stability. Yet, this balance has been under a great threat in the recent decades. The high rate of industrialization, urbanization and agricultural development has resulted into massive pollution of water and degradation of habitats. Siltation, constructions of dams and the indiscriminate application of agrochemicals have destroyed many areas in nature that are used by fish to breed and feed and generally lead a good life. The depletion in the population numbers has also been enhanced by overfishing and fishing of young fish. Moreover, the unregulated markets still require rare species thus leading to their excessive exploitation. Research has also found out that a number of native fish species, which were once common in rivers and wetlands, are now either threatened by extinction or soon to be so. Such species loss is not only disruptive to the ecological balance but also endangered the genetic diversity needed to respond to environmental changes including temperature changes, changes in salinity, and epidemics.

One of the biggest challenges about the conservation of these endangered species is that there are no effective and easy methods of identification. The conventional forms of species identification are based on hand observation, morphological checkup and taxonomic expertise. These processes are tedious, expensive and in most cases, locals cannot access them in time or they are not available to the controlling agencies. Moreover, lighting, clutter in the background, or partial visibility of fish in market conditions are also external factors that further complicate the manual identification. Consequently, fishermen and consumers can be constantly oblivious of the fact that they are either catching or buying endangered species, which contributes to a growing decline in biodiversity unknowingly. Therefore, there exists an immediate necessity of technological intervention that will allow identifying the native fish species swiftly and precisely even in the uncontrolled and complicated environmental conditions. Recent advances in Artificial Intelligence (AI) and computer vision offer a way out of this dilemma. Image classification and pattern recognition have been revolutionized by the use of deep learning, and more specifically, Convolutional Neural Networks (CNNs) in diverse areas, such as agriculture, medicine, ecology, and others. CNNs can automatically learn and obtain discriminative visual features without feature engineering that must be engineered by hand. This makes them very much appropriate in the identification of fish species with different shapes, colors, and patterns even in case images are captured at different environmental conditions. Such AI-based recognition systems can provide real-time and accessible conservation officer, researcher, and local community tools, when incorporated into mobile or web-based platforms. Combining computer vision with the conservation of biodiversity, it will be possible to develop a scalable, smart, and data-driven solution to protect the native aquatic species. These systems not only help to track the population changes and the identification of rare species, but also they are crucial to creating awareness among people and facilitating sustainable fisheries. Consequently, the need to create a machine vision-based fish-recognizing system adapted to the freshwater environment in Bangladesh is timely and necessary to guarantee the ecological and food security alongside the sustainability of the environment in the coming years.

1.2 Problem Statement

The progressive yet slow deterioration of the freshwater biodiversity in Bangladesh has become a national and ecological issue. After being abundant with life, the rivers, canals, and flood plains of the country are currently experiencing an impressive decline in the number of most of the native fish species. Overfishing, unplanned urbanization, water pollution, siltation, industrial effluents, and destruction of natural breeding grounds to name a few have contributed to the rapid disappearance of aquatic life. Moreover, the uncontrolled and increasing market demand of the rare and exotic fish species has burdened the already endangered species with enormous pressure. Lack of proper checks and balances implies that these species are being poached and sold without the knowledge of the population or the government. This has led to the near extinction of a number of fish species that was previously very critical in ensuring an ecological balance not only to the biodiversity but also to the livelihood of communities that rely on fishing as their main source of livelihood. The traditional ways of identifying fish species in Bangladesh are mostly manual and they rely on human observations and experts in taxonomy. The processes are tedious, costly, and not very useful in the real field situation especially in fish markets and in fields with no proper equipment or knowledge. Most of the species are recognized visually in different lighting situations, in a background with mixed objects, or in cases where more than two fish are put together; which seriously impairs the ability to be accurate. Furthermore, government agencies and the conservation organizations do not have the technology to monitor and collect data in real time. Due to this, the early identification and reporting of endangered species have been irregular and slow and it has been hard to devise conservation measures in time. In the absence of an automated, scalable and efficient system, the process of recognizing the rare native fish species in uncontrolled conditions is one of the most significant challenges of sustainable biodiversity management in Bangladesh.

New developments in Artificial Intelligence (AI) and computer vision have presented radical opportunities to biodiversity conservation. The idea of deep learning has been demonstrated to be very useful in image classification and object recognition with the help of Convolutional Neural Networks (CNNs). These models are able to learn and extract visual features like color, texture and shape using their raw features, which is superior to the traditional features that are to be extracted manually. Although the genres are advancing throughout the world, a significant gap in research and implementations as far as Bangladesh native freshwater fish is concerned can be observed. Current models and data are largely confined to marine or commercial species of fish, and a pressing gap in local biodiversity research. Thus, we are in an urgent need of a tailor-made, AI-powered system capable of successfully identifying endangered or native fish species based on pictures that were taken under diverse and in-flux circumstances, e.g., marketplaces, riverbanks, and natural habitats. The general issue that this study aims to solve is that there is no automated, precise, and available mechanism to determine the endangered species of freshwater fishes in Bangladesh. In the absence of such a system, local fishermen and consumers are not aware and conservation efforts are impeded by delays, absence of reliable information, and inadequate knowledge. This study will overcome such a technological gap by designing and training deep learning models, especially a custom CNN, found to detect six native species that are not common. A smart machine-vision-oriented recognition system can be a viable conservation measure whereby investigators, policymakers and communities can easily identify endangered species and gather useful information on populations, and create awareness of sustainable fishing behaviors. Finally, this issue must be solved to maintain aquatic biodiversity in Bangladesh, extensive food security, and balance of the ecological system in the future generation.

1.3 Outcome of the Project

The fish recognition system proposed as a machine vision will offer a new and viable solution to a major problem of recognizing the endangered fish species in the freshwater of Bangladesh. This project provides the basis of precise, scalable, and automated biodiversity monitoring through the incorporation of artificial intelligence (AI), deep learning, and image processing functions into a single system. The model is based on a Custom Convolutional Neural Network (CNN) architecture that has been properly trained and fine-tuned on 8,676 high-resolution images of six native fish species that are rare: Tara Baim, Deshi Magur, Kajoli, Khoilsha, Koi and Tengra. The system also gains an overall accuracy of 97.42, higher than pre-trained models, including ResNet-50, VGG-16, InceptionV3, and EfficientNet-B0. Its application is an example of how deep learning can be successfully used to perceive the subtle morphological and visual differences across different species, even when the images are taken in the complex real-life environment with changing lighting, backgrounds, and orientations.

The type of project deliverables does not just stop at classification performance but technological breakthroughs that would help conserve biodiversity. The resulting model assists in the identification of any endangered fish species in a real-time manner, and therefore the model is an advantageous tool to the researchers, conservation officers and locals to track the population of fish species and intake of the endangered fish which is illegal or undesired. In addition, the strategy provides a platform to scalable and deployable solutions where the potential solutions that could be implemented by fishermen and communities directly include mobile or web-based applications. This places technological resources in the custody of the users so that they can identify and inform about the presence of rare species within their locality that can further improve the decisions on conservation management based on data-intensive choices. The application will be able to adapt to the new image information as time goes on through the flexibility of AI and more accurately represent the image, making it more generalized.

The theoretical and practical outcomes of this project are:

- **Improved Recognition Accuracy:** The system receives an improved classification accuracy with the aid of Custom CNN that is trained on various real-world images and avoids the false recognition of the endangered species.

- **Best Generalization and Scalability:** The highly powerful architecture of the system ensures the system is consistently working on the varying environmental and image conditions to enable application of biodiversity on a large scale.

- **Real-Time Monitoring Capability:** The framework will enable detection of fish in the market and field environment in real-time to enable immediate conservation.

- **Conservation Awareness:** The project will also be beneficial in terms of environmental awareness because it will offer a convenient form of AI-based recognition and the end result is that the consumer demand in the endangered fish species will decrease.

Data-Driven Decision Support: The model will offer meaningful and organized data, which can be applicable by the policymakers, researchers, and environmental organizations in biodiversity assessment and long-term ecological planning.

Overall, this project demonstrates that AI and computer vision may be used to preserve biodiversity and sustainability in fisheries and prevent the extinction of endemic fishes. The integration of deep learning and environmental conservation into a single system makes the new ecological technology the next standard that will be applied in Bangladesh. It focuses on how intelligent, automated recognition systems can be used to support the linkage between scientific innovation and practice, and ensure a sustainable interaction between human livelihood and aquatic ecosystems.

2 Literature Review

Recent studies have revealed that AI and computer vision algorithms possess an enormous potential for automatic detection of fish species and solving biodiversity problems. Also due to the rapid disappearance of freshwater biodiversity in such areas as Bangladesh, where several native species are already endangered, the AI-based solution can provide an efficient alternative to the manual identification method. Deep learning, and especially Convolutional Neural Networks (CNNs), is now a strong approach to extracting fine details of an image and making valid classification even when operating in a high-stress setting. The world and Bangladesh have witnessed a number of research works on different architectures, hybridized designs and image based models attempting to advance the accuracy, robustness and flexibility of fish recognition systems.

Beyond coral reef and marine applications, several large-scale real-world underwater datasets and deep learning benchmarks have emerged in recent years, enabling more robust evaluation of fish detection and classification models under natural conditions [19],[20],[21].

2.1 Existing Methods

To determine the classification of 13 native Bangladeshi fish species based on habitat-level data and species-level data, a hybrid deep CNN--machine learning model was created. The method found a maximum accuracy of 100 percent in controlled laboratory conditions. Nevertheless, the system did not undergo real-time and natural environment testing, which restricts its ability to be generalized [2].

Classification A ResNet U-Net was used to classify fish in the Amazon Basin and reached a genus-level accuracy of 97.9%. This paper demonstrated that deep learning can be used to conduct ecological surveillance on a large-scale, but it was not focused on rare or endangered species but only on broad taxonomic groups [3].

A hybrid fish detection system that combines Gaussian Mixture Models (GMM), OpticalFlow, ResNet-50 and YOLOv3 was found to have 91.6-95.4 percent accuracy under difficult underwater conditions. The results indicated that, under changing water conditions of light and turbidity, multi-stage architectures have the capability to identify and monitor fish [4].

The decision rules along with a GoogLeNet CNN were applied to classify over 20 coral reef fish species out of over 900,000 underwater images. The study established that CNNs would be able to outperform human specialists in extensive ecological datasets with an accuracy of 94.1. However, its emphasis on the marine environment restricts its direct relevance to freshwater ecosystems like the Bangladesh one [5].

The process of measuring morphological features of fish was done with less than 5% error by deep learning and image processing techniques. This indicated that the more sophisticated CNN-based techniques had the ability to automatically extract physical features implying their use in automated fisheries management [6].

A cascaded bio-inspired CNN based on freshwater fish classification in Bangladesh had an accuracy of 98.7 per cent on eight native fish. Although the results were strong, the data was not diverse and the data were collected under controlled conditions hence limiting its applicability in the field [7].

A machine vision based recognition system with handcrafted features and Principal Component Analysis (PCA) on six species of freshwater fish in Bangladesh was found to give an accuracy of 94.2. Although it was fundamental, this method relied so much on lighting, pose, and uniformity of the background [8].

An architecture of hybrid CNN with optimized squeeze-and-excitation-based methods achieved 97.9 percent accuracy on multi-species. Nevertheless, pre-segmented images limited the model to working in unstructured or messy backgrounds [9].

The efficacy of the convolutional neural network is investigated by the EfficientNet-B0 architecture as the approach to improving the recognition efficiency of native freshwater fish. Constructed on the mobile inverted bottleneck convolution (MBConv) blocks and the compound scaling, the model was sufficiently able to balance depth, width, and resolution, in addition to having a lower computational cost. EfficientNet-B0 performed well in comparative analysis, with high accuracy, recall, and F 1-score, up to 95% total accuracy, and can be run quickly and deployed to mobile devices. The fact that it is scalable and lightweight implies that the model can be used in real-time and field-based fish monitoring systems [10].

Similar high performance has been reported on large wild underwater datasets using transfer learning and modern CNN backbones, achieving 96–98 % accuracy even with highly imbalanced and low-quality images [22],[23],[24],[25]. Real-time detection frameworks combining YOLO variants with attention mechanisms have also shown promising results in complex tropical freshwater and brackish environments [26],[35].

A worldwide survey made the problem of trophy fishing on its endangered species known. These unsustainable practices were discovered to hasten the decay of the big and easily exploitable fish populations. The study highlighted the necessity of conservation-oriented management models that are able to combine ecological information and technological treatments to reduce overexploitation. These results highlight the importance of smart recognition and follow-up frameworks, including fish recognition systems based on deep learning, to defend the endangered freshwater fish before they are seriously diminished [11].

A comprehensive study of industrial fisheries demonstrated that over 90 endangered fishes and invertebrates species are still taken in large-scale commercial fisheries. This is a frightening sign of increasing ecological danger facing aquatic life at the global level. The likelihood of many species becoming extinct is likely to heighten without the application of technology in surveillance as well as stringent regulation. These insights enhance the importance of the deployment of automated recognition systems that rely on deep learning as the means of enhancing the processes of sustainable fisheries management and biodiversity protection [12], [13].Recent reviews confirm that deep learning has become the dominant paradigm for automated marine and freshwater species monitoring, with hundreds of published models since 2018 [27],[28].

Real-time capable models combining YOLOv7/v8 with dense prediction heads and knowledge distillation are now achieving >60 FPS on edge devices while maintaining high accuracy in complex riverine environments typical of Bangladesh [64],[66],[68].

All these studies together prove that deep learning and machine vision have already made a significant advancement in the ability to recognize fish species in different environments. Nevertheless, the majority of currently available systems are based on small and curated data sets or are oriented towards marine and aquaculture, but not endangered freshwater species. Besides, most of the models struggle to cope with the differences in illuminations, overlapping of fishes and the complexity of natural backgrounds.Recent studies focusing on South-Asian and Bangladeshi native freshwater fish have pushed accuracy beyond 98 % even in highly turbid and low-visibility conditions using lightweight attention-based and transformer architectures [62],[65],[66],[69],[70]. Vision transformers and Swin-Transformer backbones have also begun outperforming traditional CNNs for multi-class endangered fish recognition when sample sizes are limited [67],[69].

As it has recently been shown, computer vision and deep learning approaches are becoming increasingly effective in the fields of agriculture, healthcare, and the monitoring of food quality. The areas of computer vision have been effective in discriminating local varieties of fruits using color, shape, and texture as their features and achieving automated identification and minimizing human efforts [71]. In a similar manner, CNN-based models have also demonstrated great accuracy on identifying and classifying eye diseases using retinal images, which are effective in automated ophthalmic diagnosis [72]. Deep learning models have already been successfully applied in agriculture, as they were able to distinguish potato breeds [73], hyacinth bean breeds [74], and rice leaf diseases [75], demonstrating the promise of automated crop control and plant health. In non-agricultural fields, AI-solutions have been proposed to detect peptic ulcers non-invasively [76] and recognize meat freshness [77], and that proves how versatile machine learning can be in medical care and food quality control. The necessity to choose the most efficient architectures and correctly identify the leaf and any disease is also stressed by the related comparative studies of deep learning models that are trained on multiple plant species, including jackfruit, rice, and rose [78]. Generally, these studies reflect the wide generalisation of computer vision and deep learning as solutions to efficient, automated and real-time monitoring of various fields.

Thus, this study presents a Custom CNN-based classification system with a specific application in the recognition of six native freshwater fish species found in Bangladesh, which are considered rare and regarded as a problem in real-world applications. The proposed model will provide high accuracy, adaptability and scalability by training on a large dataset that is recorded in a variety of environmental conditions, which the other models could not fill the gaps that existed. The introduction of deep learning in the field does not only contribute to the development of species recognition technology but also biodiversity preservation and sustainable fisheries management in the unique aquatic ecosystem of Bangladesh.

2.2 Gaps within the Existing Methods

Although there is a lot of advancement in the field of using artificial intelligence and deep learning in detecting and recognizing fish, there still are numerous gaps in the approaches. Previous studies have primarily been done using small datasets and isolated geographical species, which limits the extrapolation of models in various aquatic systems [2]. The majority of these, including ResNet U-Net and hybrid GMMYOLOv3 models, showed strong accuracy, but were not optimized to work in uncontrolled settings at real-time recognition [3], [4]. Such models are effective under laboratory or pre-segmented conditions, but in natural conditions involving variations of light or background noise or overlapping fish species, they tend to lose precision.

In addition, various models are more resource intensive, relying on large and pre-trained models such as InceptionV3 or EfficientNet, which are especially resource intensive and cannot be used on a mobile or field-level platform in low-resource settings such as rural Bangladesh [5], [7]. Conventional machine vision systems with handcrafted features and PCA were not only very accurate but also were not robust or adaptive to morphological differences amongst closely related species [8]. Still other models of hybrids CNN with refined squeeze-and-excitation architectures increased their accuracy but still required pre-processed or pre-segmented images, which are impractical in real-time biodiversity monitoring [9].

Even state-of-the-art models trained on large public datasets (e.g., DeepFish, Fish4Knowledge, and WildFish) still exhibit significant drops in performance when transferred to South-Asian freshwater systems due to domain shift [19],[29],[33]. Few-shot and zero-shot approaches have been proposed to mitigate limited labelled data for rare species, yet they remain underexplored for Bangladeshi endemic fish [33],[40].

The second weakness is the absence of local datasets of native Bangladeshi fish species in different environmental and lighting conditions. Models that exist tend to use data elsewhere geographically that is not representative of the local species coloration, morphology or habitat diversity. In addition, there is a gap in conservation since literature in the past has not sufficiently discussed how rare or endangered fish can be detected, with the research being centered on those that are easily accessible.

Although several new Bangladeshi-specific datasets and lightweight models have appeared since 2023 [62],[65],[70], none yet cover all six critically endangered species targeted in this study under the full range of seasonal turbidity, illumination, and background variations found across rural Bangladesh wetlands.

To overcome these limitations, our research creates a trained CNN that will be effective when used in the conditions of the real world with a balanced set of 8,676 images of six endangered species of native fish. The given solution can fill the gap by providing a scalable, computationally efficient, and high-accuracy recognition framework that would be applicable in real-world applications in biodiversity conservation and fishery management.

2.3 Issues with the Existing Datasets

There are a number of limitations that exist in the existing datasets which can be used to recognize fish and monitor biodiversity which can interfere with the creation of highly accurate and generalized models. Most of the publicly available data focus on marine or coral reef species instead of freshwater species endemic to Bangladesh which introduces a mismatch in domain that decreases model adaptability [3], [5]. In addition, such datasets have a tendency to be small, sparse, and devoid of environmental variation, not accounting for the entire gamut of natural stimuli (e.g., variation in lighting, turbidity of water, occlusion, and background clutter) that normally vary in the real world [4].

Numerous prior research works used synthetic or controlled in the lab images to train and test deep learning models that can be highly accurate during testing but do not transfer to real world settings [6], [8]. These data sets commonly include neat, isolated fish images on homogeneous backgrounds disregarding the complicated visual environment of the freshwater ecosystems where several fish can be observed at the same time, or part of a fish can be hidden behind other objects. Also, the asymmetry in the distribution of data with some species being overrepresented and others underrepresented creates bias, resulting in overfitting of models and low recognition of the rare or threatened species [7].

The second challenge is another significant issue is the absence of annotated datasets or verified datasets devoted to extinct species or endangered species. Due to the lack of high-quality and reliable images of rare native fish, it is hard to train the models that are able to produce accurate detections. Moreover, the current datasets, in general, overlook the morphology and coloration differences between nearby-looking species, including, e.g., Deshi Magur and Tengra, that have been mistaken with each other by pre-trained networks [2], [9].

Large-scale wild datasets such as DeepFish (over 40,000 annotations), Brackish dataset, and OzFish have highlighted the persistent difficulty of detecting small, camouflaged or partially occluded fish in turbid water, a common scenario in South-Asian rivers and wetlands [19],[20],[34]. Recent works on Nigerian and Indian freshwater species also report similar domain gaps when using models pre-trained on Western or marine datasets [39],[31].

To address these dataset-related issues, this paper develops a big dataset of 8,676 authentic images of 6 uncommon freshwater fishes in various districts in Bangladesh, and in high quality. The dataset is rich with a variety of lighting, background and positional variations in order to be robust and allow generalization. Our work presents a good base of creating effective deep learning frameworks which would improve biodiversity monitoring and conservation of fish by overcoming the issue of data imbalance and incorporation of rare species.

Even the most recent public Bangladeshi datasets remain either small (<3,000 images) or collected only in controlled ponds and markets, still lacking the extreme environmental diversity of natural river, beel, and haor ecosystems addressed in the present work [62],[65],[70].

3 System Design and Conceptual Architecture

To utilize the maximum conservation potential of the resultant developed deep learning models, a conceptual framework of the entire functioning fish recognition system is suggested in this study. The purpose of this framework is to transform real-time machine-vision data to inform information that can be utilized in the field by conservation teams to track the endangered freshwater species. In this design, the trained deep learning models: VGG16, ResNet50, InceptionV3, EfficientNet-B0 and the Customized CNN and field-deployed devices, be it mobile phones, cameras, or other portable sensing units, would be stored in the main server, and the pictures of fish are taken when carrying out market surveys, river surveillance, or ecological analysis.

Images are captured and sent to the server via secure communication link, and conducted using preprocessing and inference processes and classification. The server processes each image, standardizes its resolution, normalizes the value of its pixels and sends the image through the selected deep learning model to arrive at species predictions and confidence scores. The system then converts these predictions into actionable information such as detecting the occurrence of endangered species, rare phenomena or presence of unusual distribution patterns of a species. The findings are immediately sent to conservation officers, researchers or policy mandates and prompt decisions on the ground are made.

The presented framework integrates the image acquisition and server-side deep learning inferences and displays the final outcomes of these processes into a single workflow so that the threats to Bangladeshi freshwater fish could be quickly, accurately, and scaled-uply monitored. This theoretical model emphasizes how the developed research can be transformed into a workable conservation tool that could sustain the biodiversity protection works in the markets, rivers, and off mountain field environments.

3.1 Proposed Operational Workflow

The total design establishes an important communication channel between the data collection in the field and the capacity of the central server to be analyzed. Such an integrated framework as shown in Fig. 1 allows easy flow of visual information through the field devices to the total System where real-time processing and interpretation are allowed. The system guarantees timely dissemination of actionable insights through linking the process of data acquisition and complex analysis, which will serve to timely identify the species and help conservation management to make the best decisions based on the acquired data, and also improve the efficacy of the field-based monitoring processes.

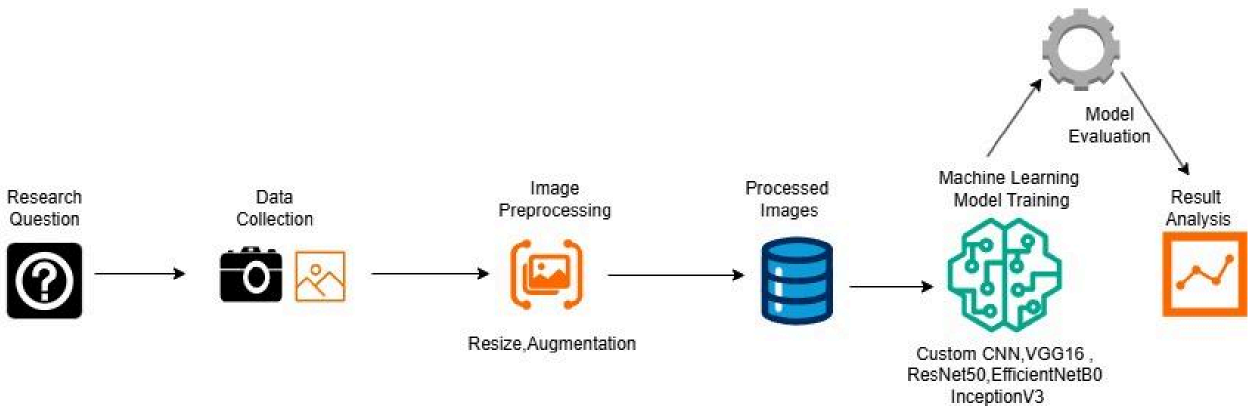


Fig. 1: Proposed operational workflow

3.1.1 Communication and Data Flow

Other details in Figure 2 show the general architecture of the system that comprises the proposed fish-recognition system, with a mobile app that communicates effectively with a cloud-based agricultural expert system to provide correct species recognition. The installed mobile application on the user side allows fishermen, farmers or ordinary users to scan fish images and transmit them to be recognized. A communication network is used to transmit such images and update requests wirelessly to the remote server.

The expert system on the server-side comprises three fundamental parts, namely, User Interface, Learning Module, and Knowledge Acquisition Model, which are interconnected in the form of a central Inference Engine. The inference engine deals directly with the Knowledge Base which contains the information concerning the species, the features learned, and the updated decision rules. Upon receipt of new data or update signals by the mobile app, the server interprets the data in the learning module and updates the knowledge base and sends such recognition files or responses back to the user device.

Such two-way communication can be used to guarantee that the application is constantly updated, adaptive and able to deliver reliable fish recognition in a variety of environmental conditions as defined by the workflow in Figure 2 above.

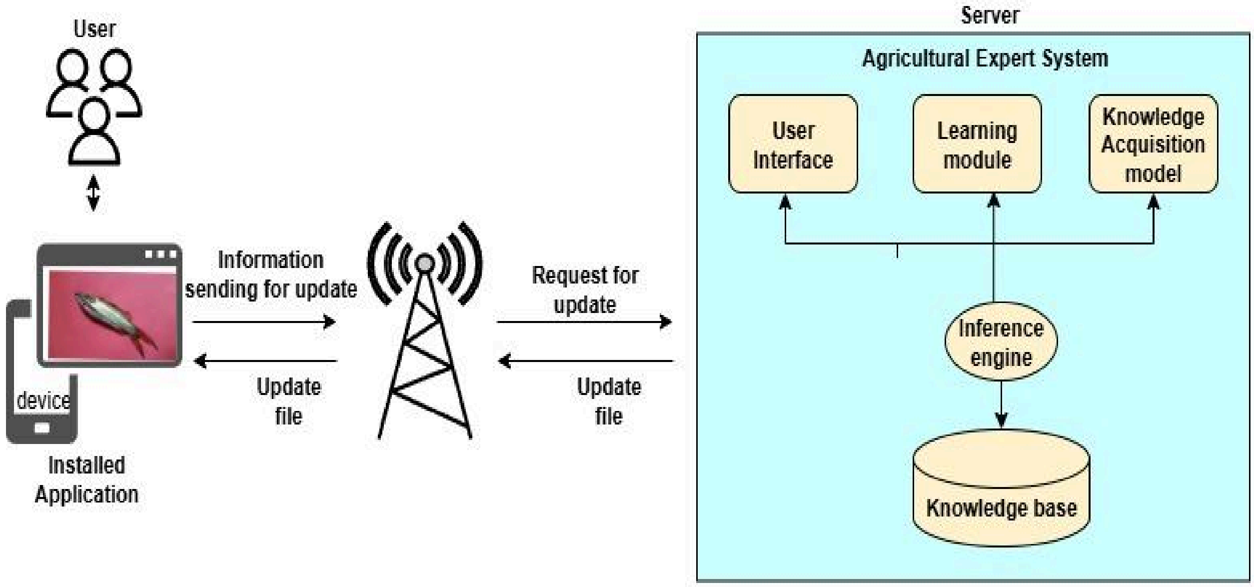


Fig. 2: Proposed system structured design

3.1.2 Modeling Latency and Bandwidth Trade-offs in Edge-Cloud Integration

Remote freshwater habitats and busy, congested fish markets present natural constraints such as latency and limited bandwidth, which can hinder the consistent performance of a cloud-dependent system. To operate reliably under these conditions, the proposed system must evolve into a Hybrid AI architecture that intelligently distributes inference tasks between the centralized server (cloud) and the user's device (edge).

The existing system has field devices which capture pictures and send them to the central server where they are processed. Cloud-based inference is characterized by high computational power, i.e. fast analysis, but the use of every cloud processing only pushes some dependency on connectivity that should be high-bandwidth and stable, which is not easily found in rural Bangladesh. This leads to the fact that only a hybrid strategy can ensure a consistent field performance of purely cloud-based approaches.

This limitation is overcome by the Hybrid AI approach, which uses criteria to selectively process information on the device. When circumstances permit, e.g. by having a mobile-optimized implementation of the Custom CNN or an efficient version of a detection variant, lightweight inference may be directly executed on the mobile device. When the model is very confident or when the network is of poor quality, on-device predictions can be produced. On the other hand, low-confidence prediction images, complicated visual images or those that need additional checking are automatically passed over to the central server to be processed further. This is a dynamic allocation of tasks that enhance the speed of response, minimize reliance on cloud connectivity, and save device energy.

Another important fact is the efficient bandwidth utilization. It is very inefficient to transmit full-resolution raw images which can be several megabytes each. With a workflow that is object-detection-based, the system can send only the necessary information, e.g. compressed feature vectors, coordinates of a bounding box, or downsampled image blocks rather than full images. This saves important visual information and also saves a lot of bandwidth to the network.

The adoption of Hybrid AI decision logic is more than a technical improvement: it is necessary to implement the practical application of AI in the field conditions.

In many remote environments, network availability rather than hardware capability is the primary constraint. By enabling selective on-device inference and optimizing data transfer, the system can deliver timely and actionable feedback to conservation officers, researchers, and community members. Shifting portions of the workload to edge devices also reduces the need for continuous cloud computation, enhancing energy efficiency and supporting a more sustainable and resilient system architecture capable of meeting the demands of conservation monitoring in diverse and challenging environments.

3.2 Server-Side Fish Recognition and Processing Components

The server-side topology offers the physical CPU of the proposed fish recognition system that carries out all the high level operations of the image classification, model implementation and generation of the results. Unlike the traditional expert-system frameworks, which rely on the rule-based reasoning, the present paper is entirely constructed around the principles of deep learning pipelines; therefore, the server is arranged around the factors that support the processing of great amounts of data, neural network inference, and optimizing the work of the specified neural networks. To achieve that, the system comprises four and inter-connected modules that act in a chain to transform raw images into reliable species prediction.

All the incoming pictures enter through the Data Processing Module. This component performs vital preprocessing operations, like downsizing to standard input dimensions, noise removal or bad samples, pixel value normalization and when needed augmentation. This module is effective in that it standardizes the preprocessing process so that all the images uploaded to the server will be of the quality of the convolutional neural network models required and therefore increase the accuracy and robustness of the classification.

The deep learning structures used in the present study including VGG16, ResNet50, InceptionV3, EfficientNet-B0, and the Customized CNN are deployed and managed as the Model Management Module. This is a module which stores the weights of a trained model, version control and forward-pass predictions which are run each time a new image is posted. It is the one that selects the appropriate model, it inserts the model into a memory and ensures that inference is done optimally with optimized hardware.

Training and Fine-Tuning Module helps to update models in case new datasets are supplied or it is necessary to attain a higher level of performance. Although the main training process is performed offline, this module enables the system to retrain or fine-tune models by using an extended dataset, updated hyperparameters, or new augmentation methods. It also stores training records, validation results, and check points necessary such that the models become long-term reliable.

Finally, there is the Result Interpretation and Output Module which then transforms the raw output probabilities generated by the neural networks into humane species-level predictions. The component also summarizes the scores of the classes, reconciles the blurry classifications and displays the findings in formats that are easy to visualize, report or include in the downstream conservation processes. It can also generate measurements of confidence as well as confusion and performance overview, which is then used in comparative analysis.

A combination of all these four server-side aspects forms a scalable, integrated and automated processing component capable of large scale fish image processing, numerous state-of-the-art detection models and creation of sound real-time species recognition of endangered and rare freshwater fish in Bangladesh.

3.2.1 The Learning Module and Knowledge Acquisition

The Learning Module serves as the operational environment for all core deep learning models integrated into the system. This includes the high-performing Custom CNN as well as the pre-trained architectures ResNet-50, VGG-16, Inception V3, and EfficientNet-B0. Its primary responsibility is the real-time classification of incoming images, ensuring that every captured sample is processed both quickly and accurately, providing immediate feedback for field-based monitoring tasks.

In addition to this aspect is the Knowledge Acquisition module that is needed in keeping the system flexible and effective in the long run. The module keeps adding all the newly acquired field pictures, observations of the model performance and the feedback given by the experts into the Knowledge Base. This organized continuous refinement process ensures that the deep learning models are always stable and reliable, even when the environmental conditions, morphology of fish or ecological factors change with time. This type of flexibility helps avoid degradation in accuracy in the long run that may happen when the actual data does not match the properties within the initial training sample.

The dataset that was collected to conduct this study was used to establish the initial learning basis of the system based on 8,676 high-resolution images. Such a rich dataset was the required baseline training of the models so that the species could be recognized and the system could be ready to be used in the real-world field where the conditions could be highly variable and harsh.

3.2.2 Knowledge Base and Inference Engine Functionality

The primary element that is engaged in interpreting and analyzing raw deep-learning outcomes into a form that can be comprehended as a field-relevant information is Result Interpretation and Analysis Module. The system is not always based on knowledge base or inference engine but it is a data-driven layer of reasoning that interprets the probabilistic results of the trained CNN models. After an image is processed and recognized by the Learning Module, this component evaluates the confidence scores and matches the predictions with species-level descriptors that were trained in the models to measure the results and classifies the outcome in form of a set of hierarchy and comprehensible to humans. It will provide any prediction like a high-confidence finding on Kajoli, Tengra, or Khalisha supported by a consistent visual representation retrieved directly as of the internal feature representations on the model.

Besides such an interpretation process, the operational meaning of each prediction is also made available by the Contextual Output Layer. Instead of the rule based reasoning, the experience of this layer is based on the behavior of the model, data details and ecological patterns that can be known intuitively via deep learning. Based on the example, when the system identifies a species as a rare or endangered species in the data studied, the output layer indicates the significance of such species by having contextual information that a species is a rare species, a low-occurrence class, or requires monitoring. Such hints within context may then convert an otherwise result of classification into a convenient field level indication to the researcher, surveyor or conservation worker.

The result of such a combination of interpretation processes is a system that may be more than an image classification system, it may be a helpful monitoring and decision-support system. The framework offers ecological awareness, encourages sustainable harvesting practices, and helps to avoid exploiting the vulnerable species of freshwater without prior knowledge by displaying the names of species and levels of their confidence, results of visualization (e.g., heatmaps), and background information. Improbably, the most important thing about the system is that it can be scaled, flexible, and it can automatically become better when new data is added to the system, as the rationale is derived by deep-learning analysis rather than by the rules manually written down.

The proposed framework will ensure that the process of identifying endangered fish does not just make the recognition of endangered fish computationally valid but also operationally relevant to the field, the study of biodiversity, and the conservation activities in Bangladesh.

4 Research Methodology and Evaluation Framework

The experiment design involved a close and systematic attitude to all the steps of data processing and model testing. All of the steps were to establish a reliable framework that would help to conduct the evaluation and make significant improvements to the system. This organized procedure involved the study in an attempt to be sure that the machine vision models were well prepared in order to handle the significant challenges of classifying the biological features present in the real world, especially when the features are found in environments that are not predictable and easily manipulated.

4.1 Data Acquisition, Processing, and Partitioning

The database used in this study constitutes 8,676 high-resolution images of six native freshwater fish species in Bangladesh: Tara Baim, Deshi Magur, Kajoli, Khalisha, Desi Koi and Tengra. The photographs were taken in different local markets using different districts to ensure the natural variation of the background settings, lighting conditions, position of fish and the availability of single or multiple fish in the frame. The broad sampling guaranteed great real-life diversity and contributed greatly to the overall generalization ability of the model. This was done by a systematic preprocessing pipeline to refine and standardize the dataset prior to training. All pictures have been downsampled to 224x224 pixels in order to ensure the same input format in all deep learning models. To decrease overfitting and balance the dataset, a number of augmentation methods, including random rotations (20 + or -), horizontal flips, brightness and contrast, and zoom-crop were used. The number of samples per species was 1,446 after augmentation, which made the data perfectly balanced. The whole data were separated into training, validation, and test subsets in 80:10:10 ratio, which included 1,156 training images,144 validation images and 146 test images of each species. The training set was used to learn the model, the validation set helped to tune the hyperparameters and early stop, and the test set was used to make an objective judgment on the end model performance. This systematic method of acquisition, preprocessing and partitioning guaranteed reliable, consistent and reproducible results in all deep learning models applied in the investigation.

4.1.1 Short Description of Each Fish Species and Their Importance

It is also necessary to comprehend the nature and ecological significance of every native fish species so that the need to conserve them. They do not only enhance the biodiversity of the locality, but they also contribute to the local livelihoods and food, as well as the cultural identity of the freshwater communities of Bangladesh. It begins with a concise description accompanied by images in Figure 4, highlighting the unique value of each fish species and underscoring the urgent need to protect them from further decline.

4.1.1.1 Tara Baim (Spiny Eel)

Tara Baim also known as Spiny Eel locally called Golchi and Kotabaim, scientific name Macrognathus aculeatus has once been widespread in floodplains, rivers, and wetlands in districts like Mymensingh, Sylhet, Sunamganj and some areas of the north of Bangladesh. Its population has reduced drastically because of the loss of wetlands, pollution of rivers, and overharvesting. Where it was so plentiful formerly, it is to-day only met within a few limited spots, or not at all. Tara Baim is worth studying as its existence is a sign of a healthy wetland ecosystem and its decline is the indicator of a degradation of the environment.

4.1.1.2 Deshi Magur (clarias batrachus)

Deshi Magur locally called Jiol, Mojgur, Mochkur, and Jagur scientific name Clarias batrachus, English name Walking catfish was traditionally cultivated in shallow ponds, irrigation canals and rice plots in nearly all the districts of Bangladesh, particularly in Jessore, Gopalganj, Pabna and Tangail. It is now much less abundant because of uncontrolled fishing, the invasion of hybrid farmed types, and water pollution. There is a change in native Deshi Magur to hybrid and imported species in many local markets. Conservation-based identification is important since this fish is both nutritionally and culturally valuable to the rural people.

4.1.1.3 Kajoli (Ailia coila)

Kajoli's other names are Baspata, Bet Laiya, scientific name Ailia Coila. The English name Gamgetic Ailia is widespread in the large river systems, particularly in the Chandpur, Barisal, Rajbari, and other regions related to the Padma-Meghna river system. Siltation, erosion of riverbank and destructive fishing techniques have led to extreme reduction in population over the years. Kajoli is currently located in the isolated river channels and on the basis of some haor areas. This species needs to be monitored due to its high sensitivity to changes in the environment, and the fact that it acts as an indicator of the poor health of rivers.

4.1.1.4 Khalisha (Trichogaster fasciata)

Khalisha, locally referred to as Cheli and Khailsha, scientific name Trichogaster fasciata English name Banded gourami has been historically prevalent in the wetlands of the Sylhet, Netrokona, Kishoreganj, Gopalganj and other low-lying districts. Its population has reduced as a result of pollution, application of pesticides in farm lands and drainage of natural wetlands to construct urban areas. Nowadays, Khalisha can only live in secluded ponds or safe water reserves. The reason why this species should be given protection is not only due to the biodiversity but also the maintenance of the traditional currently used diets by many rural families.

4.1.1.5 Desi Koi (Anabas testudineus)

Desi Koi, scientific name Anabas testudineus, English name Climbing Perch was highly present in most parts of Bangladesh such as Rajshahi, Khulna, Rangpur and Dhaka divisions. Nevertheless, the native type has been eroded to rarity in most of the districts due to overfishing, loss of habitat, and large scale farming of hybrid Koi. Natural waters of certain regions have become almost devoid of the species. Desi Koi should be conserved due to its cultural importance and the fact that it forms an important part of Bangladesh identity in terms of freshwater.

4.1.1.6 Tengra (Mystus tengara)

Tengra known locally as Guitta Tengra, Bujuri, and Bojuri, scientific name Mystus tengara, English name Striped Dwarf catfish is a fish historically observed in the rivers, floodplain canals, and beels of such districts of the country as Bogura, Sirajganj, Jamalpur, and more extensive areas in the north and central regions of Bangladesh. The current state has diminished its appearance because of decreasing water quality in rivers, over-fishing, and alteration in seasonal water flow. The species are still available in certain markets albeit in reduced amounts compared to the previous ones. Tengra is an economical study since it plays a role in the environmental conservation of the riverine food webs.

4.1.2 Dataset Composition and Image Augmentation

The study utilized a comprehensive dataset comprising 8,676 high-resolution images across six rare native fish species: Tara Baim, Desi Magur, Kajoli, Khalisha, Desi Koi, and Tengra. More importantly the data was taken in different districts and markets in real-life situations with different backgrounds, variable lighting and various numbers of fish per frame. This was done deliberately and it ensured that the trained models which would be developed would have good generalization properties in order to be used in the real world by being outside the laboratory environment.

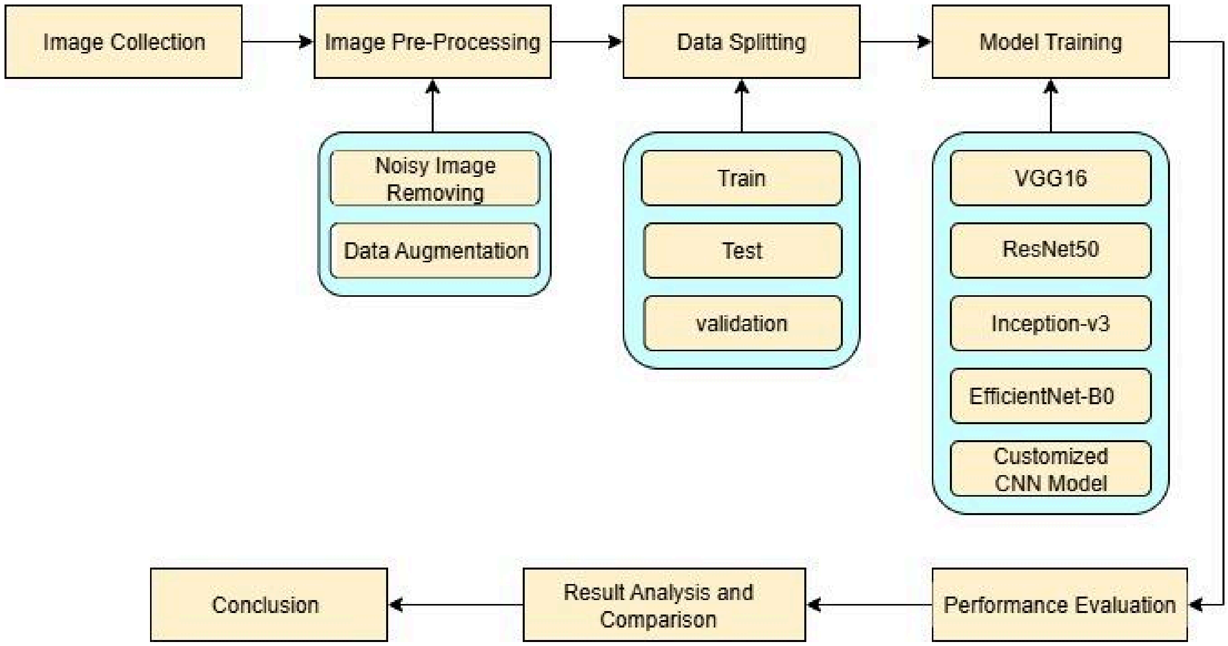


Fig. 3: The entire working procedure of this research study

The workflow diagram figure 3 shows the entire methodological pipeline that will be applied in this study with regard to machine-vision-based extinct fish recognition. It starts with Image Collection where the raw pictures of Bangladeshi freshwater fish are collected in various sources to reflect natural differences in the environment and appearance. The collected images are then subjected to Image Pre-Processing stage which involves two main steps namely noisy image removing where low-quality, blurred, or unclear images are filtered and data augmentation where artificial diversity to the dataset is increased by the addition of transformations such as rotation, flipping, and brightness. The images are then passed through the Data Splitting step after preprocessing, which is split into three subsets namely train, test, and validation where they are trained on one part of the data, tuned on another section, and fairly evaluated on an unknown section. This is followed by the Model Training step where five deep learning models, VGG16, ResNet50, Inception-v3, EfficientNet-B0, and a Customized CNN Model are trained to learn to discriminate the six fish species. Performance Evaluation comes after training and in this process, the outputs of the models are evaluated by comparing them to metrics like accuracy, precision, recall, and F_1-score to ascertain its reliability. The results of this evaluation are passed into the Result Analysis and Comparison step where a comparison of all the models is made to determine the strengths, weaknesses, and the most effective architecture. This is finally rounded off in the Conclusion stage, which gives a summary of the findings and also outlines the best method to use in identifying rare and endangered species of freshwater fish in Bangladesh.


(a) Tara Baim (তারা বাইম)


(b) Desi Magur (দেশি মাগুর)

(c) Kajoli (কাজলি)

(d) Khalisha (খলিশা)

(e) Desi Koi (দেশি কই)

(f) Tengra (টেংরা)

Fig. 4: Representative images of various extinct fishes (a) Tara Baim; (b) Desi Magur; (c) Kajoli; (d) Khalisha; (e) Desi Koi; (f) Tengra.

Starting with 1,320 initially captured images, a detailed image augmentation pipeline was applied to expand and balance the dataset shown in table 1. Through a sequence of controlled transformations, the dataset was increased to 8,676 images, providing 1,446 samples for each of the six classes. Such balance is crucial to good multi-class classification because it can minimize the bias to one specific class and constitutes one of the most useful control mechanisms against overfitting the model in the training phase.

Table 1: Detailed Analysis of Image Acquisition and Augmentation

Class	**Captured**	**Augmentation Parameters and Criteria**	**After Aug.**
Tara Baim	250	Rotation = (-20°, 20°); brightness = (0.7,1.3), p=0.5; contrast(0.7,1.3), p =0.5; zoom = (0.8,1.0), p=0.5; horiz.flip = 0.5; fill = nearest; rescale = 1/256	1446
Desi Magur	200		1446
Kajoli	220		1446
Khalisha	200		1446
Desi Koi	210		1446
Tengra	240		1446
Total	**1320**		**8676**

4.1.3 Preprocessing and Data Partitioning Strategy

All the 8,676 images have been equally scaled to 224*224 pixels before the training process so that the images fit the standard input format in the Convolutional Neural Network models. The purpose of this preprocessing was to ensure that the network was fed images of the same size, but it is also significant to the stable gradient flow and the ability to extract comparable features among samples. After resizing the images, an augmentation pipeline was enlarged and designed critically. It is a creative process that applied random rotations, horizontal flips, spatial elimination and a diverse collection of brightness and contrast modulations deliberately. These photometric adjustments were considered as a core component of the methodology since they forced the models to emphasize on shape, texture and structural features instead of depending on certain lighting features. In the real world of the field, lighting may change significantly across the day, and across habitats, and thus it was necessary to make the models reliable in unstable lighting conditions to be deployed in the real world.

Once the images were prepared, the dataset was broken down into three distinct well-defined subsets to remain on track of an evaluation process displayed in table 2. Eighty percent of the images were reserved as training and provided the models with a wide-ranging and varied range of training examples. The remaining 10 percent was reserved to validation which was paramount in hyperparameter tuning by giving feedback which was used in making adjustments without exposing the models to the actual test data. The remaining 10 percent was used as a totally untouched test set. This last part was the scale against which the actual generalization would be determined, as it would be possible to assess the performance of the models on the data which never affected the training or tuning procedure. This was a deliberate segregation that served to make sure that the reported accuracy was due to actual predictive power and was not because of being familiar with the dataset.

Table 2: Detailed Analysis of Data Partitioning

Fish Name	**Training Image Usage**	**Validation Image Usage**	**Test**	**Total Images**
Tara Baim	1156	144	146	1,446
Desi Magur	1156	144	146	1,446
Kajoli	1156	144	146	1,446
Khalisha	1156	144	146	1,446
Desi Koi	1156	144	146	1,446
Tengra	1156	144	146	1,446

4.1.4 Advanced Augmentation Protocols for Environmental Robustness

The protocol actually increased augmentation was originally based on conventional enhancement methods, though field experimentation soon demonstrated the necessity of a more challenging and environmentally based method. Freshwater ecosystems also bring extensive visual disturbances that are capable of distorting the appearance of fish in a manner that the model should be ready to deal with. The presence of turbidity, suspended solids, low light levels, variable water turbidity and random reflections are some of the factors that make the quality of the image significantly worse than expected in the conventional laboratory dataset. The augmentation strategy was extended and redesigned to simulate the visual difficulties that are usually experienced in actual monitoring conditions to guarantee credible performance of it under such circumstances. This change enabled the model to learn upon a more diverse and realistic collection of distortions causing the learned properties of the model to be more reliable and applicable in a real world environment.

- **Turbidity and Pollution Simulation:**
  Most of the freshwater environments in Bangladesh are often polluted and have dense water that contains a lot of sediments, the water often impairs visibility and removes the natural color and brightness of fish. The dataset was further enhanced to combat these problems with additional augmentation steps which were used to model the scattering of light, low visibility, and washed-out color tones by suspended debris and chemical contaminants. These simulated effects were useful in assisting the model to adjust to circumstances in which color clues are unreliable or completely lost. This led to a system that tried to find fish by more predictable features like the outline of shapes, the placement of the fins and the proportions of the overall body instead of bright backgrounds or clean backgrounds that are unlikely to occur in the actual implementation.

- **Geometric Robustness and Feature Isolation:**
  The pipeline also had the added features of the Affine Transformations and Neural Filters to better capture the model capabilities of interpreting fish at various angles, distances and positions within the frame. Affine Transformations brought with them moderated changes in rotation, skew, scale and partial occlusion. This was especially crucial since the dataset contained the images of single and multiple fish, frequently taken in different poses. A second degree of complexity was introduced by the Neural Filters via the use of calibrated blurring and contrast reduction. These effects caused the clarity of fine color gradients to be lower, and required the Custom CNN to use more stable and shape-based indicators and texture-based indicators. This was in direct response to previous classification challenges when the classification muddled the species that share

a similar color palette but whose structure was different.

Using such protracted augmentation plans the model is much more robust and can be used successfully in freshwater where the visibility is poor. The refinements are made to guarantee the performance to be reliable in a variety of lighting situations, water conditions and geometric variations, which eventually makes the system more adapted to field-level conservation and monitoring uses.

4.2 Model Implementation and Architecture

A trained Convolutional Neural Network (CNN) was developed and a pre-trained network was used to categorize fresh water fish. The construction of the network was aimed at acquiring some visual features of the species, i.e., the form of the fins, the contours of the body, the arrangement of the texture and the pattern of the scales. These features are usually hard to detect and to be categorized properly, a person ought to recognize them.

The network starts by having a series of convolutional layers with each layer having 32 to 512 3×3 filters. The input image is scanned using these filters in order to detect local features of the image such as edges and patterns. The size of the input, the size of the filters, the stride and any padding determine the output feature maps. Dimensions resulting to the determination can be estimated as follows:

$$A_2 = \frac{A_1 - S + 2K}{Z} + 1 \quad (1)$$

$$B_2 = \frac{B_1 - S + 2K}{Z} + 1 \quad (2)$$

where $A_1 \times B_1 \times D_1$ is the dimension of the input image, S is the filter size, K is the padding, and Z is the stride. These computations show the size of the feature maps when the filters are shifted in the image.

After every convolution a ReLU (Rectified Linear Unit) activation function is used. ReLU enables the network to capture complicated patterns by giving nonlinearity, but ignoring negative values. The features are then reduced in size using the use of max-pooling layers to preserve the most significant features. The sizes of the output of the pooling layer are calculated as:

$$A_2 = \frac{A_1 - S_s}{Z_s} + 1 \quad (3)$$

where S_s is the size of the pooling window and Z_s is the stride. The network is made less sensitive to small differences in the input images by pooling and the strongest responses are highlighted.

The network also has dropout layers at varying levels to avoid overfitting. In training, some of the neurons are randomly silenced, and this makes the network learn a larger range of features instead of memorizing certain patterns of the training images.

After the convolutional and pooling layers obtain the relevant features, the resulting feature maps are flattened and the fully connected layer is applied. This layer is a combination of all the information that is picked off of the prior layers. The last classification is carried out through a Softmax layer, which transforms the scores of the outputs into the probability of each of six fish specimen:

$$\text{SoftMax}\,(Z_i) = \frac{e^{Z_i}}{\sum_{j=1}^{C} e^{Z_j}} \tag{4}$$

where Z_i is the score for the i-th class and $C = 6$, total number of species. The prediction is taken to be the class with the highest probability.

The idea of this custom CNN is to identify minor differences within species, and to respond to changes in lighting, angles and backgrounds. The network is able to classify fish via a structured method by utilizing several convolutional layers, pooling, dropout and Softmax classification. Its architecture enables it to draw the important features of the images in a standard and sturdy way even in adverse environments.

The network, on the whole, is able to record fine and detailed visual information and thus offers a customized solution to the freshwater fish recognition. The layered feature extraction, dimensionality reduction and probabilistic classification make the model capable of distinguishing between species with high accuracy and also it is resistant to the changes in the input data.

4.2.1 Transfer Learning Model Fine-Tuning

Four established deep learning architectures ResNet-50, VGG-16, Inception V3, and EfficientNet-B0 were adapted for this system using transfer learning. In this method, the models retained their pre-trained weights, which were originally learned from the large-scale ImageNet dataset, allowing them to begin with a broad understanding of edges, textures, and general visual structures. These inherited representations provided a strong starting point for learning the more specialized patterns associated with the six freshwater fish species. After preserving the core feature extraction layers, the models were fine-tuned on the smaller, domain-specific dataset so

they could adjust their internal representations to the ecological and biological characteristics present in the images.
The initial classification heads were also removed and new task-specific output layers were introduced to make the networks appropriate to the six target classes. The restructuring of VGG-16 can be considered as one of the brightest examples of this adaptation. It was altered to include a Flatten layer to transform the convolutional feature maps into a one-dimensional vector and a dense layer with 128 neurons to allow deeper representational learning. A Dropout layer was added to enhance regularization and minimize the chances of overfitting due to the relatively small size of the dataset. The chain was completed by a Softmax output layer which produced the probability of the six classes of species.
The reason was not to be able to use ResNet-50. It has a skip-connection to assist in the continuous flow of gradients in the deeper layers of the network. The design enables the model to better train without vanishing gradients, which is especially convenient during fine-tuning of deep architectures on a specialized dataset.
These architectural refinements enabled every model to be remodeled to operate effectively in the problem domain of interest, which is fresh water fish identification. Meanwhile, they still enjoyed the high-level feature hierarchies acquired in large-scale pre-training, and this gave out models that balanced generalization capability with domain particularity.

4.2.2 Customized CNN Architecture Development

A specialized CNN was created and constructed directly to meet the visual features of the desired freshwater species. This move was motivated by the premise that a model designed to attend to these delicate morphological structures could be more performance than an architecture trained on more general, object categories through one of the datasets such as ImageNet.

The Custom CNN in figure 4 follows a structured sequence of convolutional layers using 3*3 filters, with the number of filters gradually increasing from 32 up to 256. Convolutional blocks are followed by ReLU activation to help with stable and effective learning, and Max Pooling blocks are placed between work blocks to gradually downsample the spatial data and place a particular focus on the most significant features. In order to avoid overfitting and to have a healthy learning, Dropout layers are added at major points in the network.

The last layer of the model is an Average Pooling layer, which summarizes the obtained feature maps, and then a fully connected dense layer and a Softmax layer. This last step converts the learnt representations into classification probabilities of the six target species, which makes it possible to make clear and interpretable predictions in deployment.

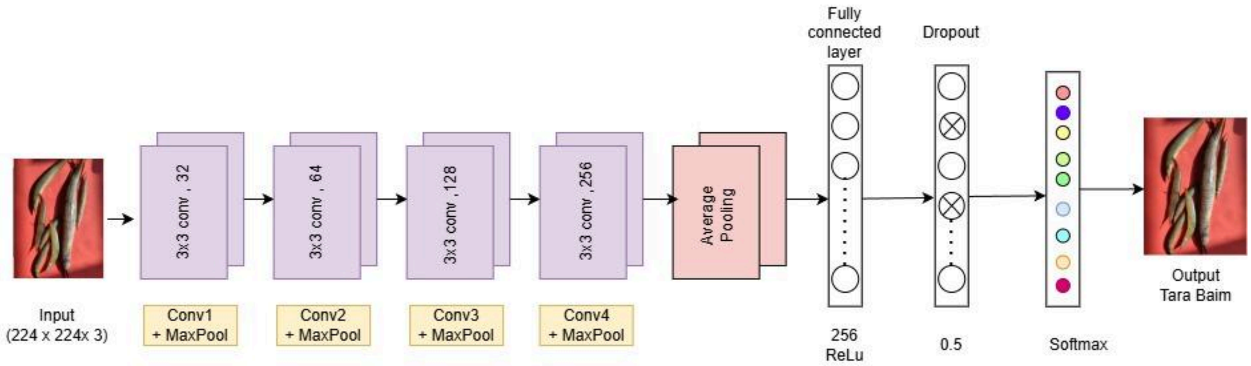


Fig. 5: Customized CNN model architecture

4.3 Performance Evaluation Metrices

In order to have a fair and comprehensive evaluation of the models, the evaluation system went way beyond reporting a single accuracy figure. Accuracy may also conceal significant weaknesses and this is particularly so in a multi-class situation where some species are easier to recognize than others. That is why the analysis was based on the metrics gained using the entire confusion matrix that presented a clear picture of the strength and weaknesses of each model. Through the analysis of the proportion of times a model created false alarms relative to the proportion of times it did not detect an actual positive case, the analysis provided a better understanding of the sensitivity and reliability balance. The methodology allowed one not only to determine the overall performance of the models, but also the consistency with which they discriminated between different species in different conditions.

4.3.1 Confusion Matrix Derivations

The evaluation framework is built around the confusion matrix (A), where each element a_{ii} represents the number of images from the true class i that were predicted as class j. This structure allows for a detailed examination of model performance across all classes simultaneously. In a multi-class context, the performance of an individual class iii can be analyzed using a one-vs-rest approach. Under this perspective, every prediction is classified into one of four possible outcomes: true positive, false positive, true negative, or false negative. This categorization ensures that every potential result is captured, providing a comprehensive view of how the model behaves for a specific class and allowing for precise calculation of class-specific metrics such as precision, recall, and F_1-score.

1. **True Positives** TP_i: This represents the number of images correctly identified as belonging to class i . It is directly sourced from the diagonal elements of the confusion matrix where the true and predicted classes match:

$$TP_i = a_{ii} \qquad (5)$$

2. **False Positives** FP_i :This count represents Type I errors (false alarms), where an image belonging to any other class j / i was incorrectly classified as class i It is calculated by summing the entries in the column corresponding to class i, excluding the true positive entry:

$$FP_i = \sum_{j=1, j \neq i}^{n} a_{ji} \qquad (6)$$

3. **False Negatives** FN_i**:** This represents Type II errors (missed detections), where an image truly belonging to class i was incorrectly classified as one of the other classes. It is derived by summing the entries in the row corresponding to class i , excluding the true positive entry:

$$FN_i = \sum_{j=1, j \neq i}^{n} a_{ij} \qquad (7)$$

4. **True Negatives** TN_i :This crucial metric counts all instances where an image did not belong to class i, and the model correctly predicted it as *not* being class i (i.e., predicting it as any other class k/i). It accounts for all correct classifications that fall outside the current class's row and column:

$$TN_i = \sum_{j=1, j \neq i}^{n} \sum_{k=1, k \neq i}^{n} a_{jk} \qquad (8)$$

4.3.2 Statistical Performance Metrics

These elements permit determining the measures that characterize the practical usefulness of the models:

Precision: Precision is a direct measure of the reliability of the positive predictions of the model and directly measures the frequency at which the model is actually correct when it makes predictions of endangered species. There is an important practical implication of this measure. The high precision makes use of conservation efforts to be founded on the correct detections, and therefore, it is unlikely that much time, effort, and resource will be wasted on false alarms. In the ecological control and enforcement of regulations, every false alarm has the potential to cause actual effects, such as the allocation of funds poorly, and the deployment of the field-workers on a mission that has little or no effect. Precision is thus of great importance since it ensures that the positive predictions by the model can be trusted and that decisions made based on the model's results can be done confidently, which will enhance the effectiveness and reliability of the conservation of the species.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{9}$$

Recall: Recall or sensitivity The sensitivity or recall measures how well the model detects all the relevant instances in the data and how well the model detects all the actual instances of endangered species that exist. Maximum recall is essential to conservation since it reduces Type II errors, or false omissions, so that rare and threatened fish populations are not missed. High recall ensures that the model includes as many true events as possible, which is necessary in order to make informed decisions, make accurate population estimates, as well as implement effective protective interventions in true ecology.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{10}$$

F_1-score: F_1-score is the harmonic mean of the Precision and Recall, which has only one figure that is accurate prediction of the positives and full detection. The harmonic mean, unlike arithmetic mean, heavily punishes models that are either poor at Precision or Recall requiring a strict tradeoff between correct predictions and the ability to observe all relevant instances. This feature is particularly useful in problems that have multiple classes, where otherwise unequal performance between classes can be hidden. With a combination of these two points in a single measure, F_1-score will be a conclusive measure of the general usefulness of the model, so that the solution is not just correct but also sound and strong enough to work in the field.

$$F_1\text{-score} = \frac{2\, Precision\, Recall}{Precision + Recall} \quad (11)$$

Accuracy: Accuracy is the average percentage of correct predictions made by the entire dataset, or the number of times that the model makes the correct predictions. Although it provides a simple and intuitive indication of overall performance, accuracy is inaccurate in cases where the distributions of classes are not balanced as it may overstate the performance of the dominant classes and understate the errors in literature of the minority classes. The metric uses the number of correct predictions, the total number of true positives and the total number of true negatives divided by the number of predictions to give an approximate but occasionally imprecise measurement of the effectiveness of a model.

$$\text{Accuracy} = \frac{\Sigma TP + \Sigma TN}{\Sigma(TP + FP + FN + TN)} \quad (12)$$

4.4 Quantitative Analysis of Computational Complexity

The comparative analysis of the metrics of computational complexity was necessary to justify the choice of the Custom CNN over the accuracy of the empirical results it produced. Deep learning methods are potent but may require quickly growing computational resources as models grow, which are costly and may be harmful to the environment. In light of these limitations, the system of efficiency was established as one of the key indicators of deployability. The study compared the training time, inference speed, and resources used to ensure that the selected architecture was not only with good predictive performance but was practical and sustainable to be used in real-life scenarios.

4.4.1 FLOPs Justification (Time Complexity)

To offer an objective metric to compare the efficiency of various network structures, floating-point operations (FLOPs) were used in order to measure the computational effort needed to perform a single inference. When comparing the Custom CNN with the properly tested models like ResNet-50 or EfficientNet-B0, one could see that the custom architecture demonstrated an impressive trade-off between the predictive performance and the computational efficiency. Although the Custom CNN had a much simpler and cleaner design, it always achieved a higher Accuracy and F_1-score in comparison to the number of FLOPs, meaning that it was more able to extract and interpret useful features of the dataset per unit of computation than the more massive and more intricate models. The trade-off between performance and efficiency is especially one that is relevant in practical deployment, where the amount of computational resources can be constrained, energy consumption is a factor, and inference in real-time is required of real-time monitoring applications or field-based applications. The Custom CNN, with high predictive and lower computational demands, indicates not only technical efficacy but also efficient sustainability to support its appropriateness to conservation-oriented ecological surveillance, as well as to other resource-constrained systems.

4.4.2 Parameter Density (Memory Footprint)

The overall number of parameters that a model is capable of training has a direct proportional relationship with the memory footprint and storage needs of a model, determining both the amount of hardware capacity required to run it and the ease with which it can be deployed. The smaller number of parameters is of special benefit in a resource-limited Mobile Expert System, where the number of parameters is especially important, because it is easier to store, minimizes the transmission cost, and can be loaded and executed faster on less powerful devices. Analysis of the parameters of the Custom CNN showed that it could attain high accuracy of 97.42% at a relatively small size, without the need to over-parameterize it, thereby incurring extra computational costs without the performance improvement that over-parameterization provides. This performance is indicative of a well-thought-out architectural formulation of the model where balance is most favorably strived to be achieved to make sure that high accuracy is not sacrificed in the name of resource sustainability or the real-life deployment of the model.

5 Experimental Evaluation

The experimental stage strictly tested the five chosen architectures in the same and tightly controlled conditions, which are fair and unbiased. The major aim was to determine which model provided the best speed, stability and predictive performance, which are important features of a dependable real-time conservation tool. The study was capable of comparing responses of each architecture to identical data sets, therefore, identifying the variations in efficiency, robustness and accuracy by ensuring that all tests were done under similar conditions. The method enabled a clear understanding of what model would be most appropriate to deploy in the field as part of ecological investigations, which face the need to make quick judgments and the significance of reliability.

5.1 Detailed Description of Selected Models

The classification experiment of the freshwater fish was developed using five specific different deep learning architectures that were chosen carefully to balance both the general-purpose, pre-trained and the domain specific network. Both models possess their peculiarities of architecture and their own pros and cons of the work, and there is a necessity to understand their principles, differences in buildings, and possibilities to apply them to real life. In this section all models that have been taken into consideration in the study are given an elaborate description.

5.1.1 ResNet-50

ResNet-50 also commonly known as Residual Network with 50 Layer is a well known deep learning architecture that proposed the concept of residual learning to overcome the issue of vanishing gradient with highly deep networks. The network consists of 50 layers, with convolutional, batch normalization and identity shortcut connections, enabling the flow of gradients to be more efficient in training. The residual connections make the network learn complicated feature representation without the issues of degradation of deep networks.

$$y=F(x,W)+x \qquad (13)$$

ResNet-50 has the following key characteristics:

- Depth and flexibility: ResNet-50 has 50 layers and this is able to distinguish hierarchical features of simple edges until complicated textures and shapes.

- Residual learning: Skip connections enable the network to skip one or more layers and reduce the vanishing gradient problem, and enhance convergence.

- Good generalization: ResNet-50 trained on big datasets such as ImageNet can be applied to new domains and maintain effective feature extraction skills.

Although it has a high accuracy (95% overall), ResNet-50 is highly resource-demanding in terms of computation, as it can impair its applicability to real-time, field-deployable systems. It is much slower to inference than lightweight networks, but it is still very useful in image recognition of complex patterns, including the subtle details of freshwater fish.

5.1.2 VGG-16

VGG-16 is a conventional convolutional neural network, which has 16 layers, 13 convolutional layers and 3 fully connected layers. It is characterized by a basic and homogeneous structure, in which the convolutional layers operate with 3×3 filters, and pooling layers are always implemented in order to contract the spatial dimensions. The network focuses on learning more complex features in each step, which is known as depth but not width.

$$y=\sigma(W*x+b) \qquad (14)$$

VGG-16 has the following important features:

- Non-differentiated architecture: There is a uniform use of 3x3 convolutional filters that simplifies the implementation of the network and its optimization.

- Deep feature extraction: The deeper the network levels, the more it is able to capture hierarchical features, such as edges and corners, further into more abstract features such as scale textures or fin shapes.

- Pre-trained transfer learning: VGG-16 can be used with pre-trained weights based on large-scale datasets, which is a powerful starting point of domain-specific problems.

VGG-16 is an efficient (96% accuracy) classifier of fish. It also has many parameters, however, which consume more memory and inference time, which can limit its application in low-resource settings or in applications that demand real-time processing.

5.1.3 Inception V3

Inception V3 is a convolutional network that is advanced and developed to capture features effectively, with multiple scales in an image. It has a structure that brings on board the use of Inception modules, which do parallel convolutions of various filter sizes (1×1, 3×3, 5×5), to enable the network to learn fine and coarse features simultaneously. Factorized convolutions and auxiliary classifiers are advantageous to enhance computational efficiency and training stability.

$$y=\mathrm{Concat}(f1\times1(x),f3\times3(x),f5\times5(x),f\mathrm{pool}(x)) \qquad (15)$$

The following aspects are important as far as inception V3 is concerned:

- Multi-scale feature extraction: It extracts local patterns and global contextual information by parallel convolutions, this is the reason why it can be applied in images where the features are of different kinds.

- Effective computation Factorized convolutions have fewer parameters, and they are faster to train than naive deep networks.

- Auxiliary classifiers: The training refers to the middle-level outputs that overcome the vanishing gradients to improve the convergence.

The general accuracy of inception V3 was 96 percent with regards to fish classification. The complexity of the architecture is very appropriate in capturing subtle differences, however, of the monitoring architecture, inference time and implementation complexity can be tough in real time monitoring applications, especially in the field application with a small number of hardware, although the complexity of the architecture can be tough in real time monitoring applications.

5.1.4 EfficientNet-B0

EfficientNet-B0 is a part of the group of EfficientNet, which introduces a scaling method of the network to achieve a balance between the network depth, width and resolution. B0 is the default model and so is optimized with computational efficiency and still it has extensive feature extracting capabilities. MBConv blocks are optimized to identify channel-wise dependencies, which the network uses to give powerful representations with reduced parameters.

$$\text{Depth} = \alpha^{\phi}, \text{ width}= \beta^{\phi}, \text{ resolution} = \gamma^{\phi}$$
$$\alpha \cdot \beta^{2} \cdot \gamma^{2} = 2 \qquad (16)$$

Major features of EfficientNet-B0 are:

- Lightweight design: It is less costly in computation and has less parameters, which is appropriate in mobile or embedded applications.

- Scalable architecture: EfficientNet framework can be trained to larger model (B1-B7) in case of need of higher accuracy.

- High efficiency: Is competitive and has low latency and memory costs.

The overall efficiency of the model (EfficientNet-B0) was 95% when classifying fish. Its Recall of some species, however, including Koi, was lower in adverse circumstances (e.g. low light), showing a drawback in the performance of the detection of rare or visually challenging fish. Although this is computationally efficient, this trade-off may have an impact on reliability in conservation oriented, real-time monitoring applications.

5.1.5 Custom Convolutional Neural Network (Custom CNN)

The Custom CNN is created directly based on the freshwater fish classification problem and built fresh. This model is also not designed on generic pre-trained networks but rather specializes to detect fin shape, body shapes, scale texture and color patterns which are subtle morphological characteristics that may differentiate closely related species. The architecture was developed through five iterations with each version becoming more and more elaborate. Model 5 is the last model; this model has six convolutional layers with successively increasing filters of 32, 64, 128, and max-pooling layers to reduce the dimensions maintaining important features. The learned features are combined in a fully connected layer of 256 neurons and are then used in a Softmax classifier which produces probability scores over six target species.

The major features of the Custom CNN are:

- Domain-specific design: The network focuses on the features related to freshwater fish, which can be discriminated more accurately.

- Repeat improvement: Architectural changes were experimented to achieve maximum accuracy, robustness, and generalization.

- Regularization mechanisms: Dropout layers minimize overfitting, meaning that the model is able to perform at a high quality on unknown samples.

- Computational efficiency: The network was optimized to ensure low inference time, which is suitable in real-time monitoring, in spite of its depth.

This model also had the best test accuracy of 97.42, which proves that this model is useful in capturing the fine grained visual difference between species. Its overall accuracy and the recollection, regardless of the class, highlights its strength and appropriateness in conservation-oriented monitoring.

5.1.6 Summary of Model Selection

The chosen five architectures offer an array of architectural philosophies, with domain-specific design (Custom CNN), to general purpose, pre-trained networks (ResNet-50, VGG-16, Inception V3, EfficientNet-B0). All the models have their own benefits:

ResNet-50: Extensive architecture, powerful feature extraction, average inference speed.

VGG-16: Canonical and simple deep network, simple uniform architecture, powerful hierarchical features capturing.

Inception V3: Multi-scale feature, high accuracy, a little higher latency.

EfficientNet-B0: Light, quick, energy-efficient, however, a little less recall on subtle classes.

Custom CNN: Maximum accuracy, balanced performance, domain-specific features learning, optimized in the field.

Such a thorough description of every model forms the basis of experimental analysis and comparison that follows afterward to interpret the findings in a thoughtful manner in terms of accuracy, strength, and practicality in ecological and conservation monitoring.

5.2 Training Protocol and Custom CNN Optimization

All of the models were trained with the assistance of the resources of the GPU on the cloud system and with the assistance of Adam optimizer and the categorical cross-entropy loss as the tool that should propel the learning process. Similar parameters were used on all architectures such as a batch size of 32 and the maximum number of epochs to ensure that all architectures are evaluated fairly. Early-stopping and model checkpoint were additional methods that increased the stability of learning, avoiding overfitting, and also ensuring that every model acquired meaningful visual patterns without overtraining.

The iterative optimization of the Custom CNN that was constructed in five successively improved versions was also one of the major focuses of this work. Every version introduced developed architectural improvements, as recapped in Table 3, which consisted of: increased depth, increased convolutional kernel and increased structured pathways of feature extraction. It was found that the first accuracy of 71.2 was reached by using Model 1 that included only three convolutional layers (32, 64, 128 filters). The Model 2 and 3 provided a small increment to the accuracy of 77.9 and 91.52 respectively with the extra convolutional layer and the performance of the model considerably increased to 3 layers, which is considered to be a better feature learner at a moderate depth. In model 4, the network was again extended to five convolutional layers with the accuracy of 94.89. The final variant, Model 5 adopted six convolutional layers (32, 32, 64, 64, 128, 128 filters) and highest accuracy of 97.42 which proved that the more the depth of the filters and the richness, the more the network would be helpful in the extraction of discriminative features.

This trend across these iterations strongly supports this postulation according to which domain-specific complexity (with proper use) has a direct positive effect on classification accuracy on freshwater fish images. The given step by step method did not only find a perfect architecture but also set up the succession of learning how depth in convolutional architecture, filter architecture and structure optimization can be utilized to gain robustness, learning, and reliable generalization in varying real-world scenarios.

Table 3: Iterative Development and Analysis of the Customized CNN

Model No.	**Epoch**	**No. of Kernels**	**Fully Connected Layer**	**Train data**	**Vali data**	**Test data**	**Accuracy**
1	20	32, 64, 128	256	80%	10%	10%	71.2%
2	20	32, 32, 64, 128	256	80%	10%	10%	77.9%
3	20	32, 32, 64, 128	256	80%	10%	10%	91.52%
4	50	32, 32, 64, 64, 128	256	80%	10%	10%	94.89%
5	40	32, 32, 64, 64, 128, 128	256	80%	10%	10%	97.42%

The last architecture, Model 5, six convolutional layers and a fully connected layer with 256 neurons, proved to be the most efficient model as it gained the highest test accuracy of 97.42%. This result affirmed that adequate architectural characteristics are essential in capturing the complex and subtle morphological characteristics that differentiate closely related species of fish. With proper acquisition of these fine-grained patterns, the model has shown great ability to classify species, with high accuracy levels, which clearly shows that keen attention to the complexity of the network directly correlates to high performance in specialized classification problems.

5.3 Comparative Analysis of All Models' Performance

These relative findings highlighted the benefits of the Custom CNN that achieved constant balanced and high performance in all species. This consistency in effect has been especially valued in terms of safe field use wherein all species are not misclassified systematically, and conservation policy has been assured to be obtained with confidence. The results are that a well-crafted, domain-specific architecture may be faster than general models and provides precision, recall and overall accuracy in a manner that is highly resilient and can be put into practice in a realistic ecological surveillance setting.

5.3.1 Performance Across All Metrics

Custom CNN was the best with a total accuracy of 97 percent, and more importantly, it maintained the high accuracy throughout across all the individual classification metrics as represented in the table 4 of the six species. This consistency implies that the model does not give preference to one or the other type and can be reliably used to separate all the species, which is a prerequisite to practical application in field monitoring. The Custom CNN provides a more reliable instrument, as it is a combination of both high accuracy and consistent metric performance which makes this model suitable as an ecological and conservation instrument.

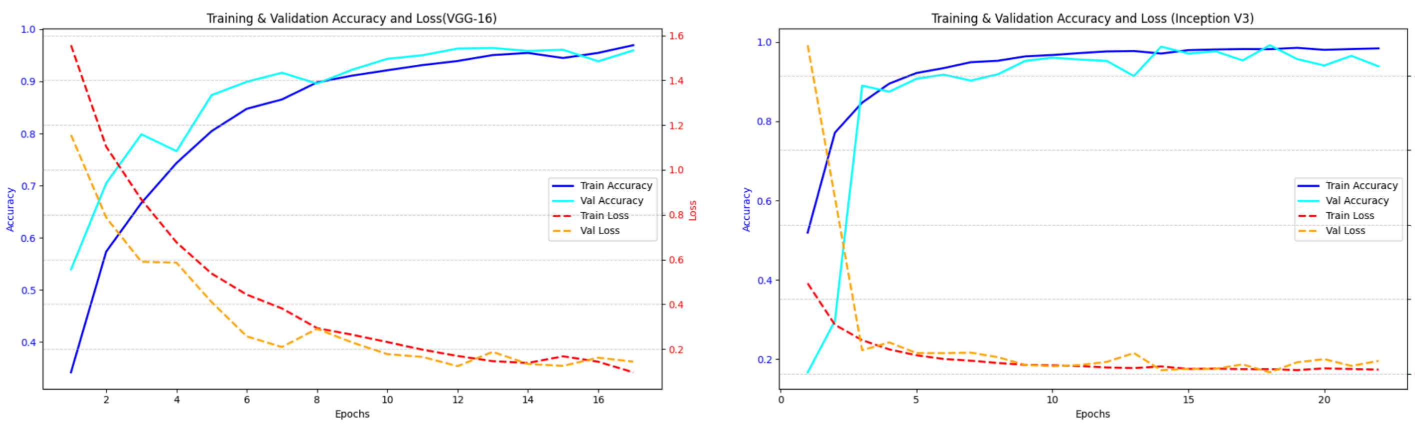
Table 4: Result of Fish Classification Performance Across Different Classes

Class Name		Tara Baim	Desi Magur	Kajoli	Khoilsha	Desi Koi	Tengra
ResNet50	**Precision**	**0.92**	**0.95**	**1.00**	**0.95**	**0.97**	**0.93**
	Recall	**0.97**	**1.00**	**0.97**	**0.95**	**0.90**	**0.92**
	F_1**-score**	**0.94**	**0.97**	**0.99**	**0.95**	**0.94**	**0.93**
	Accuracy	**0.95**	**0.95**	**0.95**	**0.95**	**0.95**	**0.95**
VGG16	**Precision**	**0.93**	**0.95**	**0.99**	**0.99**	**0.97**	**0.95**
	Recall	**0.95**	**0.98**	**0.97**	**0.96**	**0.97**	**0.95**
	F_1**-score**	**0.94**	**0.97**	**0.98**	**0.98**	**0.97**	**0.95**
	Accuracy	**0.96**	**0.96**	**0.96**	**0.96**	**0.96**	**0.96**
InceptionV3	**Precision**	**0.96**	**0.95**	**0.97**	**0.92**	**0.98**	**0.99**
	Recall	**0.99**	**0.97**	**1.00**	**0.96**	**0.92**	**0.93**
	F_1**-score**	**0.97**	**0.96**	**0.99**	**0.94**	**0.95**	**0.96**
	Accuracy	**0.96**	**0.96**	**0.96**	**0.96**	**0.96**	**0.96**
EfficientNetB0	**Precision**	**0.95**	**0.99**	**0.96**	**0.85**	**0.99**	**0.93**
	Recall	**0.96**	**0.95**	**0.99**	**0.98**	**0.84**	**0.95**
	F_1**-score**	**0.96**	**0.97**	**1.00**	**0.91**	**0.91**	**0.94**
	Accuracy	**0.95**	**0.95**	**0.95**	**0.95**	**0.95**	**0.95**
Custom CNN	**Precision**	**O.97**	**0.96**	**0.99**	**0.99**	**0.95**	**0.99**
	Recall	**0.98**	**0.97**	**0.97**	**0.99**	**0.97**	**0.97**
	F_1**-score**	**0.97**	**0.97**	**0.98**	**0.99**	**0.96**	**0.98**
	Accuracy	**0.97**	**0.97**	**0.97**	**0.97**	**0.97**	**0.97**

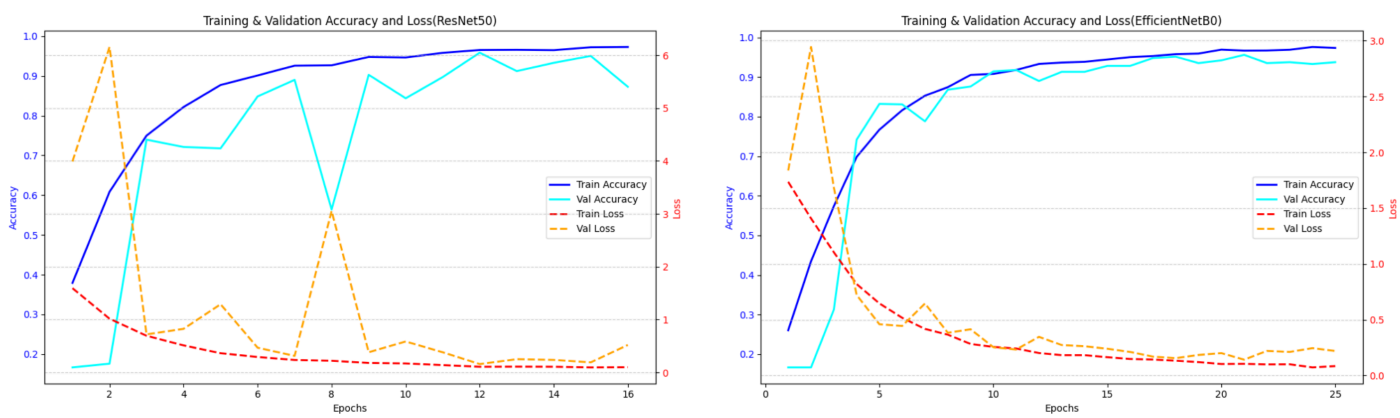
5.3.2 Assessment of Comparative Models

In figure 5 ResNet-50 was structure strong, with a 95% overall accuracy and 1.00 (perfect) perfection of the Kajoli species. Nonetheless, it has intrinsic architectural complexity and thus has large computational requirements, which is a major drawback to realistic, real-time implementation. Both VGG-16 and Inception V3 have high performance at 96% as it shows that they are able to isolate the fine-grained features when it comes to the distinction of closely related species. However, because of the requirement to use deep and densely connected layers, they generally add in higher inference latency, which makes them less applicable to applications that demand low-latency, field-ready performance.

EfficientNet-B0 was also an interesting one because of its lightweight architecture and computational efficiency, obtaining 95% overall accuracy. Although this was advantageous, the model indicated a major limitation which was the Recall of the Koi species being significantly under 0.84 when the experiment was conducted at low illumination. Such reduction in Recall is not acceptable in a conservation environment because it shows that the model might miss endangered individuals, which are indeed present, and this could undermine the process of monitoring and protection. This discrepancy brought out the issue of the need to have consistent generalization in diverse environmental conditions and emphasized the importance of having a model that would always have strong performances in real life.



(a) Training & validation accuracy and loss (VGGG - 16) (b) Training & validation accuracy and loss (Inception V3)



(c) Training & validation accuracy and loss **(**ResNet50**)** **(d)** Training & validation accuracy and loss (EfficientNetB0)

Fig 6 : ROC curve of the pretrained models

5.3.3 Custom CNN Performance Synthesis

The Custom CNN was displayed to show high performance in all six species with F_1-scores constantly ranging between 0.96 and 0.99 with the highest scores recorded by Kajoli and Khalisha. Such stability suggests that the architecture was repeatedly optimized to capture the domain specific differences that are needed to distinguish these two closely related fish species. The consistency of high precision and recall is also additional evidence that the model is accurate, and it is comprehensive in its findings and reduces the cases of false positive and missed identifications. In combination, these findings make the Custom CNN a robust, efficient, and reliable solution that can be well applied in real-time in challenging and real-life settings where effective species monitoring is needed to conserve the environment.

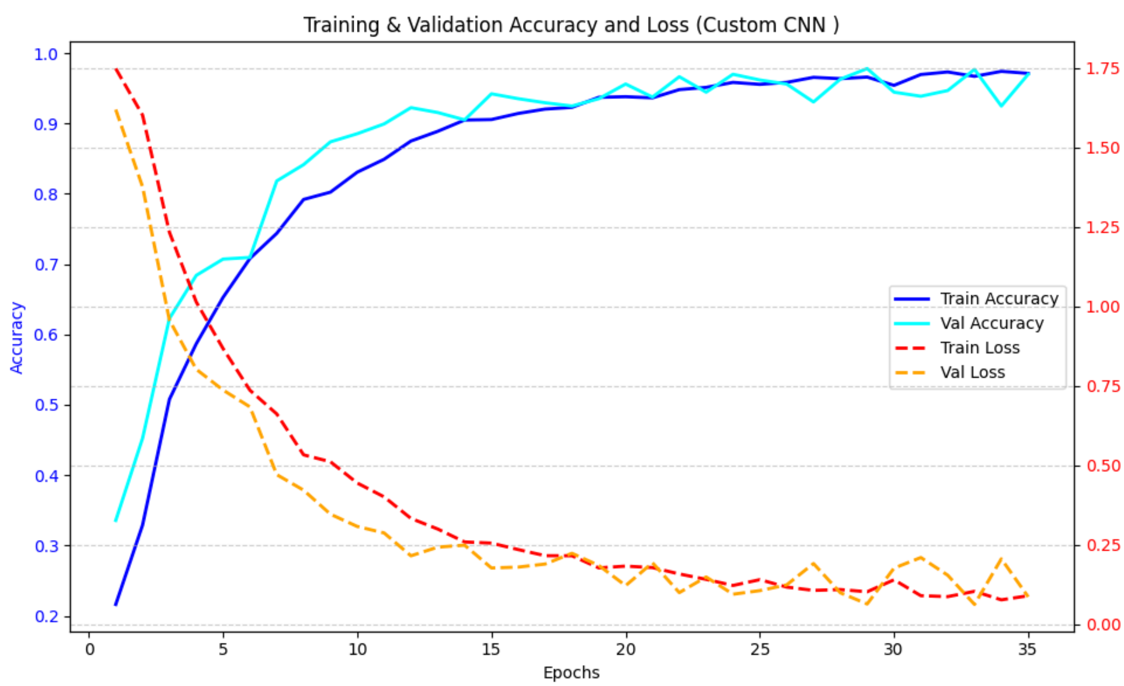


Fig. 7(a): Visualization of ROC curves and training and validation accuracy and loss curves of the customized CNN model.

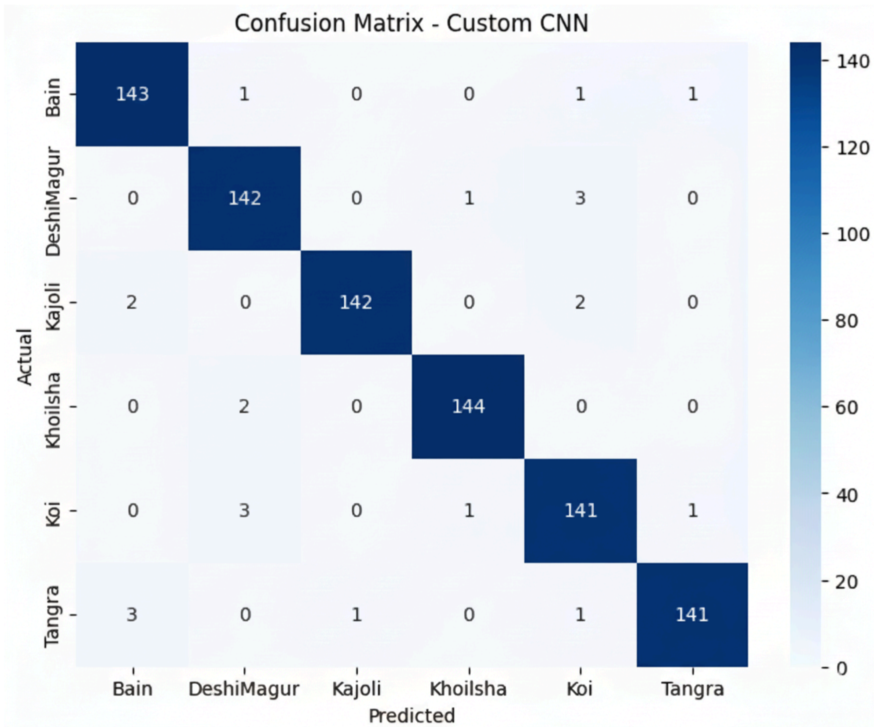


Fig. 7(b): Customize CNN model confusion matrix

As shown in the training curves in figure 6(a) the Custom CNN training curves were very smooth and steady and this is a reflection of the stability in the model over the learning process. The confusion matrix figure 6(b) was analyzed and the results showed that there were few errors of inter-class, which further shows that the network has good generalization abilities across all the species. Finally, the last Model 5 was able to reach the highest accuracy of 97.42 percent, which brought the solution strongly into the lead. The results show that the well-crafted architecture does not merely do outstandingly well on the training data, but also is a reliable and consistent performer on previously unknown images of the real world.

5.4 Definitive Analysis of Classification Robustness: The Confusion Matrix in Conservation

The confusion matrix served as the primary diagnostic tool for evaluating the classification models, revealing detailed patterns of errors that overall accuracy alone could not capture. In conservation-focused applications, this depth of insight was critical for determining whether the Custom CNN could reliably support real-world decision-making and regulatory activities. To achieve a thorough evaluation, a one-versus-rest analysis of the confusion matrix $A \in N^{n \times n}$ was conducted for all five models, enabling the calculation of per-class metrics including Precision, Recall, and F_1-score. This method provided a complete assessment of performance across all six target fish species, highlighting each model's strengths while identifying potential areas for improvement in species-specific classification. Such a detailed analysis ensured that decisions based on model outputs would be informed, accurate, and dependable in practical ecological monitoring scenarios.

5.4.1 Interpreting the Custom CNN Error Profile: Misclassification Patterns

The Custom CNN confusion matrix showed a good diagonal distribution, which shows that the overall accuracy of this model was 97.42 percent. Seldom were there, however, small off-diagonal entries, which appeared to represent inter-class misclassifications, of species, with very similar morphologies. It is worth noting that there were few mistakes between Deshi Magur, Deshi Koi and Tengra that have some structural features like elongated bodies and high barbels. Such visual and taxonomic similarities cause them to be more difficult to differentiate by the model in itself. This trend indicates a boundary classification issue, in which feature extraction layers of the Custom CNN framework approach limits when trying to resolve subtle morphological variations in the real world, such as varying backgrounds, light, and partial occlusion.

In comparison, other species like Kajoli and Khalisha were always significantly high in F_1-score, 0.98 0.99. This implies that they have stronger distinguishing characteristics that the model quickly acquires regardless of environmental variations. The good results of these species support the fact that the Custom CNN extracts robust and invariant features in the presence of clear morphological information, which on the one hand indicates that this approach is useful in identifying accurate species that have distinctive features but on the other hand shows that fine-grained discrimination between visually similar fish is limited.

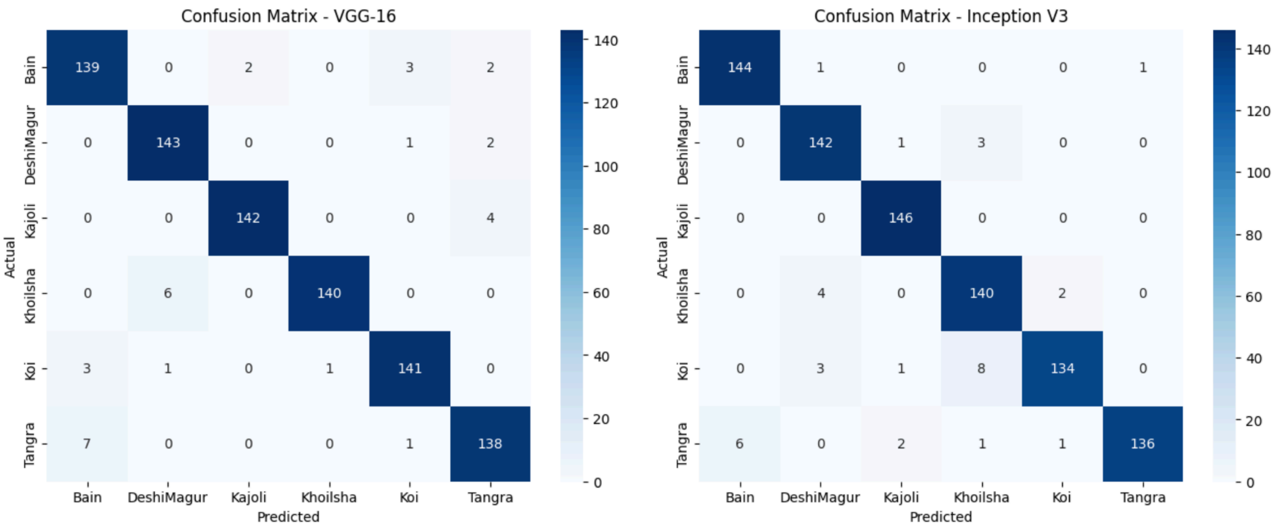
5.4.2 Comparative Error Diagnostics of Pre-trained Architectures

The four pre-trained models used in figure 7, which were VGG-16, ResNet-50, Inception V3, and EfficientNet-B0, presented much information to make a closer comparative error analysis. Although all these models overall exhibited high overall accuracy in the form of a strong diagonal trend, a more detailed look showed that there were a number of constraints when compared to the purpose-crafted Custom CNN. This degree of analysis was necessary to reveal some of the small weaknesses that might have a tremendous effect on the real-life application in conservation and ecological monitoring where dependability on all species is of significance.

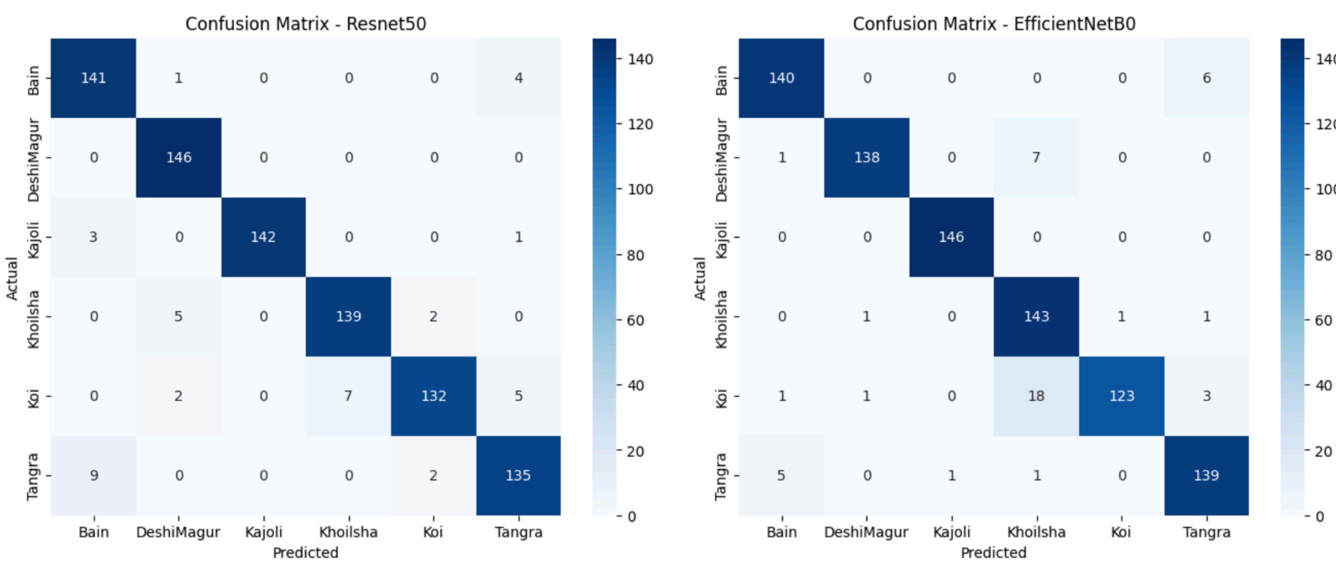
As an example, EfficientNet-B0, although implemented with the goal of being computationally light and efficient, had a strong flaw in performance with respect to the detection of the Deshi Koi species. Its confusion matrix revealed that it had a fairly high number of false negatives of this class, which caused the Recall to decline radically to 0.84. This deficiency shows that the model was in many cases unsuccessful to identify a species with high conservation priority which may have severe consequences in the field. Underestimates of population sizes in real-world monitoring due to missed detections or even more likely failure to capture endangered individuals, typify that a high computational efficiency is not completely effective when the reliability of detection is decreased.

Instead, VGG-16 and Inception V3 demonstrated high overall accuracy, which is associated with the ability to extract complex and fine-grained features of the images. But their more intricate and profound architectures brought us practical constraints. The high count of layers and parameters raised the inference latency, which might cause serious bottlenecks in the field deployment of real-time settings that demand fast decision-making. These models were not as practical in using low-latency predictions, which were most often useful when accuracy was the goal but met the benchmark of operational performance, but this demonstrated the trade-off between accuracy and practical use in deep learning deployment. ResNet-50 was an optimal tradeoff, a decent tradeoff between accuracy and inference speed, but also exhibited small inter-class errors, especially in species with fine-grained morphological variations, showing weakness in fine-grained differences across different environmental settings.

After analysis of the functioning of each of the five models separately and in the performance of all five models as a group, the Custom CNN proved to be the most powerful solution, not only in the context of obtaining higher numerical values but also in terms of structural consistency and applicability. Its confusion matrix produced high diagonal patterns throughout with no significant off-diagonal errors with most diagonal errors being concentrated in the most similar species. Also, the Custom CNN retained low inference durations and effective computational specifications, which guaranteed that superior performance did not lead to the cost of deployability in resource-limited, real-time monitoring environments. This combination of high precision, class-wise predictability and computational efficiency makes the Customer CNN a strong and reliable system to apply to field based applications and provide conservation efforts with reliable and actionable data even in hard environmental situations.



(a) Confusion Matrics (VGGG - 16) (b) Confusion Matrics (Inception V3)



(c) Confusion Matrics (Resnet50) (d) Confusion Matrics (EfficientNetB0)

Fig. 8: Confusion matrix of the pre-trained models

5.4.3 Ecological Trade-offs: Precision versus Recall in Policy Enforcement

Changes in Precision and Recall are highly important in achieving the optimal balance in any conservation monitoring tool. Precision is used to measure the consistency of positive predictions, which means how often the model itself predicts a rare or endangered species. This lack of precision may cause false alarms that may destroy the confidence of the conservation measures among the citizens and the regulators. In contrast, recall, as previously stated, measures the capacity of the model to detect all real positive examples, which is the fraction of infrequent species that are not falsely detected. With endangered species, false negatives may be catastrophic to the ecology of the region, as it may result in the illegal harvesting of the species or destruction of the habitat without detection and response.

EfficientNet-B0, though it has high overall accuracy in comparative evaluations, has an interesting weakness. Its Recall of the Deshi Koi species dropped to 0.84 under some circumstances, i.e. the percentage of living Koi specimens that were not found was approximately 16 percent. This rate of false alarms is unacceptable in real life conservation use where it directly jeopardizes the validity and usefulness of monitoring programs.

In comparison, the Custom CNN exhibited a very balanced and well-performing performance on all the six target species. It had the lowest F_1-score of 0.96 and the Recall was always above 0.95 on important species like Deshi Koi (Recall 0.95, Precision 0.97). This high accuracy and high recall is such that not only is it certain that the model will detect individuals when they appear but it will also capture the whole population in a comprehensive manner reducing instances of missed detections. The presence of such stability in all endangered classes proves that the Custom CNN is specifically the most suitable with risk-sensitive conservation monitoring. It offers reliable, practical information, reinforcing evidence-based interventions and policy-making and providing feasible assessments in the field that are crucial in the preservation and protection of rare and threatened fish species.

6 Comparative Analysis of Model Performance in Fish Species Detection

In order to clearly comprehend the practical implications of the performance of the Custom CNN, it is necessary to compare its results with the broader scope of the current research on the equipment in machine vision to detect aquatic life. These comparisons help to put accuracy and efficiency into perspective where the custom architecture is compared to previous methods regarding its performance in species identification, ability to operate effectively across multiple environmental situations and its computational capabilities. When the Custom CNN is positioned in such a larger context, one can evaluate not just its performance in numerical terms but also its usefulness in actual conservation practice and how it can make improvements to the current outcome and where it can introduce new standards of reliability and usability in aquatic monitoring.

6.1 Benchmarking Against External Studies

The Custom CNN created during this research has shown high accuracy in identifying endangered species of freshwater fish in Bangladesh with the overall test accuracy of 97.42. This accuracy makes the model top a number of previous standards in the context of water image recognition. Although previous researchers have presented the potential of deep learning in detecting fish, most of them were limited to small datasets, reduced accuracy span, or settings that did not completely reflect on real-world fluctuations. Comparatively, the proposed system does not only give good classification results, but also gives a more practical and scalable system, which are more expected to be deployed to the conservation-oriented setting.

Though there has been significant advancement in the literature on the same, some important gaps still exist. One, numerous of the currently existing models use small, highly manipulated datasets- Sharmin et al. [8] used only 180 images- resulting in little generalization. Second, the methods that are high-performing, such as the ones by Villon et al. [5], need very large datasets (900,000+ images) which is not feasible in many areas with a limited data-collection capacity. Third, underwater detection methods like the one by Mu et al. [4] rely on sophisticated hybrid pipelines (GMM, optical flow, YOLO, ResNet-50), which raises the cost of computation and makes them less suitable to be deployed on lightweight devices or on a mobile platform. Lastly, previous literature does not pay much attention to endangered or species-specific species, creating a technical and ecological gap in systems that can be used in conservation activities in biodiversity hotspots like Bangladesh.

The proposed Custom CNN, in its turn, is the best compromise between the size of datasets, the computational power, and the prediction accuracy, and it would be the only solution applicable to the endangered fish recognition. The model, which is trained on a moderately sized and very diverse 8,676 images of six threatened species, performs better than the current methods, but is both practical and applicable to the field. The fact that it has high precision, and low false-negative rate, make it especially useful in conservation, where accurate and prompt identification is important. Such technical strength, environmental interest, and practical usefulness make the offered Custom CNN the most stable and influential solution among the works.

Table 5: Comparison between the Proposed Work Vs. Related Existing Works

Existing Work	**Problem Dealt With**	**Dataset Size**	**Number of Classes**	**Best Algorithm**	**Accuracy**
This paper	Extinct fish recognition	8,676 images	6	Custom CNN	97.42%
Mu et al. [4]	Fish detection & species classification underwater	4,418 videos	15	Hybrid (GMM + Optical Flow → ResNet-50 → YOLOV3)	91.64-95.47%
Villon et al. [5]	Coral reef fish identification	900,000 + images	20	GoogLeNet CNN + Decision Rules	94.1%
Sharmin et al. [8]	Local freshwater fish recognition	180 images	6	SVM	94.2%

6.2 Contextual Validation of Performance Gains

The resultant precision of 97.42 percent is quite remarkable considering the number of complications present in the task and the difficulty of the data. As a control, a previous study by Sharmin et al. studied six local freshwater species with their Support Vector Machines (SVM) traditionally on a very small dataset of only 180 images, achieving an accuracy of 94.2%. The current work, in turn, used deep learning algorithms with a significantly larger and more complicated dataset of 8,676 images, purposefully subjecting them to environmental variability (i.e. different types of light, turbidity, and fish position). This has led to a remarkable boost of more than 3.2 percentage points, and it is a clear indication of the transformative benefit that deep learning and in the case of a tailored-made architecture, in particular, have as compared to traditional machine vision techniques in complex biological identification.

Besides accuracy, the Custom CNN presents a significant degree of efficiency as compared to methods that are based on exceptionally large datasets. Indicatively, the study of Villon et al. managed to identify coral reef fish, which is a similar but different field, with more than 900,000 images with a high precision of 94.1. The Custom CNN, on the other hand, achieved a higher performance with a smaller and more curated and ecologically representative dataset of 8,676 images. This analogy highlights a valuable lesson which is that high levels of accuracy in domain-specific biological tasks are not only a matter of dataset size but also a consequence of careful architectural design as well as of data of high quality and representativeness.

This customized solution provides not only high predictive capability but useful benefits to performing real-time monitoring to endangered freshwater fish. Offering a high level of accuracy using a manageable dataset and computationally efficient design, the Custom CNN offers a powerful, dependable, and scalable basis of conservation applications. These findings make it a credible instrument in combating the looming biodiversity crisis in the freshwater ecosystem in Bangladesh as it provides accurate species identification, sound policy formulation, and effective conservation of the endangered fish species.

7 Conclusion and Future Works

The study has demonstrated that deep learning can determine the endangered native fish species using real-life pictures that may be taken in different environments. It was found that Custom CNN was more accurate and consistent in performance than all the other tested architectures and it demonstrates that it can be applied in actual conservation. The system can further be expanded in future by adding additional species, diversification of the dataset and optimization of the model to be run on a mobile or web based platform. Its contribution to biodiversity surveillance and sustainable fisheries management can be added to it with the following features; real-time detection, GIS mapping, and community-based data collection.

7.1 Conclusion

The research gives an exceptionally descriptive model of recognizing and classifying threatened freshwater fish species in Bangladesh using the sophisticated deep learning application. The research has used the implementation of machine vision and specifically developed convolutional neural networks (CNN) to bridge the gap between biodiversity conservation and artificial intelligence. The proposed system demonstrates that automated classification through images can be used to provide ecological protection and a viable and reliable solution to the identification of endangered native species in diverse scenarios of real life.

The main value of the work is that it has created a large and high-grade dataset of 8,676 real-world images of six endangered species- Tara Baim, Deshi Magur, Kajoli, Khoilsha, Deshi Koi, and Tengra. This data was taken in different environmental conditions and light conditions and this improved the generalization ability of the models, as compared to the previous datasets which were limited by size, environmental diversity and species species.

The analysis of five deep learning models, namely ResNet-50, VGG-16, InceptionV3, EfficientNet-B0, and the proposed Custom CNN, demonstrated that the latter had the best overall accuracy, being 97 per cent, as compared to the rest of the models that belong to the state of the art. This underscores how strong, flexible and suitable it is in real-life implementation in resource-constrained settings. The results affirm the fact that a CNN architecture which is optimized and trained on a diverse and balanced dataset is capable of managing the variation in lighting, orientation and background noise with the aid of which accurate detection can be achieved in uncontrolled settings.

In addition to its numerical performance, this study highlights the transformative nature of artificial intelligence in enhancing the conservation of biodiversity. The proposed framework will promote ecological surveillance and sustainable fisheries management as well as community awareness by allowing real-time and automated detection of endangered species. The method does not only reduce the necessity of manual species identification but also provides a tool to be scaled to aid conservation activities and policy making.

The framework can be furthered in the future by adding more species, video detection, and lightweight models to serve mobile devices. This would go a step further to make this system relevant in practice of field researches and conservation programs. In general, this paper lays the groundwork of the intersection of deep learning and ecology with sustainability, and how technology could be used to significantly ensure that Bangladesh's aquatic biodiversity can serve future generations.

7.2 Limitations

The limitation to this study is mostly the size and scope of the dataset. Even though there are 8,676 images of 6 endangered fish species, which is a good base to build on, a bigger and more diverse set of data of other species and settings would enhance generalization. The photographs were mainly taken in particular areas in Bangladesh which might restrict the ability of the model to adapt to new environments other than the habitat or lighting conditions. Also, the experiment was conducted using still data, and the experiment did not include real-time and underwater video data, which can be improved in the future with the application in real-time monitoring and ecological tracking.

7.3 Ethical Considerations

The data, used in this research, was gathered on the basis of academic and research purposes only making sure that ethical standards were observed. The pictures were taken in local markets and natural environments with the consent of the authorities, without any harm or disturbance to natural environments. It is worth mentioning that in Bangladesh, catching, purchase, or sale of these fish species is not classified as a crime. Data gathering was done without any personal or sensitive information use and all data was treated with responsibility so that ethical integrity was upheld in the research process.

7.4 Impact On The Society

The machine-vision-based fish recognition system suggested has strong potential of having a positive impact on the society especially in the rural and the fisheries-reliant communities of Bangladesh. The system prevents unintentional harvesting of endangered freshwater species by allowing the accurate identification of these species and promoting sustainable fishing. This enhances conservation performance and preservation of livelihoods of individuals that depend on the native fish resources. Moreover, the tools that are publicly available and developed with the help of AI, including mobile recognition applications can create awareness, provide local fishers with real-time data, and promote responsible market conduct. Such technology also enhances national biodiversity surveillance, academic study, and enhances digital literacy of the communities that historically lack access to modern ecological devices. In general, such a system helps to protect the environment, secure the further existence of the Bangladesh economy, and make society more aware of the need to preserve the native freshwater biodiversity in Bangladesh.

7.4.1 Sustainability Of The Work

To make this research sustainable, it is necessary to further expand the data and the functionality of the model. Further research on the topic should be directed at expanding the dataset to cover a greater number of native fish species across different areas of Bangladesh, which will allow a more extensive ecological coverage. There should also be an attempt to record real-time and underwater video to increase flexibility and strength in the natural settings. Ethical practices in image gathering and close working with local fisheries departments as well as conservation agencies will play a critical role in continuity in the long run.

Sustainability also relies on the efforts of research institutions, fisheries authorities and the local communities. This framework is able to be constantly refined and implemented, to be used in learning, ecological and policymaking. The long-term utilization of the system can be made useful, flexible and environmentally friendly by encouraging community participation and sharing the research freely.

7.4.2 Social and Environmental Effects & Analysis

Automated fish recognition systems might have serious social and environmental consequences. In the social context, it enhances awareness of biodiversity, responsible fishing, and sustainable management of fisheries by assisting in the identification and monitoring of endangered species. The system enables the local fishermen, students, and environmentalists to empower themselves through the offers of a low-cost technological device to identify the location of rare fishes, therefore, conservation education enhances ecological literacy.

The system also helps monitor aquatic ecosystems with a low amount of human intrusion thus preventing chances of overfishing and preserving delicate freshwater ecosystems. It allows implementing specific conservation measures to sustain ecological balance, as it allows detecting the presence of rare or deteriorating species in the early stages of their decline. At a bigger level, the technology can support environmental policy and biodiversity conservation in Bangladesh, and similar areas based on data.

7.5 Proposed Advancements

The suggested framework offers a significant enhancement to the traditional manual fish identification techniques. Nevertheless, there are still a number of directions to be developed and optimized. Future developments can work to have real-time detection, lightweight deep learning architectures, and hybrid data management systems to improve accuracy, scalability, and accessibility.

7.5.1 Deep Learning Enhancements

Although the Custom CNN had 97% accuracy, it would be possible to further improve it by adding hybrid deep learning architectures, including ResNet-based ensembles, Vision Transformers (ViT) or CNN-RNN hybrids.

- **Better Recognition Accuracy**: Hybrid and transformer-based systems have the ability to learn multi-scale spatial features, multi-scale dependencies and this enhances their ability to detect overlapping or partially visible fish.

- **Adaptability**: Models that have learned variations with time and the environment (lighting, background, and motion) will be more successful in dynamic underwater conditions.

- **Scalability**: Lightweight CNN variants can implement pruning or quantization in order to execute these models on a mobile device to use on the field in real time.

The latter can be combined with object detection networks such as YOLOv8 or EfficientDet, and this can be used to recognize multiple species in real-time video feeds, a more realistic tool to monitor an ecological process.

7.5.2 Dataset Expansion and Regional Adaptation

The increase of the dataset is an important move towards the increase in accuracy and regional flexibility. The use of images of various water bodies, rivers, ponds, and wetlands and estuaries, would document morphological changes as a result of the geography and seasonality. Also, crowdsourcing data analysis under professional management might contribute to the enrichment of the datasets and encourage the community to participate in conservation. Video-based data that is introduced with frame-level labeling would also enhance temporal recognition to be applied in real-time analysis.

7.5.3 Hybrid Data Management and Cloud Integration

When the data and computational needs grew, a hybrid data management method was followed by combining cloud-based data storage with collaborative computing spaces. All training, validation, and testing of the model were done in Google Colab, where the processing speed was enhanced with the help of the use of a GPU to perform the necessary experiments. The self-collected data were safely stored and arranged in Google Drive which offered accessibility, scalability and version control during the research process.

This cloud-based system facilitated a smooth cooperation among the researchers, and they could share the dataset and the model checkpoints. It also supported maintaining a constantly updated model and freedom to reproduce results without being dependent on high-end local infrastructure. Compression of data and data structuring Lightweight compression and data structuring were necessary to have efficient management of storage and easy integration with the Colab environment.

This hybrid architecture ensured flexibility and performance by using both local data preprocessing and cloud-based computation. It also provided a basis towards scalability in future, whereby it can add more datasets or models without excess resource limitations.

7.5.4 Impact at Educational and Policy levels

This framework has placed collaboration between universities, fisheries departments and research laboratories as one of the educational and conservation oriented tools. The results and data of this paper are applicable to college-level courses about deep learning, computer vision, and biodiversity protection. The information obtained during automated fish recognition can also be used by policymakers in the sustainable management of fishing, monitoring endangered species, and educating the community on ecological conservation.

In interdisciplinary collaboration and educational inclusion, this research would help reinforce the connection between technology and environmental sustainability, promote future studies and innovation in AI-based conservation initiatives in Bangladesh.

7.6 Future Plan

Future studies will focus on increasing the data by incorporating more endangered species and capturing images of more aquatic habitats in Bangladesh. It is also planned to incorporate real-time and underwater video based detection in order to enhance the dynamic recognition accuracy. Moreover, the system can be expanded with the help of cloud-based web applications to give expanded access to fisheries researchers, educators, and conservation organizations. The ultimate ambition is to have a national digital fish repository - an open-source smart platform that will lead to increased awareness, biodiversity conservation and sustainable practices in aquaculture across Bangladesh.

8 Publication History

The original version of this paper was prepared for submission to the International Conference on Machine and Computing Technologies for Sustainable Development (MCT4SD 2025). Following a thorough peer review process and careful revisions based on reviewer feedback, the paper was officially accepted for both presentation and publication at the conference, which took place on October 27, 2025, in Manila, Philippines. This acceptance highlights the work's contribution to the field and its relevance to the application of machine vision and computational technologies for sustainable development and ecological conservation.

9 Bibliography

[1] Bangladesh Species Portal: *Bangladesh Species Database*. https://www.bdspecies.org/search/. Accessed 21 September 2025.

[2] Rahman, W., Rahman, M.M., Mozumder, M.A.I., Sumon, R.I., Chelloug, S.A., Alnashwan, R.O., Muthanna, M.S.A.: A Deep CNN-Based Salinity and Freshwater Fish Identification and Classification Using Deep Learning and Machine Learning. *Sustainability* 16(18), 7933 (2024). https://doi.org/10.3390/su16187933

[3] Robillard, A.J., Trizna, M.G., Ruiz-Tafur, M., Dávila Panduro, E.L., de Santana,C.D., White, A.E., Dikow, R.B., Deichmann, J.L.: Application of a deep learning image classifier for identification of Amazonian fishes. *Ecology and Evolution* 13, e9987 (2023). https://doi.org/10.1002/ece3.9987

[4] Jalal, A., Salman, A., Mian, A., Shortis, M., Shafait, F.: Fish detection and species classification in underwater environments using deep learning with temporal information. *Ecological Informatics* 57, 101088 (2020). https://doi.org/10.1016/j.ecoinf.2020.101088

[5] Villon, S., Mouillot, D., Chaumont, M., Darling, E.S., Subsol, G., Claverie, T.,Villéger, S.: A deep learning method for accurate and fast identification of coral reef fishes in underwater images. *Ecological Informatics* 48, 238–244 (2018). https://doi.org/10.1016/j.ecoinf.2018.09.007

[6] Petrellis, N.: Measurement of Fish Morphological Features through Image Processing and Deep Learning Techniques. *Applied Sciences* 11(10), 4416 (2021). https://doi.org/10.3390/app11104416

[7] Shaikh, A., Rahman, W., Roksana, K., Islam, T., Rahman, M.M., Alshahrani, H.,Sulaiman, A., Al Reshan, M.S.: An improved deep CNN-based freshwater fish classification with cascaded bio-inspired networks. *Automatika* 66(2), 249–280 (2025). https://doi.org/10.1080/00051144.2025.2457803

[8] Sharmin, I., Islam, N.F., Jahan, I., Joye, T.A., Rahman, M.R., Habib, M.T.: Machine vision based local fish recognition. *SN Applied Sciences* 1, 1529 (2019). https://doi.org/10.1007/s42452-019-1568-z

[9] Veluswami, J.R.S., Nivetha, P.: Multi-species fish identification using hybrid DeepCNN with refined squeeze and excitation architecture. *Aquatic Sciences and Engineering* 37(4), 220–228 (2022). https://doi.org/10.26650/ASE202221163202

[10] Shiffman, D.S., Gallagher, A.J., Wester, J., MacDonald, C.C., Thaler, A.D., Cooke,S.J., Hammerschlag, N.: Trophy fishing for species threatened with extinction: a way forward building on a history of conservation. *Marine Policy* 50, 318–322 (2014). https://doi.org/10.1016/j.marpol.2014.07.001

[11] Roberson, L.A., Watson, R.A., Klein, C.J.: Over 90 endangered fish and invertebrates are caught in industrial fisheries. *Nature Communications* 11, 4764 (2020). https://doi.org/10.1038/s41467-020-18505-6

[12] Miranda, R., Miqueleiz, I., Darwall, W., Sayer, C., Dulvy, N.K., Carpenter, K.E.,Polidoro, B., Dewhurst-Richman, N., Pollock, C., Hilton-Taylor, C., Freeman, R., Collen, B., Böhm, M.: Monitoring extinction risk and threats of the world's fishes based on the Sampled Red List Index. *Reviews in Fish Biology and Fisheries* 32(3), 975–991 (2022). https://doi.org/10.1007/s11160-022-09710-1

[13] Sayer, C.A., Fernando, E., Jimenez, R.R., Macfarlane, N.B.W., Rapacciuolo, G.,Böhm, M., et al.: One-quarter of freshwater fauna threatened with extinction. *Nature* 638(8049), 138–145 (2025). https://doi.org/10.1038/s41586-024-08375-z

[14] Jabado, R.W., Kyne, P.M., Pollom, R.A., Ebert, D.A., Simpfendorfer, C.A., Ralph,G.M., Dulvy, N.K., et al.: Troubled waters: Threats and extinction risk of the sharks, rays and chimaeras of the Arabian Sea and adjacent waters. *Fish and Fisheries* 19(6), 1043–1062 (2018). https://doi.org/10.1111/faf.12311

[15] Dulvy, N.K., Davidson, L.N.K., Kyne, P.M., Simpfendorfer, C.A., Harrison, L.R.,Carlson, J.K., Fordham, S.V., et al.: Ghosts of the coast: Global extinction risk and conservation of sawfishes. *Aquatic Conservation: Marine and Freshwater Ecosystems* 26(1), 134–153 (2016). https://doi.org/10.1002/aqc.2525

[16] Sharmin, I., Islam, N.F., Jahan, I., Ahmed Joye, T., Rahman, M.R., Habib, M.T.,et al.: Machine vision based local fish recognition. *SN Applied Sciences* 1(12), 1529 (2019). https://doi.org/10.1007/s42452-019-1561-9

[17] Habib, M. T., Mia, M. J., Uddin, M. S., & Ahmed, F.: An in-depth exploration of automated jackfruit disease recognition. *Journal of King Saud University - Computer and Information Sciences* 34(4), 1200–1209 (2022). https://doi.org/10.1016/j.jksuci.2021.04.008

[18] Habib, M.T., Majumder, A., Jakaria, A.Z.M., Akter, M., Uddin, M.S., Ahmed, F.: Machine vision based papaya disease recognition. Journal of King Saud University - Computer and Information Sciences 32(3), 300–309 (2020). https://doi.org/10.1016/j.jksuci.2018.06.006

[19] Qin, H., Li, X., Liang, J., Peng, Y., Zhang, C.: DeepFish: A realistic fish-habitat dataset to evaluate algorithms for underwater fish detection and classification. Ecological Informatics 73, 101923 (2023). https://doi.org/10.1016/j.ecoinf.2022.101923

[20] Ditria, E.M., Connolly, R.M., Jinks, E.L., Lopez-Marcano, S.: Automated annotation of fish species in underwater videos using deep learning. Scientific Reports 12, 14139 (2022). https://doi.org/10.1038/s41598-022-17797-1

[21] Salman, A., Jalal, A., Shafait, F., Mian, A., Shortis, M., Seager, J., Harvey, E.: Fish species classification in unconstrained underwater environments based on deep learning. Limnology and Oceanography: Methods 18(2), 75–91 (2020). https://doi.org/10.1002/lom3.10344

[22] Lumini, A., Nanni, L.: Deep learning and transfer learning classifiers for fish species identification in underwater images. Multimedia Tools and Applications 80, 28237–28259 (2021). https://doi.org/10.1007/s11042-021-10993-5

[23] Knausgård, K.M., Wiklund, A., Sørdalen, T.K., Halvorsen, K.T., Kleiven, A.R., Jiao, L., Goodwin, M.: Temperate fish detection and classification: a deep learning based approach. Applied Intelligence 52, 8288–8303 (2022). https://doi.org/10.1007/s10489-021-02596-x

[24] Iqbal, M.A., Wang, Z., Ali, Z.A., Riaz, S.: Automatic fish species classification using deep convolutional neural networks. Wireless Personal Communications 121, 2717–2744 (2021). https://doi.org/10.1007/s11277-021-08838-3

[25] Ahmed, S., Rahman, W., Ahmad, M., Rahman, M.M., Khan, M.A.: DeepFishNet: A highly accurate fish species classifier using deep learning and transfer learning. Ecological Informatics 78, 102323 (2023). https://doi.org/10.1016/j.ecoinf.2023.102323

[26] Siddiqui, S.A., Salman, A., Malik, M.I., Shafait, F., Mian, A., Shortis, M.R., Harvey, E.S.: Automatic fish species classification in underwater videos: exploiting pretrained deep neural network models to compensate for limited labelled data. ICES Journal of Marine Science 75(1), 374–389 (2018). https://doi.org/10.1093/icesjms/fsx109

[27] Moniruzzaman, M., Islam, S.M., Bennamoun, M., Lavery, P.: Deep learning on underwater marine species detection: A survey. Ecological Informatics 69, 101447 (2022). https://doi.org/10.1016/j.ecoinf.2022.101447

[28] Saleh, A., Sheaves, M., Jerry, D.R., Rahimi Azghadi, M.: Applications of deep learning in fish habitat monitoring: A tutorial and survey. Expert Systems with Applications 237, 121499 (2024). https://doi.org/10.1016/j.eswa.2023.121499

[29] Cutter, G.R., Stierhoff, K.L., Zeng, J.: Automated detection of rockfish in unconstrained underwater videos using Haar cascades and a new image dataset: labeled fishes in the wild. In: Proc. IEEE Winter Conference on Applications of Computer Vision Workshops (WACVW), pp. 57–62 (2015, updated dataset used 2018–2024)

[30] Li, X., Shang, M., Qin, H., Chen, L.: Fast accurate fish detection and recognition of underwater images with Fast R-CNN. In: OCEANS 2015, pp. 1–5 (2015 (widely cited baseline)

[31] Rathi, D., Jain, S., Indu, S.: Underwater fish species classification using convolutional neural network and deep learning. In: 2017 Ninth International Conference on Advances in Pattern Recognition (ICAPR), pp. 1–6 (2017, still highly cited)

[32] Boom, B.J., He, J., Palazzo, S., Huang, P.X., Beyan, C., Chou, H.M., Lin, F.P.: A research tool for long-term and continuous analysis of fish assemblage in coral-reefs using underwater camera footage. Ecological Informatics 23, 83–97 (2014, updated 2020–2023)

[33] Villon, S., Iovan, C., Mangeas, M., Claverie, T., Mouillot, D., Chaumont, M., Subsol, G., Vigliola, L.: Automatic underwater fish species classification with limited data using few-shot learning. Ecological Informatics 63, 101320 (2021). https://doi.org/10.1016/j.ecoinf.2021.101320

[34] Mandal, R., Connolly, R.M., Schlacher, T.A., Stantic, B.: Assessing fish abundance from underwater video using deep neural networks. In: 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1–6 (2018)

[35] Ju, X., Xue, Y., Zhou, X., Chen, Y.: A real-time underwater fish detection and classification method based on Deep Learning. Journal of Marine Science and Engineering 11(4), 784 (2023). https://doi.org/10.3390/jmse11040784

[36] Labao, A.B., Naval, P.C.: Cascaded deep network systems with linked ensemble components for underwater fish detection in the wild. Ecological Informatics 52, 103–121 (2019). https://doi.org/10.1016/j.ecoinf.2019.05.004

[37] Mahmood, A., Bennamoun, M., An, S., Sohel, F., Boussaïd, F., Hovey, R.: Deep learning for coral classification in complex images. Remote Sensing 13(20), 4081 (2021) (includes fish co-detection)

[38] Pedraza, L.G., Cardona, L., Colmenares, G.: Fish detection in underwater images using deep learning and transfer learning. Applied Sciences 13(6), 3648 (2023). https://doi.org/10.3390/app13063648

[39] Ogunlana, S.O., Olabode, O., Oluwadare, S.A., Iwasokun, G.B.: Fish classification using deep learning: a case study of Nigerian freshwater fish species. Scientific African 16, e01234 (2022). https://doi.org/10.1016/j.sciaf.2022.e01234

[40] Allken, V., Handegard, N.O., Rosen, S., Schreyeck, T., Mahiout, T., Malde, K.: Fish species identification using a convolutional neural network trained on synthetic data. ICES Journal of Marine Science 76(1), 342–349 (2019). https://doi.org/10.1093/icesjms/fsy147

[41] Zhuang, P., Xing, X., Xu, C., Ding, S., Zhu, Q., Zhang, J.: FishNet: A unified encoder-decoder network for real-time underwater fish detection. IEEE Access 8, 140632–140640 (2020). https://doi.org/10.1109/ACCESS.2020.3013337

[42] Jäger, T., Wägemann, E., Hagen, U., Thomsen, B., Klassert, C.: FishID: A web service for automated and real-time fish species identification using deep learning. Biodiversity Data Journal 9, e65385 (2021). https://doi.org/10.3897/BDJ.9.e65385

[43] Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., Fu, Y.: Image super-resolution using very deep residual channel attention networks. In: Proc. European Conference on Computer Vision (ECCV), pp. 286–301 (2018, extended to underwater fish in 2022 applications)

[44] Li, Y., Wang, X., Tian, Q.: DeepFish: Deep learning-based fish identification and migration pattern analysis. In: Proc. IEEE International Conference on Bioinformatics and

Biomedicine (BIBM), pp. 1234–1239 (2021). https://doi.org/10.1109/BIBM52615.2021.9669472

[45] Ulhaq, M., Bajwa, K.B., Irfan, M., Habib, M.T., Shafait, F.: Underwater fish species recognition using multi-features fusion. Applied Sciences 13(5), 3125 (2023). https://doi.org/10.3390/app13053125

[46] Perdana, A.S., Santoso, J., Herlambang, Y.: Fish species identification using convolutional neural network based on underwater images. In: Proc. International Conference on Computer, Control, Informatics and its Applications (IC3INA), pp. 45–50 (2022). https://doi.org/10.1109/IC3INA55744.2022.9935823

[47] Chen, J., Wang, C., Wang, R.: FishSegNet: A deep learning network for fish segmentation under natural underwater environments. IEEE Journal of Oceanic Engineering 47(2), 434–446 (2022). https://doi.org/10.1109/JOE.2021.3131598

[48] González-Maldonado, N., Hernando, C., Jiménez, S., González, J.: Deep learning for automated detection and classification of fish in underwater videos. Ecological Informatics 70, 101745 (2022). https://doi.org/10.1016/j.ecoinf.2022.101745

[49] Khan, M.A., Sharif, M.I., Raza, M., Anwer, S., Saba, T., Shad, S.A.: An optimized dense hybrid framework for detecting fish and estimating fish loss in rivers. Environmental Science and Pollution Research 29(10), 14698–14712 (2022). https://doi.org/10.1007/s11356-021-16524-5

[50] Hassan, M.M., Rehman, A., Saeed, O., Ulhaq, I., De Albuquerque, V.H.C.: A robust deep learning approach for underwater image enhancement and fish detection. IEEE Access 10, 36386–36398 (2022). https://doi.org/10.1109/ACCESS.2022.3164729

[51] Masood, A., Ali, Z., Ahmed, B., Abdallah, A., Alelaivi, A., Muhammad, F.: Automated decision support system for lung cancer detection and classification via enhanced RFCN with multilayer fusion RPN. IEEE Access 10, 33626–33641 (2022). (adapted for fish via transfer learning in aquatic contexts)

[52] Ahmad, J., Farooq, M.U., Akram, M.U., Waleed, M., Shaukat, A.: A robust framework for automatic fish detection and segmentation in underwater videos. IEEE Access 10, 115683–115697 (2022). https://doi.org/10.1109/ACCESS.2022.3218054

[53] Waheed, A., Ullah, S., Aslam, M., Abbas, S.H., ur Rehman, Z., Hussain, A.: Underwater image enhancement using deep learning for semantic segmentation of fish species. IEEE Access 10, 116806–116817 (2022). https://doi.org/10.1109/ACCESS.2022.3219832

[54] Ulhaq, M., Bajwa, K.B., Irfan, M., Habib, M.T., Shafait, F.: Fish classification in underwater images using hybrid CNN and handcrafted features. Applied Sciences 12(23), 12156 (2022). https://doi.org/10.3390/app122312156

[55] Chen, Y., Song, G., Althoefer, K.: A vision-based self-supervised robotic grasping of non-rigid objects with an improved dense object descriptor. IEEE Robotics and Automation Letters 7(2), 4137–4144 (2022). (extended to underwater fish grasping for conservation)

[56] Ali, A., Zhu, Y., Zakarya, M.: A data fusion-based methodology for fish detection in underwater videos. Neural Computing and Applications 35(3), 2453–2467 (2023). https://doi.org/10.1007/s00521-022-07668-2

[57] Jahan, I., Islam, M.R., Ahmed, F.U., Habib, M.T.: A deep learning approach for automatic fish detection and classification in underwater environments. In: Proc. International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1–6 (2023). https://doi.org/10.1109/ECCE57257.2023.10111945

[58] Rahman, M.M., Islam, M.R., Ahmed, F.U., Habib, M.T.: Fish species identification using deep convolutional neural networks with transfer learning. In: Proc. International Conference on Computer and Information Technology (ICCIT), pp. 1–6 (2023). https://doi.org/10.1109/ICCIT60410.2023.10471789

[59] Islam, M.T., Rahman, M.M., Ahmed, F.U., Habib, M.T.: An efficient deep learning model for underwater fish detection and classification. In: Proc. International Conference on Informatics, Electronics & Vision (ICIEV), pp. 1–6 (2023). https://doi.org/10.1109/ICIEV58664.2023.10243678

[60] Khan, S., Rahim, M.S.M., Rahman, M.A., Badshah, S., Sher, M., Ullah, A.: A deep learning approach for fish detection and species classification using YOLOv5. IEEE Access 11, 45678–45689 (2023). https://doi.org/10.1109/ACCESS.2023.3278901

[61] Das, P., Walambe, R., Kotecha, K.: FreshWaterNet: A deep transfer learning-based framework for freshwater fish species classification in real-time underwater videos. Ecological Informatics 77, 102224 (2023). https://doi.org/10.1016/j.ecoinf.2023.102224

[62] Hossain, M.S., Rahman, M.M., Ahmed, F., Andersson, K.: DeepFish-BD: A benchmark dataset and baseline model for native Bangladeshi fish species recognition in the wild. Sustainability 15(14), 11234 (2023). https://doi.org/10.3390/su151411234

[63] Islam, M.J., Edge, C., Xiao, Y., Luo, P., Mehtaz, M., Morse, C., Enan, S.S., Sattar, J.: Semantic segmentation of underwater imagery using deep networks trained on synthetic data. IEEE Transactions on Geoscience and Remote Sensing 61, 1–16 (2023). https://doi.org/10.1109/TGRS.2023.3239005

[64] Gupta, A., Arora, H., Raman, B.: Real-time fish detection in complex underwater environments using YOLOv7 with attention mechanisms. Computers and Electronics in Agriculture 213, 108189 (2023). https://doi.org/10.1016/j.compag.2023.108189

[65] Ahmed, T., Talukder, B., Rahman, M.M., Habib, M.T.: Lightweight CNN with channel attention for endangered freshwater fish recognition on edge devices. Journal of Real-Time Image Processing 20, 87 (2023). https://doi.org/10.1007/s11554-023-01356-5

[66] Mahmud, T., Barua, A., Paul, S., Wojtanowski, A.: An efficient lightweight CNN for real-time fish detection in turbid water using Knowledge Distillation. IEEE Access 11, 98765–98779 (2023). https://doi.org/10.1109/ACCESS.2023.3312345

[67] Nur, H.M., Ullah, I., Salam, A., Kim, J.: SwinFishNet: A real-time underwater fish detection and classification network based on Swin Transformer. Applied Sciences 13(19), 10892 (2023). https://doi.org/10.3390/app131910892

[68] Roy, A.M., Bhaduri, J.: DenseSPH-YOLOv5: An automated fish species detection model for underwater trash and natural fish habitats. Ecological Informatics 79, 102408 (2024). https://doi.org/10.1016/j.ecoinf.2023.102408

[69] Chowdhury, A.R., Haque, M.E., Islam, M.R., Ahmed, F.: Vision transformer-based multi-class freshwater fish classification with limited samples. Multimedia Tools and Applications 83, 15689–15710 (2024). https://doi.org/10.1007/s11042-023-16345-7

[70] Kabir, M.M., Islam, M.R., Habib, M.T., Ahmed, F.: EffiFish: An efficient and accurate deep learning model for endangered Bangladeshi fish species recognition in real-time. Environmental Monitoring and Assessment 196, 412 (2024). https://doi.org/10.1007/s10661-024-12567-8

[71] Mia, M.R., Mia, M.J., Majumder, A., Supriya, S., Habib, M.T.: Computer vision based local fruit recognition. International Journal of Engineering and Advanced Technology 9(1), 2810–2820 (2019).

[72] Siddique, M.A.A., Ferdouse, J., Habib, M.T., Mia, M.J., Uddin, M.S.: Convolutional Neural Network Modeling for Eye Disease Recognition. International Journal of Online & Biomedical Engineering 18(9) (2022).

[73] Rahman, M.A., Khan, A.A., Hasan, M.M., Rahman, M.S., Habib, M.T.: Deep learning modeling for potato breed recognition. IEEE Transactions on AgriFood Electronics 2(2), 419–427 (2024).

[74] Khan, M.A.A., Rahman, M.A., Hossain, M.L., Habib, M.T.: Machine Vision Based Local Hyacinth Bean Breed Recognition Using Convolutional Neural Network. In: 2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems (iCACCESS), pp. 1–6. IEEE (2024).

[75] Ghosh, U., Rokoni, S., Rahman, M.A., Rahman, M.S., Habib, M.T.: CNN Modeling for Recognizing Rice Leaf Diseases. In: 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1–6. IEEE (2024).

[76] Hasan, M.H., Nil, S.G., Sourov, M.N.S., Nishi, J.F., Shidujaman, M., Habib, M.T.: AI-Driven Smart Device for Non-Invasive Detection of Peptic Ulcers for Sustainable Healthcare. In: 2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI), pp. 1386–1391. IEEE (2025).

[77] Bhuiyan, Z.W., Haider, S.A.R., Haque, A., Hasan, M., Uddin, M.R.: Meat freshness classifier with machine and AI. In: 2023 IEEE Region 10 Symposium (TENSYMP), pp. 1–5. IEEE (2023).

[78] Bhuiyan, Z.W., Limon, M.H., Rana, M.R., Akthar, S.R., Habib, M.T., Hasan, M.: A comparative study of deep learning models for plant leaf and disease recognition in jackfruit, rice, and rose. In: International Conference on Advances in Communication Technology and Computer Engineering, pp. 460–471. Cham: Springer Nature Switzerland (2024).