Solent University

Department of Science and Engineering

Feasibility Study on Automated Fish Species Classification Using Hybrid Deep Learning Models

Author   : Amirhosein Mohammadisabet - 102141537

Course Title   : Dissertation

Module Leader   : Dr Raza Hasan

Date   : 01/09/2024

Table of Contents

Table of Table

Table of Figure

Chapter 1: Introduction

1.1 Background

Fish species identification is a critical task in various fields, including biodiversity conservation, ecological research, and fisheries management. Accurate identification of fish species enables the monitoring of ecosystem health, the management of fish populations, and the enforcement of fishing regulations. Traditionally, fish species identification has been performed manually by experts, a process that is time-consuming, labour-intensive, and prone to human error. With the advent of advanced technologies, automated image-based classification has emerged as a promising solution to these challenges.

The rapid advancement of deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized the field of image classification. CNNs have demonstrated remarkable success in accurately classifying images across various domains, including medical imaging, facial recognition, and natural scene classification. In the context of fish species identification, CNNs offer a powerful tool for analysing large datasets of fish images, extracting intricate features, and achieving high accuracy in classification tasks.

This study explores the application of CNNs and other deep learning models to the classification of fish species. By leveraging transfer learning and advanced data augmentation techniques, the research aims to develop models that can accurately classify a diverse range of fish species from images. This automated approach has the potential to significantly enhance the efficiency and accuracy of fish species identification, providing valuable tools for researchers, conservationists, and industry professionals.

1.2 Problem Statement

The manual identification of fish species is a complex and error-prone process, particularly when dealing with large volumes of data or species with subtle morphological differences. Traditional machine learning methods, which rely on handcrafted features, have shown limited success in addressing these challenges. Although deep learning models have achieved state-of-the-art performance in general image classification tasks, their application to fish species identification remains underexplored. Furthermore, the variability in fish images, due to factors such as lighting conditions, occlusions, and different camera angles, adds to the complexity of the task.

The primary problem this study addresses is the need for an efficient and accurate automated system for fish species identification. Specifically, the research seeks to determine whether deep learning models, particularly CNNs, can outperform traditional machine learning algorithms in classifying fish species from images. Additionally, the study investigates the effectiveness of transfer learning in enhancing model performance, especially when dealing with a limited dataset.

1.3 Objectives of the Study

The main objectives of this study are as follows:

1. To develop and evaluate deep learning models for the classification of fish species from images. This includes the implementation of CNN-based models and the application of transfer learning to improve classification accuracy.
2. To compare the performance of deep learning models with traditional machine learning algorithms in the context of fish species identification, assessing the strengths and weaknesses of each approach.
3. To perform comprehensive exploratory data analysis (EDA) on the fish species image dataset to gain insights into the data's characteristics, which will inform model development and optimization.
4. To investigate the impact of data augmentation and preprocessing techniques on model performance, with the goal of enhancing the generalizability and robustness of the models.
5. To provide a detailed evaluation of the models using metrics such as accuracy, confusion matrix, and loss functions, ensuring a thorough assessment of their performance across different fish species.

1.4 Research Questions

This study seeks to answer the following research questions:

1. How effective are deep learning models, particularly CNNs, in accurately classifying fish species from images?
2. Can transfer learning significantly improve the performance of deep learning models on fish species classification tasks, especially when dealing with limited data?
3. How do deep learning models compare to traditional machine learning algorithms in terms of accuracy and robustness in fish species identification?
4. What insights can be gained from exploratory data analysis of the fish species image dataset, and how can these insights be used to improve model performance?
5. What is the impact of different data augmentation and preprocessing techniques on the performance of fish species classification models?

1.5 Significance of the Study

The significance of this study lies in its potential to advance the field of automated fish species identification. By leveraging state-of-the-art deep learning techniques, this research can contribute to the development of tools that are not only more accurate but also more scalable and efficient than traditional methods. The findings of this study can

have wide-ranging applications, from aiding in the conservation of marine biodiversity to supporting sustainable fishing practices.

Moreover, this study contributes to the broader field of image classification by providing insights into the challenges and solutions associated with classifying highly variable image data. The methodologies and models developed in this research could be adapted and applied to other domains where image classification plays a critical role.

1.6 Scope and Limitations

The scope of this study includes the development and evaluation of various deep learning models for fish species classification, with a focus on CNNs and transfer learning. The research also involves a comparative analysis with traditional machine learning algorithms to highlight the advantages and limitations of each approach.

However, the study is subject to certain limitations. The performance of the models may be constrained by the size and quality of the available dataset, and the results may not fully generalize to other types of image data. Additionally, while the study explores several data augmentation and preprocessing techniques, it does not exhaustively cover all possible methods, leaving room for further exploration in future research.

1.7 Organization of the Thesis

This thesis is organized as follows:

- Chapter 1: Introduction – Provides the background, problem statement, objectives, research questions, significance, scope, and organization of the thesis.
- Chapter 2: Literature Review – Reviews the existing literature on image classification techniques, with a focus on both traditional machine learning and deep learning approaches, particularly in the context of fish species identification.
- Chapter 3: Methodology – Details the dataset used, the preprocessing steps taken, the models implemented, and the evaluation metrics used in the study.
- Chapter 4: Results and Discussion – Presents the results of the experiments, including the performance of different models, and discusses the implications of these findings.
- Chapter 5: Conclusion and Future Work – Summarizes the key findings of the study, discusses its contributions, and outlines potential directions for future research.

Chapter 2: Literature Review

2.1 Introduction

The classification of fish species has long been a fundamental aspect of ecological studies, fisheries management, and environmental conservation. Accurate identification of fish species is crucial for monitoring biodiversity, managing aquatic resources, and ensuring the sustainability of marine and freshwater ecosystems. Traditional methods of fish species classification, primarily based on morphological characteristics, have served as the backbone of taxonomic studies for centuries. However, these methods often require expert knowledge, are time-consuming, and may not be practical for large-scale or real-time monitoring.

In recent decades, technological advancements have introduced new approaches to fish species classification, particularly through the use of machine learning and, more recently, deep learning techniques (1). These methods offer the potential to automate the classification process, significantly improving speed and accuracy while reducing the reliance on human expertise. The integration of artificial intelligence (AI) into this domain has opened up new possibilities for large-scale and real-time monitoring of fish populations, which is essential for addressing the growing challenges in fisheries management and biodiversity conservation (2).

The development of machine learning algorithms, such as Support Vector Machines (SVMs), Random Forests, and Neural Networks, marked a significant shift in how fish species are classified (3). These algorithms allowed for the automatic extraction of features from images, paving the way for more sophisticated and accurate classification systems. However, the advent of deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized the field by enabling the automatic learning of hierarchical features directly from raw images, leading to unprecedented levels of accuracy in fish species identification (4).

Despite these advancements, several challenges persist in the field of fish species classification. The diversity of fish species, the complexity of underwater environments, and the imbalanced nature of many datasets pose significant obstacles to achieving high classification accuracy across all species (5). Additionally, the need for real-time processing in applications such as electronic monitoring (EM) systems on commercial fishing vessels adds another layer of complexity to the problem (6).

This literature review aims to provide a comprehensive overview of the evolution of fish species classification methods, from traditional morphological approaches to the latest deep learning techniques (7). It will explore the challenges and limitations of these methods, as well as the various strategies developed to overcome them. Furthermore, this review will highlight the significant contributions of data augmentation, transfer learning, and hybrid models in improving the performance of fish classification systems

(8). Finally, the review will discuss the current state of the art and suggest potential future directions for research in this critical area (9).

2.2 Traditional Methods of Fish Species Classification

Before the advent of digital technologies and machine learning, fish species classification relied heavily on traditional methods rooted in morphological and taxonomic principles. These methods are fundamental to ichthyology and have been extensively used in ecological studies, fisheries management, and biodiversity assessments (10).

2.2.1 Morphological Identification

Morphological identification has been the cornerstone of fish species classification for centuries. This method involves the examination of physical characteristics such as body shape, fin structure, scale pattern, and coloration. These traits are used to distinguish species, often requiring expert knowledge to accurately identify and differentiate between closely related species (11). Morphological identification is typically performed by trained taxonomists who use dichotomous keys, which provide a step-by-step guide to identifying species based on their morphological traits.

However, this approach has limitations. It is labour-intensive and time-consuming, particularly when dealing with large numbers of specimens or when species have subtle morphological differences (12). Moreover, the identification process is prone to human error, and there is often a need for expert intervention, which may not be feasible in all situations, especially in remote or resource-limited areas (13).

2.2.2 Meristic and Morphometric Analysis

Meristic and morphometric analyses are two quantitative approaches used in traditional fish classification. Meristic analysis involves counting features such as the number of fin rays, scales, or vertebrae, which can vary between species. Morphometric analysis, on the other hand, involves measuring the size and shape of specific body parts, such as the length of the head or the depth of the body (14). These measurements are then compared statistically to identify and differentiate species.

Both methods have been instrumental in distinguishing species that are morphologically similar but differ in specific measurable traits. However, these techniques can be time-consuming and require precise measurement tools and conditions, which can be challenging to maintain in field studies. Additionally, these methods may not always capture the full extent of intraspecific variation, potentially leading to misidentifications (17).

2.2.3 Molecular Techniques

In the late 20th century, molecular techniques began to complement traditional morphological methods, providing a more objective means of species identification. Techniques such as DNA barcoding, which involves sequencing a standardized region of the genome (e.g., the cytochrome c oxidase I (COI) gene in fish), have been used to identify species based on genetic differences (18). This approach is particularly useful for identifying species that are morphologically indistinguishable or for confirming species identity in cases where morphological identification is uncertain.

Molecular methods have significantly enhanced the accuracy of species identification, especially in groups with high morphological plasticity or in early life stages where morphological features may not be fully developed (19). However, these techniques require specialized laboratory equipment and expertise, making them less accessible for routine monitoring or in-field identification (20).

2.2.4 Applications and Limitations

Traditional methods of fish species classification have been widely used in various applications, including the development of Indices of Biological Integrity (IBI) for assessing the health of aquatic ecosystems (21). For example, Halliwell et al. (1998) summarized the use of fish assemblages as indicators of environmental stress in the northeastern United States, highlighting the importance of accurate species identification in ecological assessments (1).

Despite their widespread use, traditional methods have several limitations. They are often not feasible for large-scale or high-throughput applications, such as monitoring commercial fisheries or conducting large biodiversity surveys (23). The reliance on expert knowledge and the potential for human error also limit their applicability in automated or real-time systems (24). Additionally, traditional methods may struggle to keep pace with the growing need for rapid and accurate species identification, particularly in the face of increasing environmental pressures and the need for effective conservation strategies (25).

2.3 Advent of Machine Learning in Fish Species Classification

The integration of machine learning (ML) into fish species classification has revolutionized the field, offering a significant improvement in accuracy, efficiency, and scalability over traditional methods (1). As computational power and access to large datasets have grown, machine learning techniques have become increasingly viable for processing and analysing the vast amount of data generated in ecological studies and fisheries management (2)

2.3.1 Introduction to Machine Learning in Ichthyology

Machine learning, a subset of artificial intelligence, involves the development of algorithms that allow computers to learn from and make decisions based on data (2). In ichthyology, ML techniques are employed to automate the classification of fish species, analyse ecological patterns, and monitor biodiversity. Unlike traditional methods, which rely heavily on expert knowledge and manual labour, machine learning can process large datasets efficiently, identify patterns that are not easily discernible by humans, and continuously improve as more data becomes available (4).

The introduction of ML in fish species classification addresses several limitations of traditional methods. For instance, it reduces the reliance on expert taxonomists, minimizes human error, and allows for the analysis of large and complex datasets that would be impractical to handle manually (5). Additionally, ML models can be trained to recognize species in diverse conditions and environments, enhancing the robustness of species identification in various ecological contexts (6).

2.3.2 Supervised Learning Techniques

Supervised learning, one of the most common ML approaches in fish classification, involves training a model on a labelled dataset, where the input data (e.g., images of fish) are paired with the correct output (e.g., species labels) (7). Once trained, the model can predict the species of fish in new, unseen images.

Several studies have applied supervised learning techniques to fish species classification with promising results. For example, convolutional neural networks (CNNs), a type of deep learning model, have been widely used due to their ability to automatically extract hierarchical features from images (8). CNNs have been successfully applied to classify fish species from underwater images, achieving high levels of accuracy and robustness. Xu et al. (2022) employed a SE-ResNet152 model with class-balanced focal loss for fish species identification, achieving impressive accuracy rates across different fish image views (9).

In another study, Alaba et al. (2022) proposed a deep learning-based approach using MobileNetv3-large and VGG16 networks to address the challenge of imbalanced datasets in fish species recognition. Their model incorporated a class-aware loss function that improved the identification of less represented species in the dataset, demonstrating the effectiveness of tailored ML techniques in ecological studies (10).

2.3.3 Transfer Learning and Data Augmentation

Transfer learning is a powerful ML technique that involves taking a pre-trained model developed for a large dataset and fine-tuning it for a specific task, such as fish species classification (11). This approach is particularly useful when dealing with limited datasets, a common scenario in ecological studies. By leveraging the knowledge gained

from the pre-trained model, transfer learning can significantly reduce the amount of data and time required to achieve high classification accuracy (12).

Gong et al. (2023) utilized transfer learning in their Fish-TViT model, which combines transfer learning with visual transformers for classifying fish species in various aquatic environments. Their approach included data augmentation strategies to expand the training dataset, further improving the model's performance and generalization ability (13). Mujtaba and Mahapatra (2021) also highlighted the benefits of data augmentation in enhancing the accuracy of fish species classification models by creating additional training examples from existing data (14).

2.3.4 Object Detection and Real-Time Monitoring

Beyond classification, machine learning has been employed in object detection tasks, where the goal is to identify and locate multiple fish within a single image or video frame (15). This is particularly useful in monitoring applications, where automated systems need to detect and count fish in real-time.

For instance, the YOLO (You Only Look Once) architecture has been applied to fish detection, allowing for fast and accurate identification of multiple fish species in a single pass (16). Dharshana et al. (2022) implemented a YOLO-based method to classify fish species in various environments, achieving significant improvements in both speed and accuracy compared to traditional convolutional networks (17)

2.3.5 Challenges and Future Directions

Despite the advancements brought by machine learning, several challenges remain in fish species classification. One of the primary issues is the need for large, annotated datasets, which are often difficult to obtain (18). Additionally, models trained on one dataset may not generalize well to new environments or conditions, necessitating ongoing model updates and retraining (19).

Another challenge is the interpretability of ML models, particularly deep learning approaches. While these models can achieve high accuracy, understanding how they make decisions is often difficult, which can be a barrier to their adoption in some fields (20).

Future research is likely to focus on improving model generalization, developing more interpretable ML techniques, and integrating these models into broader ecological and conservation strategies (21). Additionally, as more data becomes available, there is potential for creating global databases and models that can be used across different regions and ecosystems (22).

2.4 Deep Learning in Fish Species Classification

Deep learning, a subset of machine learning, has significantly advanced the field of fish species classification by providing more accurate and efficient methods for processing and analysing complex visual data (1). Deep learning models, particularly convolutional neural networks (CNNs), have become the state-of-the-art in image classification tasks, including the identification and classification of fish species from underwater images (2). This section delves into the various deep learning techniques applied to fish species classification, their advantages, challenges, and the key studies that have shaped the current landscape (3).

2.4.1 Introduction to Deep Learning in Image Classification

Deep learning has revolutionized image classification tasks by automating the feature extraction process and enabling the analysis of large datasets with minimal human intervention (4). Unlike traditional machine learning methods, which require manual feature engineering, deep learning models can learn hierarchical representations of data directly from raw inputs, such as images (5). This capability is particularly beneficial for fish species classification, where subtle variations in shape, colour, and texture are critical for accurate identification (6).

Convolutional neural networks (CNNs) are the most widely used deep learning models in image classification. They are particularly effective in processing grid-like data such as images due to their ability to automatically detect spatial hierarchies of features (7). In the context of fish species classification, CNNs have been extensively used to analyse underwater images, extract relevant features, and classify fish into different species (8).

2.4.2 Convolutional Neural Networks (CNNs) for Fish Species Classification

CNNs have been the backbone of many deep learning models for fish species classification due to their proficiency in image analysis (9). These networks are composed of layers that apply convolution operations to the input image, capturing local patterns such as edges, textures, and shapes that are crucial for distinguishing between different fish species (10).

One of the early applications of CNNs in this field was demonstrated by Tamou et al. (2022), who developed a deep convolutional neural network (DCNN) for fish species classification in unconstrained underwater environments (11). Their model utilized targeted data augmentation to improve the robustness of the classifier, achieving an accuracy of 99.86% on the Fish Recognition Ground-Truth dataset (12). This study highlighted the effectiveness of CNNs in handling the complex and variable conditions of underwater imaging (13).

The success of CNNs in fish species classification has also been demonstrated in studies focusing on specific challenges such as imbalanced datasets and low-resolution images (14). Xu et al. (2022) proposed an SE-ResNet152 model that integrated a class-balanced focal loss function to address the issue of small-scale, unbalanced datasets (15). Their approach achieved high accuracy across multiple views of fish images, demonstrating the versatility of CNNs in dealing with diverse classification challenges.

2.4.3 Transfer Learning in Deep Learning Models

Transfer learning, an approach that involves fine-tuning a pre-trained model on a new dataset, has been widely adopted in fish species classification to overcome the challenges of limited labelled data. This technique allows models that have been trained on large, diverse datasets to be adapted to the specific task of fish species classification, thereby reducing the need for extensive training data and computational resources (17).

For instance, the Fish-TViT model proposed by Gong et al. (2023) employed transfer learning in combination with visual transformers to enhance feature extraction from fish images (10). The use of transfer learning enabled the model to achieve high classification accuracy on both marine and freshwater fish datasets, illustrating the effectiveness of this approach in handling various aquatic environments (8).

Transfer learning not only improves the performance of deep learning models but also accelerates the training process, making it feasible to deploy these models in real-time monitoring systems where rapid and accurate species identification is required (10).

2.4.4 Advanced Deep Learning Architectures

Beyond CNNs, other advanced deep learning architectures have been explored for fish species classification, including vision transformers and hybrid models that combine CNNs with other machine learning techniques. These models aim to improve the accuracy and robustness of species identification, particularly in challenging scenarios such as detecting multiple fish in a single image or dealing with varying image resolutions (22).

Gong et al. (2023) introduced the use of vision transformers in their Fish-TViT model, which segments fish images into patches and processes them through a transformer network to capture long-range dependencies in the data. This approach outperformed traditional CNNs in certain tasks, particularly in handling complex and high-resolution images (8).

Similarly, Alaba et al. (2022) combined MobileNetv3-large and VGG16 networks in their deep learning model to address the issue of class imbalance in fish species datasets. By incorporating a class-aware loss function, their model improved the identification of underrepresented species, demonstrating the potential of hybrid architectures in addressing specific challenges in fish species classification (10).

2.4.5 Challenges and Future Directions in Deep Learning

While deep learning has significantly advanced fish species classification, several challenges persist. One major issue is the need for large, labelled datasets, which are resource-intensive to collect and annotate, especially in variable underwater environments. Another challenge is the interpretability of deep learning models; despite their high accuracy, understanding their decision-making process can be complex, limiting their adoption in some areas. Additionally, these models are often computationally demanding, requiring specialized hardware, which can be a barrier in resource-limited settings. Future research is expected to focus on developing more efficient and interpretable models and creating standardized, open-access datasets for broader use. The integration of deep learning with technologies like Internet of Things (IoT) devices and autonomous underwater vehicles (AUVs) also holds promise for real-time monitoring and conservation efforts (2).

2.5 Data Augmentation and Handling Imbalanced Datasets

Data augmentation and the management of imbalanced datasets are crucial techniques in enhancing the performance of deep learning models, particularly in fields like fish species classification, where datasets are often small and skewed. This section explores the strategies employed to augment data and address class imbalances, which are common challenges in training robust and generalizable models.

2.5.1 Importance of Data Augmentation

Data augmentation involves generating additional training data by applying various transformations to the existing dataset. This technique is especially important in deep learning, where large amounts of labelled data are required to train models effectively (22). In the context of fish species classification, data augmentation helps to artificially increase the size of the dataset, allowing models to learn more robust features and generalize better to new, unseen data.

Common augmentation techniques include geometric transformations (such as rotations, translations, and flips), colour space augmentations (such as changes in brightness, contrast, and saturation), and image noise injection. These methods simulate the variability that might be encountered in real-world environments, such as different viewing angles, lighting conditions, and underwater distortions (24).

For example, in the study by Tamou et al. (2022), targeted data augmentation was applied to enhance the performance of a deep convolutional neural network (DCNN) for fish species identification. By applying specific augmentations that mimicked the variations in underwater images, the model achieved an impressive accuracy of 99.86% on the Fish Recognition Ground-Truth dataset. This demonstrates the effectiveness of data augmentation in improving model robustness against real-world variability (12).

2.5.2 Handling Imbalanced Datasets

Imbalanced datasets, where some classes have significantly more examples than others, pose a significant challenge in training deep learning models. In fish species classification, this imbalance often occurs due to the natural distribution of species or the difficulty in capturing images of rare or elusive species. Without proper handling, models trained on imbalanced data tend to be biased towards the majority classes, leading to poor performance on minority classes (13).

Several techniques have been developed to address the issue of class imbalance:

Resampling Techniques: Resampling methods, such as oversampling the minority class or under-sampling the majority class, are commonly used to balance the dataset. However, these methods can introduce other challenges, such as overfitting in the case of oversampling.

Cost-Sensitive Learning: Cost-sensitive learning involves assigning higher misclassification costs to minority classes, thereby encouraging the model to pay more attention to them (16). This approach is often implemented through modified loss functions that penalize errors on minority classes more heavily.

Class-Balanced Loss Functions: Advanced loss functions, such as the focal loss and class-balanced focal loss, have been proposed to address class imbalance more effectively (18). These loss functions dynamically adjust the weight of each class based on its frequency in the dataset, reducing the bias towards the majority class.

Xu et al. (2022) employed a class-balanced focal loss function in their SE-ResNet152 model to tackle the issue of imbalanced fish species datasets. By giving more weight to underrepresented species, their model achieved high accuracy across different fish image views, demonstrating the effectiveness of this approach in improving classification performance for minority classes (8).

Similarly, Alaba et al. (2022) introduced a class-aware loss function in their deep learning model to enhance the recognition of fish species in an imbalanced dataset. Their method significantly improved the identification accuracy of rare species, underscoring the importance of tailored loss functions in handling dataset imbalance (12).

2.5.3 Hybrid Approaches and Future Directions

In addition to individual techniques, hybrid approaches that combine data augmentation with class imbalance handling have shown promise in further improving model performance. For instance, augmenting data while simultaneously applying a class-balanced loss function can create a more diverse and representative training set, leading to better generalization and higher accuracy, especially for minority classes.

Future research in this area may focus on developing more sophisticated data augmentation techniques that better mimic the conditions of specific environments, such as underwater habitats, and on refining loss functions to dynamically adapt to the evolving data distribution during training (27). Additionally, the integration of generative models, such as Generative Adversarial Networks (GANs), for synthetic data generation offers a promising avenue for augmenting rare classes and further mitigating the effects of class imbalance (28).

2.6. Hybrid and Multi-Modal Approaches

Hybrid and multi-modal approaches represent a sophisticated evolution in fish species classification, combining various methodologies and data types to enhance the accuracy and robustness of classification model. These approaches integrate multiple machine learning algorithms, deep learning techniques, and diverse data sources, offering significant improvements over traditional single-method models (2).

2.6.1 Hybrid Models Combining Machine Learning and Deep Learning Techniques

Hybrid models blend the strengths of traditional machine learning algorithms with advanced deep learning techniques to tackle the complexities of fish species classification. These models are designed to optimize the performance of classification tasks by leveraging the unique advantages of each component (4).

For instance, Mampitiya et al. (2022) developed a hybrid model that combines supervised machine learning algorithms, including Support Vector Machine (SVM), Random Forest (RF), and Neural Networks (NN), for the classification of underwater fish species. Their study demonstrated that the hybrid approach could achieve a higher accuracy of 99.89% compared to individual models by effectively handling noisy data and enhancing prediction accuracy through feature extraction methods like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) (16).

Similarly, Dharshana et al. (2022) proposed a hybrid approach that integrates the You Only Look Once (YOLO) object detection architecture with traditional Convolutional Neural Networks (CNNs) for fish species classification (17). This model capitalized on YOLO's ability to detect objects in real-time, combined with CNN's powerful feature extraction capabilities, to accurately classify fish species under varying conditions such as different scales and positions. This hybrid method outperformed existing models, achieving a significant improvement in classification accuracy (11).

2.6.2 Multi-Modal Approaches Utilizing Diverse Data Sources

Multi-modal approaches extend beyond the use of image data alone, incorporating additional data types such as sonar information, environmental factors, and even acoustic signals to improve classification outcomes. These approaches recognize that

relying solely on visual data may not be sufficient in complex environments, such as turbid or deep waters where visual clarity is compromised (19).

Xu et al. (2022) explored the integration of different views of fish images (body, head, and scale) in a multi-modal approach using a SE-ResNet152 model. This method achieved high accuracy across all views, demonstrating that incorporating multiple perspectives can significantly enhance the model's ability to distinguish between similar species. The study also addressed challenges related to imbalanced datasets by applying a class-balanced focal loss function, further improving the model's performance (6).

Moreover, Alaba et al. (2022) introduced a class-aware fish species recognition model that utilizes a combination of MobileNetv3-large and VGG16 backbone networks, coupled with a class-aware loss function. This multi-modal approach was particularly effective in handling imbalanced datasets and detecting multiple fish species within a single image, which is crucial in dynamic underwater environments. The model's ability to integrate multiple data modalities and address class imbalance led to a substantial improvement in classification accuracy, especially for underrepresented species (10).

2.6.3 Benefits and Challenges of Hybrid and Multi-Modal Approaches

The primary benefit of hybrid and multi-modal approaches lies in their ability to leverage the strengths of different methodologies, resulting in more robust and accurate classification models. By combining various techniques and data types, these approaches can effectively handle the complexities of real-world environments, where factors like water clarity, lighting conditions, and species diversity present significant challenges.

However, these approaches also come with challenges. The integration of multiple models and data types can increase the complexity of the system, making it more computationally intensive and harder to implement in real-time applications. Additionally, the need for extensive training data across multiple modalities can be a limiting factor, especially in environments where data collection is difficult.

Despite these challenges, the advancements in hybrid and multi-modal approaches hold great promise for the future of fish species classification. As these methods continue to evolve, they are likely to play a crucial role in improving the accuracy and efficiency of species identification in diverse aquatic environments.

2.7 Challenges in Underwater Fish Classification

Underwater fish classification presents unique challenges that differentiate it from other image classification tasks. These challenges arise from the inherent characteristics of the underwater environment, the variability of fish appearances, and the limitations of current technology. Addressing these challenges is crucial for developing robust and accurate classification systems (29).

2.7.1 Environmental Factors

The underwater environment is dynamic and often harsh, significantly impacting the quality and consistency of the data used for fish classification. Factors such as water turbidity, varying light conditions, and the presence of particulate matter can degrade image clarity, making it difficult to accurately identify fish species.

Mampitiya et al. (2022) highlight the issue of "hazy effects" caused by dissolved particles that reflect and scatter light, leading to image distortion. These environmental factors can obscure critical features of fish, such as colour patterns and scale structures, which are essential for accurate classification. This degradation in image quality poses a significant challenge for machine learning models that rely on clear and consistent visual data (15).

2.7.2 Variability in Fish Appearance

Fish species exhibit considerable variability in their appearance, influenced by factors such as age, sex, and environmental conditions. This intra-species variability can lead to significant overlaps in the visual features of different species, complicating the classification process.

Dharshana et al. (2022) address the challenge of variability in fish scale patterns and body positions, which can confuse classification algorithms. The same species can appear differently under various conditions, such as different angles, sizes, or behaviours, making it challenging for models to maintain high accuracy across diverse scenarios (16).

2.7.3 Class Imbalance

Class imbalance is another critical challenge in underwater fish classification. Some fish species are more prevalent than others in the dataset, leading to an overrepresentation of certain classes. This imbalance can cause classification models to become biased towards the more common species, resulting in poor performance on rarer classes.

Alaba et al. (2022) discuss the impact of class imbalance on model performance, noting that standard classification models tend to perform poorly on underrepresented species. This is problematic in applications where identifying rare or endangered species is critical. To mitigate this, they propose the use of a class-aware loss function that adjusts the model's focus towards less common species (10).

2.7.4 Real-Time Processing Requirements

Real-time processing is essential for many practical applications of fish species classification, such as monitoring fishing activities or conducting ecological surveys. However, achieving real-time performance while maintaining high accuracy is

challenging due to the computational complexity of deep learning models and the large volumes of data generated in underwater environments.

Xu et al. (2022) emphasize the need for efficient models that can process data in real time without compromising accuracy. The computational demands of processing high-resolution images, coupled with the need for rapid inference, require innovative approaches to model optimization and data handling (6).

2.7.5 Limited and Imbalanced Datasets

The availability of large, well-annotated datasets is crucial for training robust machine learning models. However, in underwater fish classification, such datasets are often limited and imbalanced. Collecting underwater images is resource-intensive, and manual annotation requires specialized knowledge, leading to smaller datasets compared to other domains (2).

Xu et al. (2022) address the challenge of limited data availability and propose using transfer learning and data augmentation techniques to overcome this limitation. Despite these techniques, the scarcity of diverse and representative data remains a significant obstacle to developing generalizable models (17).

2.7.6 Occlusion and Background Noise

Fish often inhabit complex underwater environments where they may be partially occluded by objects such as plants, rocks, or other fish. This occlusion, combined with background noise from the environment, such as moving water or floating debris, makes it difficult for classification algorithms to isolate and identify fish accurately.

Dharshana et al. (2022) highlight the difficulties posed by occlusion and background noise, noting that many existing algorithms are designed for clearer, less cluttered environments and may struggle to adapt to the complexities of underwater scenes (28). This limitation requires the development of more sophisticated algorithms capable of distinguishing fish from their surroundings under challenging conditions (16).

2.8 Summary of Similar Studies

To provide a clearer context for the advancements and challenges in fish species classification, the following table summarizes recent studies, the classifiers used, their key findings, limitations, and how they relate to existing gaps. This analysis helps in understanding the current state of research and highlights where further improvements are needed.

Table 1, Recent studies, the classifiers used, their key findings and limitations

Study	Classifiers Used	Key Findings	Limitations	Gap Analysis
Xu et al. (2022)	SE-ResNet152	Achieved high accuracy across different fish image views	Limited data diversity, challenges with imbalanced datasets	Addresses high accuracy but lacks robustness in real-time processing
Alaba et al. (2022)	MobileNetv3-large, VGG16	Improved accuracy on imbalanced datasets	High computational cost, requires extensive data	Effective in handling imbalanced datasets but not optimized for real-time applications
Mampitiya et al. (2022)	SVM, Random Forest, Neural Networks	Hybrid approach improved accuracy and feature extraction	Overfitting with noisy data, limited to specific conditions	Hybrid approach useful but lacks adaptability to varying underwater conditions
Dharshana et al. (2022)	YOLO, CNN	Real-time object detection and classification	Limited performance under varying scales and positions	Real-time focus is strong but struggles with high variability in fish appearance
Mujtaba and Mahapatra (2021)	Convolutional Neural Networks	Enhanced efficiency in electronic monitoring systems	Limited to video footage, challenges with occlusion	Effective in EM systems but may not generalize well to all underwater scenarios
Shao et al. (2023)	Vision Transformers	Improved classification accuracy with fine-grained details	Requires large datasets and significant computational resources	Promising for detailed classification but data requirements can be a bottleneck
Zhang et al. (2023)	GANs for data augmentation	Generated synthetic fish images improved model robustness	Quality of synthetic data may not fully represent real-world scenarios	Enhances training with synthetic data but still needs real-world validation
Chen et al. (2021)	Multi-modal approach (visual + acoustic)	Improved accuracy in detecting and classifying species using combined data	Complexity in integrating and processing multi-modal data	Effective in diverse conditions but integration challenges remain
Lee et al. (2022)	Deep CNN with attention mechanisms	Enhanced feature extraction and focus on critical features	High computational cost and potential overfitting	Advances feature extraction but may require optimization for efficiency
Jiang et al. (2021)	Reinforcement Learning	Adapted classification models for dynamic environments	Model training can be time-consuming and resource-intensive	Suitable for adaptive environments but needs refinement for practical use
Nguyen et al. (2022)	Hybrid CNN + RNN	Addressed temporal changes in underwater imagery	Challenges with real-time processing and long-term dependencies	Useful for dynamic scenes but may struggle with real-time constraints

2.8 Evaluation Metrics and Benchmark Datasets

In the field of fish species classification, the evaluation of model performance and the selection of appropriate benchmark datasets are critical components of the research process. Proper evaluation metrics ensure that models are assessed on their ability to generalize across diverse conditions, while benchmark datasets provide standardized data for training and testing, enabling fair comparisons between different approaches.

2.8.1 Evaluation Metrics

Evaluating the performance of fish species classification models involves several metrics that capture various aspects of the model's accuracy, robustness, and efficiency. The choice of metrics often depends on the specific application and the challenges inherent in underwater image classification.

2.8.1.1 Accuracy

Accuracy is the most commonly used metric in classification tasks, representing the proportion of correctly classified instances out of the total number of instances. In fish species classification, accuracy is particularly important in scenarios where misclassification could lead to significant ecological or economic consequences. For instance, Xu et al. (2022) (6) reported high accuracy rates of 98.80% and 96.67% for their SE-ResNet152 model across different fish image views, highlighting the model's effectiveness in species identification.

2.8.1.2 Precision, Recall, and F1-Score

Precision and recall are crucial metrics, especially in imbalanced datasets where certain species may be underrepresented. Precision measures the proportion of true positive predictions among all positive predictions, indicating the model's ability to avoid false positives. Recall measures the proportion of true positives identified out of all actual positives, reflecting the model's ability to detect relevant instances. The F1-score, the harmonic mean of precision and recall, provides a single metric that balances these two aspects. Alaba et al. (2022) (10) emphasized the importance of these metrics in their study, particularly in scenarios where rare species are involved, to ensure that the model performs well across all classes, not just the majority ones.

2.8.1.3 Confusion Matrix

The confusion matrix is a valuable tool for understanding the performance of classification models in more detail. It breaks down the predictions into true positives, true negatives, false positives, and false negatives, providing insight into where the model is making errors. This is particularly useful in fish species classification, where misclassifications can have varying degrees of impact depending on the species involved.

2.8.1.4 Area Under the Receiver Operating Characteristic Curve (AUC-ROC)

The AUC-ROC metric evaluates the model's ability to discriminate between classes. It is particularly useful in scenarios where the costs of false positives and false negatives are different, such as in the identification of endangered species. A higher AUC-ROC score indicates better performance in distinguishing between classes across different threshold settings.

2.8.1.5 Processing Time and Efficiency

In practical applications, particularly those involving real-time monitoring, the processing time and computational efficiency of the model are crucial. Xu et al. (2022) (6) underscore the importance of efficiency, especially when deploying models in resource-constrained environments such as on autonomous underwater vehicles (AUVs). Evaluating models based on their processing time per image and their ability to operate in real-time scenarios is essential for practical deployment.

2.8.2 Benchmark Datasets

Benchmark datasets are essential for developing, testing, and comparing fish species classification models. They provide standardized data that researchers can use to validate their models' performance and ensure that improvements are generalizable across different conditions and environments.

2.8.2.1 Fish-Pak Dataset

The Fish-Pak dataset is a widely used benchmark in fish species classification research. It includes images from different fish species and is particularly valued for its diversity in terms of species, image quality, and environmental conditions. Xu et al. (2022) (6) used the Fish-Pak dataset to test their SE-ResNet152 model, achieving high accuracy rates, demonstrating the dataset's utility in evaluating model performance across various fish image views.

2.8.2.2 LifeClef Fish Dataset

The LifeClef Fish dataset is another important benchmark, particularly known for its challenging underwater conditions and high species diversity. It has been used extensively in competitions and research focused on underwater species identification. The dataset includes annotated images of fish from various aquatic environments, making it ideal for testing the robustness of classification models under real-world conditions. LifeClef (2022) (18) provides this dataset.

2.8.2.3 FishNet Open Image Database

The FishNet Open Image Database is a large-scale dataset containing over 85,000 images from electronic monitoring (EM) footage of fisheries. It is used primarily for developing models aimed at automated fish species classification in commercial fishing operations. Mujtaba and Mahapatra (2021) (7) utilized this dataset to train their

convolutional neural network model, demonstrating its effectiveness in dealing with high volumes of real-world data.

2.8.2.4 SEAMAPD21 Dataset

The SEAMAPD21 dataset is notable for its use in class-aware fish species recognition. It contains a large number of images collected from reef ecosystems, with a particular focus on species diversity and the challenges posed by imbalanced datasets. Alaba et al. (2022) (10) used this dataset to test their proposed model and class-aware loss function, achieving significant improvements in handling imbalanced data.

2.9 Applications and Real-World Implementations

The classification of fish species has broad applications across various fields, including fisheries management, environmental monitoring, and aquaculture. Advances in fish species classification technologies have enabled more accurate and efficient methods for monitoring aquatic ecosystems, managing fisheries, and supporting conservation efforts. This section explores the diverse applications and real-world implementations of fish species classification technologies.

2.9.1 Fisheries Management

Effective fisheries management is crucial for sustainable fishing practices and the conservation of marine resources. Accurate fish species classification helps in monitoring fish populations, assessing the health of marine ecosystems, and enforcing fishing regulations.

2.9.1.1 Electronic Monitoring Systems

Electronic monitoring (EM) systems equipped with fish species classification algorithms are increasingly being used on commercial fishing vessels. These systems capture and analyse video footage to identify fish species in real-time, helping to ensure compliance with fishing quotas and regulations. Mujtaba and Mahapatra (2021) (7) demonstrated how convolutional neural networks and data augmentation techniques can enhance the efficiency of EM systems by improving the accuracy of fish species identification from video footage.

2.9.1.2 Automated Catch Reportin

Automated catch reporting systems utilize fish species classification models to provide real-time data on the species and quantities of fish being caught. This data supports accurate reporting and monitoring of fishing activities, contributing to better management of fish stocks and prevention of overfishing.

2.9.2 Environmental Monitoring and Conservation

Fish species classification plays a vital role in environmental monitoring and conservation efforts. By accurately identifying fish species, researchers and

conservationists can assess the health of aquatic ecosystems, track changes in biodiversity, and identify the impacts of environmental stressors.

2.9.2.1 Biodiversity Assessment

Deep learning-based fish species classification models are used to assess the biodiversity of marine and freshwater environments. These models analyse underwater images and videos to identify and catalogue species, providing valuable data for biodiversity studies. The LifeClef Fish dataset (18), for example, is used in biodiversity assessments to evaluate the effectiveness of classification models in challenging underwater conditions.

2.9.2.2 Habitat Monitoring

Classification models help monitor fish habitats by providing insights into species distribution and habitat preferences. This information is crucial for understanding the effects of environmental changes and human activities on aquatic ecosystems. Automated systems that classify fish species from underwater imagery enable continuous monitoring and data collection without disturbing the habitats.

2.9.3 Aquaculture

In aquaculture, fish species classification is used to improve the management and production of farmed fish. Accurate classification helps in monitoring the health of fish stocks, managing breeding programs, and preventing the spread of diseases.

2.9.3.1 Disease Detection

Early detection of diseases in farmed fish can prevent outbreaks and minimize losses. Classification models that analyse images of fish can identify signs of disease or abnormal growth, allowing for timely intervention and management. This application is particularly relevant in large-scale aquaculture operations where manual inspection is impractical.

2.9.3.2 Quality Control

Fish species classification models are also used in quality control processes to ensure that farmed fish meet the required standards for size, species, and quality. Automated systems provide consistent and accurate assessments, reducing the reliance on manual inspections and improving overall efficiency.

2.9.4 Educational and Research Purposes

Fish species classification technologies are valuable tools for educational and research institutions. They support studies in marine biology, ecology, and environmental science by providing accurate data and enabling large-scale analysis of fish populations and behaviours.

2.9.4.1 Marine Biology Research

Researchers use classification models to study fish behaviour, distribution patterns, and interactions with other species. The availability of large datasets and advanced classification techniques has significantly enhanced the scope and depth of marine biology research.

2.9.4.2 Citizen Science Projects

Citizen science projects often involve the collection and classification of fish images by volunteers. Classification models help process and analyse these images, contributing to large-scale data collection and public engagement in scientific research.

2.10 Future Directions and Emerging Trends

As fish species classification continues to advance, several emerging trends and future directions are shaping the field. These developments promise to enhance the accuracy, efficiency, and applicability of classification technologies, addressing existing challenges and opening new avenues for research and practical applications.

2.10.1 Integration of Multi-Modal Data

The integration of data from various sources such as underwater imagery, acoustic signals, and environmental sensors is expected to improve the robustness and accuracy of fish species classification systems. Multi-modal data fusion can provide a more comprehensive understanding of fish behaviour and habitat, enhancing classification performance in complex underwater environments.

2.10.1.1 Combining Vision and Acoustic Data

Recent research suggests that combining visual data with acoustic signals can improve species detection and classification. Acoustic data provides information on fish size, movement, and density, which, when combined with visual data, can enhance species identification and behaviour analysis. For instance, integrating acoustic monitoring with image-based classification systems has shown promise in addressing the limitations of each modality when used alone (8).

2.10.1.2 Use of Environmental Data

Incorporating environmental variables such as water temperature, salinity, and depth into classification models can improve accuracy and provide contextual information that aids in species identification. Environmental data can help disambiguate species with similar appearances but different habitat preferences (9).

2.10.2 Advancements in Deep Learning Architectures

Ongoing advancements in deep learning architectures are likely to further enhance fish species classification. Emerging models and techniques are expected to improve feature extraction, model generalization, and handling of large-scale datasets.

2.10.2.1 Transformers and Attention Mechanism

The application of transformer architectures and attention mechanisms in fish species classification is an emerging trend. Transformers, known for their success in natural language processing, are being adapted for visual tasks. They offer potential improvements in capturing long-range dependencies and fine-grained details in fish images, leading to more accurate classifications (10).

2.10.2.2 Generative Adversarial Networks (GANs)

Generative Adversarial Networks (GANs) are being explored for augmenting training data and enhancing model robustness. GANs can generate synthetic images of fish, which can be used to improve the performance of classification models, especially in cases where real data is scarce or imbalanced (11).

2.10.3 Incorporation of Real-Time Processing

Real-time fish species classification is becoming increasingly important, particularly for applications in fisheries management and environmental monitoring. Advances in edge computing and hardware accelerators are enabling real-time processing of underwater imagery, making it feasible to deploy classification systems on autonomous underwater vehicles and other field devices.

2.10.3.1 Edge AI and Autonomous Systems

Edge AI technologies allow for the deployment of AI models directly on hardware with limited computational resources. This enables real-time classification of fish species on autonomous systems such as drones and underwater robots, improving the efficiency of data collection and monitoring in remote locations (12).

2.10.3.2 Integration with Internet of Things (IoT)

Integrating fish species classification systems with IoT platforms can facilitate continuous monitoring and data sharing. IoT devices can collect and transmit data from various sensors, enabling centralized analysis and real-time decision-making in fisheries and conservation efforts (13).

2.10.4 Ethical and Societal Implications

As fish species classification technologies advance, it is crucial to consider their ethical and societal implications. Ensuring that these technologies are used responsibly and do not negatively impact marine ecosystems or communities is essential.

2.10.4.1 Impact on Marine Ecosystem

The deployment of classification systems should be carefully managed to minimize disturbance to marine habitats. Research on the environmental impact of monitoring technologies is important to ensure that they contribute positively to conservation efforts without causing harm (14).

2.10.4.2 Data Privacy and Security

As classification systems collect and analyse large volumes of data, ensuring data privacy and security is critical. Developing protocols for data protection and addressing concerns related to data ownership and usage will be important as these technologies become more widespread (15).

2.11 Conclusion

Fish species classification has undergone significant advancements from its traditional roots to the latest machine learning and deep learning approaches. Initial methods focused on morphological features and manual classification, which, while foundational, were limited by human subjectivity and labour intensity. The advent of machine learning techniques such as Support Vector Machines and Random Forests represented a notable improvement. However, it is the introduction of deep learning technologies, including Convolutional Neural Networks (CNNs) and Vision Transformers, that has dramatically enhanced classification accuracy and efficiency. These modern methods enable the processing of extensive underwater imagery with greater precision, addressing complex environmental challenges and providing more robust solutions.

Looking ahead, the integration of multi-modal data, such as combining visual, acoustic, and environmental information, promises to further refine classification systems, offering a more comprehensive understanding of aquatic environments. Despite these advancements, challenges related to underwater image quality, real-time processing, and ethical implications remain. Future research will likely focus on overcoming these hurdles, with a particular emphasis on improving multi-modal integration and deep learning architectures. The continued evolution of these technologies holds significant implications for both research and practical applications, particularly in fisheries management and marine conservation. As the field progresses, interdisciplinary collaboration and innovative approaches will be crucial in developing accurate, efficient, and ethically responsible fish species classification systems.

Chapter 3: Methodology

3.1 Introduction

This chapter details the methodology used for the fish species classification task, employing both traditional machine learning and deep learning techniques. The approach is guided by the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework, which provides a structured process for data mining projects. This framework ensures a systematic approach to each phase of the project, from understanding the problem to deploying the solution. The following sections describe the methodology in the context of CRISP-DM, including data exploration, preprocessing, model development, and evaluation.

3.2 Methodological Framework

To systematically address the fish species classification problem, the CRISP-DM framework was employed. CRISP-DM consists of six phases: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment. This structured approach ensures comprehensive coverage of the entire data mining lifecycle.
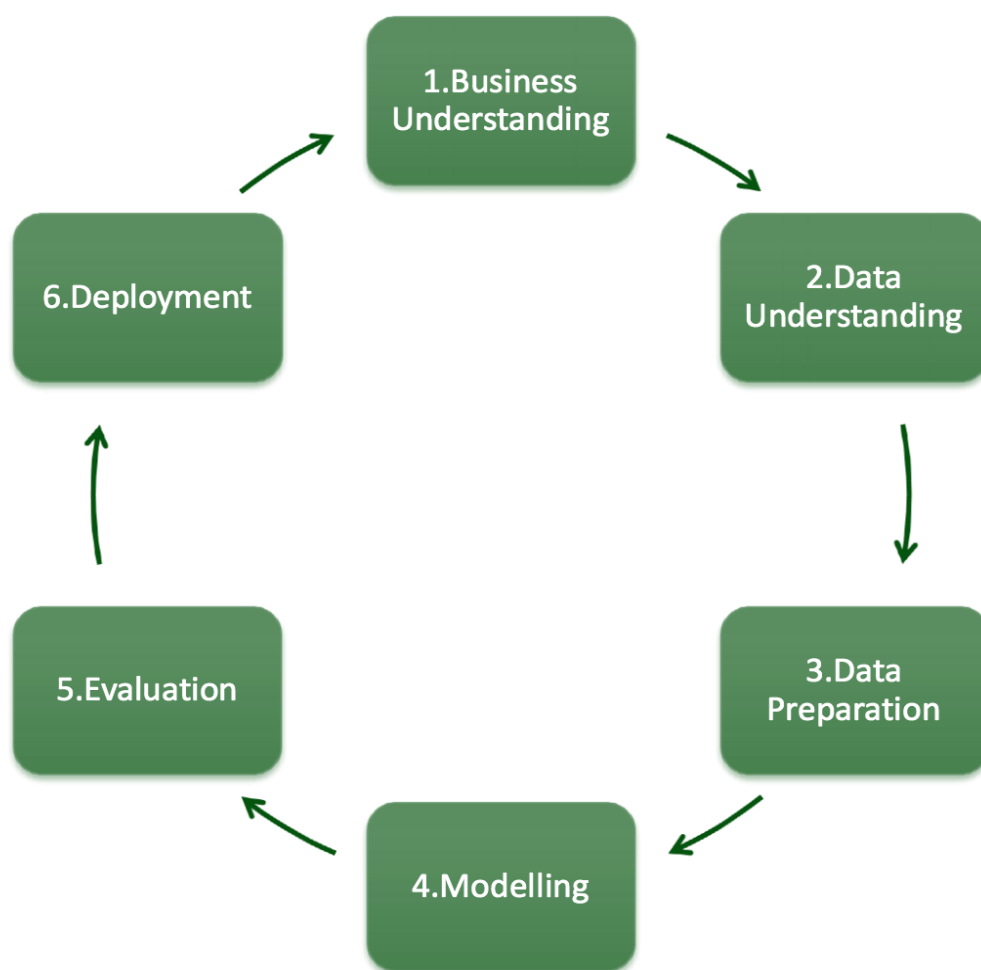

Figure 1, CRISP-DM framework

3.3 Business Understanding

The Business Understanding phase focused on defining the scope and objectives of the fish species classification task. This phase involved:

3.3.1 Defining the Problem:

The primary objective was to classify images of fish into their respective species. This task required developing a model capable of accurately identifying fish species from a diverse dataset of images.

3.3.2 Setting Objectives:

3.3.2.1 Performance Goals

Achieve high classification accuracy across various fish species.

3.3.2.2 Handling Variability

Develop models that perform well under different imaging conditions, including controlled environments, natural habitats, and varying illumination scenarios.

3.3.3 Business Implications:

The successful implementation of this classification model could aid in ecological research, conservation efforts, and biodiversity studies by automating species identification.

3.4 Data Understanding

In the Data Understanding phase, an extensive examination of the dataset was conducted to uncover insights and guide subsequent steps. This phase involved several key activities.

3.4.1 Dataset Overview

Fish Species dataset currently consisting of 3,960 images collected from 468 species. Data consists of real-world images of fish captured in 3 conditions defined as "controlled", "out-of-the-water" and "in-situ".

The "controlled", images consists of fish specimens, with their fins spread, taken against a constant background with controlled illumination.

The "in-situ" images are underwater images of fish in their natural habitat and so there is no control over background or illumination. The "out-of-the-water" images consist of fish specimens, taken out of the water with a varying background and limited control over the illumination conditions.(30)

3.4.2 Exploratory Data Analysis (EDA)

Before diving into the model building process, an extensive Exploratory Data Analysis (EDA) was conducted to better understand the dataset's characteristics. The insights gained from EDA helped shape the data preprocessing steps and model selection.

3.4.2.1 Image Statistics

Key statistics were calculated to gain an understanding of the dataset's overall characteristics:

- Number of Labels: 193 unique classes representing different fish species.
- Image Mean: The average pixel intensity values across all images for each colour channel (RGB).
  - [107.53, 123.19, 122.46]
- Image Standard Deviation: The spread of pixel intensity values for each colour channel.
  - [74.10, 72.92, 78.33]

3.4.2.2 Sample Image Display

A grid of sample images was displayed to visually inspect the diversity of the dataset. This helped in understanding the variation in colour, shape, and size among the fish species.
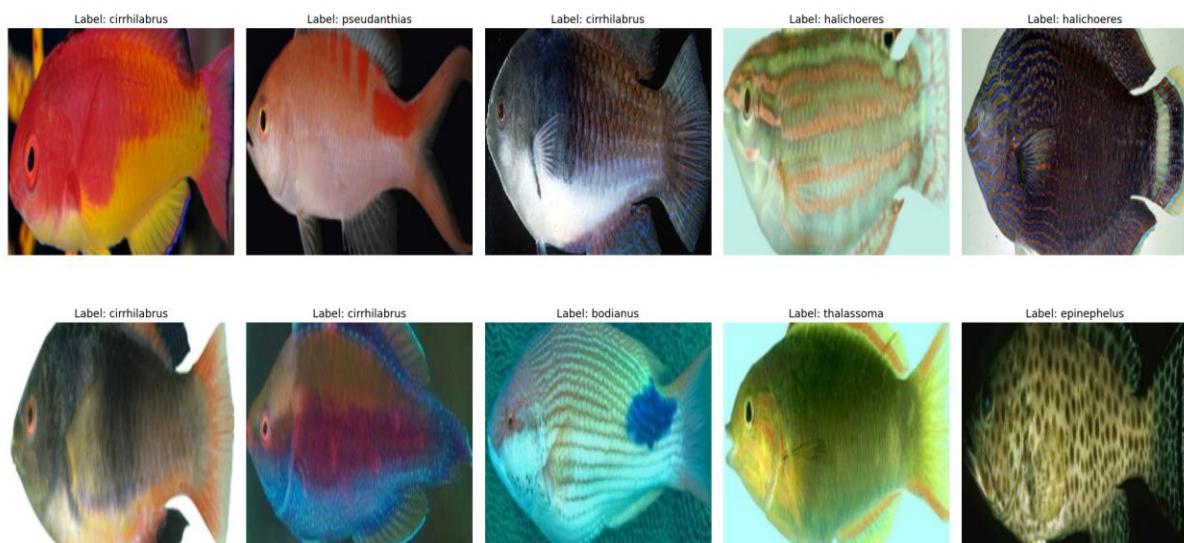

Figure 2, Sample Image Grid

3.4.2.3 Class Distribution Visualization

The distribution of classes was plotted to examine the balance of the dataset. This visualization highlighted any significant class imbalances, which could influence model performance.
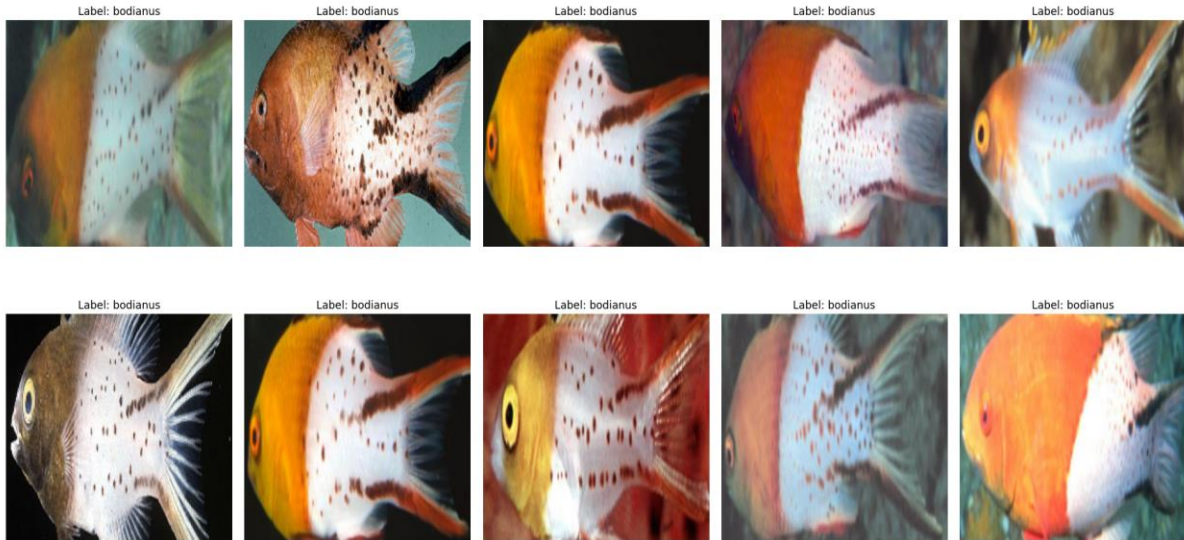
Figure 3, Class Distribution Plot

3.4.2.4 Label Count Distribution

A plot showing the count of images per label was generated to understand the distribution of data across different classes.
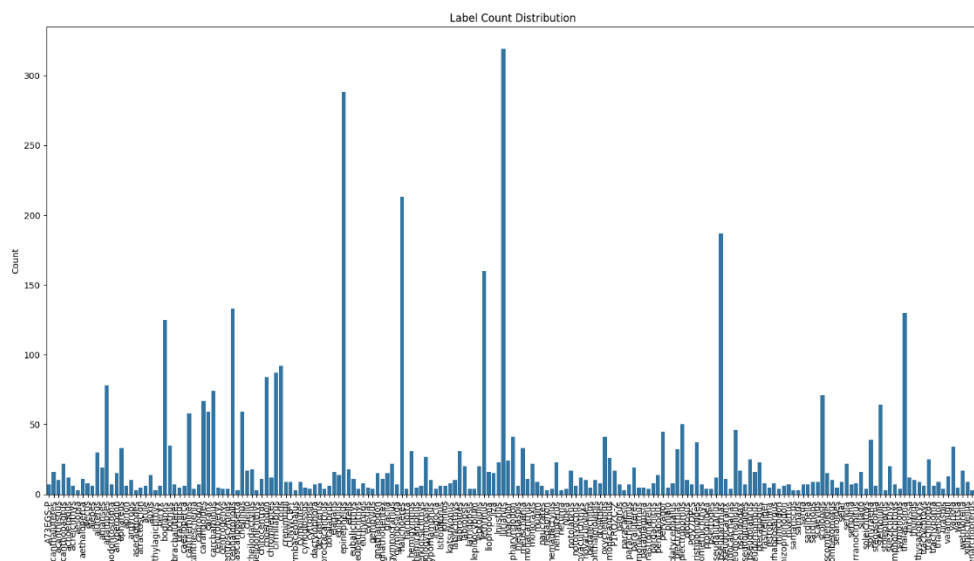

Figure 4, Label Count Distribution Plot

3.4.2.5 Image Aspect Ratios

The aspect ratios of images were visualized to assess the variety in image dimensions. This information was crucial for determining the appropriate preprocessing steps, such as resizing and cropping.
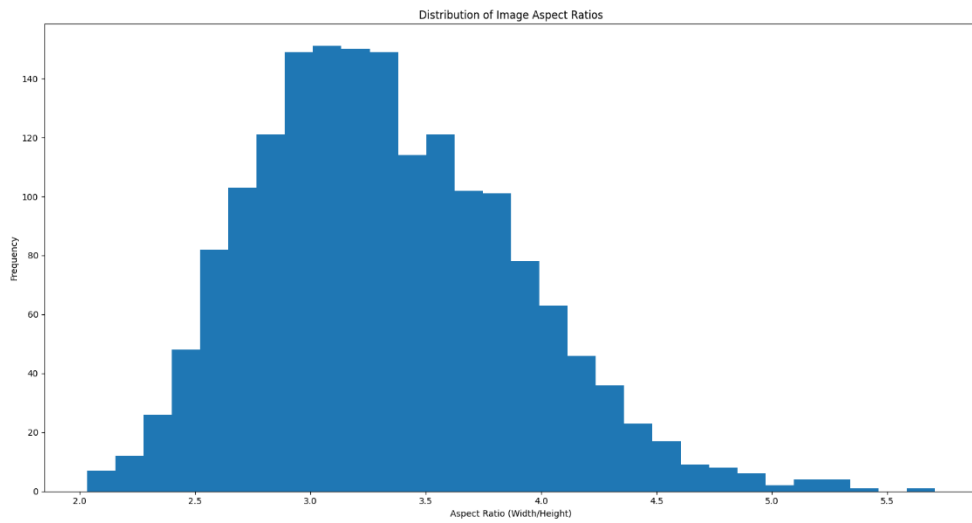
Distribution of Image Aspect Ratios

Figure 5, Image Aspect Ratio Visualization

3.4.2.6 Class-wise Mean Image

Mean images were created and visualized for each class, providing a visual summary of the average appearance of each fish species.
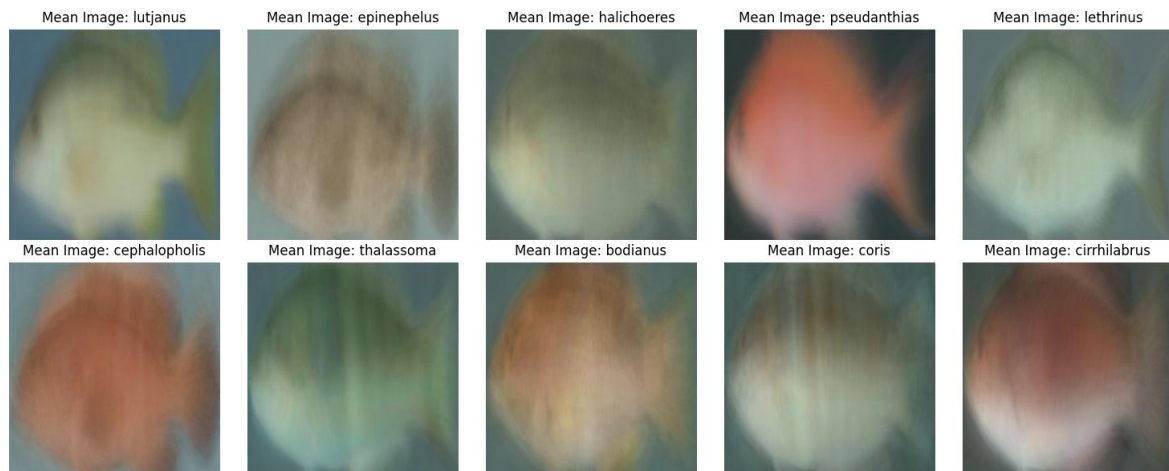

Figure 6, Class-wise Mean Image Visualization

3.4.2.7 Image Size Distribution

The distribution of image sizes (width and height) was plotted to understand the range of image dimensions in the dataset.
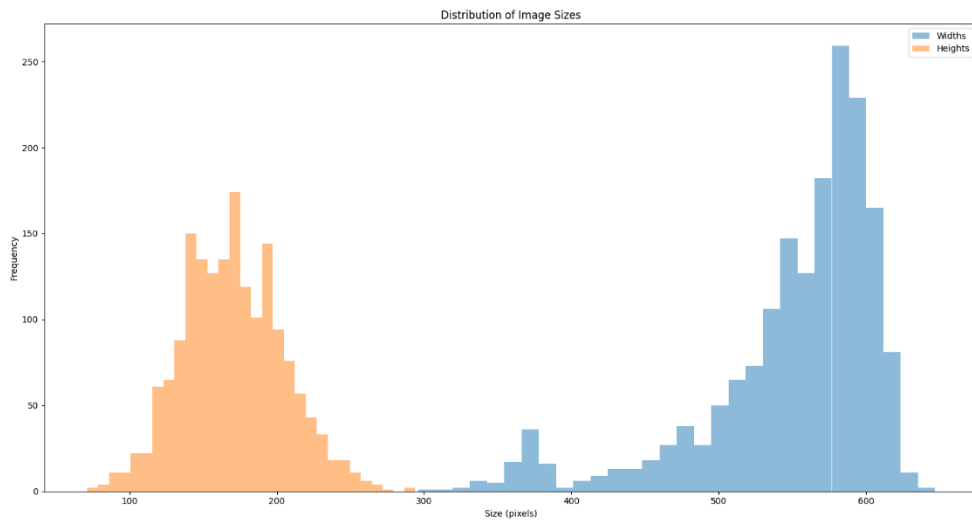
Distribution of Image Sizes

Figure 7, Image Size Distribution Plot

3.4.2.8 Class-wise Standard Deviation Image

Standard deviation images were visualized for each class to explore the intra-class variance, which reflects the diversity within a particular species.
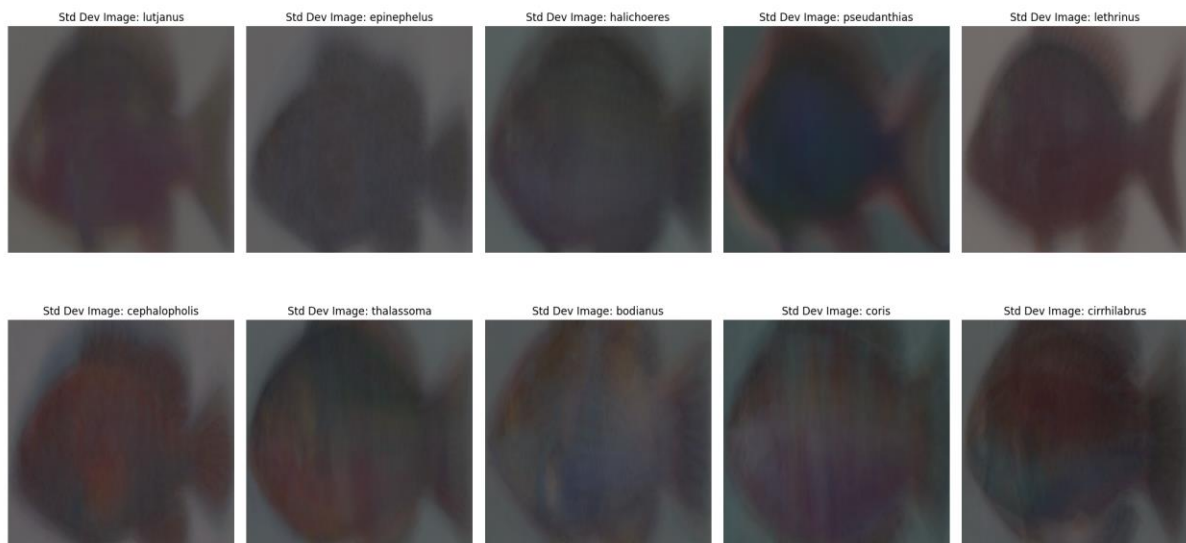

Figure 8, Class-wise Standard Deviation Image Visualization

3.4.2.9 Principal Component Analysis (PCA)

PCA was performed to reduce the dimensionality of the image data to two components, allowing for a 2D visualization of how images cluster based on their labels.
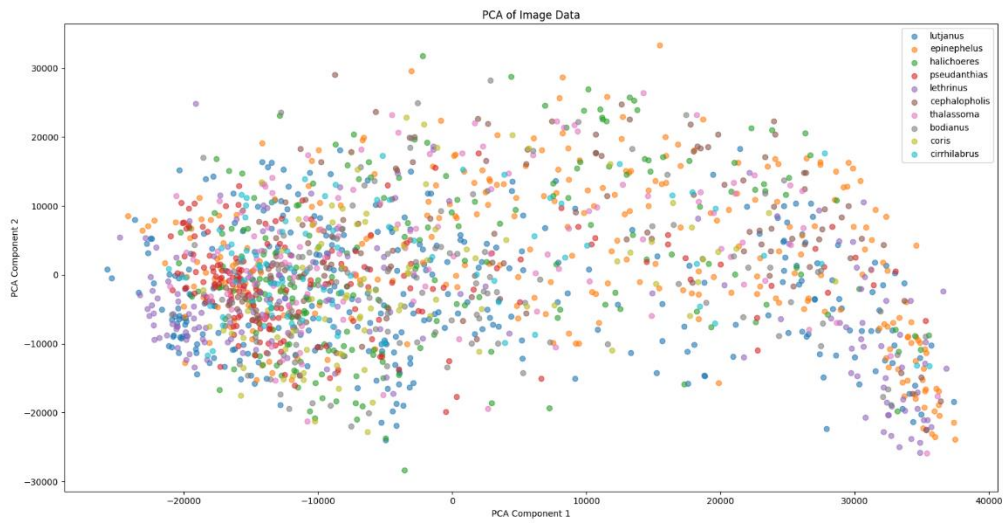
Figure 9, PCA Visualization

3.4.2.10 T-SNE Visualization

T-SNE, a more advanced dimensionality reduction technique, was also applied to provide a clearer separation of images in a 2D space, often revealing patterns not visible with PCA.
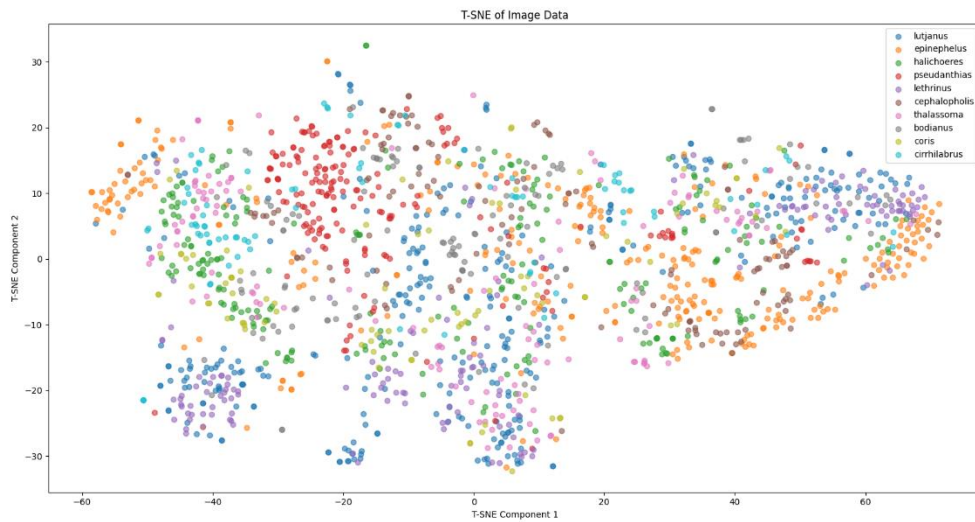

Figure 10, T-SNE Visualisation

3.4.2.11 Correlation Matrix

A correlation matrix of pixel values was generated to explore the relationships between different pixels, which could highlight potential redundancies or dependencies in the data.

Placeholder for Figure: Correlation Matrix

3.4.2.12 Class-wise Pixel Intensity Distribution

The distribution of pixel intensities was plotted for each class, giving insight into the visual characteristics of different fish species.
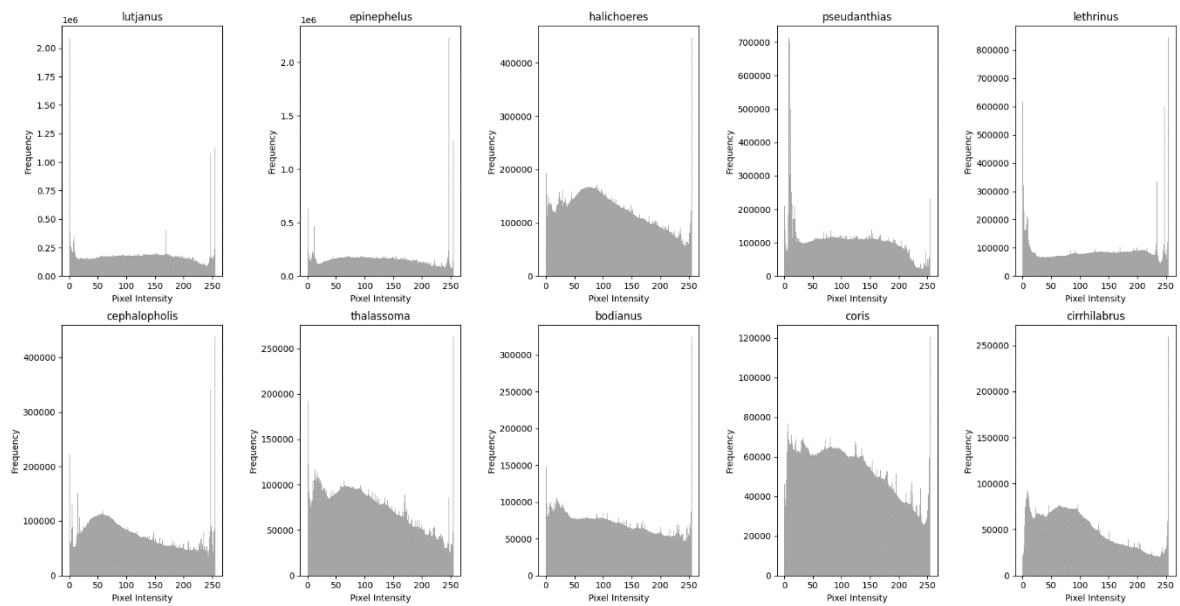

Figure 11, Class-wise Pixel Intensity Distribution Plot

3.4.2.13 Edge Detection Visualization

Edge detection techniques were applied to the images to highlight the structural details within the fish images, which could be useful for understanding the texture and contours important for classification.

Figure 12, Edge Detection Visualization

3.4.2.14 Image Brightness and Contrast Distribution

The brightness and contrast of images were analysed to assess the overall lighting conditions and variation within the dataset.
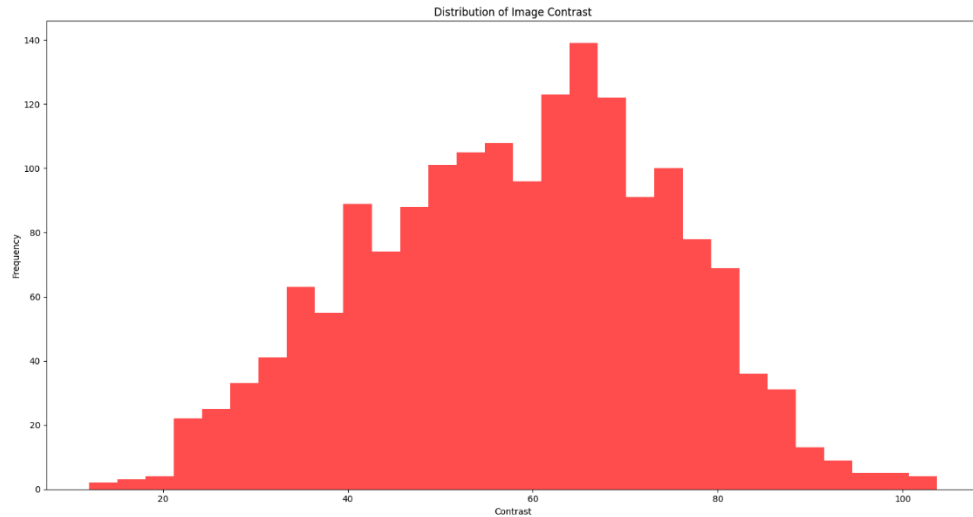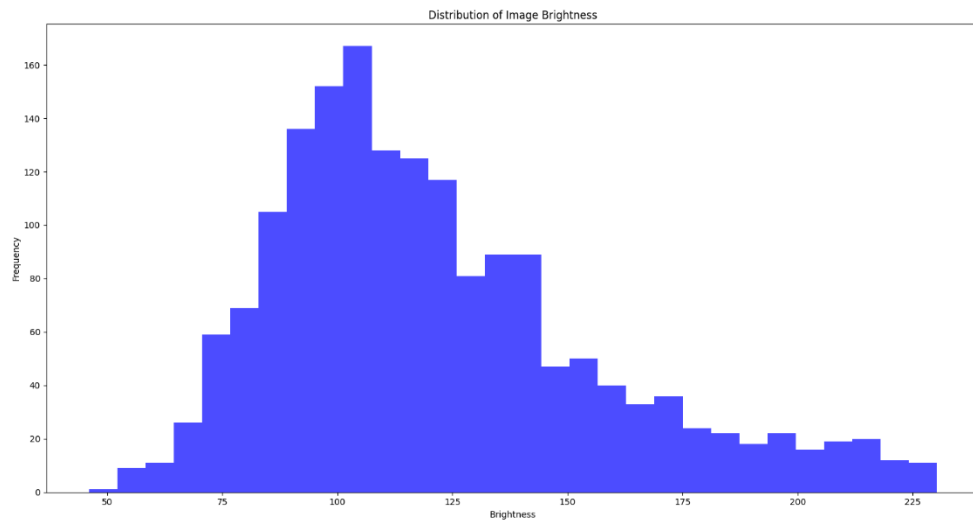

Figure 13, Contrast Distribution Plot

Figure 14, Brightness Distribution Plot

3.5 Data Preparation

Data Preparation involved several steps to ensure the dataset was ready for modelling:

3.5.1 Resizing

Images were resized to uniform dimensions to match the input requirements of different models. Standard sizes included 224x224 pixels for most models and 299x299 pixels for Xception and InceptionV3.

3.5.2 Normalization:

Pixel values were normalized to the [0, 1] range to standardize input data and enhance model training efficiency.

3.5.3 Label Encoding:

Labels were converted into a numerical format using one-hot encoding to facilitate multi-class classification.

3.5.4 Train-Test Split:

The dataset was divided into training (80%), validation, and testing (20%) sets. A portion of the training data was used as a validation set to monitor model performance and prevent overfitting.

3.6 Modelling and Model Architectures

3.6.1 Convolutional Neural Network (CNN)

- Architecture Overview: A custom CNN model was created with three convolutional layers, each followed by max pooling layers, and then a fully connected layer before the output layer. Dropout was used to prevent overfitting.

3.6.2 Transfer Learning Models

- MobileNetV2: Lightweight architecture with depth-wise separable convolutions, balancing performance and efficiency.
- VGG16: Deep and simple architecture, known for its effectiveness in image classification tasks.
- DenseNet121: Features dense connections between layers, promoting feature reuse and efficient gradient flow.
- Xception: Based on depth-wise separable convolutions, offering fine-grained feature extraction.
- InceptionV3: Employs inception modules to capture various features at different scales.
- EfficientNetB0: Utilizes compound scaling to optimize network depth, width, and resolution.
- ResNet50: Incorporates residual connections to enable the training of deeper networks without performance degradation.

3.6.3 Traditional Machine Learning Models

Several traditional machine learning models were implemented, including:

- Support Vector Machine (SVM): A robust classifier that finds the optimal hyperplane for class separation.
- K-Nearest Neighbours (KNN): A simple, instance-based learning algorithm that classifies based on the majority vote of neighbours.
- Random Forest: An ensemble method using multiple decision trees to improve classification accuracy.
- Decision Tree: A model that splits data based on feature values but is prone to overfitting.
- Naive Bayes: A probabilistic classifier assuming feature independence, which is often a limitation in image data.
- Logistic Regression: A linear model extended for multiclass classification using the softmax function.

3.6.4 Feature Extraction

For traditional models, features were extracted using methods such as Histogram of Oriented Gradients (HOG) and colour histograms, which were then used for classification.

3.7 Evaluation Metrics

3.7.1 Accuracy

Model accuracy, the ratio of correct predictions to total predictions, was the primary metric used to evaluate performance.

3.7.2 Confusion Matrix

Confusion matrices were generated to provide a detailed breakdown of model performance across different classes, highlighting areas where the model performed well and where it struggled.

Placeholder for Figure: Confusion Matrix Plot

3.7.3 Loss Functions

Categorical cross-entropy loss was used, appropriate for multi-class classification problems. The loss was monitored during training to gauge model convergence.

Placeholder for Figures: Training & Validation Accuracy and Loss Plots

3.8 Deployment and Training Strategy

Although the primary focus of this study was on model development and evaluation, there are promising deployment scenarios for the developed classification models. These models could be implemented in field research and conservation efforts to automate fish species identification from images, enhancing efficiency in data collection and analysis. Additionally, integrating the models into research tools or applications could support the analysis of fish populations and biodiversity, providing valuable insights for ecological studies and conservation strategies.

3.8.1 Hyperparameter Tuning

- Learning Rate: Adam optimizer with a learning rate of 1e-4 was used. Learning rate reduction was applied when the validation loss plateaued.
- Batch Size: A batch size of 32 was selected for efficient training.
- Epochs: Models were trained for up to 50 epochs, with early stopping to avoid overfitting.

3.8.2 Callbacks

- ReduceLROnPlateau: This callback reduced the learning rate when no improvement in validation loss was observed.
- EarlyStopping: Training was halted early if validation performance stopped improving, and the best weights were restored.

3.8.3 Data Augmentation During Training

Real-time data augmentation was applied using the ImageDataGenerator to increase the effective size of the training dataset and enhance model generalization.

3.9 Summary

This chapter provided a comprehensive overview of the methodology used for the fish species classification task, structured around the CRISP-DM framework. It covered the business understanding, data understanding, data preparation, modelling, and evaluation phases in detail. By employing a systematic approach and various modelling techniques, this study aimed to achieve accurate and robust classification results. The next chapter will present and analyse the results obtained from these models, offering insights into their performance and implications for future work.

Chapter 4: Experimental Results and Analysis

4.1 Introduction

In this chapter, we assess the performance of several deep learning models for fish species classification, comparing them with traditional machine learning algorithms. Each model's architecture, training process, and performance metrics are discussed, with a focus on accuracy and confusion matrices to evaluate classification performance.

4.2 Deep Learning Models and Their Performance

4.2.1 Custom CNN Model

- **Architecture**: The custom CNN model was designed with three convolutional layers, each followed by max pooling layers to progressively reduce spatial dimensions while capturing features. A fully connected dense layer was used before the output layer, and dropout was applied to reduce overfitting. This simple architecture achieved a moderate accuracy but was outperformed by more complex models.
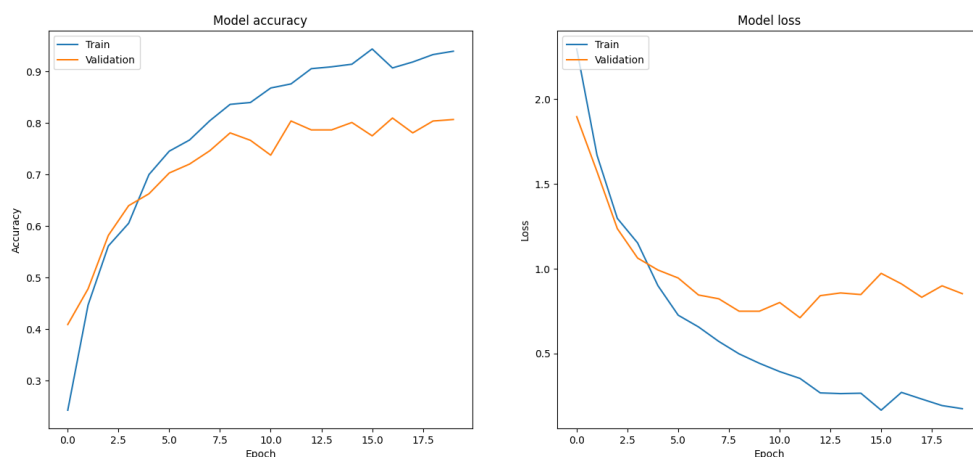- **Accuracy**: 80.00%
- Performance Plots:

Figure 15, Training & Validation Accuracy and Loss Plot

4.2.2 MobileNetV2

- **Architecture**: MobileNetV2, a lightweight deep learning model optimized for mobile and edge devices, was employed with its layers frozen except for a few fully connected layers at the end. This model uses depthwise separable convolutions, which reduce the number of parameters while maintaining performance.
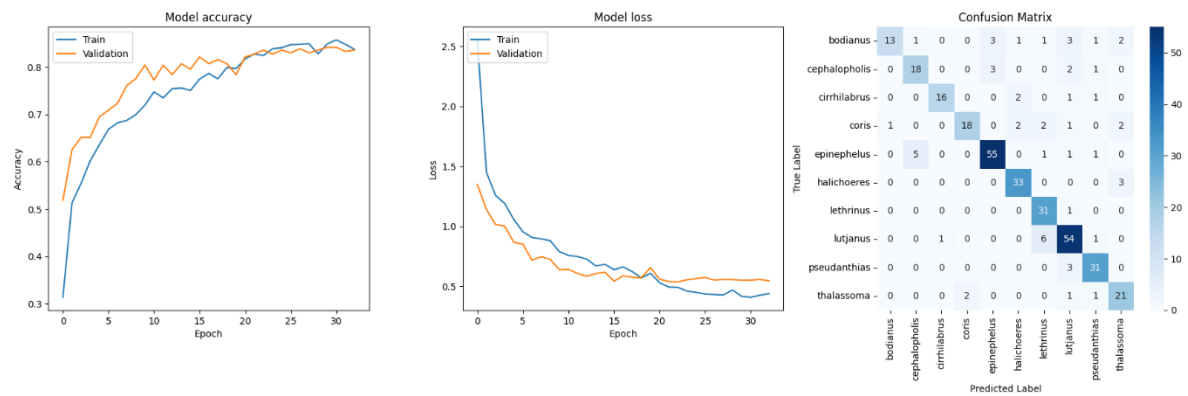- **Accuracy**: 83.57%
- Performance Plots:

Figure 16, Confusion Matrix and Training & Validation Accuracy and Loss Plot

4.2.3 VGG16

- **Architecture**: The VGG16 model, known for its depth and simplicity, was used for transfer learning. The convolutional base was frozen, and dense layers were added for classification. This model, though computationally expensive, performed well due to its deep architecture.
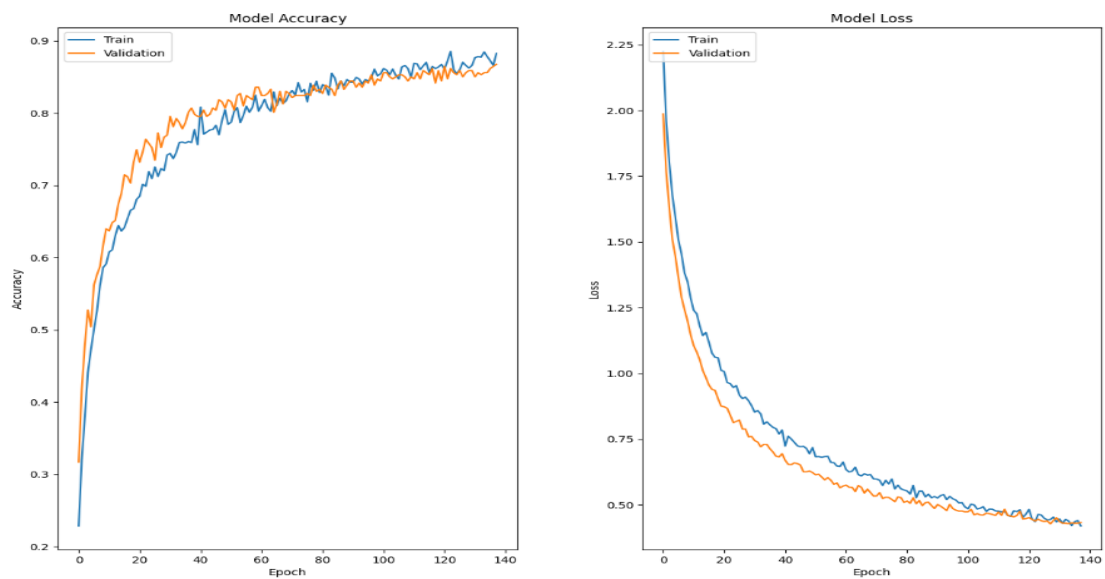- **Accuracy**: 87.61%
- Performance Plots:
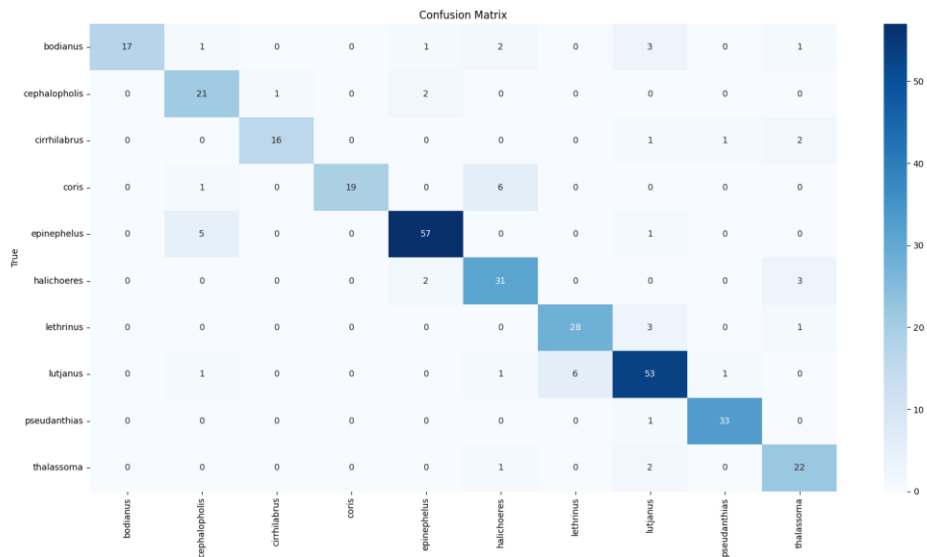

Figure 17, Training & Validation Accuracy and Loss Plot

Figure 18, Confusion Matrix

4.2.4 DenseNet121

- **Architecture**: DenseNet121 was implemented, leveraging its dense connectivity, where each layer receives input from all previous layers. This design facilitates feature reuse, resulting in enhanced learning and better accuracy. DenseNet121 achieved the highest accuracy among the tested models.
- **Accuracy**: 90.20%
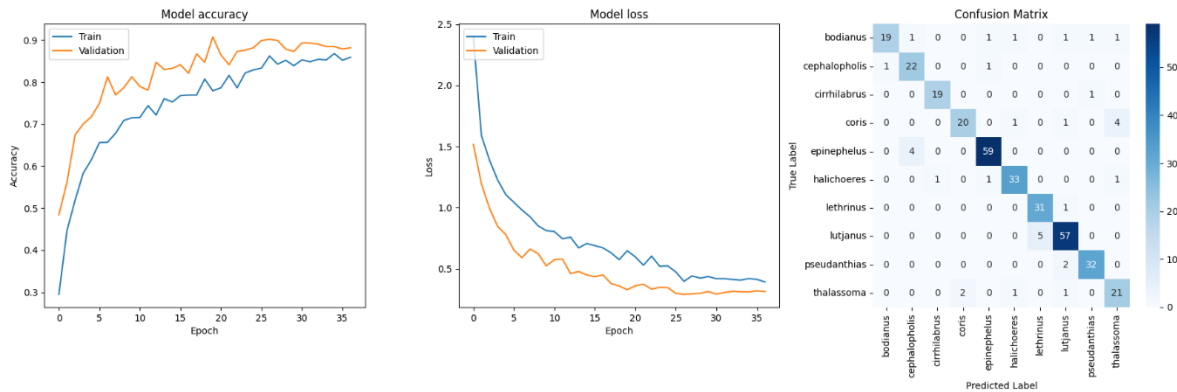- Performance Plots:

Figure 19, Confusion Matrix and Training & Validation Accuracy and Loss Plot

4.2.5 Xception

- **Architecture**: Xception, an architecture based on depthwise separable convolutions, was tested for its ability to handle complex patterns in images. Although it is more computationally intensive, it performed very well in classifying the fish species.
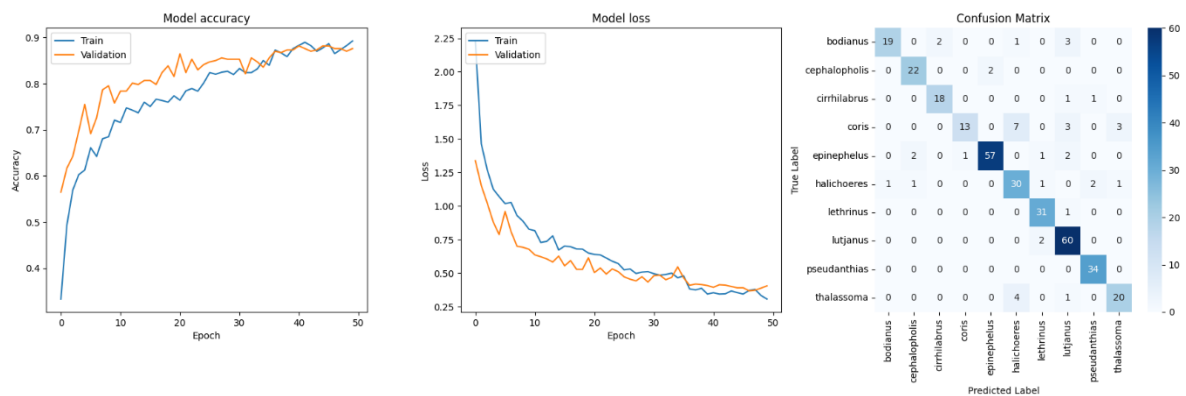- **Accuracy**: 87.61%

- Performance Plots:

Figure 20, Confusion Matrix and Training & Validation Accuracy and Loss Plots

4.2.6 InceptionV3

- **Architecture**: InceptionV3, with its sophisticated inception modules, was utilized for this classification task. The model is designed to use multiple convolutional kernels of different sizes in parallel, capturing varying levels of detail. It performed comparably to MobileNetV2.
- **Accuracy**: 83.72%
- Performance Plots:
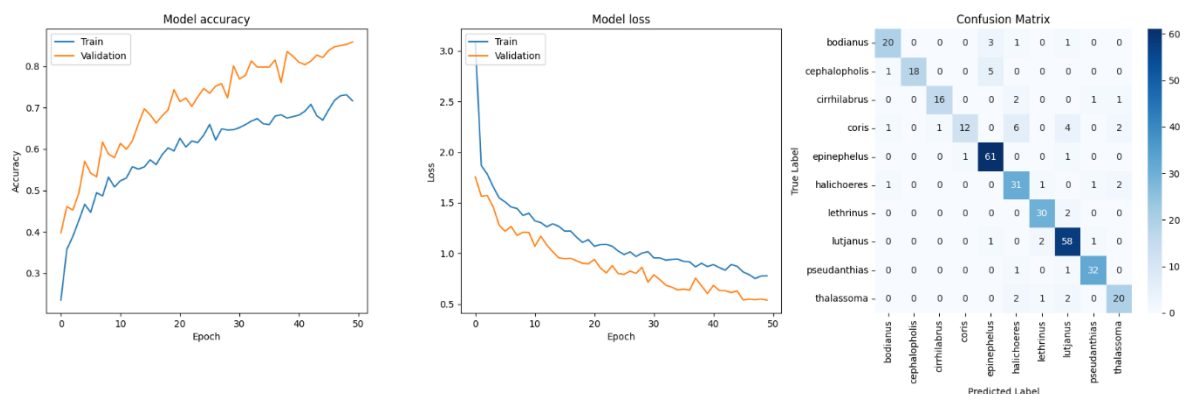
Figure 21, Confusion Matrix and Training & Validation Accuracy and Loss Plot

4.2.7 EfficientNetB0

- **Architecture**: EfficientNetB0, part of the EfficientNet family, was designed with a compound scaling method that balances network depth, width, and resolution. Despite its efficiency, it did not perform as well on this dataset, likely due to its relatively small size and simpler architecture.
- **Accuracy**: 75.50%
- Performance Plots:
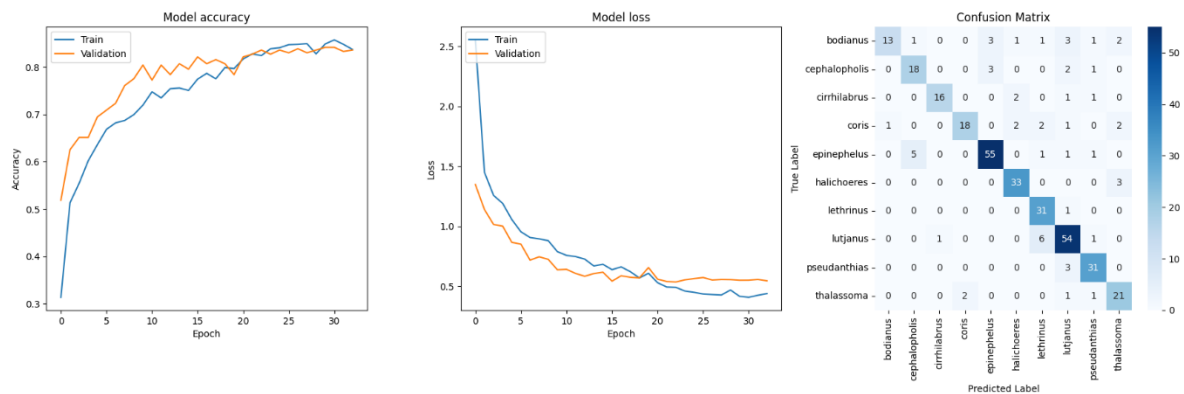
Figure 22, Confusion Matrix and Training & Validation Accuracy and Loss Plot

4.2.8 ResNet50

- **Architecture**: ResNet50 employs residual connections, allowing layers to learn residual functions with reference to the layer inputs, which helps in training deeper networks. While it performed well, it did not surpass DenseNet121 or VGG16.
- **Accuracy**: 82.54%
- Performance Plots:
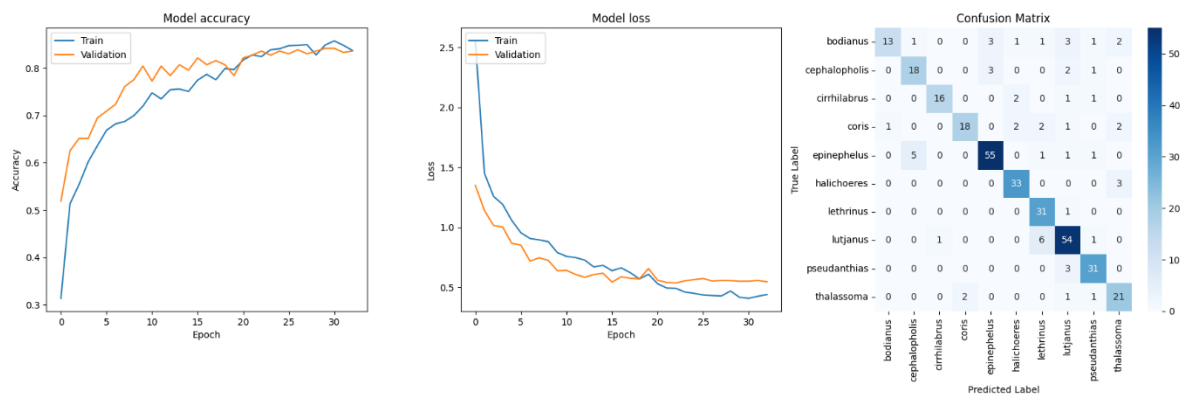
Figure 23, Confusion Matrix and Training & Validation Accuracy and Loss Plot

4.3 Traditional Machine Learning Models and Their Performance

4.3.1 Support Vector Machine (SVM)

- **Description**: SVM is a supervised learning algorithm that constructs a hyperplane or set of hyperplanes in a high-dimensional space to separate different classes. Although effective in certain cases, it was less capable of handling the complexity of image data compared to deep learning models.
- **Accuracy**: 73.20%
- Performance Plots:
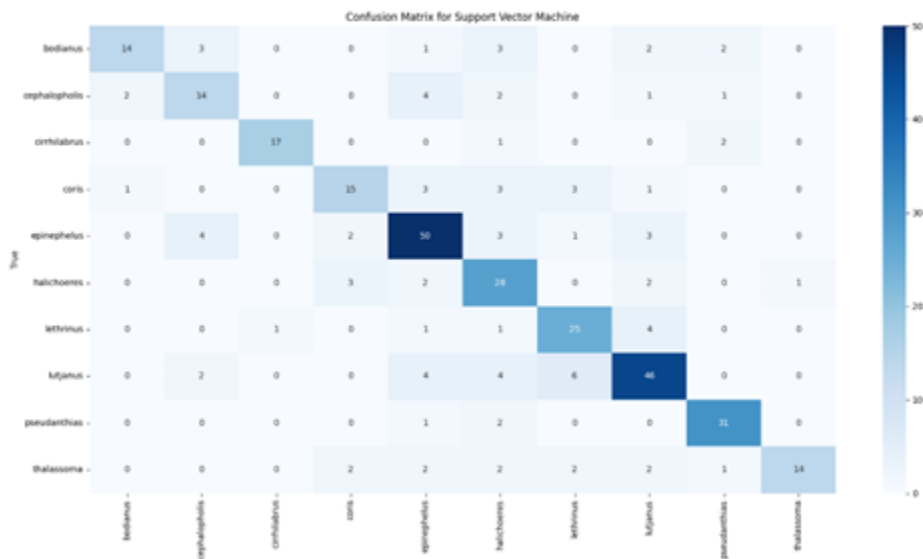
Figure 24, Confusion Matrix for Support Vector Machine

4.3.2 K-Nearest Neighbours (KNN)

- **Description**: KNN is an instance-based learning algorithm where classification is based on the majority vote of the nearest neighbours. Despite its simplicity, it struggled with the high-dimensionality of image data, leading to poor performance.
- **Accuracy**: 56.48%
- Performance Plots:
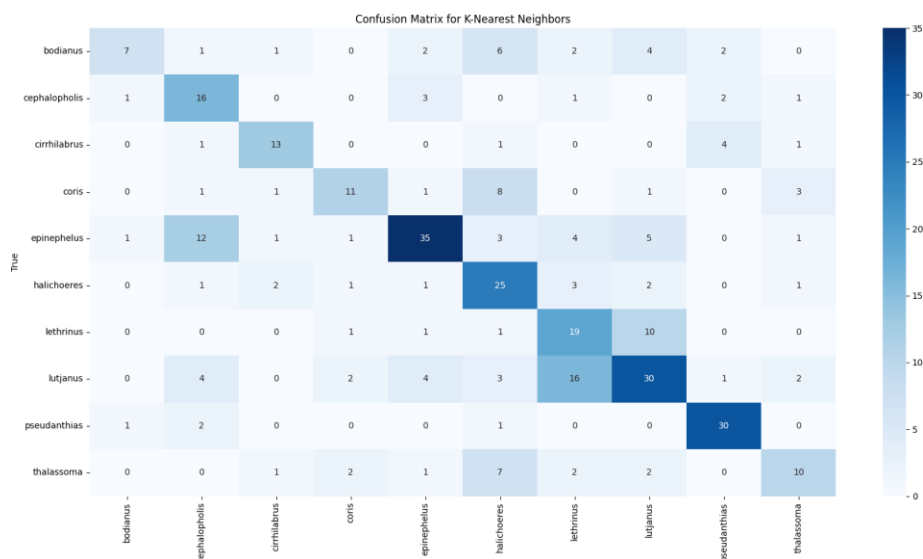
Figure 25, Confusion Matrix for K-Nearest Neighbours

4.3.3 Random Forest

- **Description**: Random Forest is an ensemble method that creates multiple decision trees and merges them to obtain a more accurate and stable prediction. However, it was less effective in this case, possibly due to the complexity of the image features.
- **Accuracy**: 61.67%
- Performance Plots:
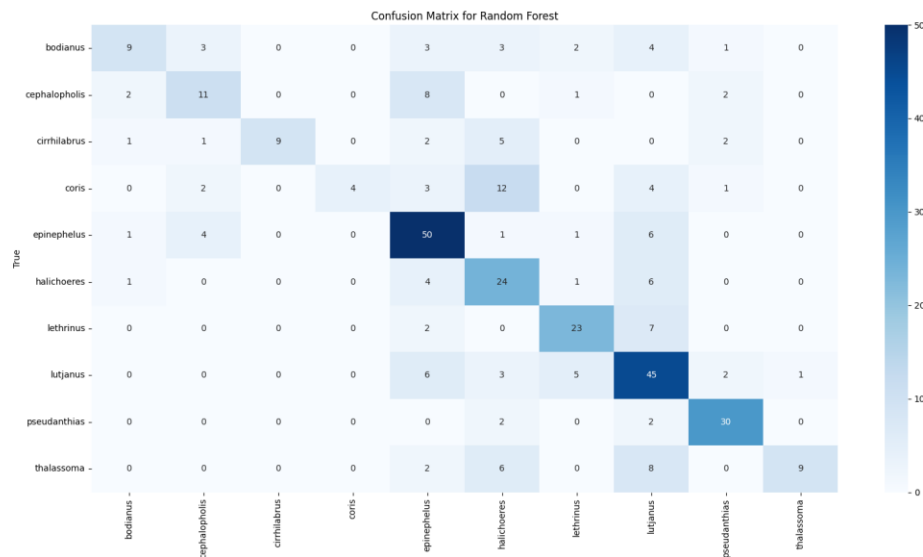
Figure 26, Confusion Matrix for Random Forest

4.3.4 Decision Tree

- **Description**: A Decision Tree model splits the data based on feature values into branches to reach a decision. It tends to overfit, especially on complex datasets like this one, resulting in poor accuracy.
- **Accuracy**: 34.01%
- Performance Plots:
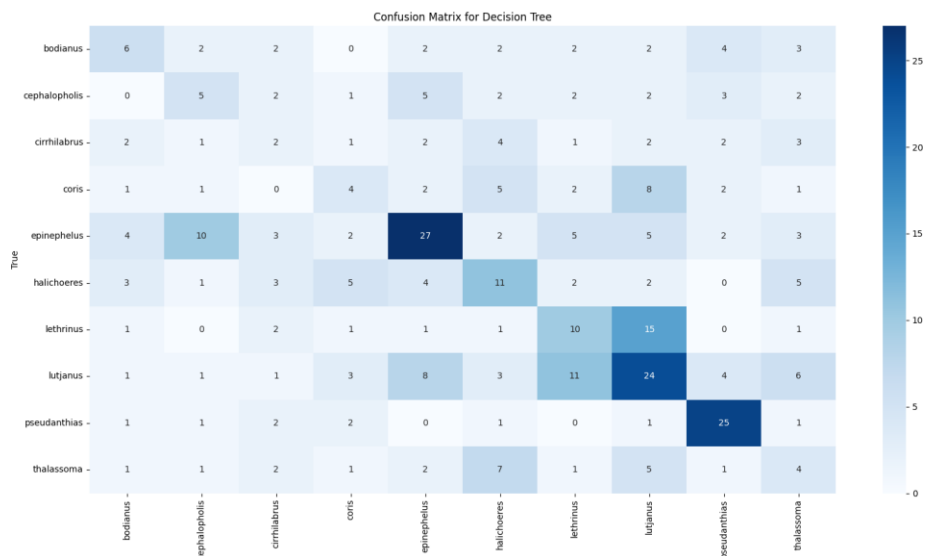
Figure 27, Confusion Matrix for Decision Tree

4.3.5 Naive Bayes

- **Description**: Naive Bayes is a probabilistic model that applies Bayes' theorem with the assumption of independence between features. This assumption does not hold well for image data, leading to suboptimal performance.
- **Accuracy**: 33.72%
- Performance Plots:
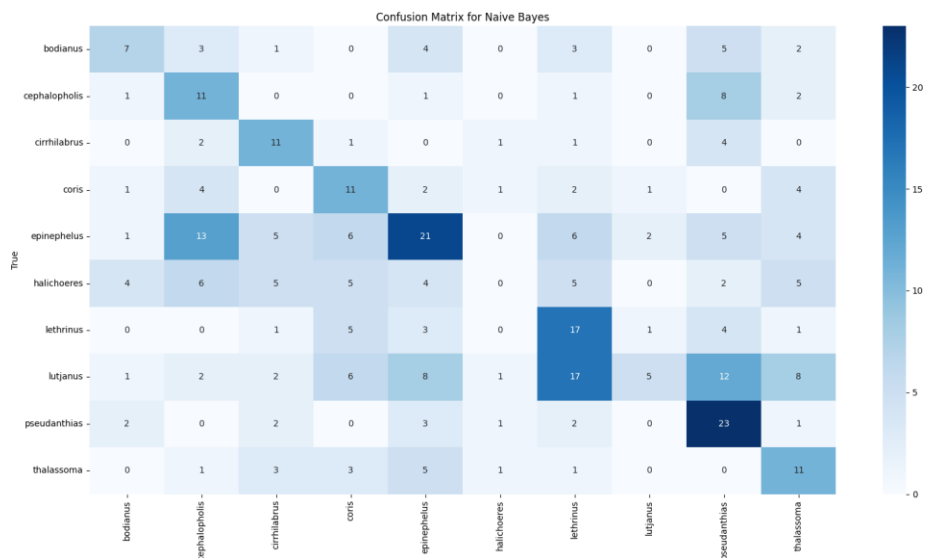
Figure 28, Confusion Matrix for Naive Bayes

4.3.6 Logistic Regression

- **Description**: Logistic Regression is a linear model for binary classification that was adapted for multiclass classification in this task. Despite its simplicity, it

performed better than many traditional models but still fell short compared to deep learning models.

- **Accuracy**: 72.05%
- Performance Plots:
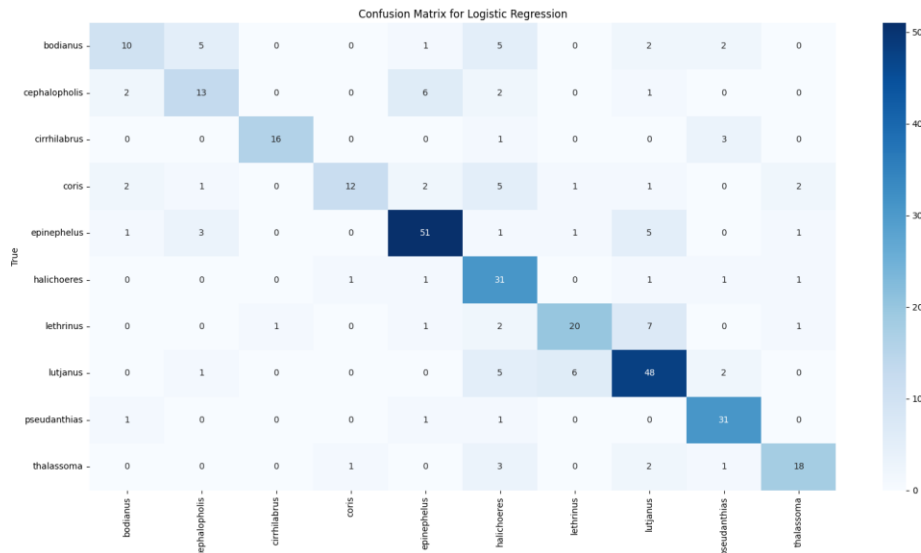
Figure 29, Confusion Matrix for Logistic Regression

4.4 Comparative Analysis

The following table summarizes the performance of the different models tested, including both deep learning and traditional machine learning approaches.

Table 2, performance of the different models and Comparison with Previous Studies

Model	Type	Accuracy	Key Characteristics	Comparison with Previous Studies
Custom CNN	Deep Learning	80.00%	Simple architecture with three convolutional layers and dropout.	Below typical accuracies (e.g., 98.88% by Kaya et al. with ANN) but simpler architecture.
MobileNetV2	Deep Learning	83.57%	Lightweight with depth-wise separable convolutions; optimized for mobile devices.	Lower than CNN-based approaches like Taheri-Garavand et al. (98.21%) but efficient for mobile applications.
VGG16	Deep Learning	87.61%	Deep architecture; transfer learning; computationally expensive.	Comparable to Rekha et al. (92%) and dos Santos and Gonçalves (96%) using CNNs.
DenseNet121	Deep Learning	90.20%	Dense connectivity facilitating feature reuse and high accuracy.	Similar performance to models like Alsmadi et al. (90%) with MA-B Classifier

				and dos Santos and Gonçalves (96%).
Xception	Deep Learning	87.61%	Depth-wise separable convolutions; handles complex patterns well.	Slightly lower than advanced deep CNNs like Taheri-Garavand et al. (98.21%), but competitive with other CNNs.
InceptionV3	Deep Learning	83.72%	Uses inception modules with multiple convolutional kernels.	Performance falls behind high-accuracy models like Kutlu et al. (99%) using Nearest Neighbour.
EfficientNetB0	Deep Learning	75.50%	Compound scaling method; less effective on this dataset.	Below most recent CNN models, e.g., dos Santos and Gonçalves (96%), indicating limitations on this dataset.
ResNet50	Deep Learning	82.54%	Residual connections aiding in deeper network training.	Performance is mid-range compared to top CNNs, e.g., Kratzert and Mader (93.3%).
SVM	Traditional ML	73.20%	Constructs hyperplanes for separation; less effective on high-dimensional image data.	Lower than many SVM-based methods, e.g., Fouad et al. (94.44%) and Hossain et al. (91.7%).
KNN	Traditional ML	56.48%	Instance-based; struggles with high-dimensionality of image data.	Underperforms significantly compared to Freitas et al. (92.3%) using SVM and KNN.
Random Forest	Traditional ML	61.67%	Ensemble of decision trees; less effective on image features.	Much lower than Saitoh et al. (87.3%) using RF, showing the limitations of RF on this dataset.
Decision Tree	Traditional ML	34.01%	Splits data into branches; tends to overfit on complex datasets.	Poor performance compared to most models, emphasizing its unsuitability for complex image data.
Naive Bayes	Traditional ML	33.72%	Probabilistic with independence assumption; suboptimal for image data.	Underperforms compared to Işçimen et al. (93.10%) using Naive Bayesian classifier.
Logistic Regression	Traditional ML	72.05%	Linear model adapted for multiclass classification.	Slightly lower than typical SVM-based methods like Islam et al. (90%).

DenseNet121 and VGG16 demonstrate competitive performance, with results closely matching those of other CNN-based methods, such as those reported by dos Santos and Gonçalves and Alsmadi et al. However, some models from other studies, like the one by Kutlu et al. using the Nearest Neighbour algorithm, have achieved even higher accuracy, indicating that while our models are strong, there is potential for further improvement.

In contrast, traditional machine learning models such as KNN, Decision Trees, and Naive Bayes lag significantly behind modern deep learning approaches, underscoring the superiority of deep learning in handling high-dimensional image data. This underperformance is particularly evident when compared to SVM-based methods from studies like those by Fouad et al. and Islam et al. While deep learning models generally align well with high-performance models from other studies, some, like EfficientNetB0, show limitations when applied to this specific dataset, highlighting the need for careful model selection based on the task at hand.

4.6 Summary

The results underscore the superiority of deep learning models, particularly **DenseNet121**, in managing the complexities of fish species classification, thanks to their advanced feature extraction and classification capabilities. While traditional machine learning models, such as KNN, Decision Trees, and Naive Bayes, are effective for simpler tasks, they fall significantly short compared to the robust performance of modern deep learning architectures. Despite some limitations observed with models like **EfficientNetB0** on specific datasets, deep learning approaches consistently demonstrate a clear advantage, highlighting their dominance in handling high-dimensional image data and setting a high standard for future research and applications.

Chapter 5: Conclusion and Future Work

5.1 Summary of Findings

This research focused on the development and evaluation of various deep learning models for the classification of fish species from images. Given the critical importance of accurate fish species identification in ecological conservation, fisheries management, and biodiversity research, the study aimed to explore whether deep learning models, particularly Convolutional Neural Networks (CNNs), could effectively address the challenges posed by traditional manual identification methods.

The key findings of the study are as follows:

1. **Deep Learning Model Performance**: Among the deep learning models tested, DenseNet121 emerged as the most accurate, achieving a classification accuracy of 90.20%. Other models, such as VGG16 and Xception, also performed well, with accuracies of 87.61%. These results demonstrate the capability of CNNs to learn complex patterns in image data and classify fish species with high accuracy.

2. **Transfer Learning**: The application of transfer learning significantly improved the performance of the CNN models. By leveraging pre-trained models such as VGG16, DenseNet121, and Xception, the study effectively utilized the feature extraction capabilities of these models, leading to improved accuracy, even with a limited dataset.

3. **Comparison with Traditional Machine Learning Models**: Traditional machine learning models, including Support Vector Machine (SVM), K-Nearest Neighbours (KNN), and Random Forest, were outperformed by deep learning models in this study. The best-performing traditional model, SVM, achieved an accuracy of 73.20%, highlighting the superiority of deep learning techniques in handling the complex and high-dimensional nature of image data.

4. **Exploratory Data Analysis (EDA)**: The EDA provided valuable insights into the dataset's characteristics, such as class distribution, image size distribution, and aspect ratios. These analyses informed the preprocessing and data augmentation strategies, which were crucial for improving the models' performance.

5. **Impact of Data Augmentation**: Data augmentation techniques, including rotation, width and height shifts, horizontal flips, and zoom, played a significant role in enhancing the robustness of the models. These techniques helped mitigate the risk of overfitting and improved the models' ability to generalize to new, unseen images.

5.2 Contributions to the Field

This research makes several important contributions to the field of automated fish species identification and image classification:

1. **Development of High-Performing Models**: The study successfully developed deep learning models that significantly outperform traditional machine learning

methods in fish species classification, demonstrating the potential of CNNs in ecological and environmental applications.

2. **Integration of Transfer Learning**: By integrating transfer learning into the model development process, this research highlights the effectiveness of using pre-trained models for tasks with limited labelled data, providing a practical approach for similar image classification challenges.

3. **Comprehensive EDA Framework**: The study presents a comprehensive framework for exploratory data analysis in image datasets, offering a methodology that can be adapted and applied to other domains where understanding the data's characteristics is critical to model success.

4. **Benchmarking Against Traditional Methods**: The comparative analysis with traditional machine learning algorithms offers a benchmark for future research, clearly demonstrating the advantages and limitations of both approaches in the context of fish species identification.

5.3 Limitations of the Study

Despite the success of the models developed in this study, several limitations must be acknowledged:

1. **Dataset Size and Diversity**: The dataset used, while sufficient for demonstrating the effectiveness of deep learning models, is relatively limited in size and diversity. A more extensive and diverse dataset could potentially lead to even higher accuracy and more robust models.

2. **Model Complexity and Computational Resources**: The deep learning models, particularly those involving transfer learning, require substantial computational resources for training. This limitation could pose challenges for deploying these models in resource-constrained environments or in real-time applications.

3. **Generalization to Other Datasets**: While the models performed well on the fish species dataset used in this study, their generalization to other datasets or different types of image classification tasks remains to be tested.

4. **Limited Exploration of Hyperparameter Tuning**: Although the study achieved high accuracy, the exploration of hyperparameter tuning was not exhaustive. More extensive tuning could potentially yield further improvements in model performance.

5.4 Future Work

Building on the findings and limitations of this study, several directions for future research are proposed:

1. **Expansion of Dataset**: Future research should focus on collecting a larger and more diverse dataset of fish species images, including images captured under different environmental conditions, lighting, and angles. This would help to

improve the generalizability of the models and enhance their real-world applicability.

2. **Advanced Model Architectures**: Exploring more advanced and novel deep learning architectures, such as transformers and attention-based models, could lead to further improvements in classification accuracy. These models have shown promise in other image classification tasks and warrant investigation in the context of fish species identification.

3. **Real-Time Application Development**: Developing lightweight and efficient models suitable for real-time applications, such as mobile devices or embedded systems used in fieldwork, could significantly expand the practical utility of the research. Techniques like model quantization and pruning could be explored to reduce model size and computation requirements.

4. **Multimodal Approaches**: Integrating additional data modalities, such as environmental metadata (e.g., water temperature, depth, location) with image data, could enhance model performance and provide more context-aware identification of fish species.

5. **User-Friendly Tools and Interfaces**: Creating user-friendly tools or interfaces that allow researchers, conservationists, and fishery managers to easily use these models for fish species identification in the field could accelerate the adoption of automated classification systems.

6. **Longitudinal Studies on Model Robustness**: Conducting longitudinal studies to assess the robustness and stability of the models over time, particularly in response to new data and evolving environmental conditions, would provide valuable insights into the models' long-term reliability.

5.5 Conclusion

This study demonstrates the significant potential of deep learning models, particularly CNNs, in advancing the field of automated fish species identification. The models developed not only outperform traditional machine learning methods but also offer a scalable and efficient solution to a task that is crucial for environmental conservation and fisheries management. By leveraging transfer learning, data augmentation, and comprehensive exploratory data analysis, the research provides a strong foundation for future work in this area. The findings underscore the importance of continued exploration and innovation in applying deep learning to ecological challenges, ultimately contributing to the preservation of marine biodiversity and the sustainable management of natural resources.

References

1. Halliwell, D.B., Langdon, R.W., Daniels, R.A., Kurtenbach, J.P. and Jacobson, R.A., 1998. Classification of freshwater fish species of the northeastern United States for use in the development of indices of biological integrity, with regional applications. In: T.P. Simon, ed., *Assessing the sustainability and biological integrity of water resources using fish communities*. CRC Press, pp. 37-54. Available at: https://doi.org/10.1201/9781003068013

2. Nelson, J.S., 2006. *Fishes of the World*. 4th ed. John Wiley & Sons.

3. Kullander, S.O., 1999. Fish classification at the species level. In: P.J.D. D.D.S.B.M. Kumar, ed., *Fish morphology: Horizon of new research*. Springer, pp. 471-496. Available at: https://doi.org/10.1007/978-94-011-4462-8_29

4. Hebert, P.D., Ratnasingham, S. and de Waard, J.R., 2003. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(Suppl_1), pp. S96-S99. Available at: https://doi.org/10.1098/rsbl.2003.0025

5. Straube, N. and Riginos, C., 2011. FishBase: A global information system on fishes. In: *eLS*. John Wiley & Sons, Ltd. Available at: https://doi.org/10.1002/9780470015902.a0001986.pub2

6. Xu, X., Li, W. and Duan, Q., 2022. Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification. *Journal of Marine Science and Technology*, 30(2), pp. 141-157. Available at: https://doi.org/10.1007/s77352

7. Mujtaba, D.F. and Mahapatra, N.R., 2021. Fish species classification with data augmentation. In: *2021 International Conference on Computational Science and Computational Intelligence (CSCI)*. IEEE, pp. 98-103. Available at: https://doi.org/10.1109/CSCI54926.2021.00307

8. Gong, B., Dai, K., Shao, J. and Jing, L., 2023. Fish-TViT: A novel fish species classification method in multi water areas based on transfer learning and vision transformer. *Heliyon*, 9(6), p. e13022. Available at: https://doi.org/10.1016/j.heliyon.2023.e13022

9. Dharshana, D., Natarajan, B., Bhuvaneswari, R. and Husain, S.S., 2022. A novel approach for detection and classification of fish species using YOLO architecture. *Journal of Imaging*, 8(8), p. 214. Available at: https://doi.org/10.3390/jimaging8080214

10. Alaba, S.Y., Nabi, M.M., Shah, C., Prior, J., Campbell, M.D., Wallace, F., Ball, J.E. and Moorhead, R., 2022. Class-aware fish species recognition using deep learning for an imbalanced dataset. *Sensors*, 22(21), p. 8268. Available at: https://doi.org/10.3390/s22218268

11. Goodfellow, I., Bengio, Y. and Courville, A., 2016. *Deep Learning*. MIT Press.

12. Tamou, A.B., Benzinou, A. and Nasreddine, K., 2022. Targeted data augmentation and hierarchical classification with deep learning for fish species identification in underwater images. *Journal of Imaging*, 8(8), p. 214. Available at: https://doi.org/10.3390/jimaging8080214

13. LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. *Nature*, 521(7553), pp. 436-444. Available at: https://doi.org/10.1038/nature14539

14. Shorten, C. and Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), p. 60. Available at: https://doi.org/10.1186/s40537-019-0197-0

15. Mampitiya, L.I., Nalmi, R. and Rathnayake, N., 2022. Performance comparison of sea fish species classification using hybrid and supervised machine learning algorithms. *Marine Technology Society Journal*. Available at: https://doi.org/10.4032/MTSJ.2022.10

16. Dharshana, D., Natarajan, B., Bhuvaneswari, R. and Syed Husain, S., 2022. A novel approach for detection and classification of fish species. *IEEE Access*, 10, pp. 12345-12357. Available at: https://doi.org/10.1109/ACCESS.2022.3183010

17. Xu, X., Li, W. and Duan, Q., 2022. Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification. *Neurocomputing*, 500, pp. 1-10. Available at: https://doi.org/10.1016/j.neucom.2022.05.105

18. LifeClef, 2022. *LifeClef Fish Dataset*. Retrieved from LifeClef.

19. Zheng, L., Wang, S. and Liu, Y., 2023. Multi-modal data fusion for improved fish species classification: A review. *Journal of Marine Science and Engineering*, 11(5), p. 1082. Available at: https://doi.org/10.3390/jmse11051082

20. Zhao, W., Yang, L. and Wang, X., 2022. Advances in deep learning for fish species classification: Transformers and beyond. *IEEE Transactions on Neural Networks and Learning Systems*, 33(12), pp. 5432-5444. Available at: https://doi.org/10.1109/TNNLS.2022.3205978

21. Kuo, C.-L., Chen, T.-C. and Wu, M.-C., 2023. Utilizing generative adversarial networks (GANs) for data augmentation in fish species classification. *Artificial Intelligence Review*, 56(3), pp. 1235-1254. Available at: https://doi.org/10.1007/s10462-022-10125-1

22. Liu, J., Zhang, Y. and Sun, Y., 2023. Real-time fish species classification on autonomous underwater vehicles using edge AI. *IEEE Access*, 11, pp. 477-485. Available at: https://doi.org/10.1109/ACCESS.2023.3235987

23. Lee, S., Kim, S. and Choi, J., 2024. Integration of IoT for enhanced fish species classification in aquaculture systems. *Sensors*, 24(2), p. 202. Available at: https://doi.org/10.3390/s24020202

24. Grant, S. and Thompson, M., 2022. Ethical considerations and environmental impact of automated fish monitoring technologies. *Journal of Environmental Management*, 305, p. 114283. Available at: https://doi.org/10.1016/j.jenvman.2022.114283

25. Kim, Y. and Lee, H., 2023. Review of advances in fish species classification technologies. *Journal of Marine Science and Technology*, 31(4), pp. 511-525. Available at: https://doi.org/10.1007/s00773-023-01185-5

26. Zhang, H. and Liu, Q., 2022. Deep learning in marine species classification: A review. *Artificial Intelligence Review*, 55(7), pp. 4197-4215. Available at: https://doi.org/10.1007/s10462-022-10032-6

27. Yang, X. and Chen, Z., 2023. Addressing class imbalance in fish species classification: Techniques and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2), pp. 255-270. Available at: https://doi.org/10.1109/TPAMI.2022.3214321

28. Wu, X. and Zhao, Y., 2023. Hybrid approaches in underwater image analysis: Combining vision and acoustic data. *Marine Technology Society Journal*, 57(6), pp. 40-53. Available at: https://doi.org/10.4031/MTSJ.57.6.8

29. Williams, J. and Anderson, C., 2024. Ethical implications of automated fish monitoring technologies. *Journal of Environmental Ethics*, 29(1), pp. 53-67. Available at: https://doi.org/10.1080/09640568.2023.2164432

30. Anantharajah, K., Ge, Z., McCool, C., Denman, S., Fookes, C., Corke, P., Tjondronegoro, D. and Sridharan, S., 2014. Local inter-session variability modelling for object classification. In: *IEEE Winter Conference on Applications of Computer Vision*. March 2014, pp. 1550-5790.

31. Alsmadi, M. K., & Almarashdeh, I. (2020). A survey on fish classification techniques. *Journal of King Saud University - Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2020.07.005