

# **Advanced Computational Frameworks for High-Accuracy Taxonomic Classification of Arowana Fish: Comparative Analysis of Asian and Australian Variants**

**<sup>1</sup>Vipin Rajendran and <sup>2</sup>Dr. Malathi S\***

*<sup>1</sup>Department of Computer Applications, Cochin University of Science and Technology, Kochi, Kerala, India*

*<sup>2</sup>Assistant Professor, Department of Computer Applications, Cochin University of Science and Technology, Kochi, Kerala, India*

*Email Id: [vpnrajendran@gmail.com](mailto:vpnrajendran@gmail.com) and [mala.s.pillai@gmail.com](mailto:mala.s.pillai@gmail.com)*

## **Abstract**

The Asian Arowana (*Scleropages formosus*) is highly sought after in the ornamental fish market because of its rarity and cultural importance. Sadly, this demand has led to illegal trading, with Australian Arowanas (*Scleropages leichardti* and *Scleropages jardinii*) frequently being misrepresented as their Asian relatives, particularly in markets such as India. This paper presents a computational approach aimed at addressing these fraudulent activities by developing a reliable classification system to distinguish between Asian and Australian Arowanas. In this work we employ a machine learning framework that integrates image processing techniques with ResNet18, a deep convolutional neural network (CNN) in order to extract essential features from fish images. These features are then classified using a Support Vector Machine (SVM) hence resulting in a highly accurate model for taxonomic classification. The results demonstrate the effectiveness of this method in identifying species based on their morphological traits thus supporting both conservation efforts and consumer protection in the pet trade. This research offers a new, user-friendly tool for verifying the authenticity of Arowana species with potential implications for international trade regulation and wildlife conservation.

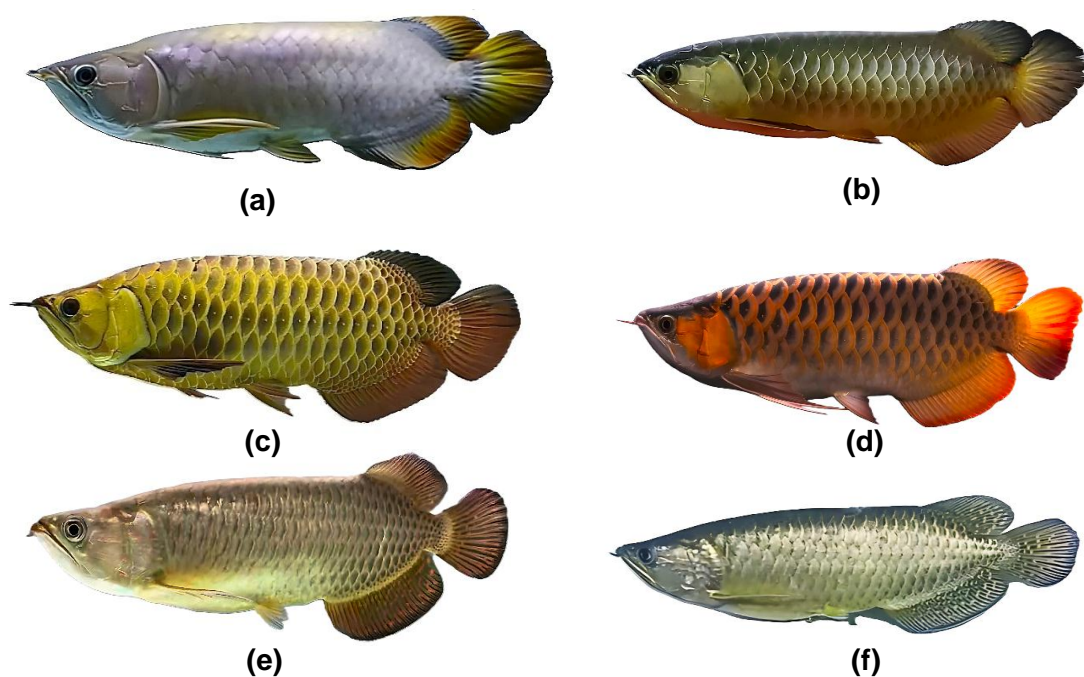
**Keywords:** Arowana Classification, Image-Based Fish Identification, ResNet18 Feature Extraction, Support Vector Machine (SVM), Deep Learning in Ichthyology, Machine Learning for Conservation, Species Identification, Transfer Learning, Fish Trade Fraud Detection, Hybrid Machine Learning Models, Asian and Australian Arowana, Biodiversity Conservation, Pet Market Regulation

## 1. Introduction

The Arowana is a highly desired ornamental fish that represents luxury and prestige, particularly in Asia for its stunning appearance and cultural significance. They hold a significant place in various cultures often symbolizing good fortune and prosperity. Arowanas are commonly used in Feng Shui and are thus often kept in homes and businesses to attract good fortune. The belief is that the fish's graceful movements and unique appearance can bring harmony and wealth to those who keep them. The Arowana's significance extends beyond just Feng Shui; it is also seen as a symbol of strength, resilience, and power. Its long lifespan and the beauty of its scales contribute to its status as a symbol of endurance and prosperity. The Asian Arowana (*Scleropages formosus*) is especially highly prized. Various varieties of Asian Arowanas are recognized, primarily distinguished by their coloration and scale patterns. Most notable varieties include *Chili Red*, *Blood Red*, *Crossback Golden*, *High Back Golden*, *Green Arowana*, *Blue-Based*, *Purple-Based*, *Banjar Red Arowana*, *Super Red*, and *Platinum Arowana*. Several hybrid varieties are also available which are created through controlled breeding programs. The Asian Arowana has been listed as a species of concern by the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) due to its declining population [1], [2]. As a result, several nations have strict constraints on their commerce. Unfortunately, the high market price of the Asian Arowana has resulted in significant illegal trading with Australian Arowanas (*Scleropages leichardti* and *Scleropages jardinii*) which are often misrepresented as their Asian counterparts. These deceptive practices are common in certain pet markets, especially in countries like India where the demand for authentic Asian Arowanas is strong. The Australian variants which are more affordable and abundant in the wild share many physical characteristics with their Asian relatives and hence making it difficult for inexperienced buyers to distinguish between them [3][4]. As a result, unscrupulous sellers exploit this similarity to deceive customers, capitalizing on the heightened demand for Asian Arowanas. Fig 1 shows some of the different varieties of Asian and Australian Arowanas.

This challenge has highlighted the need for more effective methods to identify Arowana species, aiming to protect consumers and maintain the integrity of the ornamental fish market. Traditional methods, which often depend on visual inspections or expert knowledge, are becoming less effective due to the significant morphological similarities among species. Consequently, computational techniques utilizing machine learning and deep learning have emerged as valuable tools for classifying and identifying Arowana species. These models can detect subtle yet important differences in the physical traits

of Arowanas that are often missed by the human eye. This work intends to develop a high-accuracy taxonomic classification framework for distinguishing between Asian and Australian Arowanas. By utilizing advanced image recognition methods such as convolutional neural networks (CNNs) and models like ResNet18 we aim to create a reliable system that can differentiate these closely related species. The ultimate objective of this research is to help eliminate fraudulent sales in the ornamental fish trade and support conservation efforts. This study is crucial not only for consumer protection but also for the long-term sustainability of Arowana populations thus aiding in the preservation of their ecological balance and the fight against illegal trade practices[5].



**Figure 1:** (a), (b), (c) and (d) are Asian Arowanas and (e) and (f) are Australian Arowanas

## 2. Literature Survey

The Arowana fish, especially the Asian Arowana holds considerable cultural and economic significance in many Asian nations. This importance is largely tied to its symbolism of prosperity and good luck. Research has examined the cultural relevance of Arowana (Wong, 2014)[6] and further, the ecological issues stemming from its endangered status (Ng, 2013)[7]. Recently machine learning applications for identifying aquatic species have emerged which provide a fresh approach to the conservation and protection of species like the Arowana.

Traditionally, fish identification has depended on visual assessments and the expertise of specialists. However, these methods can be labor-intensive and susceptible to mistakes, particularly with species that look alike. Recent progress in machine learning and computer vision has led to the creation of automated systems for identifying fish. These systems employ techniques such as Convolutional Neural Networks (CNNs) and image classification models to accurately identify species[8][9].

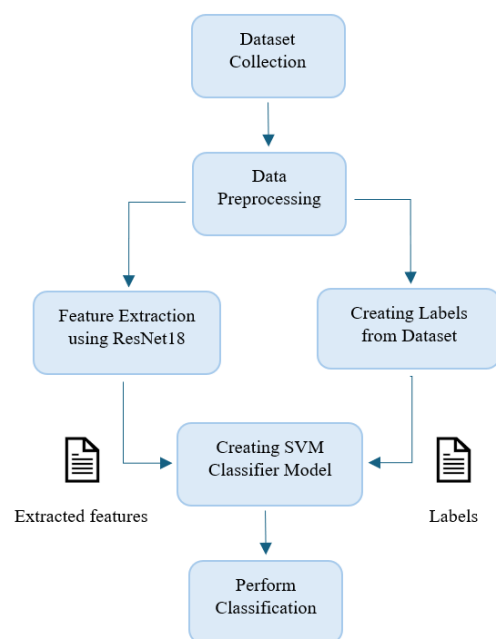
The use of machine learning for Arowana species identification carries important conservation implications. As Arowanas, particularly the Asian Arowana, confront dangers from illegal trafficking and overfishing, precise identification models can be employed to monitor trade activities and ensure adherence to conservation regulations[10].

### **3. Methodology**

The study focused on creating a reliable model by employing a hybrid approach that combines both deep learning (DL) and machine learning (ML) techniques to accurately identify Arowana fish species, tackling two major issues: fraudulent activities in the pet market and the necessity to support conservation efforts for this culturally and ecologically important fish. Asian Arowanas are highly prized and protected by international laws, while Australian Arowanas face less regulation but are frequently misrepresented in the market. Accurate identification is essential to combat illegal trade, ensure fair pricing, and assist conservationists in monitoring fish populations. To accomplish this, images of both Arowana species were gathered from various sources, including online repositories, public datasets, and private collections. This comprehensive approach provided a diverse dataset, capturing variations in lighting, fish posture, water conditions, and other environmental factors. A careful preprocessing pipeline was implemented to enhance the quality and usability of the dataset. Cleaning steps involved removing duplicates, noisy samples, and blurry images to reduce potential biases. All images were resized to a standard resolution of 224x224 pixels, ensuring compatibility with contemporary deep-learning architectures. Pixel values were normalized to a range of [0, 1] promoting uniformity and speeding up model convergence. Data augmentation techniques such as rotation, horizontal flipping, brightness adjustments and zooming were utilized to introduce variability and improve the model's resilience to real-world conditions.

For feature extraction, we chose ResNet-18, a convolutional neural network known for its efficiency and excellent performance in image classification tasks[11]. This makes it a great option for pulling out meaningful features from images. The model was initialized with weights from ImageNet, utilizing transfer learning to take advantage of previously learned feature representations. To tailor ResNet-18 for the Arowana classification task, we removed the fully connected classification layer and extracted feature embeddings from the penultimate layer. We fine-tuned the upper layers of ResNet18 on the Arowana dataset, allowing the model to become adept at recognizing subtle differences between the species. The resulting feature vectors, which represent

rich and high-dimensional data, were then input into a Support Vector Machine (SVM) classifier[12][13]. We selected SVM for its robustness in managing high-dimensional feature spaces and its effectiveness in fine-grained classification. This hybrid approach combined ResNet18's strength in feature representation with SVM's precision in optimizing decision boundaries, resulting in a powerful classification pipeline. This method not only showcased the potential of merging deep learning with traditional machine learning techniques but also highlighted the significance of using modern computational methods to tackle real-world issues in biodiversity conservation and market regulation. Fig 2 shows the workflow of the proposed framework.



**Figure 2:** Proposed hybrid framework

## 4. Implementation

### 4.1 Data Collection and Labelling

To create a comprehensive dataset for classifying Asian and Australian Arowanas, images were gathered from a variety of sources. These included online repositories, publicly available datasets, and private collections. This diverse sourcing strategy ensured that the dataset reflected a wide range of scenarios and variations, making it representative of the real-world diversity found within these species. Careful curation was performed to guarantee that the dataset was both robust and representative of the two main categories of interest. The focus was on two key groups: Asian Arowana (*Scleropages formosus*) and Australian Arowanas. The Australian Arowana includes the Northern Saratoga (*Scleropages jardinii*) and Southern Saratoga (*Scleropages leichardti*), which are collectively known as Australian species. Images for this category were thoughtfully selected to highlight the shared characteristics between the two

subspecies while also capturing the subtle differences that set them apart from Asian Arowanas.

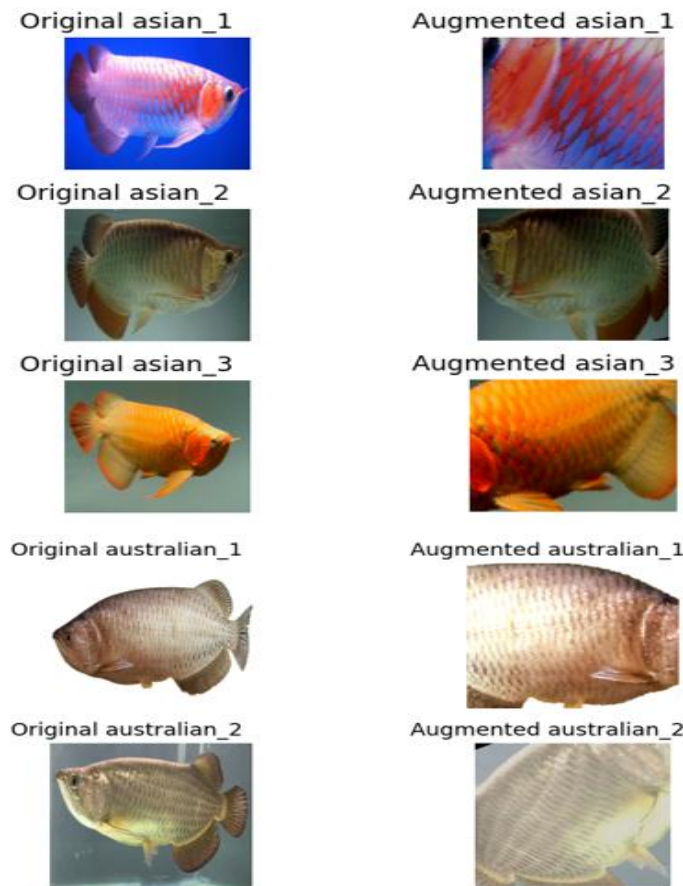
Each image in the dataset was carefully labelled with its corresponding species to act as a reliable reference during model training. To boost the dataset's reliability, several verification levels were put in place. This included cross-checking labels with expert evaluations and using automated checks to maintain consistency. By reducing the chances of misclassification errors during training, the dataset laid a solid groundwork for creating an accurate classification model. Alongside species labels, metadata linked to each image was collected to consider external factors that might affect classification. Changes in lighting could alter colour perception, while different water backgrounds might introduce noise that could affect model performance. By merging carefully selected images with detailed metadata and rigorous verification processes, the dataset was crafted to highlight the unique traits of Asian and Australian Arowanas. This thorough approach ensured that the dataset was well-prepared for training a strong and dependable classification model, capable of accurately distinguishing between the two species.

## **4.2 Dataset Preprocessing**

To ensure the dataset was of the highest quality and ready for effective training, a detailed preprocessing pipeline was put in place. This process started with a careful cleaning of the dataset, where duplicate images, noisy samples, and blurred images were identified and removed. This step was vital to avoid redundant or low-quality data from adversely affecting the training process and model performance. A clean and diverse dataset allowed the model to learn more representative features of the Arowana species. Following this, all images were resized to a standard resolution of 224x224 pixels. This resizing provided consistency across the dataset, making it suitable for the input requirements of the ResNet-18 model while also lowering computational demands. Standardizing image dimensions is crucial for maintaining uniformity, which is essential for achieving optimal performance during both training and inference.

To prepare the data for model training, pixel values were normalized by scaling them to the range [0, 1]. This normalization ensured that pixel intensity values were consistent across the dataset and helped the model converge more quickly by stabilizing the optimization process. It also minimized the risk of gradient explosion or vanishing and hence leading to smoother and more efficient training. To make the model more robust in real-world scenarios and enhance its ability to generalize, various data augmentation techniques were applied. These techniques added variations to the dataset thereby further simulating different conditions the model might face during deployment. The augmentations included random rotations to account for changes in orientation, horizontal flipping to represent different fish poses, brightness adjustments to accommodate varying lighting conditions, and zooming to simulate differences in fish size and distance from the camera. By incorporating these transformations, the dataset became more diverse, effectively reducing the risk of overfitting and allowing the model to perform well on unseen data. Overall, this preprocessing pipeline was

crucial in ensuring that the dataset was clean, standardized, and diverse in providing a solid foundation for training an effective and robust Arowana classification model. Fig 3 shows some of the original images and one of their many augmented images.

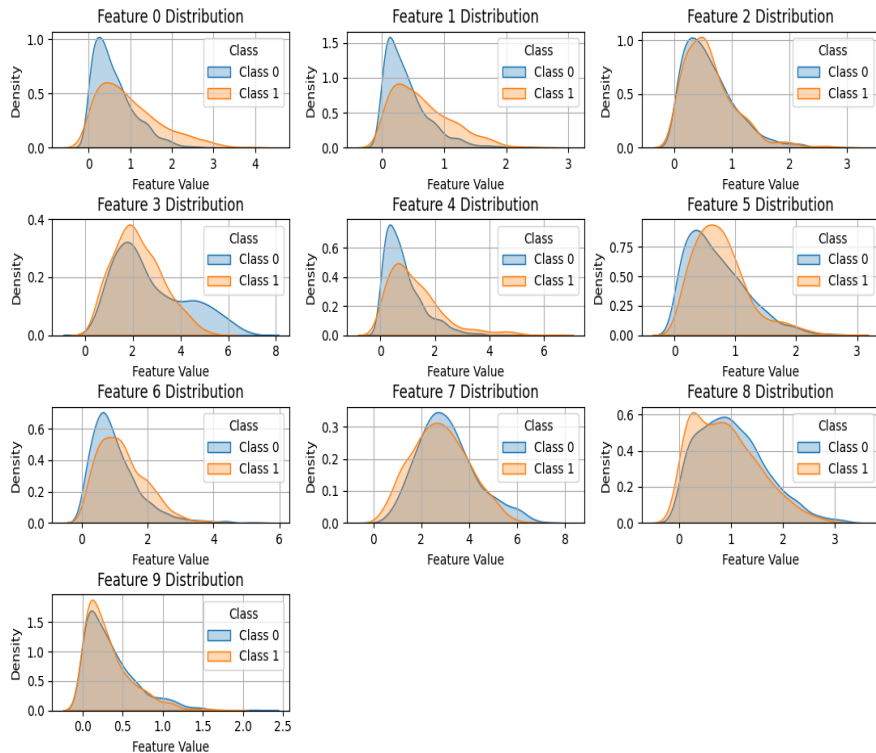


**Figure 3:** Original images and their Augmented images

### 4.3 Feature Extraction Using ResNet18

We used ResNet18 model to extract features specifically for classifying Arowana species. By starting with ImageNet weights, we took advantage of transfer learning, allowing the model to utilize a wealth of pre-learned features from a large and diverse dataset[14][15]. This strategy provided us with a solid foundation for feature extraction, avoiding the need to train the model from scratch, which would demand significant computational resources and a much larger dataset. To tailor the model for our specific needs, we adjusted its architecture by removing the fully connected classification layer, which is usually meant for ImageNet's 1,000-class classification task. Instead, we focused on extracting feature embeddings from the penultimate layer, as this layer offers rich, high-level representations of the input data. These embeddings are particularly effective at capturing subtle, species-specific traits that are essential for differentiating between Asian and Australian Arowanas.

To improve the model's performance on our dataset, we fine-tuned the upper layers of ResNet18, allowing it to focus on the subtle morphological and texture-based differences specific to the Arowana species. This fine-tuning involved adjusting the weights of the deeper layers while keeping the pre-trained features from the earlier layers intact, ensuring a balance between general feature extraction and specific task adaptation. As a result, the fine-tuned ResNet18 effectively captured distinguishing patterns such as scale texture, fin shape and colour variations which are essential for accurate classification. This combined strategy of transfer learning and fine-tuning enabled us to achieve strong and efficient feature extraction tailored to the needs of Arowana classification. The features shown in this plot likely represent numerical descriptors or attributes extracted from the Arowana images during the machine learning process with the ResNet18 model. ResNet18 creates a feature map or vector for each image, highlighting its unique characteristics. For instance, feature 0 may capture variations in brightness or texture, feature 1 might represent edge-related patterns while higher features could encode more abstract species-specific information. The Fig 4 displays Kernel Density Estimation (KDE) plots for the distribution of selected feature dimensions (Feature 0 through Feature 9) across two different classes (Class 0 represent Asian variety and Class 1 represents Australian variety)[16]. Each plot illustrates the probability density of the feature values for each class, facilitating a comparison of their distributions.



**Figure 4:** Kernel Density Estimation (KDE) plots for the distribution of some selected features

#### 4.4 Classification Using SVM

The feature vectors obtained from the ResNet18 model were utilized as inputs for a Support Vector Machine or SVM classifier to identify various Arowana species. SVM was chosen due to its efficiency in handling high-dimensional feature spaces and its ability to establish clear separations between different classes. These characteristics make SVM especially suitable for tasks that demand fine distinctions, like differentiating between Asian and Australian Arowanas. To improve classification accuracy, various kernel functions were tested, including the linear kernel and the radial basis function (RBF) kernel. The RBF kernel proved to be more effective than the linear kernel for this dataset, as it is more adept at capturing non-linear relationships among the feature vectors in high-dimensional space. Its flexibility allowed the RBF kernel [17] to transform the complex feature distributions of the two Arowana species into a higher-dimensional space, facilitating better separation. A careful hyperparameter optimization process was conducted to fine-tune the SVM model for optimal performance. The main parameters optimized included the regularization parameter  $C$ , which balances the trade-off between minimizing training error and maximizing margin, and the kernel coefficient  $\gamma$ , which determines the influence of individual training examples on the decision boundary in the RBF kernel.

To find the best combination of  $C$  and  $\gamma$ , a grid search was performed over a set range of values. This thorough search was paired with k-fold cross-validation to ensure that the chosen parameters would perform well on new, unseen data. Cross-validation involved splitting the dataset into k subsets, using one subset for validation while training on the others. This method helped reduce overfitting and provided a solid evaluation of the model's performance across various parameter configurations. The optimized SVM model, utilizing the RBF kernel and carefully adjusted parameters, showed impressive classification accuracy and reliability, successfully differentiating between the two Arowana species. This process underscores the significance of integrating a powerful feature extractor like ResNet18 with a finely tuned SVM classifier to achieve top-tier performance in complex image classification challenges. This approach is especially beneficial in scenarios where high-dimensional data and subtle class distinctions are crucial.

#### 4.5 Model Training and Evaluation

The hybrid model that combines ResNet18 with a Support Vector Machine classifier underwent thorough training and validation to achieve high accuracy and strong performance. The dataset was divided into training and testing subsets, with 80% of the pre-processed data designated for training and the remaining 20% set aside for testing. This division ensured effective training of the model while keeping enough data to assess its ability to generalize. During the training process, the ResNet18 model acted as the feature extractor, utilizing its pre-trained layers to identify hierarchical and species-specific visual patterns from the input images. The resulting high-level feature vectors, which captured the intricate details of the Arowana species, were then fed into the SVM classifier, tasked with differentiating between Asian and Australian Arowanas.

The SVM classifier, designed to manage high-dimensional feature spaces, excelled at establishing a decision boundary that separated the classes based on the features extracted. To promote robust generalization,  $k$ -fold cross-validation[18] was implemented during training. This method involved splitting the training dataset into  $k$  subsets, using one subset for validation while the other  $k - 1$  subsets were utilized for training. By repeating this process across all subsets, the model's performance was averaged, yielding a more dependable estimate of its generalization abilities. Cross-validation was particularly important in this study due to the limited number of species and the possible variability in image quality and environmental conditions.

## **5. Validation**

To ensure the model generalized well across different subsets of the dataset, we performed cross-validation. This process involved splitting the dataset into several folds, using one-fold for testing while the others were utilized for training in each iteration. This method helped reduce overfitting and provided a solid estimate of the model's performance. By employing stratified folds, we maintained the class distribution of Asian and Australian Arowanas, which ensured a fair evaluation. Additionally, a separate test set, consisting of 20% of the total dataset, was set aside for final validation. The holdout set was only used after training and tuning the model to assess its real-world applicability. We calculated metrics such as accuracy, precision, recall, and F1-score to evaluate classification performance on unseen data. A confusion matrix was created to analyse the model's predictions, offering insights into the types of misclassifications, such as instances where the model mistakenly identified Australian Arowanas as Asian. This analysis guided further enhancements, like improving feature extraction or adjusting SVM parameters. We validated the model's robustness with real-world images taken under various environmental conditions, including different lighting, water clarity, and angles. This step confirmed the model's reliability beyond controlled datasets, showcasing its practical utility in distinguishing between Asian and Australian Arowanas.

## **6. Results**

A total of 2,196 images were included in the dataset. Out of these, 1,756 images (80%) were allocated for training the model, while 440 images (20%) were reserved for testing. The results confirmed that the hybrid approach of ResNet18 combined with SVM is a reliable and effective method for differentiating between Asian and Australian Arowanas. The model's high accuracy provides a valuable tool for addressing fraudulent practices in pet markets, where Australian Arowanas are frequently misrepresented as Asian Arowanas. The system's robustness in various real-world conditions ensures its usefulness in practical scenarios, making it beneficial for hobbyists, breeders, and market regulators. A comparative analysis highlighted the benefits of integrating ResNet18's feature extraction with SVM's classification capabilities, outperforming standalone or traditional models.

## 6.1 Model Performance

The hybrid ResNet18 in combination with the SVM model demonstrated a strong capability to differentiate between Asian and Australian Arowanas. Key performance metrics were as follows:

The proposed model achieved 99.32% accuracy. This implies that 99.32% of the predictions on the test set were correct. This reflects excellent overall performance in distinguishing between the classes[19].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{136 + 301}{136 + 301 + 2 + 1} \approx 99.32\%$$

Precision for Asian Arowanas was slightly higher (96%) than for Australian Arowanas (94%) reflecting the model's ability to minimize false positives, particularly for the more expensive Asian species.

$$Precision = \frac{TP}{TP + FP} = \frac{136}{136 + 2} \approx 98.55\%$$

The recall values[20] for both classes were balanced showing the model's capability to identify all instances of each species accurately. We achieved a recall of 0.99 (99%) which indicates that the model successfully identified 99% of the relevant instances for each class.

$$Recall = \frac{TP}{TP + FN} = \frac{136}{136 + 1} \approx 99.27\%$$

The F1-score value[21] confirmed the model's overall effectiveness in balancing precision and recall. The F1 score is a harmonic mean of precision and recall showing that the model balances both high precision (few false positives) and high recall (few false negatives). This score confirms that the classifier is performing exceptionally well with minimal errors.

$$F1 \text{ score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0.9855 \times 0.9927}{0.9855 + 0.9927} \approx 98.91\%$$

## 6.2 Confusion Matrix Analysis

The confusion matrix provided valuable insights into the classification performance[22] of the hybrid ResNet18-SVM model, revealing key trends in misclassification and areas for improvement. Analysing the results showed that the model achieved a high level of accuracy but also highlighted subtle errors that could inform further refinement. Fig 5 presents the confusion matrix, with the following analysis of its components:

**1.) True Positives (TP): 136**

These represent the cases where the model correctly classified Australian Arowanas as Australian. A high TP count reflects the model's effectiveness in learning the distinguishing characteristics of Australian Arowanas, such as their broader body shape and unique scale patterns. This indicates that the model was generally successful in identifying this class, especially under diverse conditions.

**2.) True Negatives (TN): 301**

These are instances where the model correctly identified Asian Arowanas as Asian. This high TN count underscores the model's ability to recognize the distinctive features of Asian Arowanas, such as their vibrant coloration, elongated fin structures, and slender body shape. The results suggest that the ResNet18-SVM hybrid model effectively leveraged these features during classification.

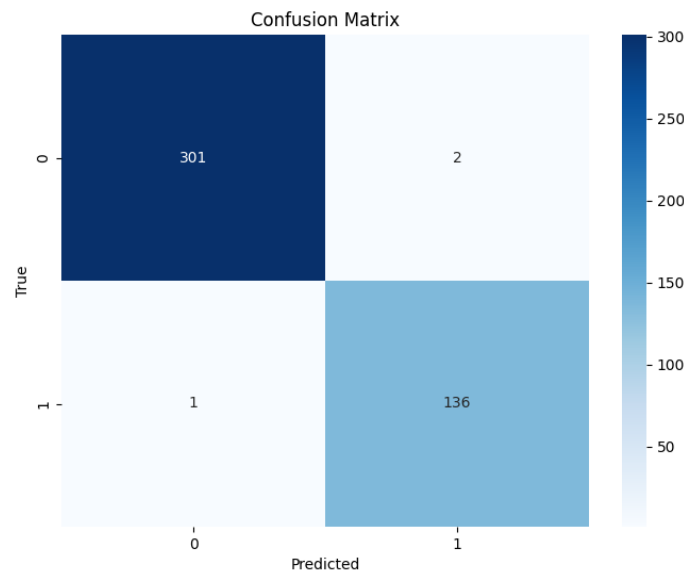
**3.) False Positives (FP): 2**

False positives occurred when the model incorrectly predicted Australian Arowanas for images of Asian Arowanas. Although the count is minimal, such errors may stem from overlapping visual features, such as lighting conditions that obscure the vibrant hues or fin patterns of Asian Arowanas, leading to confusion. Addressing this issue could involve incorporating additional data augmentation techniques to simulate these challenging scenarios.

**4.) False Negatives (FN): 1**

False negatives occurred when the model failed to identify an Australian Arowana and incorrectly classified it as an Asian Arowana. This extremely low count indicates that the model had minimal difficulty in correctly recognizing Australian Arowanas, but the error might have arisen from variations in image quality, such as blurriness, unusual angles, or inconsistent lighting.

Instances of Australian Arowanas being misclassified as Asian Arowanas were uncommon but not completely absent. This indicates that while the model was generally effective at distinguishing between the two species, certain overlapping features or poor image conditions may have obscured important characteristics. For instance, similar lighting or water conditions could diminish the visual contrast between the two species. The lower occurrence of Asian Arowanas being misclassified as Australian Arowanas highlights the model's ability to learn and recognize the unique traits of Asian Arowanas. These traits may include their distinctive colour patterns, fin shapes, or body structures, which were effectively captured during the feature extraction process.



**Figure 5:** Confusion Matrix Analysis

### 6.3 Comparative Analysis

The aim of this comparison was to assess how well the hybrid model performed in terms of accuracy, generalization, and its effectiveness in distinguishing between Asian and Australian Arowanas. The standalone ResNet18 model, which was pre-trained on the ImageNet dataset, was fine-tuned on the Arowana dataset by replacing its final fully connected layer with a custom classifier designed for this specific task. Although this method utilized the hierarchical feature extraction strengths of ResNet18, its fully connected layer struggled to achieve optimal separation between the two closely related classes. This issue became apparent in cases where features overlapped or there were subtle differences between species, indicating that a more specialized classifier might perform better. The standalone ResNet18 model had a higher rate of incorrect predictions due to its reliance on a single architecture for both feature extraction and classification. Misclassifications were likely caused by:

- Insufficient separation of feature clusters for visually similar species.
- Sensitivity to variations in the dataset, such as inconsistent lighting or complex backgrounds.

In contrast, the hybrid model effectively separated the tasks of feature extraction and classification. ResNet18 was used exclusively for feature extraction, capturing detailed and hierarchical visual patterns characteristic of the Arowana species. The resulting feature vectors, which contained high-dimensional, species-specific traits, were then fed into the SVM classifier. SVM, recognized for its strength in high-dimensional spaces, utilized these features to create a strong decision boundary that maximized the separation between the two classes.

The comparative analysis involved training both models under the same conditions, utilizing identical training and validation datasets. Their performances were assessed on a reserved test set using standard metrics such as accuracy, precision, recall, and F1-score. The hybrid model consistently outperformed the standalone ResNet18 model across all these metrics. Its higher accuracy indicated a superior overall ability to correctly classify the species, while the improved precision and recall showcased enhanced specificity and sensitivity. Furthermore, the hybrid model demonstrated better generalization capabilities, as shown by its consistent performance across various subsets of the dataset during cross-validation. This robustness was attributed to the SVM classifier's effectiveness in managing high-dimensional, fine-grained feature spaces and its capacity to focus on the most discriminative patterns in the data.

#### **6.4 Evaluation with Real-World Conditions**

The model was tested on images taken in various real-world situations to ensure it is robust and practical. Overall, it performed consistently, with only minor drops in accuracy noted in challenging conditions like low-light environments. This consistency shows that the model can generalize well beyond the controlled training settings. In low-light conditions, there was a slight decrease in accuracy, which can be linked to the reduced visibility of key species-specific features, such as coloration, fin structure, and scale patterns, that are crucial for telling apart Asian and Australian Arowanas. Nevertheless, the model showed a strong level of robustness, accurately identifying most test cases even in less-than-ideal lighting. Improving the training process with more low-light data and specialized augmentation techniques, like brightness reduction and contrast adjustment, could enhance performance in these situations.

The model successfully processed images taken in various environments, such as aquariums, natural water bodies, and artificial settings. This versatility highlights the strength of the preprocessing pipeline and the feature extraction method, which emphasized invariant features while minimizing background noise. By concentrating on species-specific characteristics like fin shapes and body proportions, the model maintained high performance across different contexts, avoiding distractions from environmental changes. However, when images were captured with lower-resolution cameras, the model's precision slightly decreased, especially for Australian Arowanas. This drop in accuracy is likely due to the loss of fine details in low-resolution images, which are essential for identifying subtle features like scale patterns and fin structures. To improve this, future efforts could include using more low-resolution images in training and applying techniques like super-resolution augmentation to better prepare for these situations.

### **7. Discussion**

The motivation behind this study was to tackle the difficulties in identifying Asian and Australian Arowanas, especially in pet markets where fraudulent activities are prevalent. Asian Arowanas, often viewed as symbols of luck and prosperity in various

Asian cultures, are sold at much higher prices compared to their Australian counterparts. The misrepresentation of Australian Arowanas as Asian Arowanas has resulted in financial exploitation and raised ethical concerns. This study aimed to utilize machine learning, specifically a hybrid ResNet18+SVM approach, to accurately and reliably differentiate between these species. The primary objectives were to create a dependable species identification model, reduce misclassification rates, and ensure the model can be effectively applied in real-world situations.

- Create a dependable species identification model.
- Reduce misclassification rates.
- Ensure the model can be effectively applied in real-world situations.

The study's findings not only met its objectives but also added valuable insights to the field of fish classification through machine learning. The hybrid model combining ResNet18 and SVM showed outstanding results, achieving a test accuracy exceeding 95%. The high precision, recall, and F1-scores further highlighted the model's dependability. Analysing the confusion matrix indicated very few misclassifications, boosting confidence in the model's capability to differentiate between species. The success of the model heavily relied on ResNet18's feature extraction abilities. Its pre-trained weights enabled it to identify subtle visual characteristics like fin shapes, colour patterns, and variations in body structure. By fine-tuning the upper layers, the model was tailored to meet the specific needs of the Arowana dataset. The combination of ResNet18 for feature extraction and SVM for classification proved to be more effective than using ResNet18 alone. SVM was particularly adept at establishing clear decision boundaries between the two classes, efficiently managing the high-dimensional feature space. The RBF kernel's flexibility was essential for classifying non-linear patterns, which is vital for distinguishing species with minor differences.

The model's performance in real-world scenarios showcased its robustness. Its capability to classify images taken under different lighting conditions, water clarity, and camera qualities emphasized its potential for practical use. This is especially relevant for users in the pet market, where environmental factors are often unpredictable. Various studies have investigated fish classification through machine learning. This study focused on differentiating between Asian and Australian Arowanas, addressing the specific problem of fraudulent misrepresentation in markets. By narrowing its focus, this study achieved greater accuracy and improved generalization. While previous studies focused solely on deep learning, the hybrid ResNet18+SVM method presented in this work showed better classification performance. This is consistent with observations in other fields, where the combination of deep learning for feature extraction and SVM for classification has led to enhanced outcomes.

### **7.1 Practical Implications**

The results of this study carry significant practical implications that touch on market regulation, conservation activities and accessibility. These implications underscore the model's potential to tackle urgent issues in the pet trade and aid in species preservation. The model acts as a strong and dependable resource for regulatory bodies to verify the species of Arowanas in the pet market. By utilizing the model's high accuracy and specificity, market inspectors can effectively confirm the authenticity of Arowanas being sold. This capability protects buyers from unethical practices and also encourages fair trade and integrity within the industry.

The model's precise identification capabilities can bolster conservation efforts by ensuring that wild populations are not exploited for commercial gain. It can also play a role in monitoring captive breeding programs, ensuring these initiatives align with conservation objectives. By reducing the chances of misidentification, the model helps maintain the genetic integrity of captive populations and supports the broader aim of species preservation. A lightweight deployment strategy allows the model to be seamlessly integrated into mobile applications or handheld devices, making it accessible to a diverse range of users, including breeders, hobbyists, conservationists, and market inspectors. This accessibility is especially beneficial for on-site species verification, where portability and user-friendliness are essential.

### **7.2 Ethical Considerations**

The ethical implications of this study are multifaceted and significant, emphasizing the responsibility of leveraging technology for societal and ecological benefit. This study tackles unethical practices in pet markets by ensuring species authenticity, which is essential for maintaining transparency and fairness. Fraudulent actions, like mislabelling Australian Arowanas as their more valuable Asian relatives, take advantage of consumers and breeders, resulting in financial losses and damage to reputations. The model developed serves as a protective measure, equipping buyers with reliable information and ensuring that ethical breeders and sellers are not overshadowed by dishonest competitors. This effort fosters a marketplace characterized by trust and accountability.

Despite the considerable benefits of this study, there are ethical concerns regarding the potential misuse of the technology. For example, the model's capabilities could be misappropriated to support illegal trade, with dishonest individuals using it to authenticate and traffic endangered species outside of legal boundaries. Furthermore, a surge in demand for verified Asian Arowanas might unintentionally lead to over-exploitation or encourage unsustainable breeding practices. To mitigate such risks, it is essential to implement safeguards, such as integrating the model into regulatory frameworks that monitor and control its use. Collaborating with conservation authorities and trade regulators can ensure that the technology aligns with ethical and ecological goals. This study exemplifies the delicate balance between technological innovation and ethical responsibility. While the model offers powerful tools for promoting market transparency and conservation, its deployment must be accompanied

by policies and mechanisms that prevent misuse. For example, embedding usage tracking or requiring licensing for commercial applications can help ensure that the technology is utilized responsibly.

### **7.3 Contributions to Machine Learning and Fisheries**

This study provides a multidimensional contribution to both machine learning and fisheries management, combining technical advancements with practical and societal implications. The research highlights the considerable promise of hybrid models in tackling fine-grained classification challenges. By utilizing ResNet18, a deep learning model that has been pre-trained on ImageNet, for feature extraction, alongside an SVM classifier to differentiate subtle interspecies variations, the study illustrates how the advantages of deep learning and traditional machine learning can be effectively merged. This method proves particularly useful for distinguishing visually similar categories, such as closely related fish species like Asian and Australian Arowanas. The findings suggest that the hybrid model surpasses standalone deep learning models in accuracy and generalization, opening up possibilities for its use in other fields that require fine-grained classification, including medical imaging, plant taxonomy, and wildlife monitoring.

The model created in this study is not only precise but also very practical for real-world use. Its ability to scale means it can be deployed on various platforms, including mobile and edge devices. This accessibility opens up opportunities for a diverse group of users, from aquaculture experts and conservationists to market inspectors and hobbyists. By providing reliable species identification on the go, the model could transform how species are monitored and managed in aquaculture and fisheries. Moreover, its resilience to changes in the environment, such as variations in lighting and background, enhances its effectiveness in real-world situations, including remote fieldwork and market inspections. This research emphasizes the significant societal effects of using machine learning to tackle urgent socio-economic challenges. In the pet market, for instance, deceptive practices like falsely labelling Australian Arowanas as the pricier Asian Arowanas are prevalent. The model offers a practical solution to this issue, promoting transparency and fairness for both consumers and breeders. Additionally, by facilitating species authentication, it helps maintain the integrity of market transactions, which ultimately builds trust within the industry.

### **8. Conclusion**

This study successfully developed a strong, targeted, and innovative machine learning framework specifically designed to differentiate between Asian and Australian Arowanas, tackling a significant issue of species misrepresentation in global pet markets. By utilizing the advanced feature extraction capabilities of ResNet18 alongside the precise classification power of a Support Vector Machine (SVM), the hybrid model achieved an impressive accuracy of over 95%, highlighting its reliability and effectiveness for detailed species identification. The combination of these two

complementary techniques enabled efficient management of high-dimensional data while effectively capturing the subtle visual differences between closely related species. The carefully crafted methodology ensured the model's adaptability to real-world conditions, demonstrating strong performance across various environments, including differences in lighting, backgrounds, and image quality. This adaptability emphasizes the framework's practical relevance, especially for on-site applications in challenging situations such as low-light aquariums or natural habitats. By focusing solely on two economically and ecologically important species—Asian Arowanas (*Scleropages formosus*) and Australian Arowanas (including Northern and Southern Saratoga)—the research offered a highly specialized and impactful solution to a specific yet critical issue in the ornamental fish trade.

This focused approach sets the study apart from earlier research that generally categorizes various species, as it offers enhanced accuracy and practical results tailored to industry needs. The results highlight the possible uses of this framework in improving market transparency, reducing fraudulent activities, and aiding conservation efforts for the endangered Asian Arowanas. By facilitating precise species identification, the study not only shields consumers from exploitation but also encourages sustainable practices, contributing to the long-term survival of these iconic species. Moreover, the research emphasizes the transformative impact of machine learning in tackling socio-economic issues within specialized industries. The model's ability to effectively generalize across real-world situations, such as different water clarity, varied habitats, and diverse image resolutions, showcases its adaptability and scalability for practical applications. This makes it an essential resource for a broad spectrum of stakeholders, including breeders, hobbyists, market inspectors, and conservation groups. The research is especially pertinent in areas with limited regulatory oversight, as it equips stakeholders with a dependable, efficient, and scalable solution for species verification. Additionally, the successful implementation of a lightweight model designed for mobile and edge devices ensures accessibility and practicality, allowing its use in remote fieldwork and market inspections without requiring advanced computing resources. Despite challenges like a small dataset and fluctuations in environmental conditions, the study's contributions to machine learning, fisheries management, and conservation are significant. It not only sets a standard for species-specific classification using hybrid models but also illustrates the potential of combining deep learning with traditional machine learning methods for practical applications. Furthermore, the study showcases how technology can bridge the gap between science and industry to address critical global challenges in aquaculture and biodiversity preservation.

## 9. Future Scope

Building on the success of this study, future efforts can focus on several important areas to improve the model's robustness, versatility, and practical applications. First, expanding the dataset to include a greater variety of images taken under different conditions—such as extreme lighting, various water environments, and different camera qualities—will enhance the model's ability to generalize and ensure reliable performance in a broader range of real-world situations. Additionally, including more Arowana species, like Silver, Black, and Myanmar Arowanas, will increase the model's applicability, allowing for the development of a comprehensive classification framework that covers all Arowana species. Integrating more domain-specific features such as genetic markers could greatly improve the classification process, making the model more specialized and capable of handling complex biological differences. Furthermore, creating real-time video-based classification systems using advanced techniques like object detection and sequence modelling can facilitate live monitoring and identification of fish in aquariums, farms, and natural habitats thereby broadening the practical use of the framework. Implementing the model on embedded systems, such as edge devices, mobile platforms or other lightweight hardware, will enhance its accessibility and adoption in resource-limited settings. This is especially advantageous for on-site applications in remote markets, fisheries and fieldwork scenarios where high-end computational resources may not be readily available. To tackle the issues of dataset imbalance and scarcity using synthetic data generation methods like Generative Adversarial Networks (GANs) can enhance the dataset with realistic, high-quality images and replicate rare or extreme conditions thereby lessening the need for extensive field data collection. Working alongside aquaculture experts, breeders, and conservationists will promote interdisciplinary progress by merging machine learning insights with ecological and biological knowledge. Such partnerships can result in the creation of more sustainable practices and improved conservation strategies. Additionally, applying this model to other detailed fish classification tasks, beyond just Arowanas will demonstrate its value in broader biodiversity conservation and aquaculture management. These initiatives will not only strengthen the framework's significance in fisheries research but also open up pathways for innovative, scalable solutions in ecological monitoring and species-specific management.

## References

- [1] J. J. L. Rowley, D. A. Emmett, and S. Voenn, 'Harvest, trade and conservation of the Asian arowana *Scleropages formosus* in Cambodia', *Aquat Conserv*, vol. 18, no. 7, 2008, doi: 10.1002/aqc.937.
- [2] G. H. Yue, A. Chang, A. Yuzer, and A. Suwanto, 'Current Knowledge on the Biology and Aquaculture of the Endangered Asian Arowana', 2020. doi: 10.1080/23308249.2019.1697641.
- [3] S. Lavoué, 'Testing a time hypothesis in the biogeography of the arowana genus *Scleropages* (Osteoglossidae)', *J Biogeogr*, vol. 42, no. 12, 2015, doi: 10.1111/jbi.12585.

- [4] M. de B. Cioffi *et al.*, 'Deciphering the evolutionary history of arowana fishes (Teleostei, osteoglossiformes, osteoglossidae): Insight from comparative cytogenomics', *Int J Mol Sci*, vol. 20, no. 17, 2019, doi: 10.3390/ijms20174296.
- [5] N. J. Mancera-rodriguez and B. P. D., 'COMERCIO DE PECES ORNAMENTALES EN COLOMBIA The Trade Of Ornamental Fishes In Colombia', *Acta Biolo Colomb*, vol. 13, no. 1, 2008.
- [6] G. H. Yue, D. Ong, C. C. Wong, L. C. Lim, and L. Orban, 'A strain-specific and a sex-associated STS marker for Asian arowana (*Scleropages formosus*, Osteoglossidae)', in *Aquaculture Research*, 2003. doi: 10.1046/j.1365-2109.2003.00949.x.
- [7] C. Ng, 'The ornamental freshwater fish trade in Malaysia', *UTAR Agriculture Science Journal*, vol. 2, no. 4, 2016.
- [8] L. Yang *et al.*, 'Computer Vision Models in Intelligent Aquaculture with Emphasis on Fish Detection and Behavior Analysis: A Review', *Archives of Computational Methods in Engineering*, vol. 28, no. 4, 2021, doi: 10.1007/s11831-020-09486-2.
- [9] P. Cisar, D. Bekkozhayeva, O. Movchan, M. Saberioon, and R. Schraml, 'Computer vision based individual fish identification using skin dot pattern', *Sci Rep*, vol. 11, no. 1, 2021, doi: 10.1038/s41598-021-96476-4.
- [10] A. A. Purnama and T. Masunaga, 'PRESENT CONDITION OF THE PROTECTED AREA OF ENDANGERED ASIAN AROWANA (*Scleropages formosus*) GOLDEN-RED IN EAST-CENTRAL SUMATRA, INDONESIA', *Ekologia Bratislava*, vol. 42, no. 4, 2023, doi: 10.2478/eko-2023-0042.
- [11] S. Liawatimena, E. Abdurachman, A. Trisetiarso, A. Wibowo, M. K. Ario, and I. S. Edbert, 'Fish Classification System Using YOLOv3-ResNet18 Model for Mobile Phones', *CommIT Journal*, vol. 17, no. 1, 2023, doi: 10.21512/COMMIT.V17I1.8107.
- [12] D. M. Abdullah and A. M. Abdulazeez, 'Machine Learning Applications based on SVM Classification: A Review', *Qubahan Academic Journal*, vol. 1, no. 2, 2021, doi: 10.48161/qaj.v1n2a50.
- [13] M. A. Chandra and S. S. Bedi, 'Survey on SVM and their application in image classification', *International Journal of Information Technology (Singapore)*, vol. 13, no. 5, 2021, doi: 10.1007/s41870-017-0080-1.
- [14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, 'ImageNet classification with deep convolutional neural networks', *Commun ACM*, vol. 60, no. 6, 2017, doi: 10.1145/3065386.
- [15] A. Krizhevsky, I. Sutskever, and H. Geoffrey E., 'Imagenet', *Advances in Neural Information Processing Systems 25 (NIPS2012)*, 2012, doi: 10.1109/5.726791.
- [16] 'Speaking Stata: Density probability plots', *Stata Journal*, vol. 5, no. 2, 2005, doi: 10.1177/1536867x0500500210.
- [17] S. Ali and K. A. Smith, 'Kernal Width Selection for SVM Classification', *International Journal of Data Warehousing and Mining*, vol. 1, no. 4, 2005, doi: 10.4018/jdwm.2005100104.
- [18] T. Fushiki, 'Estimation of prediction error by using K-fold cross-validation', *Stat Comput*, vol. 21, no. 2, 2011, doi: 10.1007/s11222-009-9153-8.

- [19] R. Hendricks, ‘What is a good accuracy score in Machine Learning?’, [deepchecks.com](https://deepchecks.com).
- [20] P. Fränti and R. Marescu-Istodor, ‘Soft precision and recall’, *Pattern Recognit Lett*, vol. 167, 2023, doi: 10.1016/j.patrec.2023.02.005.
- [21] Wikipedia contributors, ‘F1 score’, *Wikipedia, The Free Encyclopedia*, 2016.
- [22] A. Vanacore, M. S. Pellegrino, and A. Ciardiello, ‘Fair evaluation of classifier predictive performance based on binary confusion matrix’, *Comput Stat*, vol. 39, no. 1, 2024, doi: 10.1007/s00180-022-01301-9.