Computational Biology and Bioinformatics
2023; 11(2): 33-48
http://www.sciencepublishinggroup.com/j/cbb
doi: 10.11648/j.cbb.20231102.12
ISSN: 2330-8265 (Print); ISSN: 2330-8281 (Online)

Automated Fish Measurement and Classification Using Convolutional Neural Networks (CNNs)

Soad El Hiak*, Xu He

College of Mechanical and Electrical Engineering, Harbin Engineering University, Harbin, China

Email address:
soad@hrbeu.edu.cn (Soad El Hiak)
*Corresponding author

To cite this article:
Soad El Hiak, Xu He. (2023). Automated Fish Measurement and Classification Using Convolutional Neural Networks (CNNs). *Computational Biology and Bioinformatics*, *11*(2), 33-48. https://doi.org/10.11648/j.cbb.20231102.12

Received: October 21, 2023; **Accepted**: November 25, 2023; **Published**: December 5, 2023

Abstract: Managing fisheries requires regular monitoring and assessment of fish populations. Traditional methods of evaluating fish stocks, particularly their size, can be time-consuming, labor-intensive, and inaccurate. Recently, digital image processing (DIP) and machine learning (ML) have emerged as promising technologies to automate fish measurement and classification. In this study, we aim to develop deep learning models to predict, and classify shape and size of the fish using convolutional neural networks (CNNs) and DIP techniques. The study utilizes publicly available fish datasets and evaluates the efficiency of the proposed models using metrics such as precision, recall, and F1 score. The developed models utilize Python programming language with TensorFlow and Keras libraries. The regression component investigates the intricate relationship between various physical attributes of fish, uncovering the connections between body length, height, and weight. This analysis provides valuable insights into the correlations among these attributes, enhancing our understanding of fish characteristics. Simultaneously, the classification segment introduces an innovative approach to fish classification, incorporating shape and size attributes. Through a combination of classifiers and ensemble learning with stacking, exceptional accuracy is achieved in identifying distinct fish classes. This integration of techniques facilitates a more nuanced classification process, allowing for comprehensive categorization based on visual attributes. Our study establishes a robust framework for fish analysis and classification, Utilizing the combined strengths of digital image processing (DIP) and machine learning (ML). The developed models not only enhance the accuracy and efficiency of size classification but also contribute to the broader goal of sustainable fisheries management. This research sets a foundation for future endeavors in automating fish stock assessments, contributing to the advancement of fisheries science and management practices.

Keywords: Machine Learning, Convolutional Neural Networks (CNNs), Fish Measurement and Classification

1. Introduction

Classifying fish serves as a valuable tool for tasks such as assessing fish populations, tallying their numbers, monitoring ecosystem health, and documenting fish groupings [13]. Fish measurement and classification are essential tasks in fisheries science and management. They help in understanding the distribution and abundance of fish populations, determining legal and sustainable catch limits, and identifying trends in species health and productivity. Traditional methods of fish length measurement and classification, such as manual measurement or visual identification, can be time-consuming, labor-intensive, and prone to measurement errors. These limitations may lead to inaccuracies and errors in fish stock assessments, which can have significant implications for the management and conservation of fish populations. Automated fish measurement and classification systems based on computer vision and machine learning techniques have shown tremendous potential in addressing these limitations, providing faster, more consistent, and objective methods for measuring and classifying fish. Deep learning algorithms, such as CNNs, have been particularly effective in image recognition and classification tasks, making them well-suited for automated fish length measurement and classification systems. However, developing these systems can be challenging due to factors such as variations in fish morphology and morphology changes over time and under varying environmental conditions. Furthermore, to develop an accurate and robust automated fish length measurement and classification system, it is necessary to have sufficient training data that represent a wide range of fish species and sizes.

Traditional methods of fish size measurement are often prone to a range of limitations and potential sources of errors. These include

1. Human Error: Inaccuracies in the measurement process can occur due to the lack of precision by the person collecting the data and the measurement tools used: [10]
2. Bias: Measurements may be influenced by the observer's perception and subjective interpretations of the collected data.
3. Inter-Observer Variability: Different observers may obtain different results when measuring the same specimen, leading to variability in the data collected.
4. Lack of Standardization: Traditional methods often lack standardization, making it challenging to compare data between different studies or researchers.

New advanced methods and technologies have been developed to address the limitations of traditional methods of fish size measurement. These technologies include:

1. Image Processing Techniques: Image processing techniques use algorithms and software to analyze digital images of fish and obtain length and weight measurements. These techniques include geometrical and texture-based methods, which use image features such as color and shape, as well as machine learning algorithms. These methods rely on the quality of the obtained images. [4]
2. Stereoscopic Image Analysis: Stereoscopic image analysis technologies use two or more synchronized cameras to capture fish images and extract their three-dimensional features. These methods are useful in accurately estimating fish length and volume and can be improved when combined with AI techniques. [16]
3. Computer Vision Approaches: The precision of fish length measurement and weight estimation has been notably enhanced by recent developments in computer vision techniques, particularly through the utilization of deep learning algorithms and convolutional neural networks (CNNs) [3].
4. Automated Machine Vision Systems: Automated machine vision systems use cameras and sensors to capture real-time images of fish and calculate their size and weight. These systems can provide accurate, non-invasive, and quick measurements. [7]

The limitations of traditional methods for fish size measurement highlight the need for innovative new technologies and methodologies Lately, there have been notable breakthroughs and advancements in the fields of computer vision, image processing, and machine learning, among others, have shown promising results in improving the accuracy and efficiency of fish size measurement and classification. These technologies can help address issues such as human error, bias, and inter-observer variability, and their implementation can significantly improve data quality for fisheries management, aquaculture, and research purposes.

2. Methodology

2.1. Convolution Neural Network CNN

In the context of Deep Learning, feature extraction is abstracted, but the intricate design of the neural network significantly influences the automatic extraction of these features. One prominent design in this context is the Convolutional Neural Network (CNN), extensively utilized in the field of image classification challenges and [2] Comprising several purpose-driven layers, the CNN architecture exhibits a structured elegance that unlocks its formidable capabilities. At the core of the CNN architecture lie multiple layers, each meticulously crafted to serve a distinct purpose in the feature extraction process. Figure 2 illustrates this intricate arrangement, capturing the essence of how information flows through these layers, gradually unveiling the essential characteristics of the input data. The true brilliance of CNNs lies in their adeptness at capturing hierarchical representations of features. By sequentially stacking convolutional and pooling layers, the network progressively learns and consolidates features, forming a multi-layered representation that becomes increasingly perceptive and discriminative. Due to their ability to automatically learn essential patterns from raw data, CNNs have emerged as a powerhouse in the domain of image classification, revolutionizing the way machines perceive and interpret visual information. As research and advancements in Deep Learning continue to flourish, CNNs stand tall as a beacon of innovation, reshaping the landscape of Artificial Intelligence with their transformative potential.

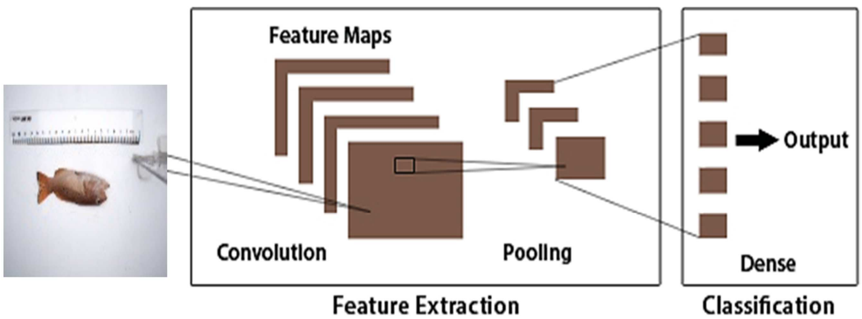


Figure 1. *A typical CNN architecture [11].*

2.2. Regression Model

In the realm of statistical analysis, regression is utilized to construct models that illustrate the relationship between a dependent variable and one or more independent variables. In our research, the objective is to harness regression analysis to predict fish attributes by utilizing data obtained from their measurements and images. [17]

Training the model: After clean the data and split our dataset into training (80%) and to testing (20%) then in our Training we used an implementation of a deep learning model of CNN using the TensorFlow and Keras libraries.

The model architecture is defined using the Sequential API of Keras. The model begins with a Conv2D layer, which performs convolutional operations on the input data. The layer has 32 filters, each with a size of 3x3, and uses the 'Relu' activation function.

Following the Conv2D layer, a The MaxPooling2D layer is introduced to reduce the resolution. The output from the preceding layer. This layer reduces the spatial dimensions of the input by taking the maximum value in each 2x2 region. Another Conv2D layer with 64 filters and a size of 3x3 is added, followed by another MaxPooling2D layer.

The output of the second MaxPooling2D layer is then flattened into a 1D array using the Flatten layer. This prepares the data for the fully connected layers of the model. [5]

Next, two Dense layers are added as fully connected layers. The first Dense layer has 64 units and uses the 'Relu' activation function. The second Dense layer has 4 units, which corresponds to the number of output classes or labels in the dataset. In this case, there is no activation function has been specified for the output layer, allowing the output values to be any real number.

After defining the model architecture, the compilation step is performed using the 'compile' function. The optimizer is specified as 'Adam', which is a popular optimization algorithm for training deep learning models. The choice for the loss function is 'mean_squared_error,' employed for quantifying the disparity between the predicted and observed values. Furthermore, the 'mean_absolute_error' metric has been selected for monitoring the mean absolute error throughout the training process.

Subsequently, the model undergoes training using the 'fit' function. The training dataset, denoted as X_train and y_train, is provided, along with the number of epochs and batch size. In this instance, the model is trained for 100 epochs, with a batch size of 1. During the training process, the model's parameters are iteratively adjusted using the designated optimizer and loss function to minimize the discrepancy between predicted and actual outputs.

Prediction: Next, the trained model is loaded into the prediction section by loading the saved model file generated during the training phase, then we utilized a machine learning model to make predictions on the test and training datasets. The predictions were obtained using the 'model. predict' function on both datasets, generating 'y_pred_test' for the test dataset and 'y_pred_train' for the training dataset [8]

Evaluation: we utilized the 'sklearn' library to calculate various metrics based on the model's predictions and the true values from the training dataset. Firstly, we computed the Mean Absolute Error (MAE) by comparing the predicted values ('y_pred_train') against the true values from the training dataset ('y_train'). The MAE represents the average absolute difference between the actual and predicted values. Next, we calculated the Root Mean Absolute Error (RMSE) by taking the square root of the MAE [1]. The RMSE is a commonly used metric that provides a measure of the square root of the average squared differences between the predicted and true values. Then we computed the Mean Squared Error (MSE), which involves taking the average of the squared differences between the predicted and true values. The MSE provides a measure of the average squared error in the model's predictions on the training data. Finally, we determined the Root Mean Squared Error (RMSE) by taking the square root of the MSE. The RMSE serves as an important metric, representing the standard deviation of the model's errors on the training data.

2.3. Classification Method

Preparation Data: the preparation of the dataset for fish classification based on size and body shape characteristics. We calculate the ratio of body length to weight to categorize fish into "Medium" and "Small" groups. Additionally, we compare body width to height to classify fish as "Narrow and tall" or "Wide and short." And combine between shape and size The consolidated "Combine" column enables in-depth analysis.

Training the model: After split datasets into Training (75%) and Testing (25%) we train the individual Machine Learning classifiers (Random Forest, KNN, SVM, Logistic Regression, Decision Tree, MLP) this six algorithms that would utilize in classification methodology and the characteristic of each model is discussed in flowing:

Logistic Regression: derives its name from the logistic function, which plays a central role in this technique. The purpose of Logistic Regression is to train a model that can decide if a new input sample belongs to one of two classes. in Equation 2.1, if the probability.

$P(y = 1|x)$ is greater than 0.5 for a given x, we answer yes; otherwise, we say no. The decision boundary is indicated by the threshold value of 0. 5.

$$decision(x) = \begin{cases} 1\ if\ P(y = 1|x) > 0.5 \\ 0\ otherwise \end{cases} \quad (1)$$

K-Nearest Neighbor: algorithm classifies data without previously constructing a model. It categorizes its classes based on "majority votes" and assigns to the most frequent class by calculating the distance between data.

Support Vector Machine (SVM): is a supervised learning algorithm that is versatile in its ability to handle both regression and classification tasks.

The role of these hyperplanes is to partition the feature space in a way that all images with feature vectors lying on one side are assigned the class label -1, whereas those on the other side receive class label +1.

The Multi-layer Perceptron (MLP): is a supervised learning algorithm that, through training on a dataset, learns a function. $f(\cdot): R^m \rightarrow R^o$. In this context, 'm' represents the number of input dimensions, and 'o' represents the number of output dimensions.

When provided with a set of features $X = x_1, x_2, \ldots, x_m$ and a target y, MLP has the capability to acquire a non-linear function approximator, making it suitable for both classification and regression tasks. [15]

Random Forest (RF): is a composite algorithm formed by creating decision trees in a randomized fashion. Each decision tree is constructed using distinct randomly chosen features and training samples. Consequently, RF delivers a substantial enhancement in accuracy when compared to a single decision tree. Simultaneously, it bolsters the resilience of the decision tree, which is susceptible to certain attacks. This attribute makes RF particularly advantageous in fish identification projects with lower security needs. [9].

Development and Evaluation: evaluate the performance of each classifier using metrics like accuracy, precision, recall, and F1-score [12]. Next Seeking to optimize and refine this performance, we turned to ensemble learning It involves combining multiple base models to make more accurate predictions than individual models could achieve alone. While there are several ensemble methods like bagging and boosting, in this section, we'll focus on an advanced technique called "stacking." is a meta-learning technique that harnesses the predictions from multiple diverse models to construct a more advanced model at a higher level [6]. Then we coupled it with Grid Search parameter tuning. This iterative approach resulted in a substantial accuracy enhancement. By combining diverse classifiers and fine-tuning hyperparameters, to optimize our model achieved higher accuracy on both training and test dataset.

2.4. Integration

Integrated regression and classification techniques to create a comprehensive solution for fish analysis and classification based on images.

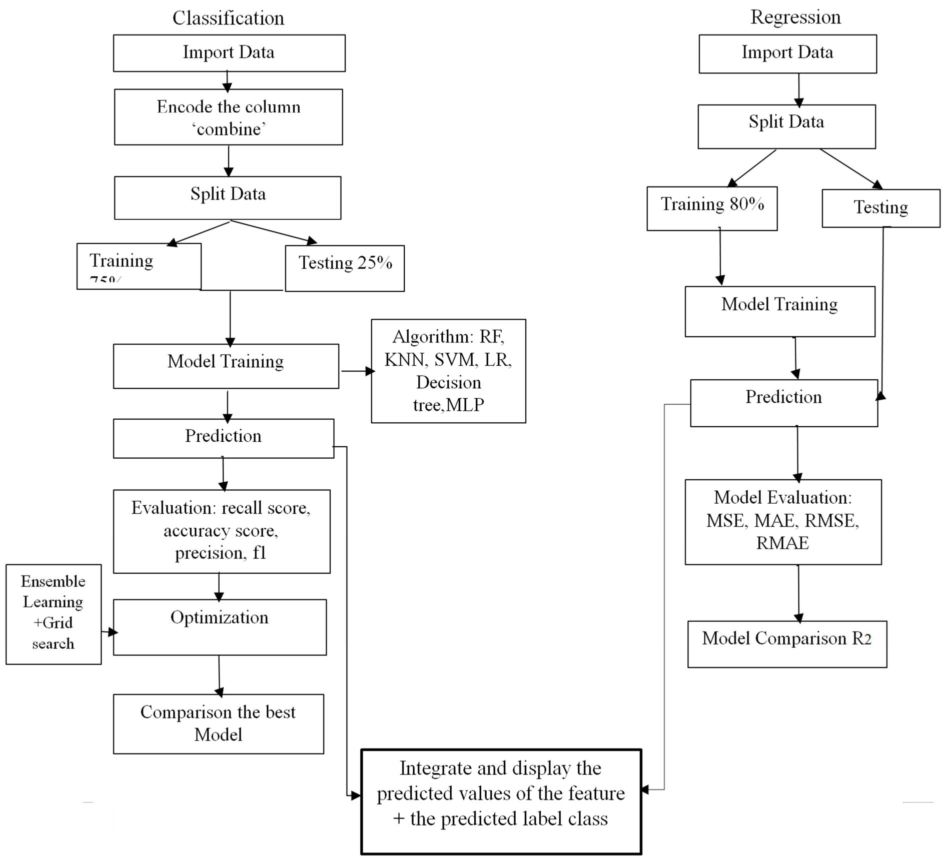


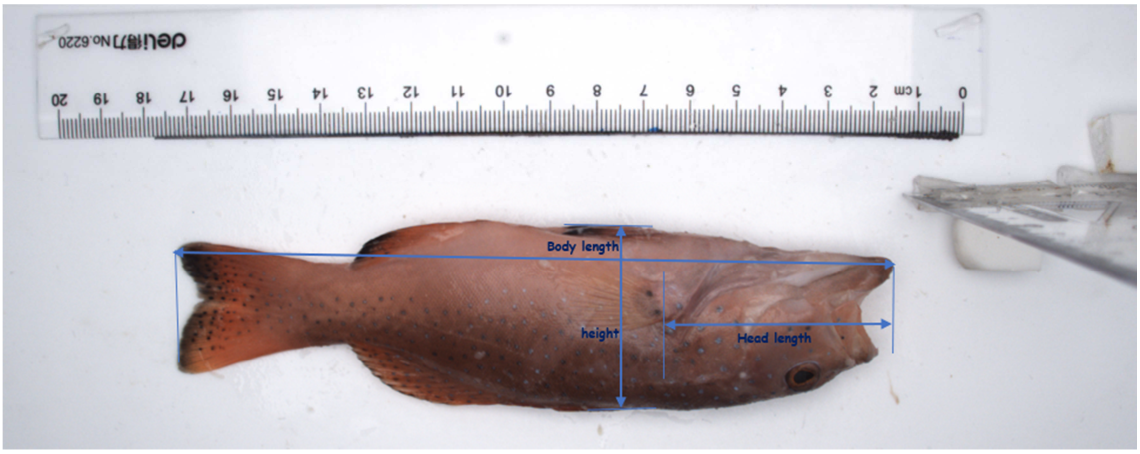
Figure 2. *General flowchart of the proposed system.*

3. Result and Discussion

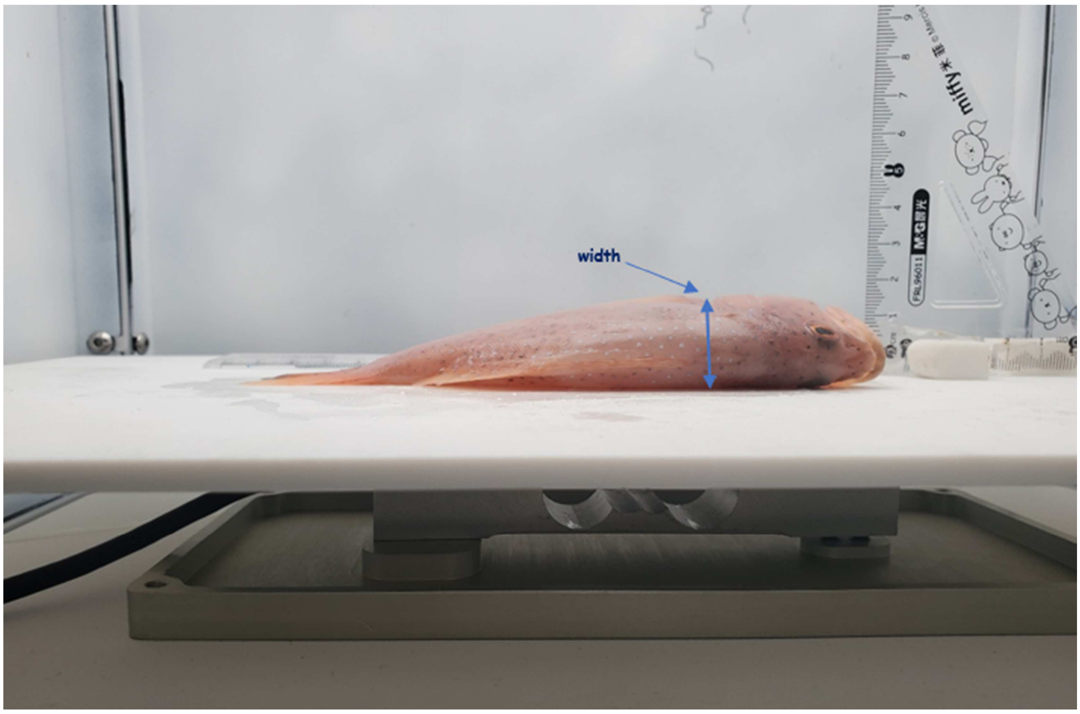
3.1. Regression Results

Import Data: In this study, deep learning techniques are applied to predict and classify different the features of the fish, with a particular focus on Plectropomus leopardus. The dataset used in this research was obtained from the Ocean University of China's laboratory, which uses specialized equipment to measure various fish sizes. The collected dataset consists of 147 Plectropomus leopardus fish, with different sizes, where each fish's size is measured using various physical attributes such as body width, height, body length, along with the fish's weight.

To capture high-resolution images of each fish, a digital camera was utilized, taking multiple images of the fish. The measurement of each fish was taken manually, using a ruler, as shown in Figure 4.



(a) The top view measurement body length, and head length



(b) The side view measurement of the body height and width

Figure 3. *The station set up for image acquisition the Plectropomus leopardus fish measurement captured by a digital camera.*

Training The Model: after we cleaned the data and split it into a training set (80%) and a testing set (20%). That mean the code approximately choose spontaneity 117 image fish for training and 30 images fish for testing,

In the training, the code snippet provided demonstrates the implementation of a convolutional neural network (CNN) using TensorFlow and Keras. The model architecture consists of several layers, including convolutional layers, max pooling layers, and dense layers. The CNN is trained using the Adam optimizer and the mean squared error loss function. The model is trained for 100 epochs with a batch size of 1, in the figure 5 curves respectively for the loss and the mean squared, In the graph, the blue line represents the loss, while the orange line represents the mean square error for the training dataset. Both lines exhibit an exponential decrease as the number of epochs increases from 0 to 100. This pattern is indicative of a well-fitted loss curve, suggesting that the model is performing effectively.

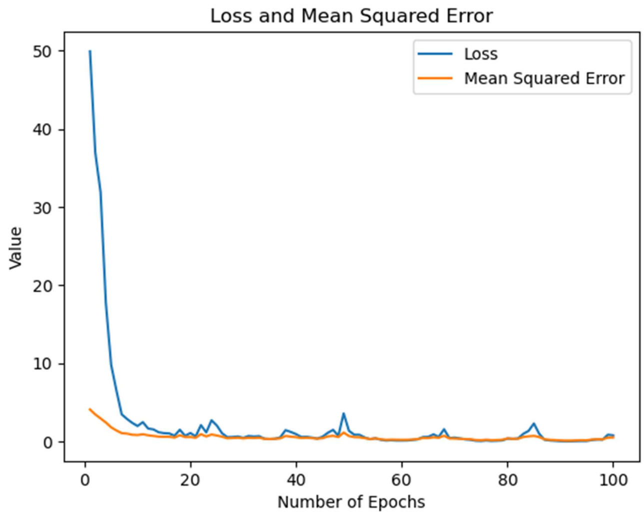


Figure 4. *Loss and mean square error for the training.*

Prediction:

As part of our experimental analysis the predictions were obtained using the 'model. predict' function on both datasets, generating 'y_pred_test' for the test dataset and 'y_pred_train' for the training dataset.

Where in the training data line is on the best fit line, and for the testing data line the majority are on the fit line.

Values and the predicted values for the training and testing sets

Evaluation:

In the evaluation we assessed the performance of a machine learning model in predicting body dimensions and weight based on four key evaluation metrics: Mean Absolute Error (MAE), Root Mean Absolute Error (RMAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). The table presented the results for each metric on both the training and testing datasets. The values in the table were expressed as percentages to provide a clearer understanding of the model's accuracy.

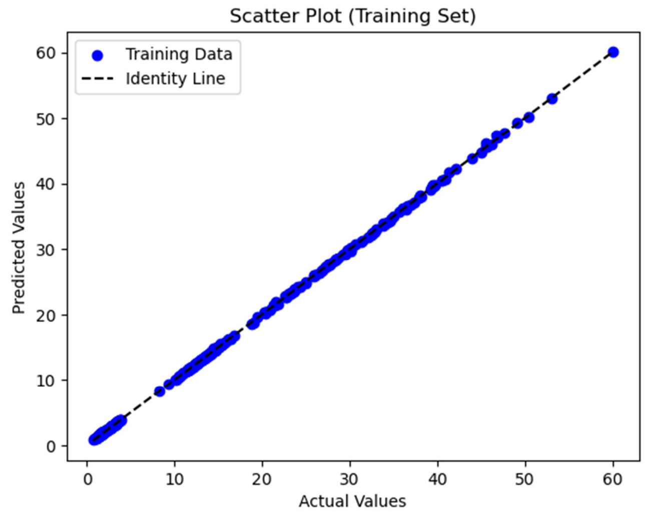


Figure 5. *The comparison between the actual.*

Table 1. *The results for each evaluation metric on both the training and testing datasets.*

Metric	Training Dataset	Testing Dataset
Mean Absolute Error (MAE)	0.152	0.911
Root Mean Absolute Error (RMAE)	0.390	0.954
Mean Squared Error (MSE)	0.0383	0.311
Root Mean Squared Error (RMSE)	0.196	0.427

Model Comparison:

The R2 Score:

The R-squared (R2) scores obtained from our experiments serve as crucial metrics to assess the model's accuracy and overall performance. [14]

For the training set, we achieved an impressive R^2 score of 0.961, indicating that approximately 96.1% of the variance in the dependent variable can be effectively predicted by our model.

However, when evaluated on the testing set, the model's performance showed a slightly lower R^2 score of 0.668. While still promising, this R^2 score indicates that the model's predictions on unseen data might have a relatively larger variance from the actual values compared to the training set.

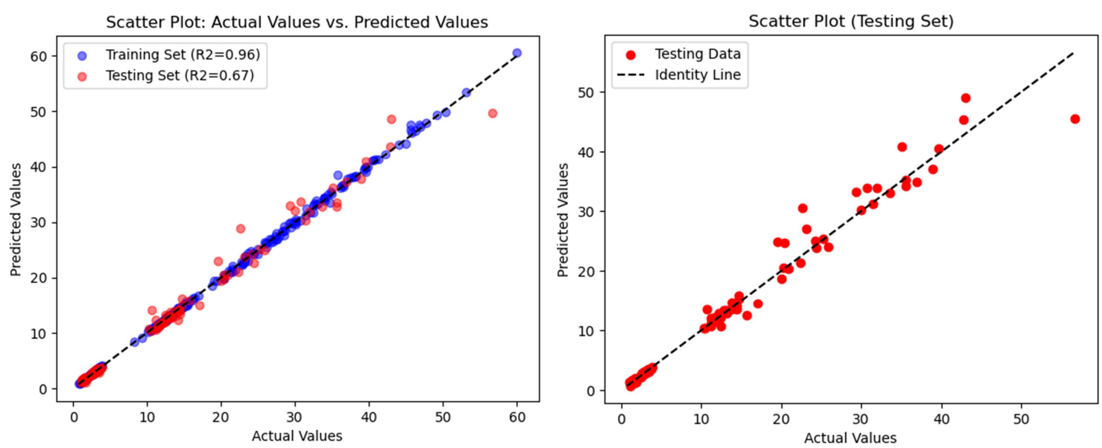


Figure 6. *Scatter Plot of Predicted vs. Actual Values with R-squared Accuracy.*

3.2. Classification Results

Preparing Data: In the process of preparing the data for my experiment, a crucial step was to calculate the size of each fish based on its body length and weight. To accomplish this, I employed a simple yet effective mathematical relationship, where the size (L/W) was determined by dividing the fish's body length (L) by its weight (W). By performing this calculation for all the fish in the dataset, I obtained a size distribution that allowed me to gain valuable insights into the population.

Upon analyzing the data, I observed that the median body length was 12.98 cm, corresponding to a weight of 29.2 g. Using these values, I calculated the median size of the fish to be approximately 0.445 (L/W ratio). With this information at hand, I made an informed decision to classify the fish into two distinct size categories: small and medium.

I established a threshold of $L/W <= 0.45$ to designate a fish as medium, while any fish with $L/W > 0.45$ was classified as small. This classification strategy was chosen based on the unique characteristics of the fish population under study, where the maximum observed body length was 17.11 cm, and the maximum weight was 60 g. Due to the limited size range of the fish in the experiment, a classification of "large" was deemed inappropriate, as there were no specimens that exceeded the chosen threshold.

In our endeavor to analyze the body shape of each fish, we employ a method centered around the interplay of body width and height. By considering the dimension perpendicular to the body length as the body width and the dimension from the bottom to the top of the fish's body as the height, we can derive valuable insights into the fish's morphology.

Our classification system involves categorizing fish into two distinct groups: "Narrow and tall" and "Wide and short." The differentiation between these groups is based on a critical threshold: if a fish's body width is less than 50% of its height, it falls into the "Narrow and tall" category. On the other hand, if the body width is equal to or exceeds 50% of its height, the fish is labeled as "Wide and short."

In the combination, we applied a comprehensive approach that combines the assessment of both the shape and size of the fish, for the purpose of classification. In the 'Combine' column, we have four unique labels with their corresponding counts:

1. SWS (Small and Wide Short): 41 occurrences.
2. MWS (Medium and Wide Short): 38 occurrences.
3. MNT (Medium and Narrow Tall): 35 occurrences.
4. SNT (Small and Narrow Tall): 33 occurrences.

These labels represent the combined classifications based on both fish shape and size. The counts indicate the frequency of each classification within the study population, providing valuable insights into the distribution of fish with different body shapes and sizes.

Analysis the labels of our classification:

The histograms provide a visual representation of the distribution of values for each feature, allowing us to observe patterns, skewness, and central tendencies in the data:

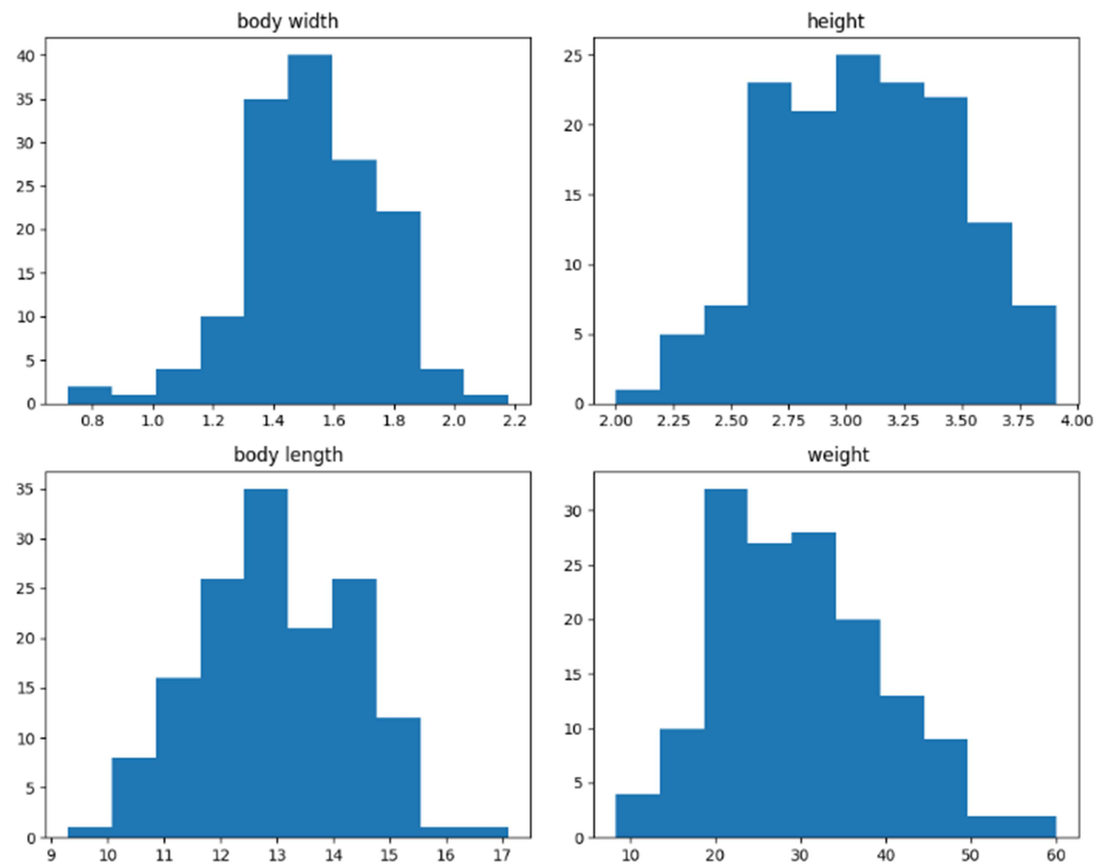


Figure 7. *Visual representation of the distribution of values for each feature.*

Creating a correlation matrix to visualize the relationships between features.

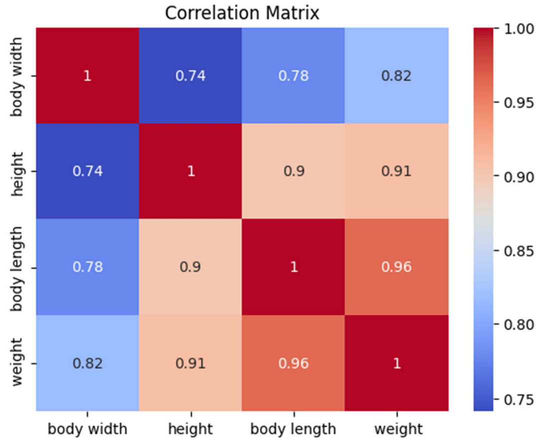


Figure 8. *Correlation matrix to visualize the relationships between features.*

The heatmap provides a visual representation of the correlations between pairs of numerical features in our data. Positive correlations are represented by warmer colors (red), while negative correlations are represented by cooler colors (blue). The numerical values displayed within the heatmap cells indicate the strength of the correlation.

Creating scatter plots to observe relationships between pairs of features the scatter plots provided a clear and intuitive representation of the relationships between features such as body width, height, body length, and fish weight. Each data point on the plot represented an individual fish, and the hue parameter was used to color-code the data points based on our combined classification of fish size and shape, and let us to understand how changes in one feature may relate to changes in another, shedding light on the interplay between body dimensions and weight.

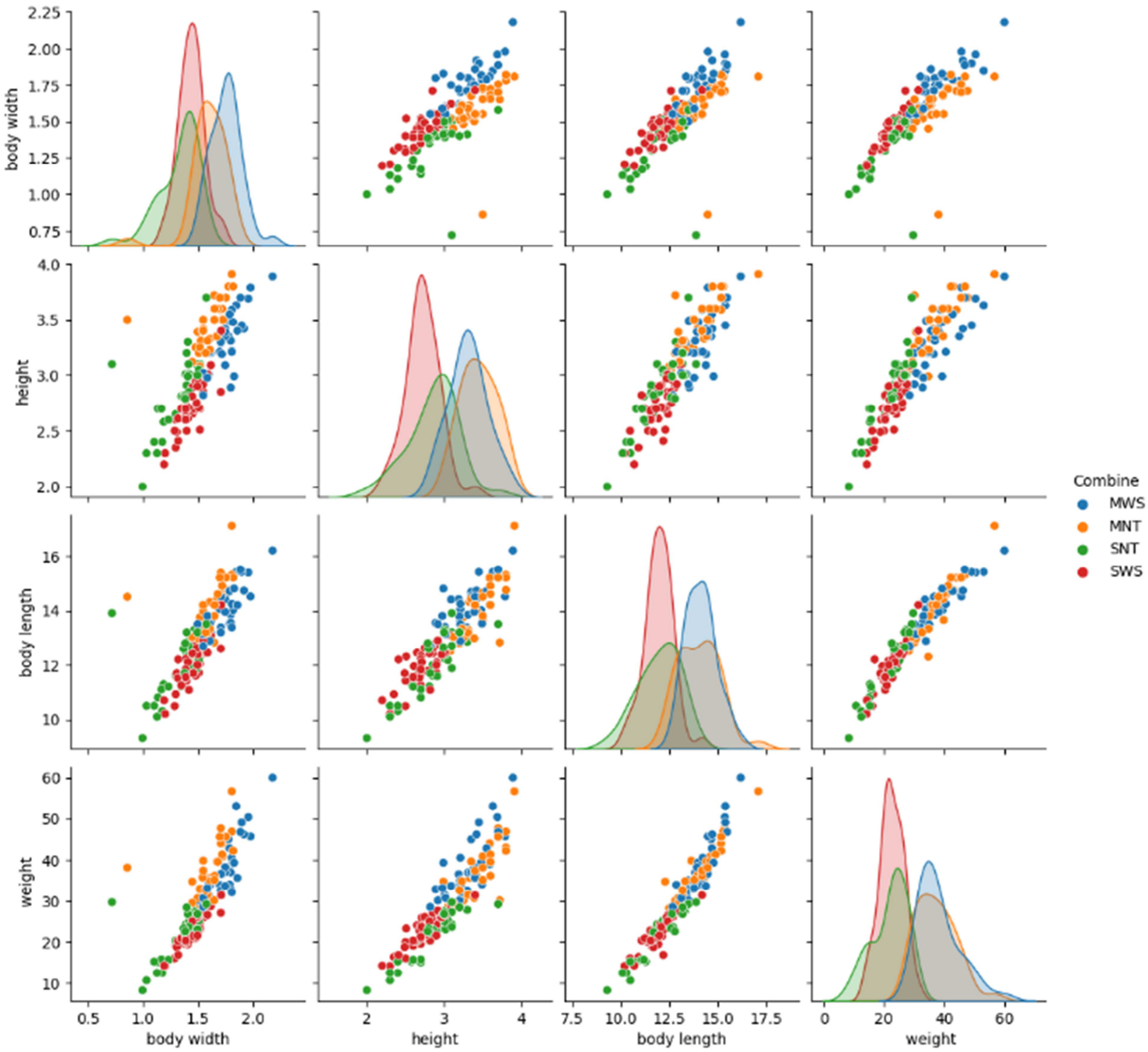


Figure 9. *Scatter plots to observe relationships between pairs of features.*

Creating box plots for each numerical feature

This plot allows to visually compare the distributions and variabilities of different numerical features in our dataset, which can help identify potential patterns, differences, or outliers in the data.

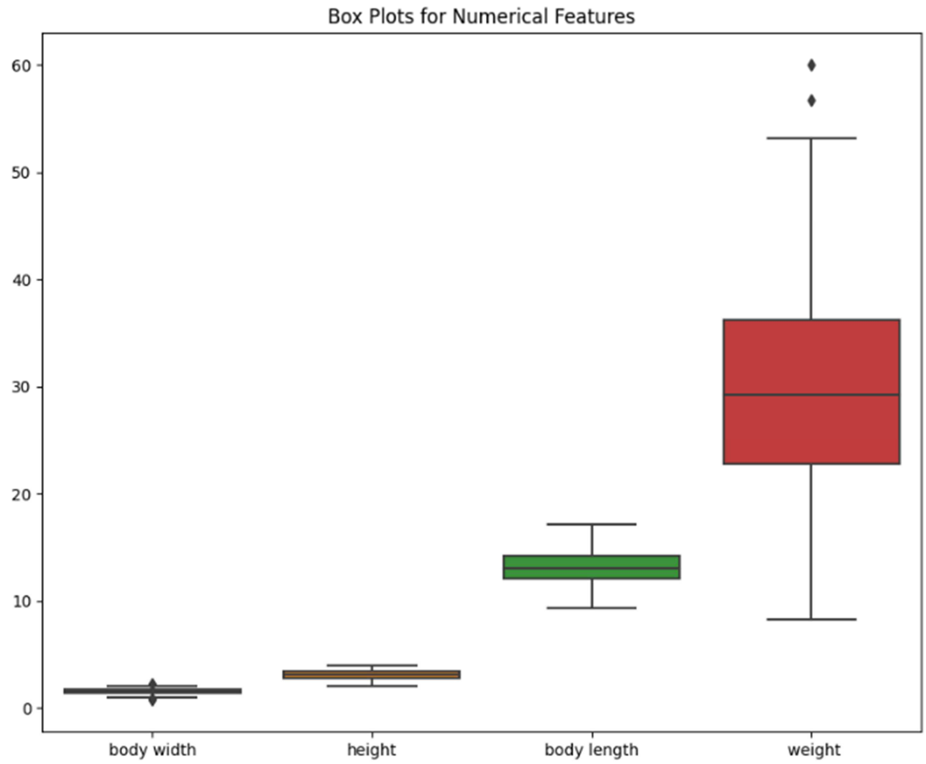


Figure 10. *Box plots for each numerical feature.*

Creating a count plot to visualize the distribution of labels

By creating the count plot using Seaborn in Python, we could visually examine the proportions of each label and understand their representation.

The count plot, with the x-axis denoting the "Combine" column, effectively showcased the number of occurrences for each label category. Through this visualization, we were able to observe the frequency of each fish size and shape classification, providing us with a clear understanding of the dataset's composition.

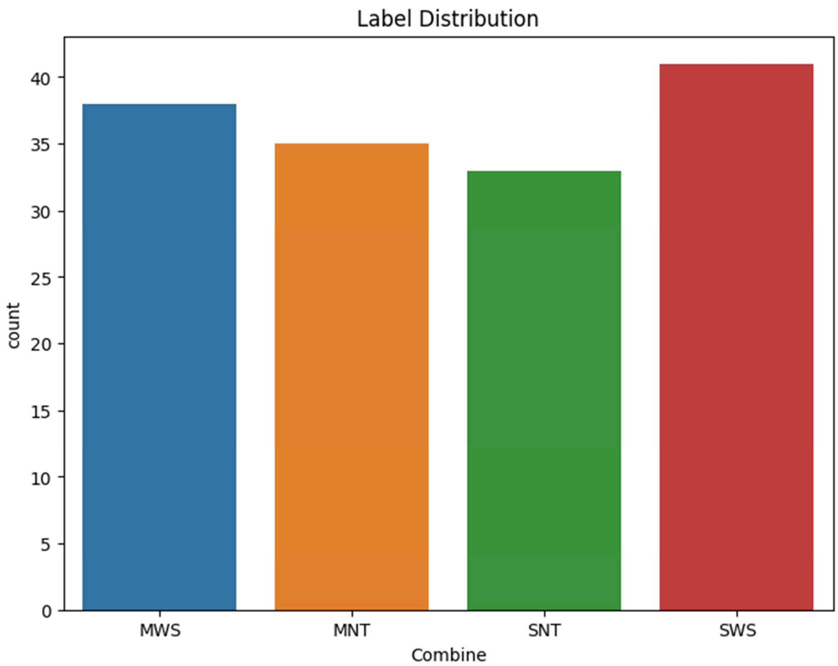


Figure 11. *Count plot to visualize the distribution of labels.*

Training the models:

In this section we performed a comparative analysis of various Machine Learning classifiers to predict the target variable in the dataset. The classifiers used in the experiment include Random Forest Classifier, K-Neighbors Classifier, SVC, Logistic Regression, Decision Tree Classifier, and MLP Classifier. The dataset underwent preprocessing, which involved encoding the target variable and scaling the features, prior to its division into training and testing sets.

Upon evaluating the classifiers, we observed distinct performance characteristics:

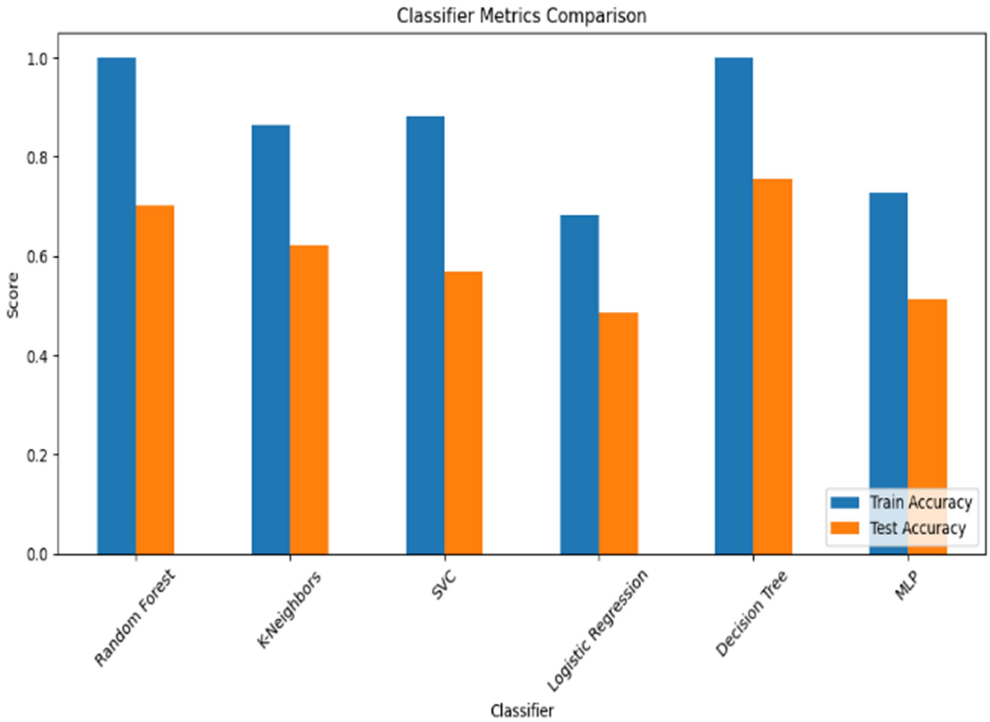


Figure 12. *Test and train accuracy for each classifier algorithm.*

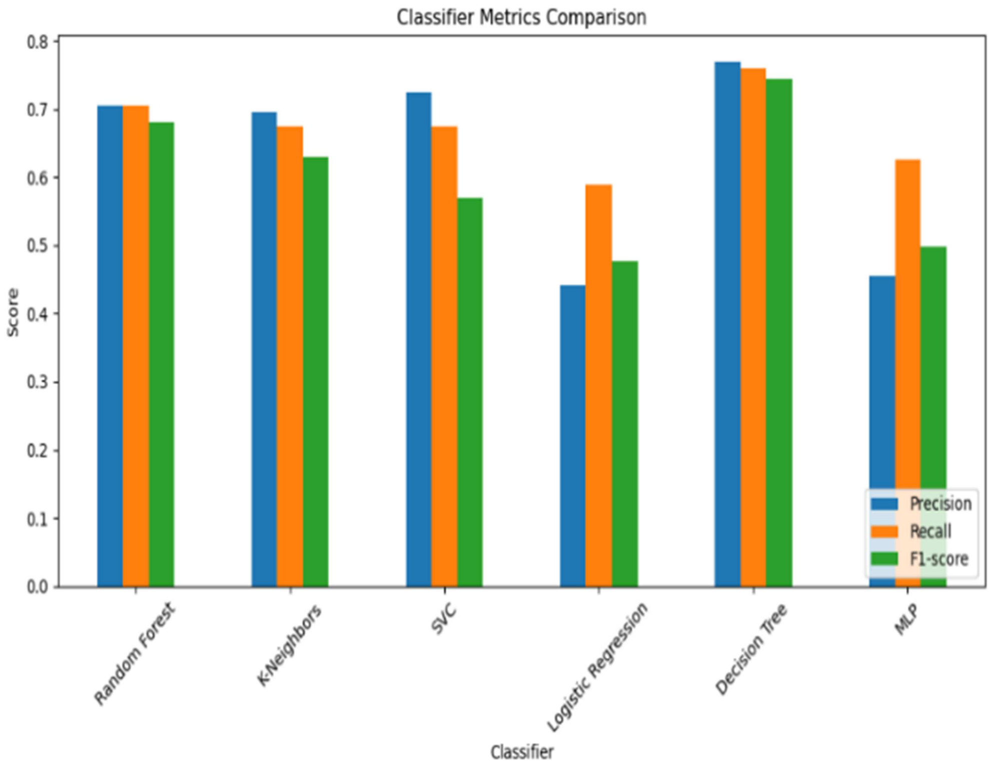


Figure 13. *The evaluation the metrics of the classifiers RF, KN, SVC, LR, Decision Tree, MLP.*

Table 2. *The train accuracy, test accuracy, precision, recall, and F1-score for each classifier.*

Classifier	Train Accuracy	Test Accuracy	Precision	Recall	F1-score
Random Forest	1.00	0.703	0.705	0.705	0.680
K-Neighbors	0.864	0.622	0.695	0.674	0.629
SVC	0.882	0.568	0.724	0.674	0.569
Logistic Regression	0.682	0.487	0.442	0.589	0.476
Decision Tree	1.00	0.757	0.769	0.759	0.745
MLP	0.727	0.514	0.455	0.625	0.498

Ensemble learning Stacking and the gird search:

We perform an optimization process to find the best combination of base classifiers and meta learner using the stacking technique. It uses a dataset containing cleaned fish data and applies feature scaling. The base classifiers considered are Random Forest, K-Nearest Neighbors, Support Vector Machine, Logistic Regression, Decision Tree, and MLP.

Then iterates over all possible combinations of classifiers, starting from combinations of two classifiers. For each combination, it creates a stacking classifier and performs a grid search to find the best meta learner. The stacking classifier is then trained and evaluated on the test set, and the accuracy is calculated. It saves the combination details, including the classifiers combination, meta learner, and accuracy, in a dataframe. The dataframe is sorted in descending order based on accuracy. Finally, the best combination of classifiers, best meta learner, and their corresponding accuracy are printed, along with the execution time. The dataframe of combination details is also displayed.

The best combination of classifiers and the optimal meta learner, Random Forest Classifier, which achieved the highest accuracy of 0.864865. This finding demonstrates the potential of ensemble learning with stacking to improve classification performance by combining the strengths of different classifiers.

The results revealed that:

1. The accuracy values range from 0.729730 to 0.864865.
2. Random Forest Classifier is the chosen meta learner for most combinations, showing its effectiveness in improving accuracy.
3. Combinations that include Random Forest Classifier as the meta learner generally achieve higher accuracy scores compared to other meta learners.
4. Combinations that include multiple base classifiers such as Random Forest Classifier, K-Neighbors Classifier, Support Vector Classifier (SVC), Logistic Regression, Decision Tree Classifier, and MLP Classifier tend to achieve higher accuracy.
5. Some combinations perform similarly, indicating that using a subset of the base classifiers can yield comparable accuracy.

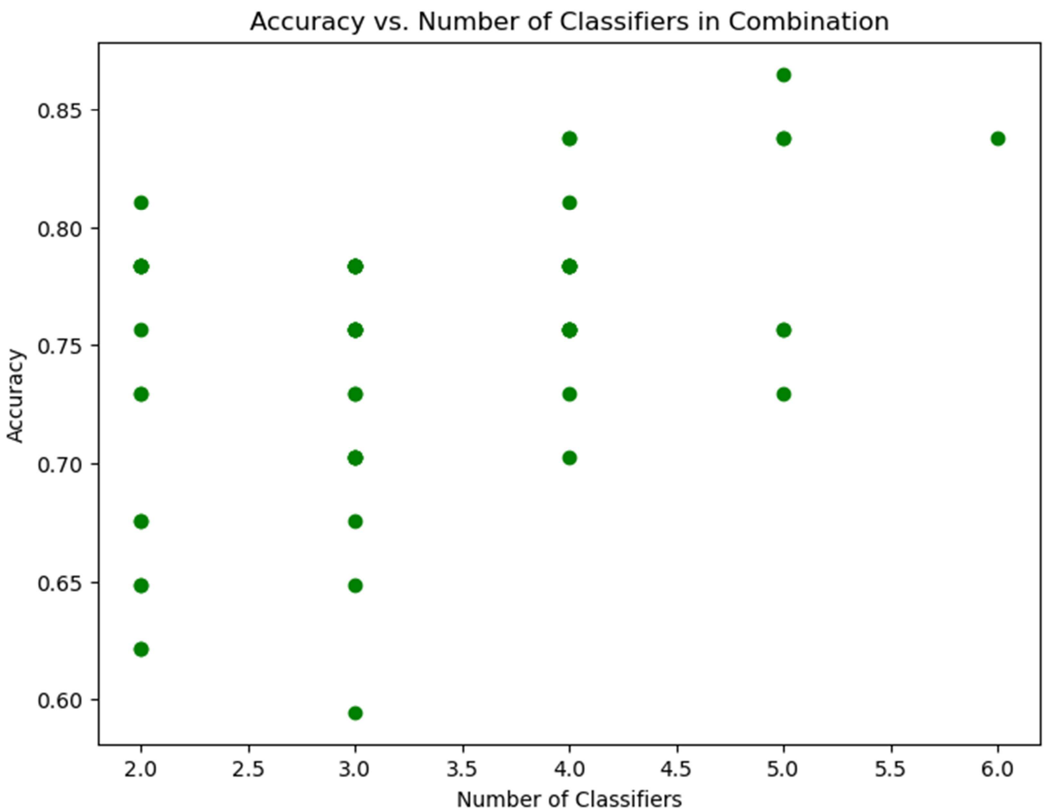


Figure 14. *Number of combination classifiers Vs Accuracy.*

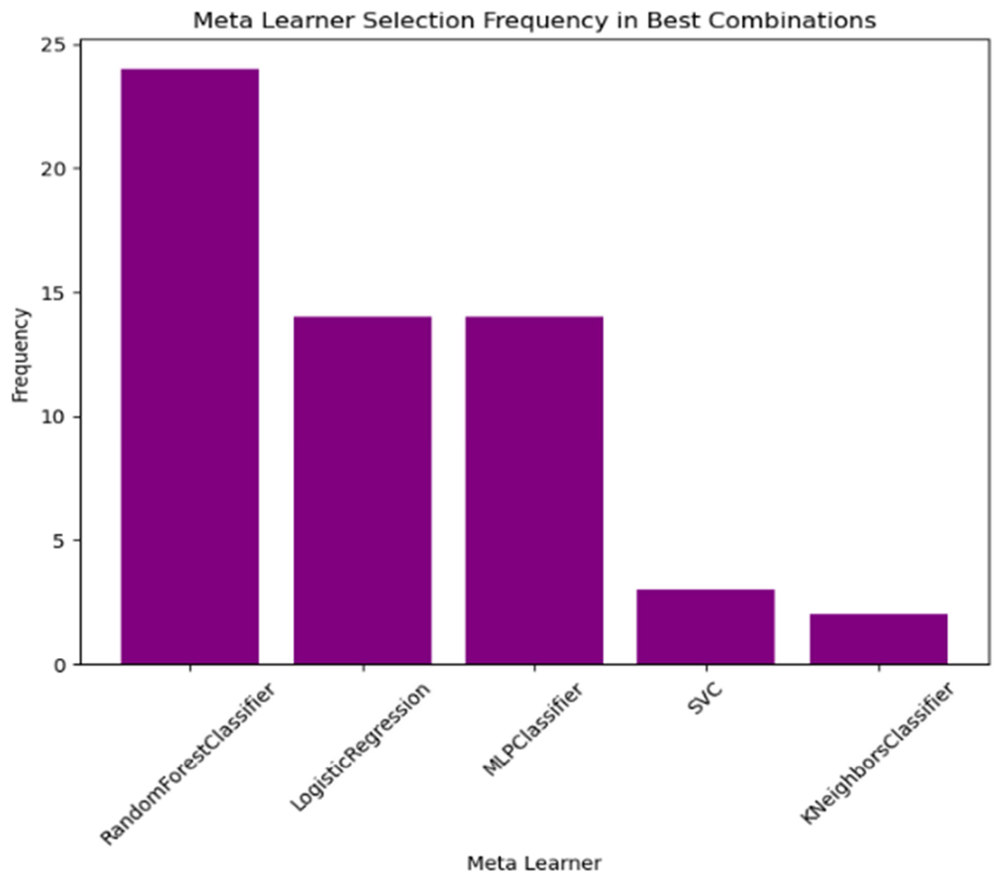


Figure 15. *Meta Learner selection Frequency in best combination.*

Train the best classifier-combination and the best meta learner found using Grid

The ensemble model consisted of five diverse base classifiers: Random Forest, K-Nearest Neighbors, Support Vector Machine, Logistic Regression, and Multi-Layer Perceptron (MLP). To ensure fair comparison and better generalization, we applied feature scaling using Min-Max scaling to bring all features to a similar range.

We used the Stacking Classifier from scikit-learn to combine the predictions of the base classifiers. The final meta-learner, which is a Random Forest classifier, was used to make the ultimate decision based on the outputs of the base classifiers.

The results of the experiment were promising, with the Stacking Classifier achieving remarkable performance on both the training and test datasets. On the training set, the model achieved an accuracy of approximately 94.55%, indicating its capability to effectively fit the training data. Moreover, on the previously unseen test set, the model demonstrated good generalization, achieving an accuracy of approximately 86.49%.

Table 3. *Train and test accuracy results.*

Accuracy (%)	
Train	94.55%
Test	86.49%

Classification Report

The classification report provides information about precision, recall, F1-score, and support for each class, as well as overall performance metrics for the model.

To classify fish features into four categories: SWS, MWS, MNT, and SNT. The classification report provides detailed information about the model's performance on the test dataset, as representing in the table 4 and the figure 16:

Table 4. *Precision, recall, and F1-score for each class.*

Classes	Precision	Recall	F1-score
SWS	0.8571	0.7500	0.8000
MWS	0.7778	1.0000	0.8750
MNT	1.0000	0.8571	0.9231
SNT	0.7778	0.8750	0.8235

Where:

SWS: small, wide, short

MWS: medium, wide, short

MNT medium, narrow, tall

SNT: short, narrow, tall

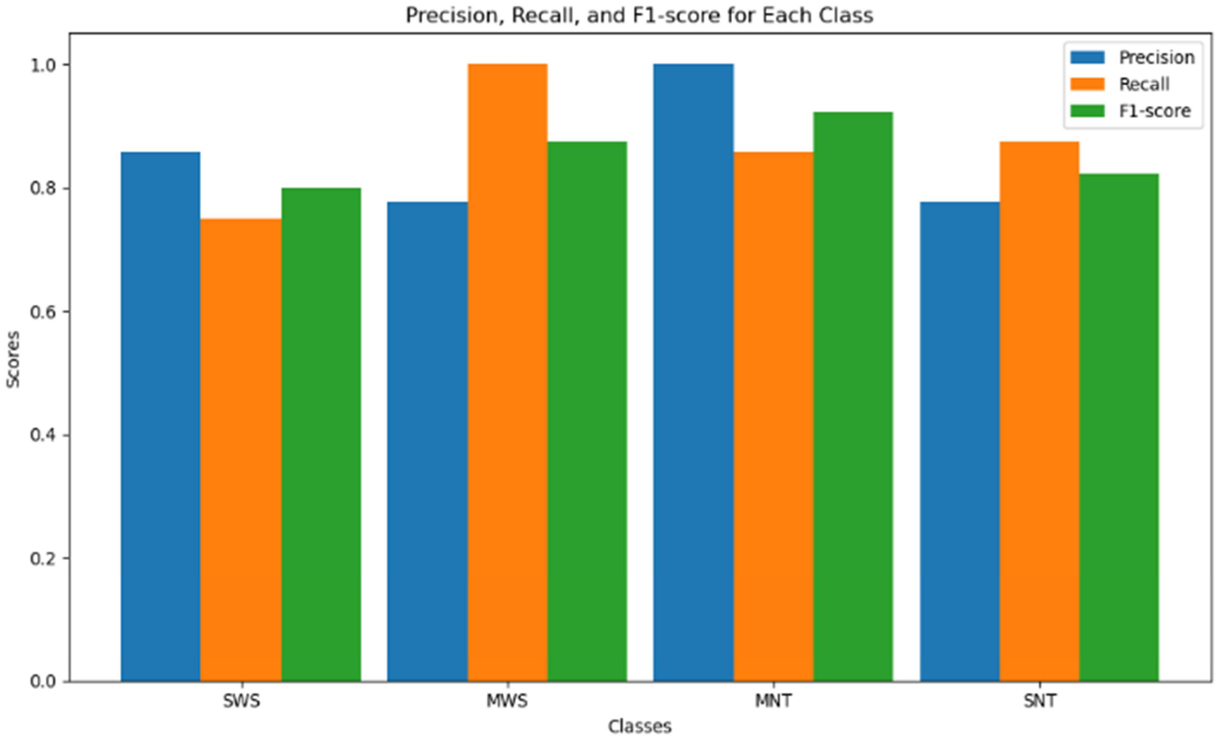


Figure 16. *Precision, recall, and F1-score for each class.*

Confusion Matrix

Diagonal elements, characterized by higher values, signify accurate predictions, where the model correctly classified instances. In contrast, non-diagonal elements indicate misclassifications. Looking at the given matrix, we can discern that our model struggled slightly with the second class, as indicated by the (2, 1) entry with a value of 1, signifying that one instance of the second class was predicted as the first class. Similarly, there is a single misclassification of the fourth class as the second class, denoted by the (4, 2) entry. On the other hand, the majority of instances have been accurately classified.

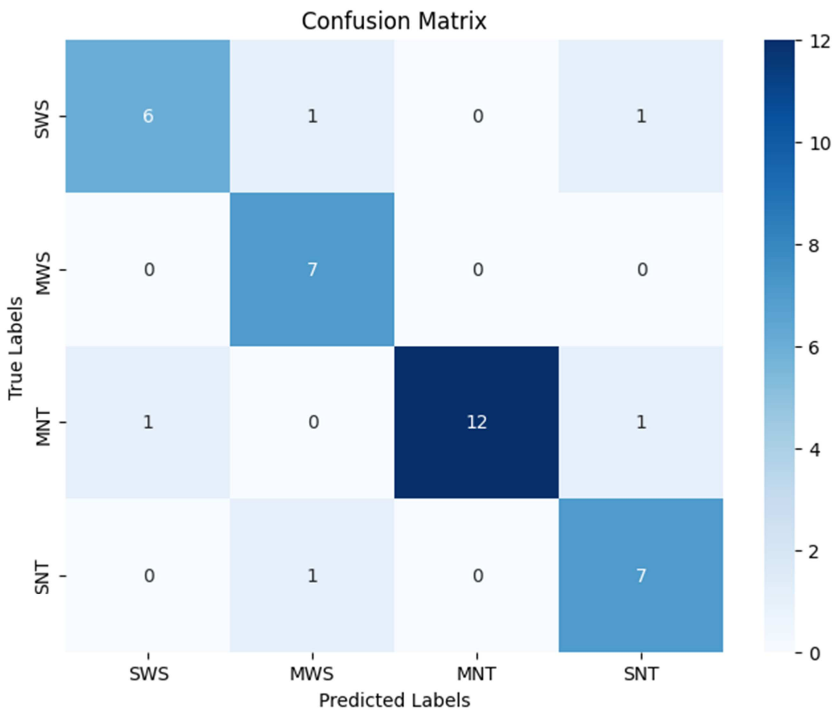


Figure 17. *Provides a visual representation of the distribution and patterns in the classification results.*

Class-wise Accuracy:

The class-wise accuracy bar plot provides a comprehensive insight into the performance of the stacking ensemble classifier for each individual class within our classification problem. Each bar on the plot represents a specific class label, ranging from "SWS" to "SNT", allowing us to assess the accuracy achieved by the classifier on each class.

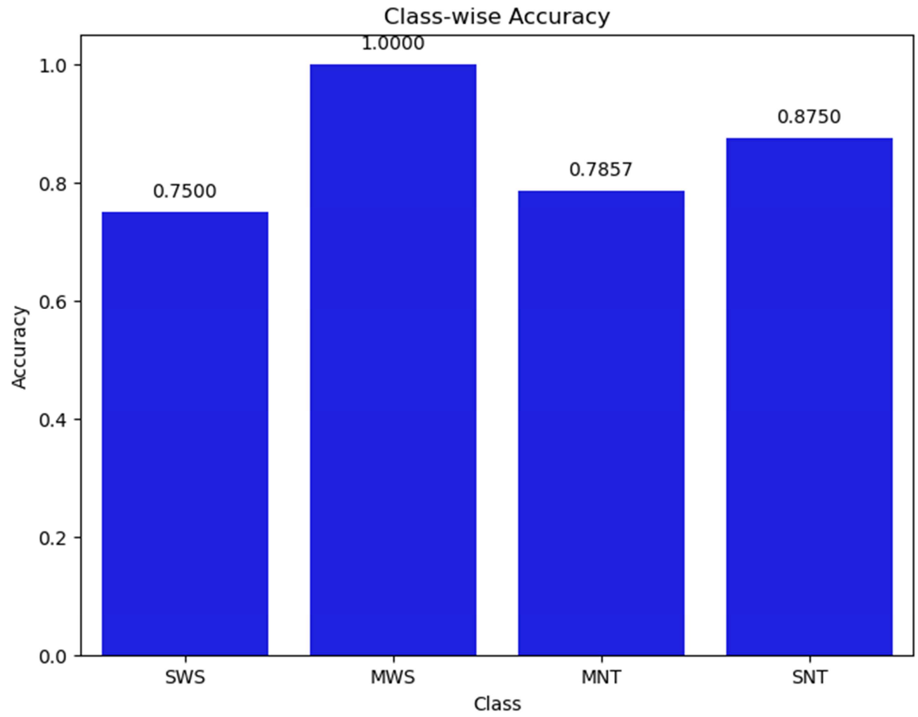


Figure 18. *Class wise accuracy.*

Micro-Averaged ROC Curve

The plot shows the micro-averaged ROC curve, which represents the overall performance of the classifier across all classes. On the y-axis, we plot the True Positive Rate (TPR), while the x-axis represents the False Positive Rate (FPR). The area under the ROC curve (AUC) serves as a single metric for assessing the classifier's performance, with a higher AUC indicating superior performance

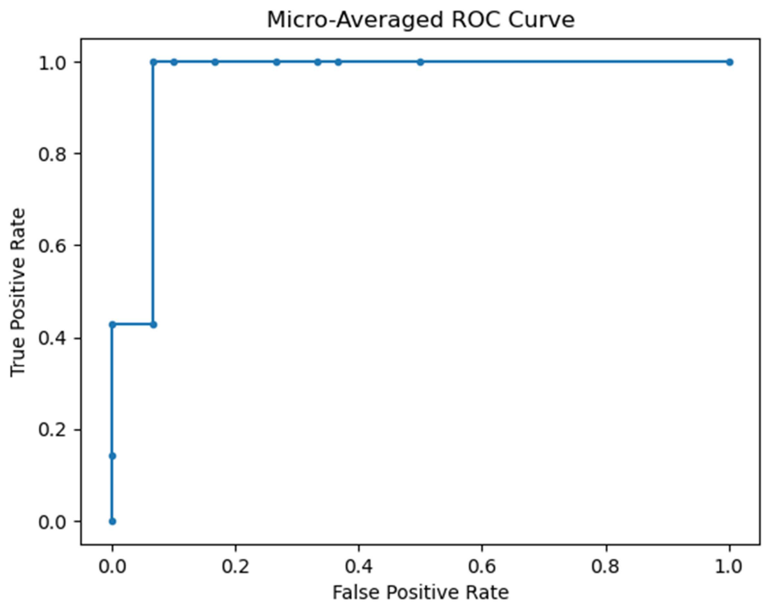


Figure 19. *Micro-Averaged ROC curve.*

The ROC Curves for each class:

This plot allows to compare the performance of the classifier for different classes. The dashed line represents the ROC curve for a random classifier.

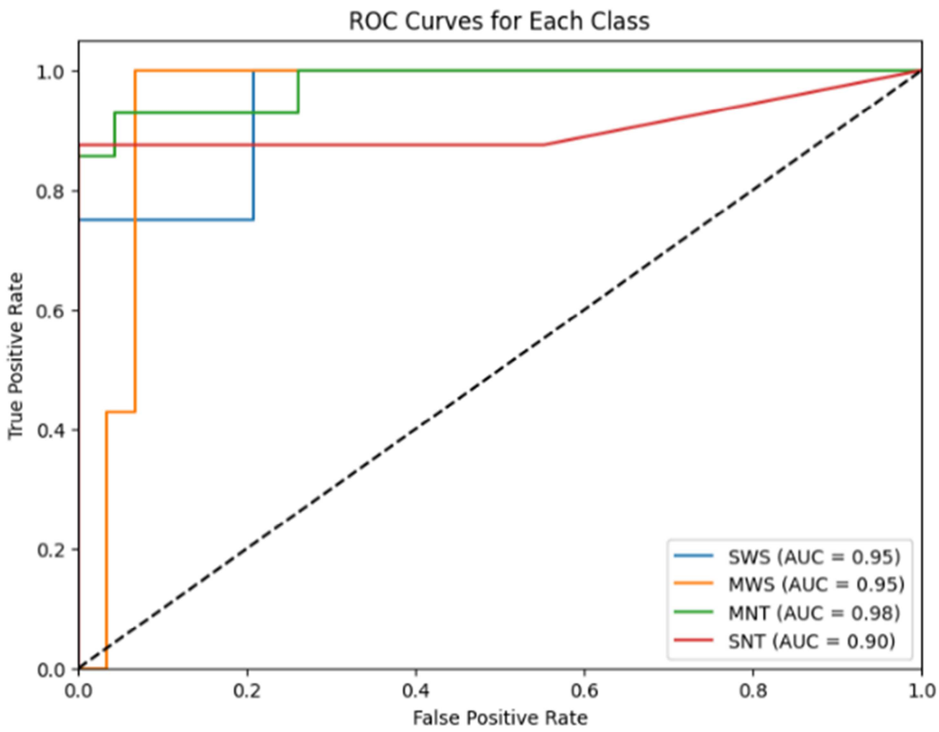


Figure 20. *The ROC curves for each class, along with the corresponding AUC values.*

3.3. Integration

In the final phase of this study, we seamlessly integrated regression and classification techniques to create a comprehensive solution for fish analysis and classification based on images. Leveraging the power of image processing and machine learning, we developed a pipeline that predicts key fish features from images and subsequently classifies them into distinct categories.

The process begins by utilizing image processing libraries to open and preprocess the fish image. After resizing and normalization, the image is reshaped to match the model's input dimensions. Employing a pre-trained model, we predict the fish's height, body width, body length, and weight from the image data. This regression step allows us to extract quantitative characteristics from the visual data, providing valuable insights into the fish's physical attributes.

With the predicted features in hand, we seamlessly transition to the classification stage. Utilizing a previously trained Random Forest Classifier, we predict the fish's category based on its extracted features. By integrating these prediction steps, we've constructed a holistic approach that combines regression and classification to offer a comprehensive understanding of fish characteristics.

4. Conclusion

The classification segment of this study introduced a novel approach to fish classification, taking into consideration both shape and size attributes. By implementing a combination of classifiers and employing ensemble learning with stacking, we achieved remarkable accuracy in identifying distinct fish classes. The integration of these techniques enabled a more nuanced classification process, providing insights beyond simple shape analysis.

Acknowledgments

This study was supported by the Natural Science Foun-dation of China under Grant 51875113, Natural Science Joint Guidance Foundation of the Heilongjiang Province of China under Grant LH2019E027. This paper is funded by the International Exchange Program of Harbin Engineering University for Innovation-oriented Talents Cultivation.

Conflicts of Interest

All the authors do not have any possible conflicts of interest.

References

[1] Akhtar Jamil, M. A. S. (2023). "Stock Price Prediction in Response to US Dollar Exchange Rate Using Machine Learning Techniques." ResearchGate.

[2] al, A. e. (2021). "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions." journal of Big Data.

[3] Barbedo, J. G. A. (2022). "A Review on the Use of Computer Vision and Artificial Intelligence for Fish Recognition, Monitoring, and Management." Fishes.

[4] Boopathi, S., et al (2023). Advances in Artificial Intelligence for Image Processing: Techniques, Applications, and Optimization. Handbook of Research on Thrust Technologies' Effect on Image Processing. e. a. Binay Kumar Pandey: 73-95.

[5] Dertat, A. "Applied Deep Learning - Part 4: Convolutional Neural Networks." From https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2.

[6] Dey, R., Mathur, R (2023). "Ensemble Learning Method Using Stacking with Base Learner, A Comparison." Springer vol 727.

[7] Guduru Dhanush, N. K., Sandeep Kumar, Praveen Kumar Shukla (2023). "A comprehensive review of machine vision systems and artificial intelligence algorithms for the detection and harvesting of agricultural produce." Scientific African.

[8] Javad Ansarifar, L. W. S. V. A. (2021). An interaction regression model for crop yield prediction.

[9] Li, D., Wang, Q., Li, X., Niu, M., Wang, H., and Liu, C (2022). "Recent advances of machine vision technology in fish classification." ICES Journal of Marine Science: 263–284.

[10] Minggang Zhou, P. S., Hao Zhu and Yang Shen (2023). "In-Water Fish Body-Length Measurement System Based on Stereo Vision." sensors.

[11] Mouhafid M, C. Y. (2022). Diagnosing COVID-19 from CT Images Based on Ensemble Learning and Transfer Learning, HEBEI UNIVERSITY OF TECHNOLOGY. Master Degree.

[12] Muntean, M., Militaru, FD (2023). "Metrics for Evaluating Classification Algorithms." Springer vol 321.

[13] Mutasem K. Alsmadi, I. A. (2020). "A survey on fish classification techniques." Journal of King Saud University – Computer and Information Sciences.

[14] Onyutha, C. (2020). "From R-squared to coefficient of model accuracy for assessing "goodness-of-fits"."

[15] scikit-learn. "Neural network models (supervised)." from https://scikit-learn.org/stable/modules/neural_networks_supervised.html.

[16] Sumei Li, Y. L., Yongtian Han (2021). "Stereoscopic image quality assessment considering visual mechanism and multi-loss constraints." Journal of Visual Communication and Image Representation.

[17] Taylor, S. (2022). "Regression Analysis." from https://corporatefinanceinstitute.com/resources/data-science/regression-analysis/.