# Image-Based Classification of Freshwater Fish Species to Support Feed Recommendation Using Random Forest

Hindayati Mustafidah[1*], Suwarsito[2], Rahmat Setiawan[3], Abdul Karim[4]

[1,3] *Informatics Engineering, Universitas Muhammadiyah Purwokerto, Indonesia*
[2] *Aquaculture, Universitas Muhammadiyah Purwokerto, Indonesia*
[4] *Department of Artificial Intelligence Convergence, Hallym University, Republic of Korea*
*corr-author: h.mustafidah@ump.ac.id

**Abstract—Accurate identification of freshwater fish species plays a vital role in aquaculture, particularly in determining appropriate feed strategies to optimize fish growth. Visual similarities among species—such as color, shape, and surface texture—often hinder novice farmers from correctly recognizing fish types. This study proposes an image-based classification system using the Random Forest algorithm to identify six freshwater fish species: pomfret (*bawal*), gourami (*gurame*), catfish (*lele*), barb (*melem*), tilapia (*nila*), and Java barb (*tawes*) and provide automated feed recommendations. A total of 120 fish images were used as the dataset, collected from various sources, including online repositories and field documentation. Feature extraction was applied to capture color characteristics (HSV), texture patterns (GLCM), and morphological features (regionprops). The model was trained on 70% of the dataset and tested on the remaining 30%. Evaluation results show that the system achieved a classification accuracy of 83.33%, with a precision of 83.53%, recall of 83.33%, and an F1-score of 82.86%. Notably, catfish, barb, and tilapia classes achieved perfect classification, while pomfret and gourami showed room for improvement due to overlapping visual features. The findings indicate that the integration of Random Forest with multi-domain image features offers an effective, affordable, and practical solution to support the digital transformation of small and medium scale aquaculture systems through intelligent species recognition and feed guidance.**

**Keywords: freshwater fish classification; image-based species recognition; random forest; intelligent feeding system; digital aquaculture**

## I. INTRODUCTION

Accurate identification of freshwater fish species is vital for achieving efficient and sustainable aquaculture, as it directly impacts feed management, fish health, growth performance, and overall operational outcomes. However, manual identification—particularly by less experienced farmers—is frequently unreliable due to morphological similarities among species, such as body shape, coloration, and fin structure, thereby necessitating automated, technology-driven solutions. In Indonesia, commonly farmed freshwater species include catfish (*lele*), tilapia (*nila*), pomfret (*bawal*), gourami (*gurame*), barb (*melem* or *nilem*), and Java barb (*tawes*), typically found in low-salinity environments such as rivers and lakes [1]. Although these species possess distinctive traits—like barbels in catfish, dorsal spines in tilapia, and oval body shape in pomfret—their visual similarities often lead to misclassification, underscoring the urgency of developing accurate and automated fish classification systems. Furthermore, precise species classification is crucial for implementing effective, species-specific feed management strategies, as nutritional requirements vary significantly across fish types. Appropriate feed formulation, tailored to each species' dietary needs, behavioral patterns, and environmental conditions, enhances growth and feed efficiency while minimizing waste, reducing costs, and preventing disease [2] [3]. Thus, integrating automated species recognition with customized feeding strategies offers a promising approach to improving productivity and sustainability, particularly in small- and medium-scale aquaculture systems in resource-constrained environments.

The advancement of artificial intelligence (AI), particularly in computer vision, has significantly enhanced automated visual classification, with the integration of digital image processing and machine learning (ML) enabling accurate object recognition under diverse conditions [4] [5]. These technologies are especially valuable in agriculture and aquaculture, where efficiency and data-driven management are crucial. In fisheries, ML supports sustainability, productivity, and monitoring, with applications including fish catch prediction [6], population modeling [7], and species

recognition based on visual traits [8] [9] [10] [11]. Studies have applied algorithms such as KNN, SVM, and Naïve Bayes—often combined with feature extraction—for image-based fish classification: ORB with KNN [12], fuzzy KNN for milkfish quality [13], SVM for classification in Bangladesh [14], and KNN for freshness detection via eye color [15]. Formalin detection has also been explored using KNN, GLCM, and multilayer perceptrons [16] [17].

Recent studies have extended image-based classification to ornamental and freshwater species, using various approaches such as betta fish identification [18] [19] [20] and marine species classification with SVM, HOG, and HSV features [21]. In freshwater fish, KNN with PCA has improved accuracy [22] [23], while SVM has been applied to African Cichlids [24]. Though deep learning methods like CNNs offer high accuracy, their reliance on large datasets and computing resources limits practical deployment. In contrast, traditional ML methods such as KNN and SVM are more lightweight but face limitations with complex data. RF, an ensemble decision tree algorithm, offers a balance of accuracy, interpretability, and computational efficiency [25] [26]. RF has demonstrated robust performance in various domains [27], including vessel detection [28], plant disease recognition [29] [30], and aquaculture diagnostics [31]. Its strengths—handling mixed data types, resistance to overfitting, and feature importance analysis—make it suitable for fish classification and water quality prediction [32] [33]. Moreover, RF has proven compatible with embedded systems, operating on low-power microcontrollers with minimal accuracy loss [34], [35], and its performance can be further enhanced with techniques like DCP, CLAHE, and GLCM-based features [36] or model tuning [5] [37], establishing it as an efficient and adaptable solution for aquaculture applications.

Based on prior research, this study proposes a freshwater fish species classification system using image analysis and the Random Forest algorithm, incorporating key visual features—texture (GLCM), color (HSV), and shape (Regionprops)—to differentiate commonly farmed species. Beyond classification, the system includes a feed recommendation module tailored to each identified fish type, thereby supporting both recognition and decision-making in aquaculture. Designed to be lightweight and accessible, the system addresses the practical needs of small- to medium-scale fish farmers with limited computing resources. By integrating technical accuracy with functional utility, it contributes to advancing digital aquaculture management [38].

## II. METHOD

### *A. Research Design*

This study adopts a quantitative experimental approach to develop and evaluate an image-based classification system for six freshwater fish species commonly farmed in Indonesia—gourami, catfish, tilapia, barb, Java barb, and pomfret—using the Random Forest (RF) algorithm. A dataset of 120 balanced images (20 per species) was collected for training and evaluation. The system aims to accurately classify fish based on visual features and deliver species-specific feed recommendations, with a focus on creating a lightweight, interpretable, and practical tool for small- to medium-scale aquaculture operations. The development process followed a structured sequence, as depicted in Fig. 1.

### *B. Data Description*

The dataset comprises 120 digital image representing six freshwater fish species commonly farmed in Indonesia: **catfish, tilapia, gourami, pomfret, barb, and Java barb**. Each class consists of 20 labeled images, captured under relatively uniform conditions. The dataset was curated to reflect typical variations in shape, color, and texture for each species while maintaining balanced class distribution. An example image of each of the six fish species is shown in Fig. 2.

In the **preprocessing** phase, all images were resized to a uniform resolution and converted from RGB into grayscale using (1) and HSV formats. Otsu's thresholding was applied to obtain binary segmentation of fish objects using (2), and normalization was performed to standardize pixel intensity distributions across the dataset.

$$Gray(x,y) = \frac{I_R(x,y) + I_G(x,y) + I_B(x,y)}{3} \quad (1)$$

with $Gray(x,y)$ is the grayish pixel value at point $(x,y)$ and $I_{R,}\, I_G, I_B$ is red, green, blue pixel values at point $(x,y)$ respectively, with $Vt(x,y)$ is the binary value at point $(x,y)$ and $T$ being the threshold value.

$$Vt(x,y) = \begin{cases} 1 & if\ Gray(x,y) > T \\ 0 & if\ Gray(x,y) \leq T \end{cases} \quad (2)$$

Next, the **feature extraction** stage was carried out to derive discriminative visual attributes essential for classification. This extraction included (1) **color features** extracted from the HSV color space, (2) **shape features** such as area, orientation, perimeter, major axis length, minor axis length, and eccentricity using the regionprops method, and (3) **texture features** computed via the Gray-Level Co-occurrence Matrix (GLCM), including dissimilarity, correlation, homogeneity, contrast, Angular Second Moment (ASM), and energy. All texture attributes were computed at four orientation angles: 0°, 45°, 90°, and 135°. In total, the texture feature extraction produced 24 data points.

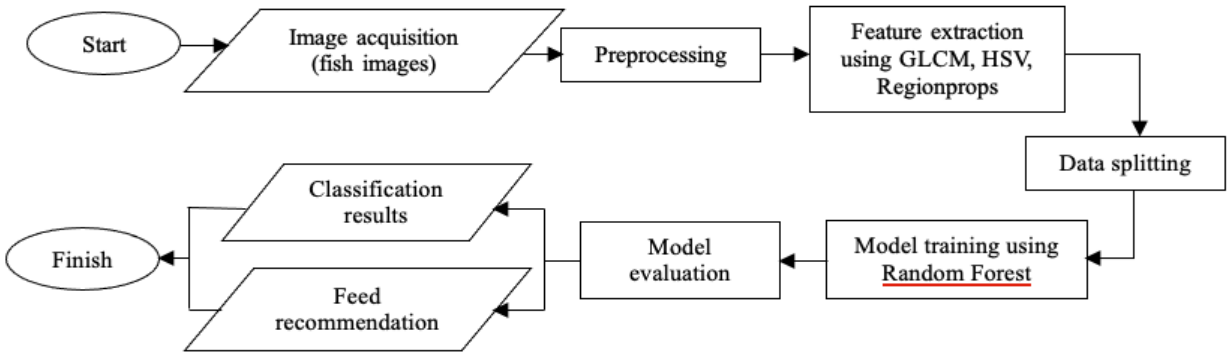



**Fig. 1 Workflow of the proposed image-based freshwater fish classification and feed recommendation model**

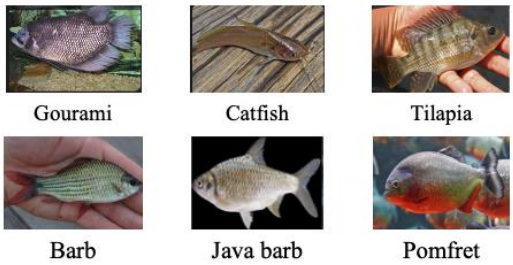



**Fig. 2 Sample images of six freshwater fish species**

*1) Color Feature:* The RGB color images were first converted to the HSV color space, which separates chromatic content (hue and saturation) from intensity (value). This separation provides better robustness to lighting variation. The transformation from RGB to HSV is defined as follows:

Let *R, G,* and $B \in [0,1]$ denote the normalized red, green, and blue values. Then, the HSV components are computed as in (3 - 5) [4]. The average values of the H, S, and V channels were used as color descriptors.

$$V = \max(R, G, B) \tag{3}$$

$$S = \begin{cases} V - \min(R,G,B); & if\ V \neq 0 \\ 0; & otherwise \end{cases} \tag{4}$$

$$H = \begin{cases} 60(G-B)/(V - \min(R,G,B)) & if\ V = R \\ 120 + 60(B-R)/(V - \min(R,G,B)) & if\ V = G \\ 240 + 60(R-G)/(V - \min(R,G,B)) & if\ V = B \end{cases} \tag{5}$$

*2) Shape features:* Shape feature were extracted from the binary images using the regionprops function from the scikit-image library [39]. Six primary geometric descriptors were selected based on their relevance to morphological analysis. The **area A** represents the total number of pixels that constitute the object. It is calculated by summing all pixel values labeled as part of the object (usually binary pixels with a value of 1), where N is the total number of such pixels as in (6) [4]. **Orientation** represents the angle (θ - in degrees) between the horizontal axis and the major axis of the ellipse that has the same second-moments as the region as in (7). The angle value ranges from −90∘−90∘ to +90∘+90∘ and describes how tilted the fish object is in the image. **Perimeter** as the length of the boundary of the object, computed as the total count of pixels outlining the object's shape as in (8) [40]. The perimeter $P$ is computed by summing the Euclidean distances between consecutive boundary pixels of the object. Each boundary point is represented by its coordinates $(x_i, x_j)$ and $M$ is the total number of boundary segments. This calculation approximates the total length around the object's edge. The **major axis length** represents the longest diameter of the ellipse that has the same normalized second central moments as the object as in (9) where $\mu_{20}, \mu_{02}, \mu_{11}$ are second-order central moments. The **minor axis length** as the shortest diameter of the fitted ellipse that reflects the **width** of the object perpendicular to the major axis as in (10). **Eccentricity** ($e$) measures the elongation of the object based on the ratio of the major axis ($a$) and minor axis ($b$) of an ellipse fitted to the object as in (11). A value close to 0 indicates a near-circular object, while a value close to 1 indicates a highly elongated shape [41].

$$A = \sum_{i=1}^{N} 1 \tag{6}$$

$$\theta = \frac{1}{2} tan^{-1}\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \tag{7}$$

$$P = \sum_{i=1}^{M} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \tag{8}$$

$$\text{Major Axis Length} = 2\sqrt{2} \cdot \sqrt{\frac{\mu_{20}+\mu_{02}}{2} + \sqrt{\left(\frac{\mu_{20}-\mu_{02}}{2}\right)^2 + \mu_{11}^2}} \tag{9}$$

$$\text{Minor Axis Length} = 2\sqrt{2} \cdot \sqrt{\frac{\mu_{20}+\mu_{02}}{2} - \sqrt{\left(\frac{\mu_{20}-\mu_{02}}{2}\right)^2 + \mu_{11}^2}} \tag{10}$$

$$e = \sqrt{1 - \left(\frac{b}{a}\right)^2} \quad (11)$$

*3) Texture features:* Texture feature were computed using the Gray-Level Co-occurrence Matrix (GLCM), which analyzes the spatial distribution of gray levels in the image. Six statistical descriptors were derived as in (12 - 17) [42]: dissimilarity, correlation, homogeneity, contrast, ASM, and energy. **Dissimilarity** measures the absolute difference between neighboring gray levels, weighted by their occurrence probability *P(i,j)* as in (12). Higher values indicate greater variation in pixel intensity. **Correlation** evaluates the linear relationship between pixel pairs. It reflects the degree to which one gray level value is linearly related to its neighbor, as in (13), where $\mu_i, \mu_j$ means of row and column indices, and $\sigma_i$ and $\sigma_j$ are the standard deviations of row and column indices. **Homogeneity** assesses the closeness of the distribution of elements in the GLCM to the diagonal, as in (14). Higher values indicate more uniform textures with minimal variation between adjacent pixels. **Contrast** measures the intensity contrast between a pixel and its neighbor, as in (15). It emphasizes texture with sharp differences; higher values suggest high local contrast or roughness. Also referred to as Energy, **Angular Second Moment (ASM)** indicates the uniformity or regularity of texture, as in (16). High ASM values correspond to homogeneous images with repeated patterns. **Energy** is the square root of ASM and similarly reflects the degree of texture uniformity as in (17). It ranges between 0 and 1, with higher values for smoother textures, where $P(i,j)$ is the normalized GLCM value at position $(i,j)$. These features provide important textural information relevant to fish scale patterns and surface irregularities.

$$\text{Dissimilarity} = \sum_{i,j} |i - j| \cdot P(i,j) \quad (12)$$

$$\text{Correlation} = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j) \cdot P(i,j)}{\sigma_i \cdot \sigma_j} \quad (13)$$

$$\text{Homogeneity} = \sum_{i,j} \frac{P(i,j)}{1 + |i - j|} \quad (14)$$

$$\text{Contrast} = \sum_{i,j} (i - j)^2 \cdot P(i,j) \quad (15)$$

$$\text{ASM} = \sum_{i,j} P(i,j)^2 \quad (16)$$

$$\text{Energy} = \sqrt{\sum_{i,j} P(i,j)^2} = \sqrt{\text{ASM}} \quad (17)$$

The data were split into 70% training and 30% testing sets with balanced classes. The Random Forest algorithm, based on [43] bagging method, trained multiple decision trees on random data subsets, with final predictions made through majority voting, with the predicted class $\hat{y}$ for an input $x$ is (18) and $h_t(x)$ denotes the class predicted by the decision tree.

$$\hat{y} = mode\{h_1(x), h_2(x), \ldots, h_T(x)\} \quad (18)$$

*C. Implementation*

The model was implemented in Python, utilizing OpenCV for image preprocessing, scikit-image and NumPy for feature extraction, and scikit-learn for model training and evaluation. All experiments were conducted on a standard laptop with an Intel Core i5 processor and 8 GB RAM.

*D. Evaluation Metrics*

The model's performance was assessed using accuracy, precision, recall, and F1-score [44] [45]. Predicted labels from test images were then used to generate species-specific feed recommendations, supporting practical aquaculture decisions.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (19)$$

$$Precision = \frac{TP}{TP + FP} \quad (20)$$

$$Recall = \frac{TP}{TP + FN} \quad (21)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (22)$$

*E. Feed Recommendation Module*

The feeding practices of freshwater aquaculture species vary according to developmental stages and species-specific requirements [46], [47], [48]. This summary outlines feed allocation across larval, fry, juvenile, and adult stages for six economically important freshwater fish: pomfret, gourami, catfish, barb, tilapia, and Java barb [49], [50].

Freshwater fish like pomfret, gourami, catfish, barb, tilapia, and baung are initially fed natural feeds such as *Tubifex*, *Moina*, *Artemia*, and rotifers, with *Tubifex* showing the best growth performance. As they grow, artificial feeds like pellets (PF500, PF788, PF800), trash fish, shrimp heads, plant matter, and organic waste are introduced. The combination of maggot and *Tubifex* enhances feed efficiency. Natural feeds are especially critical during early stages. Enriched *Moina* and *Tubifex* have achieved high survival rates, including 100% in baung larvae. A gradual transition from natural to formulated feeds, tailored to

species and growth stages, optimizes nutrition, growth, and survival [51].

## III. RESULTS AND DISCUSSION

The results of this study are presented and discussed in accordance with the objectives outlined in the methodology. The primary aim was to develop an image-based classification model for freshwater fish species using the Random Forest algorithm, supported by a rule-based feed recommendation module.

### *A. Dataset Characteristics*

The dataset consisted of 120 images of six commonly farmed freshwater fish species in Indonesia (20 images per class), captured under standardized 5600 K lighting in a controlled mini-studio at a fixed 50 cm distance. Backgrounds were segmented using semi-automatic thresholding and manual refinement to produce clean, white-background images, enabling accurate extraction of color, shape, and texture features for effective classification.

### *B. Preprocessing*

Image preprocessing involved three techniques aligned with specific feature types: grayscale conversion for texture features, Otsu thresholding for shape, and HSV transformation for color. Grayscale conversion, based on (1) and (2) and implemented using OpenCV's cv2.COLOR_BGR2GRAY, produced the grayscale image shown in Fig. 3(b) from the original in Fig. 3(a). Otsu thresholding was then applied to generate a binary image for shape analysis (Fig. 3c). Finally, the original image (Fig. 3a) was converted to HSV color space to facilitate color feature extraction, as illustrated in Fig. 4.

### *C. Feature Extraction*

Following preprocessing, all fish images were standardized to 300×300 pixels and converted to HSV color space. Segmentation using Otsu's thresholding [52] isolated the fish from the background to ensure consistency and support accurate feature extraction. Each image was then represented numerically by 33 features across three categories: color, shape, and texture. Color features—mean Hue, Saturation, and Value—were derived from the HSV model to enhance robustness under varying lighting. Shape features, extracted using regionprops, included area, perimeter, major and minor axis lengths, orientation, and eccentricity, capturing morphological traits such as size, form, and posture.

*1) Color Feature Extraction:* Color features were derived from HSV-transformed images, which separate color and brightness to enhance robustness to lighting. The mean values of H, S, and V formed a three-dimensional descriptor representing fish color composition. This method has proven effective in biological classification tasks [53], [54]. Feature computations are shown in (3–5), with examples in Fig. 5.

*2) Shape Feature Extraction:* Shape (morphometric) features were extracted from grayscale images segmented via Otsu's method using the regionprops function (skimage) to compute six geometric properties: area, perimeter, major and minor axis lengths, orientation, and eccentricity. These features quantitatively represent fish morphology—such as size, posture, and body elongation—crucial for species identification. Calculations follow (6–11), with examples shown in Fig.

(a) (b) (c)

**Fig. 3 (a) Barb image in JPG format; (b) Grayscale barb image; (c) Thresholding otsu result of barb image**

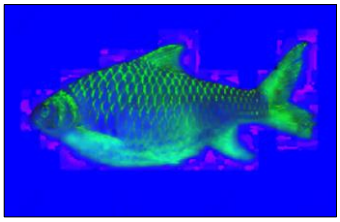


**Fig. 4 Barb image conversion to HSV format**

| | hue | saturation | value |
|---|---|---|---|
| 1 | 17.079398928783397 | 10.773332672088872 | 242.80079150962112 |
| 2 | 54.323804758032566 | 32.375972144033696 | 183.53282875669814 |
| 3 | 19.730418524053721 | 27.325451248421757 | 212.93224831868898 |

**Fig. 5 HSV extraction results**

| | area | orientation | perimeter | axis_major_len... | axis_minor_len... | eccentricity |
|---|---|---|---|---|---|---|
| 1 | 338.549132947 9769 | 0.42744538265 016185 | 44.6527212386 3954 | 8.79065026659 3833 | 3.48852076413 19353 | 0.58851911930 4384 |
| 2 | 1628.09090909 0909 | 0.79571983583 54528 | 89.3478649593 1294 | 22.3085525020 51064 | 10.5286672906 4874 | 0.69178597173 44224 |
| 3 | 495.919614147 90995 | 0.29625278906 4694 | 55.2531317653 7601 | 6.75543649938 0343 | 3.40384589100 08442 | 0.67014027984 39133 |

**Fig. 6 Regionprops extraction result**

*3) Texture Feature Extraction:* Texture features were extracted using the Gray-Level Co-occurrence Matrix (GLCM), generating 24 descriptors per image based on six metrics—contrast, dissimilarity, homogeneity, energy, ASM, and correlation—computed at four angles (0°, 45°, 90°, 135°). These features describe species-specific surface patterns such as roughness, smoothness, and texture orientation. GLCM is a well-established method in image analysis [42] and remains widely applied in fish classification using machine learning [55], [56]. Related equations (12–17) and examples are shown in Fig. 7.

*4) Combined Feature Dataset:* All extracted features—3 color, 6 shape, and 24 texture—were combined into a unified dataset comprising 33 attributes per image. With 120 processed fish images, the final dataset contained 120 rows, each representing a complete feature vector. This structured dataset served as a solid basis for subsequent classification and predictive analysis.

### D. Data Splitting

After feature extraction, the complete dataset—comprising 33 attributes per image across color (HSV), The final dataset, comprising 33 features per image (HSV, morphometric, and GLCM), was partitioned using a hold-out method: 70% (84 images) for training and 30% (36 images) for testing, with equal class distribution (16 training and 4 testing images per species). This approach supports effective model learning and fair evaluation. The RF classifier was trained on the training set, while the test set assessed generalization. The integration of color, shape, and texture features enabled species differentiation despite visual similarities, consistent with recent approaches in fish and agricultural image classification [53], [55]. The resulting dataset offers a robust basis for building accurate, generalizable models.

### E. Model Training and Parameter Optimization

The classification model was developed using the RF algorithm, which constructs multiple decision trees and aggregates their outputs through majority voting. This ensemble method enhances accuracy and reduces overfitting, making it well-suited for moderate-sized image datasets.

*1) Random Forest Mechanism:* RF uses bootstrap aggregating (bagging), training each decision tree on a randomly resampled data subset, while randomly selecting features at each node to reduce inter-tree correlation and enhance generalization. In this study, trees were built using HSV color, morphometric, and GLCM texture features, with node splits based on the Gini Impurity criterion (23):

$$Gini(t) = 1 - \sum_{i=1}^{n} p_i^2 \qquad (23)$$

where $p_i$ denotes the proportion of class $i$ at node $t$, and nn is the number of classes (Breiman, 2001). Lower Gini values reflect higher node purity, and the algorithm optimizes splits by minimizing this metric to achieve class homogeneity. A simplified example using HSV data from gourami (A), pomfret (B), and Java barb (C) illustrates tree construction (Fig. 8). For a test input (H = 20, S = 65, V = 93), the ensemble produced a majority vote for class "C" in five out of six trees, highlighting how aggregated decision paths improve classification accuracy.

| | dissimilarity_0 | dissimilarity_45 | dissimilarity_90 | dissimilarity_135 | correlation_0 | correlation_45 | correlation_90 | correlation_135 | homogeneity_0 | homogeneity_45 | homogeneity_90 | homogeneity_135 | contrast_0 | contrast_45 | contrast |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 10.02482845 7927043 | 12.274947917 859402 | 12.62718941 8562743 | 11.657906183 466906 | 0.91851916 89031444 | 0.879711124 4218389 | 0.872501085 2633807 | 0.8897091154 880908 | 0.2479412483 9907034 | 0.21421919238 423132 | 0.22061941756 53757 | 0.223041861525 0794 | 314.554487179 4869 | 446.212678281 197 | 468.680 8802 |
| 2 | 7.025086934 922613 | 8.8251821019 77337 | 9.434426229 50758 | 9.12018730 48 90989 | 0.93645260 12383669 | 0.894748523 3280523 | 0.877721615 8156594 | 0.8854351724 349184 | 0.1684384453 8844086 | 0.13193583104 624984 | 0.13358792022 709523 | 0.128701869577 2131 | 96.9262295081 9175 | 147.164932362 12608 | 171.417 88496 |
| 3 | 19.54386730 685295 | 22.512111109 063095 | 22.85052015 131681 | 23.471495146 344832 | 0.81331932 95782452 | 0.752485835 2900282 | 0.745539103 2064899 | 0.7298384908 5738 | 0.1618914446 5685967 | 0.15172639677 285416 | 0.15286206440 229555 | 0.144475358922 3876 | 831.686745235 0031 | 1097.48884672 76653 | 1150.92 32248 |

**Fig. 7 GLCM extraction result**

*2) Model Implementation and Training Results:* The RF model was implemented using scikit-learn's RandomForestClassifier() with 100 trees, Gini impurity for splitting, and bootstrap sampling enabled. Training used the .fit() method on the extracted features, with performance evaluated on the test set using accuracy, precision, recall, and F1-score. A sample node containing 60% catfish and 40% tilapia illustrates how Gini impurity guides optimal splits.

$$\text{Gini} = 1 - (p_1^2 + p_2^2) = 1 - (0.6^2 + 0.4^2)$$
$$= 1 - (0.36 + 0.16) = 1 - 0.52 = 0.48$$

This Gini value reflects moderate class impurity, prompting the algorithm to search for a feature-based split that reduces impurity further and improves class separation.

*3) Tuning Parameter*: To optimize the number of trees, n_estimators was varied from 10 to 1000. The highest accuracy (83.33%) was achieved with 500 trees, averaging $1.30 \pm 0.11$ seconds in training time (Table I). Beyond this, accuracy plateaued or slightly declined (e.g., 80.55% at 1000 trees), while training time increased, reflecting the diminishing returns principle in ensemble learning [56].

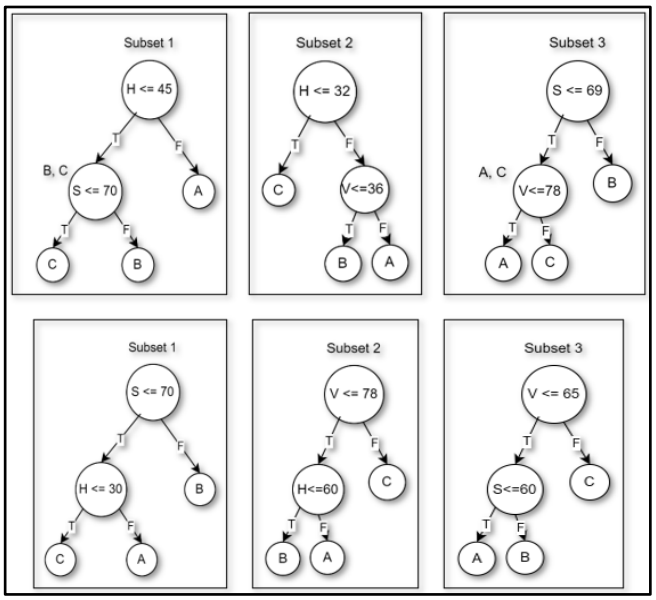



**Fig. 8 Decision tree illustration of 3 subsets**

TABLE I
ACCURACY AND RUNTIME BASED ON NUMBER OF TREES

| n_estimators | Accuracy | Average Runtime (s) |
|---|---|---|
| 10 | 63.88% | 0.07 |
| 100 | 77.78% | 0.28 |
| 500 | 83.33% | 1.30 |
| 1000 | 80.55% | 2.50 |

*4) Accuracy Convergence and Practical Considerations:* As shown in Fig. 9, model accuracy improved with increasing tree numbers, peaking at 83.33% with 500 trees, after which performance plateaued or declined slightly. Given the small dataset (120 images) and the need for computational efficiency, 500 trees were selected as the optimal configuration, balancing accuracy and resource constraints.

*F. Model Evaluation*

The Random Forest classifier (100 trees) was evaluated on a 30% test split (36 images) across six freshwater fish species. As shown in Table II, the model achieved 88.89% accuracy, 92.59% precision, 94.94% recall, and a 93.75% F1-score, based on 75 TP, 5 TN, 6 FP, and 4 FN. Class-wise evaluation (Table III) revealed perfect classification for catfish, barb, and tilapia, while misclassifications occurred mainly between pomfret and gourami due to their similar texture and color patterns. Gourami showed the lowest recall (0.50), likely affected by variations in size and orientation. The macro F1-score of 0.83 reflects strong overall performance, aligning with the "very good" category [57], [58].

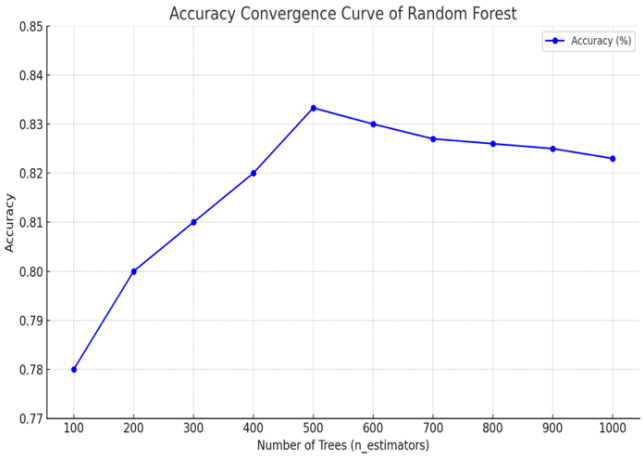



**Fig. 9 Accuracy convergence curve versus Random Forest tree count, with 500 trees (n_estimators = 500) yielding the highest accuracy**

TABLE II
MODEL PREDICTIONS VERSUS ACTUAL CLASS LABELS

| Actual \ Predicted | pomfret | gourami | catfish | barb | tilapia | Java barb |
|---|---|---|---|---|---|---|
| pomfret | 4 | 1 | 0 | 0 | 0 | 1 |
| gourami | 2 | 3 | 0 | 0 | 1 | 0 |
| catfish | 0 | 0 | 6 | 0 | 0 | 0 |
| barb | 0 | 0 | 0 | 6 | 0 | 0 |
| tilapia | 0 | 0 | 0 | 0 | 6 | 0 |
| Java barb | 1 | 0 | 0 | 0 | 0 | 5 |

TABLE III
METRIC INTERPRETATION AND CLASS-SPECIFIC PERFORMANCE

| Class | Precision | Recall | F1-score |
|---|---|---|---|
| Pomfret | 0.57 | 0.67 | 0.62 |
| Gourami | 0.75 | 0.50 | 0.60 |
| Catfish | 1.00 | 1.00 | 1.00 |
| Barb | 1.00 | 1.00 | 1.00 |
| Tilapia | 0.86 | 1.00 | 0.92 |
| Java barb | 0.83 | 0.83 | 0.83 |
| Macro Avg. | 0.83 | 0.83 | 0.83 |

The RF model developed in this study achieved an accuracy of 88.89% using image-based features, comparable to models using environmental parameters such as water temperature and pH (~88–94%) reported by [32] [59]. These results confirm that visual-based classification, when supported by effective feature extraction, can be a reliable alternative to sensor-based methods. RF also has been effectively applied in remote sensing for aquaculture classification tasks [60]. Although achieved slightly higher accuracy (91.3%) with CNN, RF offers key advantages: lower computational cost, easier interpretability, and greater feasibility for deployment on mid-range devices. In line with this, [34] showed that RF can operate on RISC-V microcontrollers with up to 90% energy savings while maintaining high accuracy, while [61] [62] reported significant improvements in speed and energy efficiency on ARM and custom hardware platforms.

In a comparison of machine learning classifiers on image-based features of rainbow trout (*Oncorhynchus mykiss*), the SVM with a radial basis function kernel achieved a correct classification rate of **82%**, demonstrating strong capability in differentiating fish based on their diet-related skin features [63]. Accuracy convergence analysis (Fig. 9) confirmed peak performance at 500 trees (n_estimators = 500), beyond which further increases yielded diminishing returns. This observation is consistent with studies combining HSV color, GLCM texture, and machine learning classifiers like RF or SVM, where F1-scores between 0.80–0.90 are commonly reported [55]. For instance, Ou et al. (2023) achieved an F1-score of 0.88 for tuna classification using a hybrid model (VGG16 + GLCM + SVM).

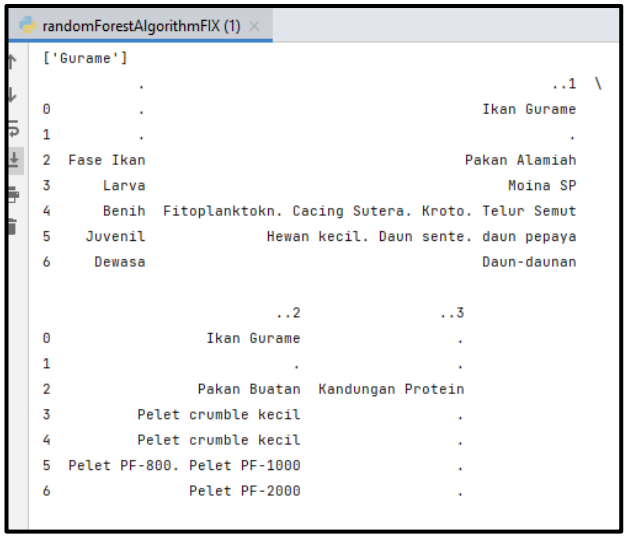
**Model Insight and Feature Importance.** The model demonstrated strong classification performance and interpretability, with visual inspection confirming accurate identification of species-specific traits such as the catfish's elongated, dark-colored body. Feature importance analysis revealed that color attributes (especially average Hue and Value in HSV) and texture descriptors (notably contrast and energy from GLCM) were key to predictive accuracy. While catfish and Java barb were classified perfectly, pomfret was often misclassified as gourami or Java barb due to similar textures and overlapping color features. Other species achieved F1-scores ≥ 0.83. Future enhancements may include data augmentation, advanced shape descriptors, and hybrid RF–CNN models for improved spatial feature extraction.

### *G. Simulation-Based Model Testing*

The model was tested using two simulation images—gourami (fish-1, Fig. 10) and Java barb (fish-2, Fig. 11)—to evaluate its ability to classify species and generate feed recommendations. As shown in Fig. 12 and Fig. 13, the system correctly identified both species and provided appropriate, species-specific dietary suggestions. These results demonstrate the model's effectiveness in combining image-based fish recognition with relevant feed recommendations.

**Fig. 10 Gourami testing image as fish-1**

**Fig. 11 Java barb testing image as fish-2**



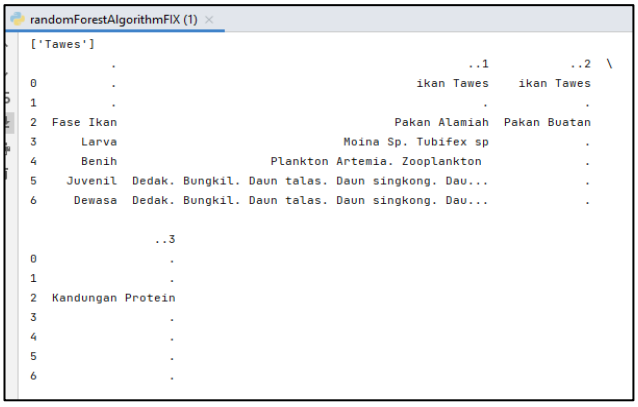
randomForestAlgorithmFIX (1)

['Gurame']

| | . | ..1 |
|---|---|---|
| 0 | . | Ikan Gurame |
| 1 | . | . |
| 2 | Fase Ikan | Pakan Alamiah |
| 3 | Larva | Moina SP |
| 4 | Benih | Fitoplanktokn. Cacing Sutera. Kroto. Telur Semut |
| 5 | Juvenil | Hewan kecil. Daun sente. daun pepaya |
| 6 | Dewasa | Daun-daunan |

| | ..2 | ..3 |
|---|---|---|
| 0 | Ikan Gurame | . |
| 1 | . | . |
| 2 | Pakan Buatan | Kandungan Protein |
| 3 | Pelet crumble kecil | . |
| 4 | Pelet crumble kecil | . |
| 5 | Pelet PF-800. Pelet PF-1000 | . |
| 6 | Pelet PF-2000 | . |

**Fig. 12 Results of fish species prediction and feed recommendations on fish-1 image**



randomForestAlgorithmFIX (1)

['Tawes']

| | . | ..1 | ..2 |
|---|---|---|---|
| 0 | . | ikan Tawes | ikan Tawes |
| 1 | . | . | . |
| 2 | Fase Ikan | Pakan Alamiah | Pakan Buatan |
| 3 | Larva | Moina Sp. Tubifex sp | . |
| 4 | Benih | Plankton Artemia. Zooplankton | . |
| 5 | Juvenil | Dedak. Bungkil. Daun talas. Daun singkong. Dau... | . |
| 6 | Dewasa | Dedak. Bungkil. Daun talas. Daun singkong. Dau... | . |

| | ..3 |
|---|---|
| 0 | . |
| 1 | . |
| 2 | Kandungan Protein |
| 3 | . |
| 4 | . |
| 5 | . |
| 6 | . |

**Fig. 13 Results of fish species prediction and feed recommendations on fish-2 image**

Gourami is a herbivorous freshwater fish that thrives on natural, plant-based feeds such as cassava leaves, papaya leaves, *Lemna sp.*, and *Hydrilla* [64]. These natural feeds align well with the species' digestive system and contribute to better water quality in aquaculture environments [65]. Although commercial pellets may be used as supplements, they should be formulated with moderate protein levels and plant-based ingredients. Studies suggest that combining 70–80% natural feed with 20–30% formulated feed improves growth performance while maintaining sustainability [66]. Java barb is an herbivorous freshwater fish whose digestive anatomy is optimized for processing plant-based feeds, such as algae and aquatic vegetation. Studies analyzing gastric contents confirm that Java barb primarily consume diatoms and green algae, supporting their natural preference for plant-derived diet [67] [68]. Supplementing with commercial pellets—often high in animal-derived protein—can disrupt digestion, reduce feed efficiency, and impair water quality. Therefore, natural feeding with leaves or aquatic plants is recommended for improved growth, health, and environmental sustainability in Java barb aquaculture.

## IV. CONCLUSION

The Random Forest model, developed using 33 integrated color, shape, and texture features, achieved a macro-averaged accuracy of 83% in classifying six freshwater fish species, with optimal performance obtained at 500 decision trees. The model demonstrated high precision in identifying catfish, barb, and tilapia (≥ 0.86), while classes like pomfret and gourami exhibited lower scores, suggesting the need for improvement. These results highlight the potential of Random Forest as a practical and cost-effective classification tool for aquaculture, particularly in settings with limited computational resources. To enhance its robustness and accuracy, especially for visually similar species, future work may involve data augmentation to improve generalization, the integration of advanced shape descriptors (e.g., Zernike moments or Fourier descriptors), and the development of hybrid ensemble models combining RF with lightweight CNNs. These enhancements would support real-time deployment and wider adoption in digital aquaculture systems.

## ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to PP Muhammadiyah for the financial support provided through the Muhammadiyah National Research Grant (RISETMU), Batch VIII of 2025, under Contract Number: 0258.574/I.3/D/2025. This support has been instrumental in the successful completion of this research.

## REFERENCES

[1] U. Zega, "Identifkasi Jenis Ikan Air Tawar di Sungai Yogi Kecamatan Fanayama," *Jurnal Education and Development*, vol. 8, no. 3, pp. 139–143, 2020.

[2] A. G. J. Tacon and M. Metian, "Feed matters: satisfying the feed demand of aquaculture," *Reviews in Fisheries Science & Aquaculture*, vol. 23, no. 1, pp. 1–10, 2015.

[3] A. Wijayanti, E. P. Madyastuti, C. P. N. Gulo, and A. F. Syarif, "KAJIAN KONSERVASI IKAN ENDEMIK TERANCAM PUNAH Betta burdigala (Kottelat & Ng, 1994) ASAL PERAIRAN BANGKA SELATAN," *Journal of Aquatropica Asia*, vol. 8, no. 2, pp. 98–102, 2023.

[4] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. Pearson, 2018.

[5] X. Zhang, H. Chen, L. Zhang, and Q. Chen, "Research on Freshwater Fish Recognition Based on Improved Random Forest," in *2023 International Conference on Intelligent Management and Software Engineering (IMSE)*, IEEE Xplore, 2023, pp. 81–84.

[6] G. A. Pradana, "A Study of Prediction Model for Capture Fisheries Production in Indonesian Sea Waters Using Machine Learning," *JISA (Jurnal Informatika dan Sains)*, vol. 6, no. 1, pp. 17–23, 2023, doi: https://doi.org/10.31326/jisa.v6i1.1553.

[7] S. Xu, J. Wang, X. Chen, and J. Zhu, "Identifying optimal variables for machine-learning-based fish distribution modeling," *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 81, no. 6, pp. 687–698, 2024, doi: https://doi.org/10.1139/cjfas-2023-019.

[8] S. Fauzi, P. Eosina, and G. F. Laxmi, "Implementasi Convolutional Neural Network Untuk Identifikasi Ikan Air Tawar," in *SEMNATI 2019*, 2019, pp. 163–167.

[9] R. M. Prasmatio, B. Rahmat, and I. Yuniar, "Deteksi dan Pengenalan Ikan Menggunakan Algoritma Convolutional Neural Network," *Jurnal Informatika dan Sistem Informasi (JIFoSI)*, vol. 1, no. 2, pp. 510–521, 2020, [Online]. Available: http://jifosi.upnjatim.ac.id/index.php/jifosi/article/view/144

[10] A. Azis, "IDENTIFIKASI JENIS IKAN MENGGUNAKAN MODEL HYBRID DEEP LEARNING DAN ALGORITMA KLASIFIKASI," *Sebatik*, vol. 24, no. 2, pp. 201–206, 2020, doi: https://doi.org/10.46984/sebatik.v24i2.1057.

[11] A. L. Unihehu and I. Suharjo, "Klasifikasi Jenis Ikan Berbasis Jaringan Saraf Tiruan Menggunakan Algoritma Principal Component Analysis (PCA)," *Jurnal Ilmiah Ilmu Komputer Fakultas Ilmu Komputer Universitas Al Asyariah Mandar*, vol. 7, no. 2, pp. 27–32, 2021, doi: https://doi.org/10.35329/jiik.v7i2.200.

[12] M. Ramadhani and D. H. Murti, "Klasifikasi Ikan Menggunakan Oriented Fast and Rotated Brief (ORB) dan K-Nearest Neighbor (KNN)," *JUTI J. Ilm. Teknol. Inf*, vol. 16, no. 2, pp. 115–124, 2018, doi: https://doi.org/10.12962/j24068535.v16i2.a711.

[13] G. M. U. Fatimah, "KLASIFIKASI KUALITAS CITRA IKAN BANDENG MENGGUNAKAN METODE FUZZY K-NEAREST NEIGHBOR BERDASARKAN DESCRIPTOR BENTUK DAN CO-OCCURANCE MATRIK," Universitas Muhammadiyah Sidoarjo, 2018. [Online]. Available: http://eprints.umsida.ac.id/1895/

[14] M. A. Islam, M. R. Howlader, U. Habiba, R. H. Faisal, and M. M. Rahman, "Indigenous fish classification of Bangladesh using hybrid features with SVM classifier," in *2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)*, IEEE, Jul. 2019, pp. 1–4. doi: 10.1109/IC4ME247184.2019.9036679.

[15] Y. Krisdiani, "Klasifikasi kesegaran ikan berdasarkan warna mata menggunakan metode K-Nearest Neighbor," Universitas Negeri Malang, 2019.

[16] A. Pariyandani, D. A. Larasati, E. P. Wanti, and M. Muhathir, "Klasifikasi Citra Ikan Berformalin Menggunakan Metode k-NN dan GLCM," in *Semantika (Seminar Nasional Teknik Informatika)*, 2019, pp. 42–47. [Online]. Available: https://semantika.polgan.ac.id/index.php/Semantika/article/view/49

[17] E. P. Wanti and M. Muhathir, "Pengidentifikasian Citra Ikan Berformalin Dengan Menggunakan Metode Multilayer Perceptron," *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, vol. 5, no. 1, pp. 491–502, 2021, doi: http://dx.doi.org/10.30645/j-sakti.v5i1.342.

[18] I. R. Diesta and W. F. Al Maki, "Klasifikasi Ikan Cupang Menggunakan Support Vector Machine," in *eProceedings of Engineering*, 2021, pp. 10556–10565.

[19] M. A. D. Akbar, A. B. Setiawan, and R. K. Niswatin, "Klasifikasi Jenis Ikan Cupang Menggunakan Metode GLCM Dan KNN," in *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, 2021, pp. 152–158.

[20] I. E. Hasym and I. Susilawati, "Klasifikasi Jenis Ikan Cupang Menggunakan Algoritma Principal Component Analysis (PCA) Dan K-Nearest Neighbors (KNN)," *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, vol. 1, no. 1, pp. 168–179, 2021, doi: https://doi.org/10.24002/konstelasi.v1i1.4242.

[21] N. Rachmat, Y. Yohannes, and A. Mahendra, "Klasifikasi Jenis Ikan Laut Menggunakan Metode SVM dengan Fitur HOG dan HSV," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 8, no. 4, pp. 2235–2247, 2021, doi: https://doi.org/10.35957/jatisi.v8i4.1686.

[22] S. Suwarsito, H. Mustafidah, T. Pinandita, and P. Purnomo, "Freshwater Fish Classification Based on Image Representation Using K-Nearest Neighbor Method," *JUITA: Jurnal Informatika*, vol. 10, no. 2, pp. 183–189, 2022, doi: http://dx.doi.org/10.30595/juita.v10i2.15471.

[23] R. Nuraini, A. Wibowo, B. Warsito, W. A. Syafei, and I. Jaya, "Combination of K-NN and PCA Algorithms on Image Classification of Fish Species," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 5, pp. 1026–1032, 2023, doi: https://doi.org/10.29207/resti.v7i5.5178.

[24] D. Yusup, S. Faisal, and A. Pratama, "Klasifikasi Jenis Ikan Hias African Cichlid Menggunakan Algoritma Support Vector Machines," *Scientific Student Journal for Information, Technology and Science*, vol. 4, no. 1, pp. 31–38, 2023, doi: http://repository.ubpkarawang.ac.id/id/eprint/2909/.

[25] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. NewYork: Springer, 2006.

[26] T. P. Trappenberg, *Fundamentals of Machine Learning*. United States of America: Oxford University Press, 2020. doi: 10.1093/oso/9780198828044.001.0001.

[27] L. Alwi, A. T. Hermawan, and Y. Kristian, "Identifikasi Biji-Bijian Berdasarkan Ekstraksi Fitur Warna, Bentuk dan Tekstur Menggunakan Random Forest," *INSYST: Journal of Intelligent System and Computation*, vol. 1, no. 2, pp. 92–98, 2019, doi: https://doi.org/10.52985/insyst.v1i2.93.

[28] N. Li, L. Ding, H. Zhao, J. Shi, D. Wang, and X. Gong, "Ship Detection Based on Multiple Features in Random Forest Model for Hyperspectral Images," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 891–895, 2018, doi: https://doi.org/10.5194/isprs-archives-XLII-3-891-2018.

[29] Y. A. Sari *et al.*, "Indonesian traditional food image identification using random forest classifier based on color and texture features," in *2019 International Conference on Sustainable Information Engineering and Technology (SIET)*, IEEE, 2019, pp. 206–211.

[30] L. S. P. Annabel and V. Muthulakshmi, "AI-powered image-based tomato leaf disease detection," in *2019 third international conference on I-SMAC (IoT in social, mobile, analytics and cloud)(I-SMAC)*, IEEE, 2019, pp. 506–511.

[31] G. Jhansi and K. Sujatha, "HRFSVM: Identification of fish disease using hybrid Random Forest and Support Vector Machine," *Environ Monit Assess*, vol. 195, no. 8, p. 918, 2023.

[32] S. M. M. Islam, M. F. Rohani, and M. Shahjahan, "Probiotic yeast enhances growth performance of Nile tilapia (Oreochromis niloticus) through morphological modifications of intestine," *Aquac Rep*, vol. 21, no. November, p. 100800, 2021, doi: https://doi.org/10.1016/j.aqrep.2021.100800.

[33] R. P. Kumar, T. Prabhu, J. Sowmya, and A. N. Kumar, "Water Quality Prediction for Smart Aquaculture," in *2024 5th International Conference on Electronics and Sustainable Communication Systems (ICESC)*, IEEE, 2024, pp. 1376–1382. doi: 10.1109/ICESC60852.2024.10690049.

[34] F. Daghero *et al.*, "Adaptive random forests for energy-efficient inference on microcontrollers," in *2021 IFIP/IEEE 29th International Conference on Very Large Scale Integration (VLSI-SoC)*, IEEE, 2021, pp. 1–6. doi: 10.1109/VLSI-SoC53125.2021.9606986.

[35] S. Koschel, S. Buschjäger, C. Lucchese, and K. Morik, "Fast inference of tree ensembles on arm devices," *arXiv preprint arXiv:2305.08579*, 2023, doi: https://doi.org/10.48550/arXiv.2305.08579.

[36] R. A. Pramunendar, F. Alzami, and R. A. Megantara, "Penerapan Random Forest Untuk Pengenalan Jenis Ikan Berdasarkan Perbaikan Citra Clahe Dan Dark Channel Prior," *Jurnal Informatika UPGRIS*, vol. 7, no. 1, pp. 67–74, 2021, doi: https://doi.org/10.26877/jiu.v7i1.8231.

[37] R. D. Shirwaikar, L. D'Souza, A. S. Bhangle, S. D. Joshi, N. A. Jathar, and J. E. Alvares, "Comparative Analysis of Machine Learning Algorithms in Fish Survival Prediction," in *2024 IEEE Bangalore Humanitarian Technology Conference (B-HTC)*, 2024, pp. 28–33. doi: 10.1109/B-HTC60740.2024.10564052.

[38] R. P. A. Yunaidi and A. Wibowo, "Aplikasi pakan pelet buatan untuk peningkatan produktivitas budidaya ikan air tawar di desa Jerukagung Srumbung Magelang," *Jurnal Pemberdayaan: Publikasi Hasil Pengabdian kepada Masyarakat*, vol. 3, no. 1, pp. 45–54, 2019, doi: https://doi.org/10.12928/jp.v3i1.621.

[39] The scikit-image team, "skimage.measure," *https://scikit-image.org/docs/stable/api/skimage.measure.html*. Accessed: Jul. 02, 2025.

[40] R. Szeliski, *Computer vision: algorithms and applications*. Springer Nature, 2022.

[41] M. Sonka, V. Hlavac, and R. Boyle, *Image processing, analysis and machine vision*. Springer, 2013.

[42] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Trans Syst Man Cybern*, vol. SMC-3, no. 6, pp. 610–621, 1973, doi: 10.1109/TSMC.1973.4309314.

[43] L. Breiman, "Random Forests," *Mach Learn*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[44] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf Process Manag*, vol. 45, no. 4, pp. 427–437, 2009, doi: https://doi.org/10.1016/j.ipm.2009.03.002.

[45] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *International Journal of Machine Learning Technology*, vol. 2, no. 1, pp. 37–63, 2011, doi: https://doi.org/10.48550/arXiv.2010.16061.

[46] Hadijah, Zainuddin, Muliani, and Farida, *Nutrisi dan Pakan Ikan*. Makassar: Azkiya Publishing, 2024.

[47] M. Aris, N. Abdullah, Yuliana, R. Labenua, A. Rumondang, and F. Muchdar, *Pakan Alami*. Ternate: Kamiya Jaya Aquatic, 2024.

[48] H. Daten, *Monograf Pembuatan Pakan Ikan Berbentuk Pelet Tenggelam*. Padang: Get Press Indonesia, 2024.

[49] mybest, "10 Rekomendasi Makanan Ikan Terbaik," *https://id.my-best.com/136458*. Accessed: Apr. 26, 2025.

[50] De Heus Animal Nutrition, "KETAHUI JENIS PAKAN IKAN ALAMI UNTUK KEBUTUHAN BUDIDAYA IKAN AIR TAWAR,"

*https://www.deheus.id/cari/berita-dan-artikel/ketahui-jenis-pakan-alami-untuk-kebutuhan-budidaya-ikan-air-tawar*. Accessed: May 26, 2025.

[51] Pemkab Buleleng, "INILAH JENIS-JENIS IKAN AIR TAWAR YANG DIBUDIDAYAKAN DI INDONESIA," *https://bulelengkab.go.id/detail/artikel/inilah-jenis-jenis-ikan-air-tawar-yang-dibudidayakan-di-indonesia-28*. Accessed: Apr. 02, 2019.

[52] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285–296, pp. 23–27, 1975, doi: 0018-9472/79/0100-0062$00.75.

[53] K. Crisannaufal and W. F. Al Maki, "Optimizing Support Vector Machine for Avocado Ripeness Classification Using Moth Flame Optimization," *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 2, pp. 330–340, 2025, doi: https://doi.org/10.35882/jeeemi.v7i2.652.

[54] J. Deka, S. Laskar, and B. Baklial, "Automated freshwater fish species classification using deep CNN," *Journal of The Institution of Engineers (India): Series B*, vol. 104, no. 3, pp. 603–621, 2023, doi: https://doi.org/10.1007/s40031-023-00883-2.

[55] L. Ou *et al.*, "Automatic classification of the phenotype textures of three Thunnus species based on the machine learning SVM algorithm," *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 80, no. 8, pp. 1221–1236, 2023, doi: https://doi.org/10.1139/cjfas-2022-0270.

[56] Q. Chen, C. Shen, H. Du, and D. Tang, "Comparing supervised classification algorithm–feature combinations for Spartina alterniflora extraction: a case study in Zhanjiang, China," *Frontiers in Remote Sensing*, vol. 6, no. July, p. 1606549, 2025, doi: https://doi.org/10.3389/frsen.2025.1606549.

[57] D. Chicco, "Ten quick tips for machine learning in computational biology," *BioData Min*, vol. 10, no. 1, p. 35, 2017, doi: 10.1186/s13040-017-0155-3.

[58] J. H. Cabot and E. G. Ross, "Evaluating prediction model performance," *Surgery*, vol. 174, no. 3, pp. 723–726, 2023, doi: https://doi.org/10.1016/j.surg.2023.05.023.

[59] M. Khan, F. Haq, and A. Awan, "Machine Learning-Based Fish Species Recommendation Using Water Quality Parameters," *International Journal of Innovations in Science and Technology*, vol. 7, pp. 110–126, May 2025.

[60] Y. Cui, X. Zhang, N. Jiang, T. Dong, and T. Xie, "Remote sensing identification of marine floating raft aquaculture area based on sentinel-2A and DEM data," *Front Mar Sci*, vol. 9, no. August, p. 955858, 2022, doi: https://doi.org/10.3389/fmars.2022.955858.

[61] S. Koschel, S. Buschjäger, C. Lucchese, and K. Morik, "Fast inference of tree ensembles on arm devices," *arXiv:2305.08579v1*, vol. May, 2023, doi: https://doi.org/10.48550/arXiv.2305.08579.

[62] Z. Takhirov, J. Wang, M. S. Louis, V. Saligrama, and A. Joshi, "Field of groves: An energy-efficient random forest," *arXiv:1704.02978v1*, vol. April, 2017, doi: https://doi.org/10.48550/arXiv.1704.02978.

[63] M. Saberioon, P. Císař, L. Labbé, P. Souček, P. Pelissier, and T. Kerneis, "Comparative performance analysis of support vector machine, random forest, logistic regression and k-nearest neighbours in rainbow trout (Oncorhynchus mykiss) classification using image-based features," *Sensors*, vol. 18, no. 4, pp. 1–15, 2018, doi: https://doi.org/10.3390/s18041027.

[64] F. Feliatra, N. A. Pamukas, D. R. Siagian, S. A. Arlisman, and R. Aulia, "Utilization of fermented chicory and cabbage waste in feed to improve growth performance of giant gourami (Osphronemus goramy)," *Aquaculture, Aquarium, Conservation & Legislation*, vol. 17, no. 5, pp. 1533–1542, 2024.

[65] F. N. Arida, F. Indira, and W. Rahma, "Evaluation of giant gouramis (Osphronemus Gouramy) growth and survival rate with additional feeding from kelakai leaves (Stenochlaena Palustris) silage," *Russ J Agric Socioecon Sci*, vol. 119, no. 11, pp. 215–219, 2021, doi: 10.18551/rjoas.2021-11.24.

[66] O. O. Olude and P. Akinduti, "Growth and Haemato-Biochemical Responses of All-Male Tilapia, Oreochromis niloticus, to Diets Containing Fermented Cassava Leaf Meal," in *Biotechnological Approaches to Sustainable Development Goals*, Springer, 2023, pp. 187–203.

[67] R. Kurnia, N. Widyorini, and A. Solichin, "Analysis of Food Competition Between Java Barb (Barbonymus gonionotus), Java Tilapia (Oreochromis mossambicus) and Nile Tilapia (Oreochromis niloticus) in Wadaslintang Reservoir, Wonosobo Regency," *Management of Aquatic Resources Journal (MAQUARES)*, vol. 6, no. 4, pp. 515–524, 2018, doi: https://doi.org/10.14710/marj.v6i4.21343.

[68] B. Ristyanadi, E. S. Prihatini, F. Mas' ud, and M. Atok, "Teknik Pembenihan Ikan Tawes (Barbonymus Gonionotus) Di Unit Pelaksana Teknis Laboratorium Kesehatan Ikan Dan Lingkungan Umbulan Pasuruan Jawa Timur," *YUME: Journal of Management*, vol. 5, no. 3, pp. 694–700, 2022, doi: https://doi.org/10.37531/yum.v5i3.3834.