

Article

A Method for Fish Feeding Intensity Assessment Based on Spatial Features and TabNet-DFWL

Lu Zhang ^{1,2,*} , Shunshun Zhou ^{1,2}, Zunxu Liu ^{1,2}, Yue Li ³, Hao Yang ^{1,2} and Wenhui Ni ^{1,2}

¹ College of Information and Artificial Intelligence, Yangzhou University, Yangzhou 225127, China; 15056721964@163.com (S.Z.); lzx_0126@163.com (Z.L.); y751923962@163.com (H.Y.); niwenhuiyzhh@163.com (W.N.)

² Jiangsu Province Engineering Research Centre of Knowledge Management and Intelligent Service, Yangzhou 225127, China

³ College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China; liyue2024@cau.edu.cn

* Correspondence: lzhang@yzu.edu.cn

Abstract

Accurate assessment of fish feeding intensity is significant for the timely understanding of feeding demands, dynamically adjusting feeding strategies, and reducing aquaculture costs. However, existing methods often rely on superficial visual features that fail to capture subtle satiety dynamics, resulting in limited reliability. To address the issue, a method for fish feeding intensity assessment based on spatial features and TabNet model with Dynamic Feature Weighting Layer (TabNet-DFWL) is proposed in this study. Fish body contours are extracted from lateral-view images through a pipeline of segmentation, enhancement, and binarization. Subsequently, spatial features highly correlated with fish feeding mechanisms are proposed to characterize behavioral changes. Based on these, an interpretable model integrating spatial features and TabNet-DFWL is constructed to achieve precise fish feeding intensity assessment. This method explores spatial features related to feeding behavior from the underlying mechanism of fish behavioral changes and establishes a feeding intensity assessment model based on TabNet-DFWL. By doing so, it avoids the black-box risk commonly associated with traditional deep learning models and significantly improves model interpretability and reliability, thereby providing a trustworthy basis for precision feeding in aquaculture. Experiments conducted on a real-world fish feeding dataset demonstrate that the proposed method achieves an accuracy of 95.96%, an average precision of 93.44%, an average recall of 93.33%, an average specificity of 98.15%, and an average F1-score of 93.38%. Compared with comparative algorithms, all evaluation metrics exhibit improvements. These results indicate that the proposed method enables accurate assessment of fish feeding intensity and can effectively support the dynamic adjustment of feeding strategies in aquaculture systems.



Academic Editors: Zhangying Ye, Chao Zhou and Weiqiang Ni

Received: 29 April 2026

Revised: 22 May 2026

Accepted: 22 May 2026

Published: 24 May 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: aquaculture; computer vision; feeding intensity assessment; spatial feature; interpretable model

Key Contribution: This study constructs a fish feeding intensity assessment model integrating spatial features with the TabNet-DFWL architecture, which avoids the black-box risk of traditional deep learning and significantly enhances model interpretability and reliability. Evaluated on a real-world dataset, the proposed method achieves 95.96% accuracy, outperforming comparative algorithms.

1. Introduction

In recent years, with the advancement and innovation of aquaculture technology, the output of aquaculture has been continuously increasing, resulting in a growing market share of farmed aquatic products. In aquaculture, feed costs account for more than 60% of the total production cost [1,2]. Therefore, realizing precision feeding is significant not only for reducing aquaculture costs and improving production efficiency, but also for minimizing feeding waste, maintaining water quality and promoting sustainable aquaculture. At present, feeding practices in aquaculture mainly rely on manual experience-based feeding or machine-based feeding with fixed timing and quantities. These approaches are unable to dynamically adjust feeding strategies according to fish feeding intensity, which easily leads to underfeeding or overfeeding [3]. When the physiological condition, nutritional requirements and environmental factors of fish change, their feeding demands also vary accordingly [4]. A decrease in feeding demands is typically accompanied by a corresponding reduction in feed intake. Consequently, feeding intensity can reflect fish appetite and satiety status, and its accurate assessment is crucial for dynamically adjusting feeding strategies in precision aquaculture systems.

Computer vision has become an important research tool in fisheries production due to its non-invasive, cost-effectiveness and high efficiency [5,6]. Computer vision-based methods for assessing fish feeding intensity generally involve image pre-processing and segmentation to obtain fish targets, followed by feature extraction and the construction of feeding intensity assessment models, thereby enabling the assessment of fish feeding intensity. At present, these methods can be categorized into machine learning-based approaches and deep learning-based approaches according to the different image features extracted.

Machine learning-based methods typically adopt a paradigm of feature engineering combined with classical classifiers. First, a set of handcrafted features is extracted from fish audio or visual streams using signal processing or statistical methods. Subsequently, machine learning models such as Support Vector Machine (SVM), Random Forest (RF), and Gradient Boosting Decision Tree (GBDT) are employed to classify and discriminate feeding intensity levels. Zhang et al. [7] extracted texture features, including inverse difference moment and correlation, from fish school images using mean background modeling and gray-level co-occurrence matrices. A Back-Propagation (BP) neural network was then used to construct a fish feeding intensity recognition model. Chen et al. [8] plotted fish swimming trajectories based on the centroid of fish schools and extracted color, shape, and texture features from the images. The eXtreme Gradient Boosting (XGBoost) algorithm was applied to select feeding evaluation factors, and optimal weights were determined using a weighted fusion strategy. On this basis, a fish feeding intensity assessment model was constructed using the fused features. Yuan and Zhu [9] employed color channel separation techniques to extract statistical features such as mean, variance, skewness and kurtosis from each channel. These features were subsequently fused using Kernel Principal Component Analysis (KPCA), and a fish feeding behavior detection model was finally established based on SVM. However, most machine learning-based methods rely on indirect evaluation using features such as texture, color, and statistical descriptors. These features are often insufficient to directly describe fish spatial distribution, posture changes, inter-individual distance, and aggregation degree [10,11]. As a result, the accuracy and reliability of the evaluation outcomes remain limited.

Deep learning-based methods for fish feeding intensity assessment typically adopt an end-to-end feature learning paradigm, which enables the automatic extraction of semantic features directly from images, videos, or audio signals to evaluate fish feeding intensity. For instance, Liu et al. [12] conducted a feeding intensity assessment based on fish feeding

image data. In this method, the Coordinate Attention mechanism was integrated into a MobileViT backbone network to construct a fish school feeding intensity recognition model. Wang et al. [13] evaluated fish feeding intensity using RGB video and optical flow videos generated by FlowNet2. A dual-stream 3D convolutional neural network was employed to establish a fish behavior recognition model, which can accurately identify fish feeding behaviors. Zhang et al. [14] performed fish feeding intensity assessment based on multimodal features, including video, audio and water surface wave data. They proposed a multilevel enhanced multimodal interaction network (MAINet) to construct a quantitative model for fish feeding intensity, achieving precise classification and quantitative evaluation of feeding intensity levels. More recently, Iqbal et al. [15] proposed LightHybridNet-Transformer-FFIA, a hybrid Transformer-based deep learning model for enhanced fish feeding intensity classification, which further improved the ability of deep models to extract feeding-related features. Wang et al. [16] developed a dual-stream spatiotemporal fusion method for fish school feeding intensity identification, in which spatial and temporal information were jointly utilized to enhance feeding intensity recognition. Deep learning models exhibit strong predictive abilities; however, as black-box models, the decision-making process often lacks transparency and interpretability. This limitation leads to reduced model credibility and restricts their further development and application in practical aquaculture scenarios [17,18].

Fish typically exhibit distinct behavioral features under different feeding intensities [19,20]. However, existing methods for assessing fish feeding intensity mainly rely on the extraction of superficial features, such as texture, color or statistical descriptors, to indirectly assess feeding intensity [21,22]. These approaches struggle to capture behavioral differences exhibited by fish at varying feeding intensities but fail to reflect the subtle and dynamic attenuation of satiety levels. Consequently, the reliability of fish feeding intensity assessment results remains limited. In contrast, spatial features have significant advantages, which can directly describe fish behavioral and accurately capture subtle behavioral changes during fish feeding. For example, as feeding intensity increases, fish schools tend to exhibit greater overall upward movement, reduced inter-individual distance and a higher degree of aggregation. Accordingly, spatial features highly correlated with feeding behavior are explored from the underlying mechanism of fish behavioral changes, and an interpretable model is constructed to assess fish feeding intensity directly and accurately in this study.

To address the limitations of insufficient behavioral feature representation and poor model interpretability in existing methods, a method for fish feeding intensity assessment based on spatial features and the TabNet-DFWL model is proposed in this study. Firstly, fish images are acquired from a lateral viewing angle. Subsequently, a series of image processing techniques are applied to obtain clear fish body contours, including image segmentation, image enhancement, image binarization and contour extraction. Then, fish spatial features that are highly correlated with fish feeding behavior are proposed. On this basis, the DFWL is introduced to optimize the TabNet to construct an interpretable model, enabling precise assessment of fish feeding intensity. The main contributions of this study are summarized as follows:

- (1) A series of spatial features highly correlated with feeding behavior is proposed from the underlying mechanism of fish behavioral changes, such as inter-individual distance, fish posture, and aggregation degree. These features can accurately reflect behavioral changes during the feeding process, enhance the effectiveness of feature representation, and improve the reliability of feeding intensity assessment.

- (2) A dynamic feature weighting layer is proposed to automatically strengthen the weights of key features within the TabNet model. This mechanism addresses the degra-

dation in assessment performance caused by the implicit expression of fish satiety and improves the model performance when processing fish feeding data.

(3) A fish feeding intensity assessment method based on spatial features and the TabNet-DFWL network is proposed. This method avoids the black-box risk of conventional deep learning models and enhances model interpretability and credibility, thereby providing a reliable basis for precision feeding in aquaculture.

The rest of this paper is organized as follows. In Section 2, the complete workflow of the proposed method is presented. The experiments and corresponding results are reported in Section 3. Finally, the conclusions are provided in Section 4.

2. Materials and Methods

2.1. Data Acquisition

The experiments were conducted in the Xuexing Building at the Yangzijin Campus of Yangzhou University in Jiangsu Province. Juvenile Asian crucian carp were selected as the experimental subjects. The fish's body lengths ranged from 15 to 25 cm and were raised in a fish tank with a water depth of approximately 40 cm. Before the experiment began, the fish had been living in the tank for more than one month and had fully adapted to the environment.

The experimental data acquisition platform is illustrated in Figure 1. It consists of a raising tank (90 cm × 65 cm × 70 cm), an oxygenation equipment (Sensen Group Co., Ltd., Zhoushan, China), a lighting system (Chihiros Aquatic Technology Co., Ltd., Guangzhou, China), a ZED 2i camera (Stereolabs SAS, San Francisco, CA, USA), and a computer (Hasee Computer Co., Ltd., Shenzhen, China). The camera was fixed at the center of the tank's side wall, 50 cm from the wall and 20 cm above the ground. It was configured to record videos at a resolution of 2208 × 1242 pixels and a frame rate of 30 frames per second. During data collection, a lateral viewing angle was adopted to record the entire feeding process to ensure class balance among images corresponding to different feeding intensity levels in the dataset.

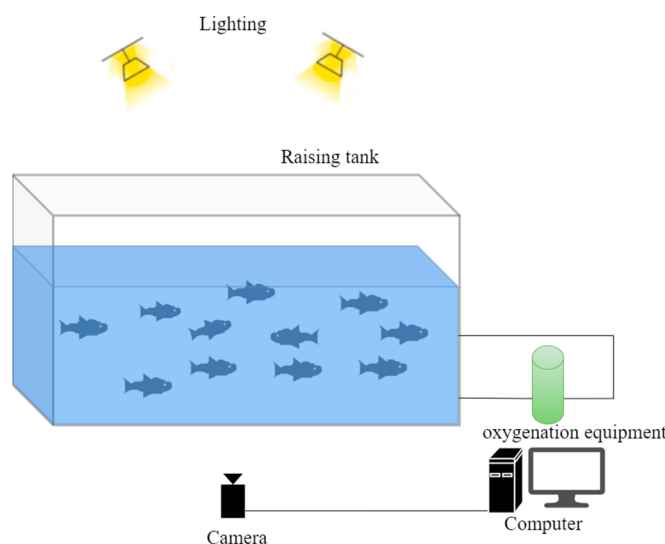


Figure 1. Experimental data acquisition platform.

The feeding scheme was implemented using fixed timing, fixed location and fixed quantity. Feeding was conducted twice daily, at 09:00 and 17:00, respectively, with 60 pellets provided at each time from directly above the center of the fish tank. To ensure that the recorded videos contained complete information on fish feeding behavior, each video acquisition was recorded for 120–150 s, including 20 s before feeding, the entire

feeding process and 20 s after feeding. The video recordings started on 25 October 2023 and ended on 13 December 2023, covering a total observation period of 50 days. During the data collection process, some data were affected by a sudden drop in temperature, human activities around the experimental area, and noise interference, which led to abnormal feeding behavior of the fish. Therefore, the recordings affected by these factors were excluded. Finally, valid fish feeding data from 39 days were retained for subsequent analysis.

To evaluate the effectiveness of the proposed method, one image was extracted every 15 frames from the recorded videos. Data cleaning was conducted to remove severely blurred or occluded samples, while no data augmentation was applied. Only the original collected images were used to preserve real-world conditions. After this preprocessing, a total of 9503 valid images were finally obtained to construct the dataset. Some representative fish feeding samples are shown in Figure 2. The dataset was divided into a training set, a validation set and a test set using a random seed (42). The training set and validation set together accounted for 75% of the total dataset, while the test set accounted for 25%. The ratio between the training set and the validation set was set to 4:1. Detailed information regarding dataset partitioning is provided in Table 1.

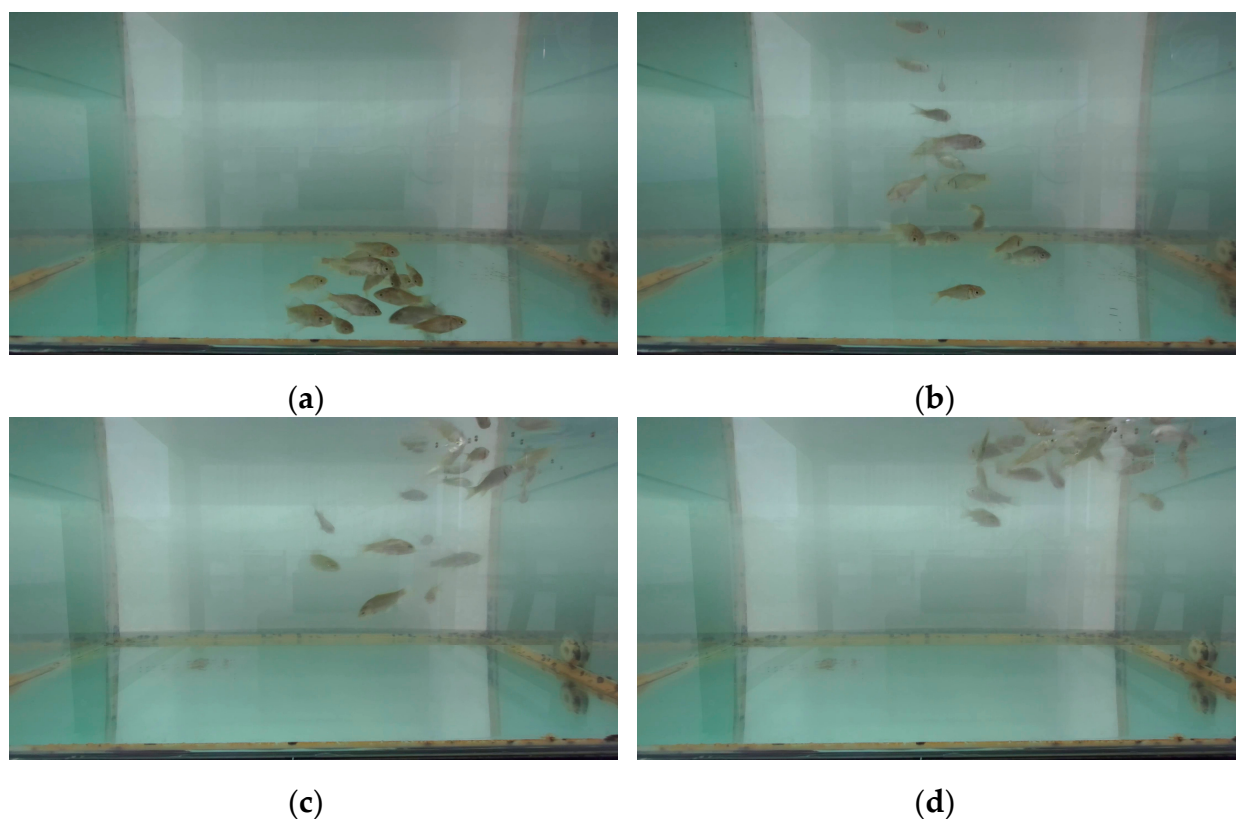


Figure 2. Fish image examples. (a) The fish feeding intensity is none; (b) the fish feeding intensity is weak; (c) the fish feeding intensity is medium; (d) the fish feeding intensity is strong.

Table 1. Dataset partitioning result.

Feeding Intensity	Training Set	Validation Set	Test Set	Total
None	1613	402	672	2687
Weak	1536	386	640	2562
Medium	1296	323	540	2159
Strong	1257	314	524	2095
Total	5702	1425	2376	9503

Fish images were classified into four standard feeding intensity levels [23], with the classification criteria presented in Table 2. Fish images at each feeding intensity are shown in Figure 2. Under the none state, fish were observed to aggregate in the lower and middle layers of the water column, with slight overlapping among individuals, and no response was exhibited towards the supplied feed. In the weak state, only nearby food was consumed, and large-scale swimming feeding behavior was not observed. Minor overlapping among individuals was present. In the medium state, fish exhibited relatively large swimming amplitudes, and most individuals presented a distinct inclined posture. In the strong state, all individuals were observed to float to the water surface for feeding. Intense competition for food was evident, accompanied by large-scale overlapping and inclination among fish individuals.

Table 2. Sample classification criteria of the dataset.

Feeding Intensity	Description
None	Fish do not respond to food
Weak	Fish eat only pellets that fall directly in front and do not move to take food
Medium	Fish move to take food, but return to original position
Strong	Fish move freely between food items and consume all food that is presented

2.2. Overall Process

A method for fish feeding intensity assessment based on spatial features and TabNet-DFWL is proposed in this study, which enhances the interpretability and credibility of feeding intensity assessment and enables precise evaluation of fish feeding intensity. The detailed workflow of the proposed method is illustrated in Figure 3.

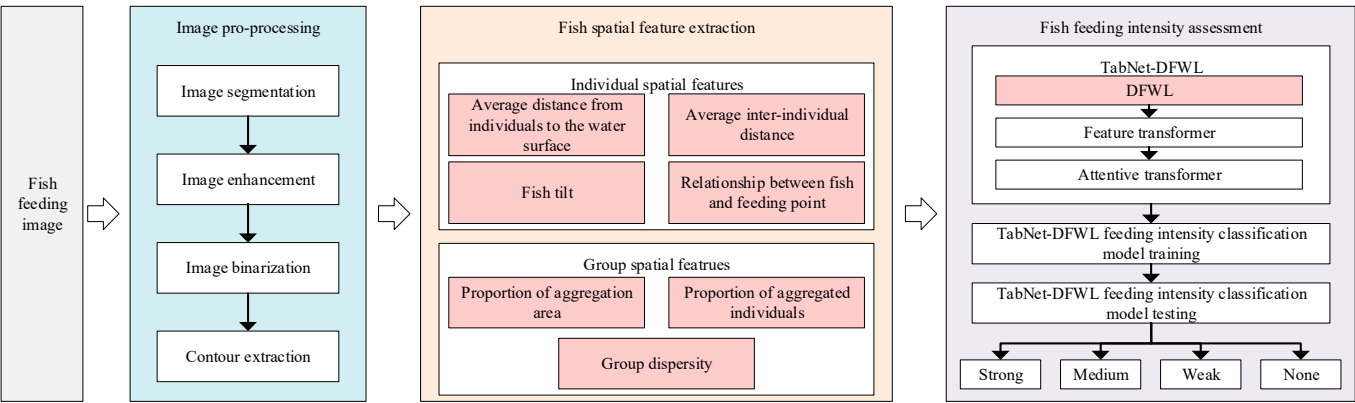


Figure 3. Overall process of fish feeding intensity assessment method.

- (1) **Data acquisition.** The camera was fixed at a distance of 50 cm from the side of the raising tank to record fish feeding videos. Video frames were extracted to obtain fish images, and a dataset was constructed accordingly.
- (2) **Image pre-processing.** Target fish bodies were extracted from the images through image segmentation. Subsequently, a series of image processing operations, including image enhancement, image binarization and contour extraction, were performed to obtain clear fish body contours.
- (3) **Fish spatial feature extraction.** Based on the extracted fish contours, centroid points of fish bodies were calculated. Then, spatial features were extracted from both individual and group perspectives.

(4) Fish feeding intensity assessment. A TabNet-DFWL model was constructed by introducing the DFWL. The mapping relationship between fish spatial features and feeding intensity was established, thereby enabling the assessment of fish feeding intensity.

2.3. Image Pre-Processing

Image pre-processing is regarded as a critical step in computer vision and image processing, aiming to improve image quality, reduce noise and enhance features, thereby providing high-quality data input for subsequent analysis and processing. However, images acquired in underwater environments are often affected by uneven illumination, distortion, and motion blur caused by water fluctuations. These issues tend to result in inaccurate feature extraction and consequently degrade the performance of feeding intensity assessment. To address these problems and improve the accuracy of image feature extraction, four pre-processing operations, including image segmentation, image enhancement, image binarization, and contour extraction, were performed in this study.

(1) Image segmentation. Image segmentation is defined as the process of dividing a digital image into multiple regions to enable image understanding and target extraction. In this study, an image differencing method was employed to segment the images and extract target fish bodies. Image differencing involves the subtraction of two images. Detection targets are obtained by calculating the differences between images, with the purpose of highlighting the regions with significant changes. The mathematical formulation of this operation is expressed as follows:

$$D(x, y) = |I(x, y) - B(x, y)| \quad (1)$$

where $D(x, y)$ denotes the differencing result at position (x, y) , $I(x, y)$ represents the pixel value of the fish image at position (x, y) , $B(x, y)$ indicates the pixel value of the background image at position (x, y) .

(2) Image enhancement. Image enhancement is defined as the process of processing image features, such as edges, contours, and contrast, through image processing techniques to improve image clarity or highlight useful information. Due to the presence of blurred target edges and significant noise caused by feces and feed residues in the differenced images, the effectiveness of image feature extraction can be easily affected. Therefore, the visual quality of images needs to be optimized through image enhancement techniques. In this study, image enhancement was achieved by applying linear transformation and histogram equalization. Linear transformation significantly improves image contrast and brightness, whereas histogram equalization increases the dynamic range of gray-level differences among pixels, thereby enhancing image clarity. After these image processing procedures, target regions and noise can be effectively distinguished in the images. The mathematical formulation of the linear transformation is given as follows:

$$I_{out}(x, y) = \alpha \cdot I_{in}(x, y) + \beta \quad (2)$$

where $I_{in}(x, y)$ denotes the pixel value of the input image at position (x, y) , $I_{out}(x, y)$ represents the pixel value of the output image at position (x, y) , α and β are defined as the gain coefficient and the bias term, respectively.

The histogram equalization formula is given in Equation (3).

$$s_k = T(r_k) = \sum \frac{n_j}{N} \cdot (L - 1) \quad (3)$$

In Equation (3), r_k denotes the k -th gray level of the input image, s_k represents the corresponding gray level of the output image, $T(r_k)$ is defined as the gray-level transformation

function, n_i indicates the number of pixels with gray level r_i , N denotes the total number of pixels in the image, and L represents the total number of gray levels in the image.

(3) Image binarization. Image binarization is defined as the process of converting a grayscale image into a binary image containing only black and white values. This operation facilitates the extraction of specific features and details of targets in the image, making them more prominent and easier to detect. In this study, image binarization was performed using an adaptive thresholding method. The calculation formula for the threshold is expressed as follows:

$$T(x, y) = \mu(x, y) + C \quad (4)$$

where $T(x, y)$ denotes the adaptive threshold at position (x, y) , $\mu(x, y)$ represents the mean pixel value within the local neighborhood centered at (x, y) , and C is defined as a constant offset. The binarization formula is expressed as follows:

$$P(x, y) = \begin{cases} 1, & \text{if } I(x, y) > T(x, y) \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

where $P(x, y)$ denotes the value of the binary mask at position (x, y) , and $I(x, y)$ represents the pixel value of the input image at position (x, y) .

(4) Contour extraction. After the binary image was obtained, although the target region had been preliminarily segmented, it still lacked structured information describing the target shape. To accurately extract spatial features that can be used for behavior analysis, the contour detection function `findContours()` was applied to detect fish contours in the image. Subsequently, the `drawContours()` function was used to visualize the detected contours, thereby achieving the extraction of fish target contours.

2.4. Fish Spatial Feature Extraction

Feature extraction is defined as the process of extracting representative information from an image to describe target objects. According to differences in research objects and objectives, corresponding features are selected and extracted to meet specific research requirements.

However, previous methods have mainly focused on local texture features or global statistical features in images. These features are unable to capture behavioral variations during the fish feeding process and fail to reflect the true characteristics of feeding behavior, which may lead to limited accuracy and reliability of assessment results. To address this issue, the spatial features of fish were extracted in this study to assess feeding intensity. Fish spatial features can be divided into individual features and group features. Individual features indicate the feeding motivation of each fish, whereas group features represent the overall behavioral tendency of the fish school from a global perspective. Such behavioral tendencies reflect the current feeding demand of the fish school. Therefore, spatial features were extracted from both individual and group perspectives to achieve an accurate assessment of fish feeding intensity in this study.

When extracting features that require determining the positional information of individuals, the centroid points of fishes were used to represent individual fish. For features directly related to the area, connected component information was adopted to characterize individuals. Since the target pixel regions are discrete in images, the centroid of the fish can be calculated using first-order moments. For a target region, its moments can be expressed as follows:

$$m_{pq} = \sum_{j=1}^N \sum_{i=1}^N i^p k^q f(i, j) \quad (6)$$

where m_{pq} denotes the $(p + q)$ -th order moment of the image, p and q are non-negative integers. i and j represent the pixel indices of the image:

$$x_c = \frac{m_{10}}{m_{00}}, y_c = \frac{m_{01}}{m_{00}} \quad (7)$$

where m_{00} denotes the zeroth-order moment of the image, which represents the area of the target region; and m_{10} and m_{01} are defined as the first-order moments of the image, corresponding to the sum of the products of the horizontal coordinates and their pixel values, and the sum of the products of the vertical coordinates and their pixel values, respectively.

2.4.1. Individual Feature Extraction

The individual features adopted in this study mainly include the average distance from individuals to the water surface, the average inter-individual distance, fish tilt, and the relationship between fish and feeding point. The extraction methods and calculation formulas for each feature are described as follows.

(1) Average distance from individuals to the water surface

In aquaculture, the throwing feeding method restricts fish to feeding by surfacing. During the feeding process, frequent upward and downward movements are observed as fish compete for dispersed feed. This phenomenon is closely related to the feeding demand of the fish. Therefore, the average distance from individuals to the water surface is regarded as an important indicator for measuring feeding intensity in aquaculture. However, this feature is difficult to obtain under a traditional top-view perspective. Consequently, a lateral viewing angle was adopted in this experiment to acquire this feature for the quantitative assessment of fish feeding intensity. The calculation formula for this feature is given as follows:

$$ADS = \frac{\sum_{1 \leq i \leq N} y_i}{N} \quad (8)$$

where ADS represents the average distance from individuals to the water surface, y_i denotes the distance from the i -th individual to the water surface, and N represents the number of fish bodies in the image.

(2) Average inter-individual distance

The average inter-individual distance is defined as the mean value of the distances among multiple individuals. When fish feed, the demand for feed intake drives fish schools to swim towards the feeding area, resulting in large-scale aggregation. In general, as feeding intensity increases, the average inter-individual distance tends to decrease. Therefore, the average inter-individual distance was selected in this study as one of the indicators for assessing the feeding intensity. The calculation method for this feature is given in Equation (9).

$$AID = \frac{2}{N(N-1)} \sum_{1 \leq i < j \leq N} d_{ij} \quad (9)$$

In Equation (9), AID represents the average inter-individual distance, d_{ij} denotes the distance between the i -th individual and the j -th individual, and N represents the number of fish bodies in the image.

(3) Fish tilt

The fish tilt is defined as the characteristic body orientation exhibited by fish during swimming. During feeding, fish are observed to raise their heads to feed near the water surface while their bodies remain submerged. As a result, fish tend to present a tilted posture when feeding. This posture differs from the horizontal swimming posture observed under none conditions and is considered an important indicator for distinguishing feeding

states of fish schools [24,25]. Therefore, this feature can be applied to feeding intensity assessment. In this study, the slope of the straight line corresponding to the fish body was selected to describe fish tilt. The straight line was obtained by connecting pixel points at the fish head and tail, and the slope of this line was calculated to characterize fish tilt. The calculation formula for the fish tilt is given in Equation (10).

$$FT = \left| \frac{y_2 - y_1}{x_2 - x_1} \right| \quad (10)$$

In Equation (10), FT represents the fish tilt, (x_1, y_1) represents the coordinates of the fish head, and (x_2, y_2) represents the coordinates of the fish tail.

(4) Relationship between fish and feeding point

Under feeding conditions, fish schools are observed to raise their heads and rapidly swim towards the feeding point. This movement pattern causes most individuals in the school to approach the feeding point, rise from positions below the feeding point, and then move upward to the water surface for feeding [26]. Based on this behavioral feature, a formula was developed in this study to describe the relationship between the distance of individuals to the feeding point and their inclination, as follows:

$$RFPR = \frac{\sum_{1 \leq i \leq N} (dist_i \times \sqrt{|r_i \times 2|})}{N} \quad (11)$$

where $RFPR$ represents the relationship between fish and feeding point, $dist_i$ represents the distance between the i -th individual and the feeding point, r_i denotes the radian value of the inclination angle of the line connecting the i -th individual and the feeding point, and N represents the number of fish bodies in the image.

2.4.2. Group Feature Extraction

The fish in the image are divided into two categories. If an individual is found to be attached to or obstructed by other individuals, these individuals are defined as aggregated individuals, and the corresponding region is considered an aggregation area. Otherwise, they are classified as non-aggregated individuals, and the corresponding area is regarded as a non-aggregation area. Based on this, the proportion of the aggregation area and the proportion of aggregated individuals were extracted. Additionally, group dispersity of fish was proposed to describe the dispersal state of the fish school.

(1) Proportion of the aggregation area

The proportion of the aggregation area is defined as the proportion of the area of the aggregation region to the total fish area in the image. When feeding, fish tend to exhibit aggregation and occlusion due to competition for food. Generally, the stronger the feeding intensity, the larger the aggregation area and its proportion. Therefore, the proportion of the aggregation area was extracted to assess the fish feeding intensity [27]. The calculation method for the proportion of the aggregation area is as follows:

$$PAA = \frac{Area_a}{Area} \times 100\% \quad (12)$$

where PAA represents the proportion of the aggregation area, $Area_a$ denotes the area of the aggregation region, and $Area$ represents the total fish area in the image.

(2) Proportion of aggregated individuals

The proportion of aggregated individuals is defined as the proportion of the number of aggregated individuals to the total number of fish in the image. The proportion of aggregated individuals serves as another indicator for measuring the degree of aggregation in fish schools. Unlike the proportion of the aggregation area (PAA), it can reduce assessment

errors that may arise when measuring area. By counting the number of aggregated and non-aggregated individuals, the degree of fish aggregation is calculated, and the proportion of aggregated individuals is obtained as one of the indicators for assessing fish feeding intensity. The calculation formula for the proportion of aggregated individuals is as follows:

$$PAI = \frac{m}{N} \times 100\% \quad (13)$$

where PAI represents the proportion of aggregated individuals, m denotes the number of aggregated individuals, and N represents the number of fish bodies in the image.

(3) Group dispersity

Group dispersity is defined as a group spatial feature used to describe the dispersal state of fish schools. The centroids of the fish in the image are treated as a set of points to construct a Delaunay triangulation. After the triangulation, the average perimeter of the triangles in the Delaunay network is taken as the measure of group dispersity to describe the dispersal state of the fish school. Under none conditions, fish schools tend to swim within the same horizontal range without aggregation. However, when feed is introduced, the fish school exhibits aggregation due to competition for food, resulting in aggregation of varying scales. As a result, the triangles in the Delaunay triangulation become smaller. In the Delaunay triangulation, the size of each triangle reflects the distance between the vertices, which indicates the dispersal degree of the fish school. Therefore, the group dispersity can be used as an evaluation indicator of fish feeding behavior. The calculation formula for it is as follows:

$$GD = \frac{\sum_{i=1}^t L_i}{t} = \frac{\sum_{i=1}^t (L_{i1} + L_{i2} + L_{i3})}{t} \quad (14)$$

where GD represents the group dispersity, t denotes the number of Delaunay triangles, L_i represents the perimeter of the i -th triangle, and L_{i1} , L_{i2} and L_{i3} denote the lengths of the three sides of the triangle. Generally, the smaller the value of GD , the lower the dispersal degree, and the stronger the corresponding feeding intensity. A partial display of the Delaunay triangulation is shown in Figure 4.



Figure 4. Triangular section of fish in different states. (a) Triangulation result of feeding ($GD = 461$); (b) triangulation result of un-feeding ($GD = 943$).

Figure 4a shows the triangulation results of the fish image under feeding conditions. The fish school exhibits large-scale aggregation and mutual occlusion. In this case, group dispersity is 461, indicating a low dispersity and suggesting that the fish school is in a feeding state with relatively high feeding intensity. Figure 4b shows the triangulation results under unfeeding conditions. The fish school mainly swims near the bottom, with a group dispersity of 943, indicating a higher dispersity and a lower feeding desire.

2.5. Fish Feeding Intensity Assessment Model Based on TabNet-DFWL

Based on the acquired spatial features of fish, such as distance, posture and aggregation degree, a classification model is constructed to assess fish feeding intensity. TabNet is a deep learning model specifically designed for tabular data, combining the advantages of both tree models and neural networks. It is known for its superior classification performance and strong interpretability [28]. Therefore, this model was selected in this study as the assessment model for assessing fish feeding intensity. However, since fish do not exhibit a clear “stop eating” behavior like mammals, their most significant behavioral change as they become satiated is reflected in the subtle attenuation of a series of key behavioral features, such as gradually moving away from the water surface, posture returning to horizontal, and a decrease in aggregation. These changes are very subtle, and their expression is relatively implicit, making it difficult to define their range. As a result, the model’s learning and decision-making may be influenced by these features, leading to potential assessment errors.

To address the issues, TabNet-DFWL was proposed in this study by introducing the DFWL. The TabNet-DFWL is able to perform adaptive feature weighting based on context, as well as identify and strengthen key features with strong discriminative power during training automatically. It solves the problem of model mis-assessment caused by the weak signals of satiety-related features and easy submergence in fish feeding intensity scenarios, as well as achieves high sensitivity in capturing subtle feature changes in fish feeding behavior and improves the model performance in the fish feeding intensity assessment task. The network structure of TabNet-DFWL is shown in Figure 5.

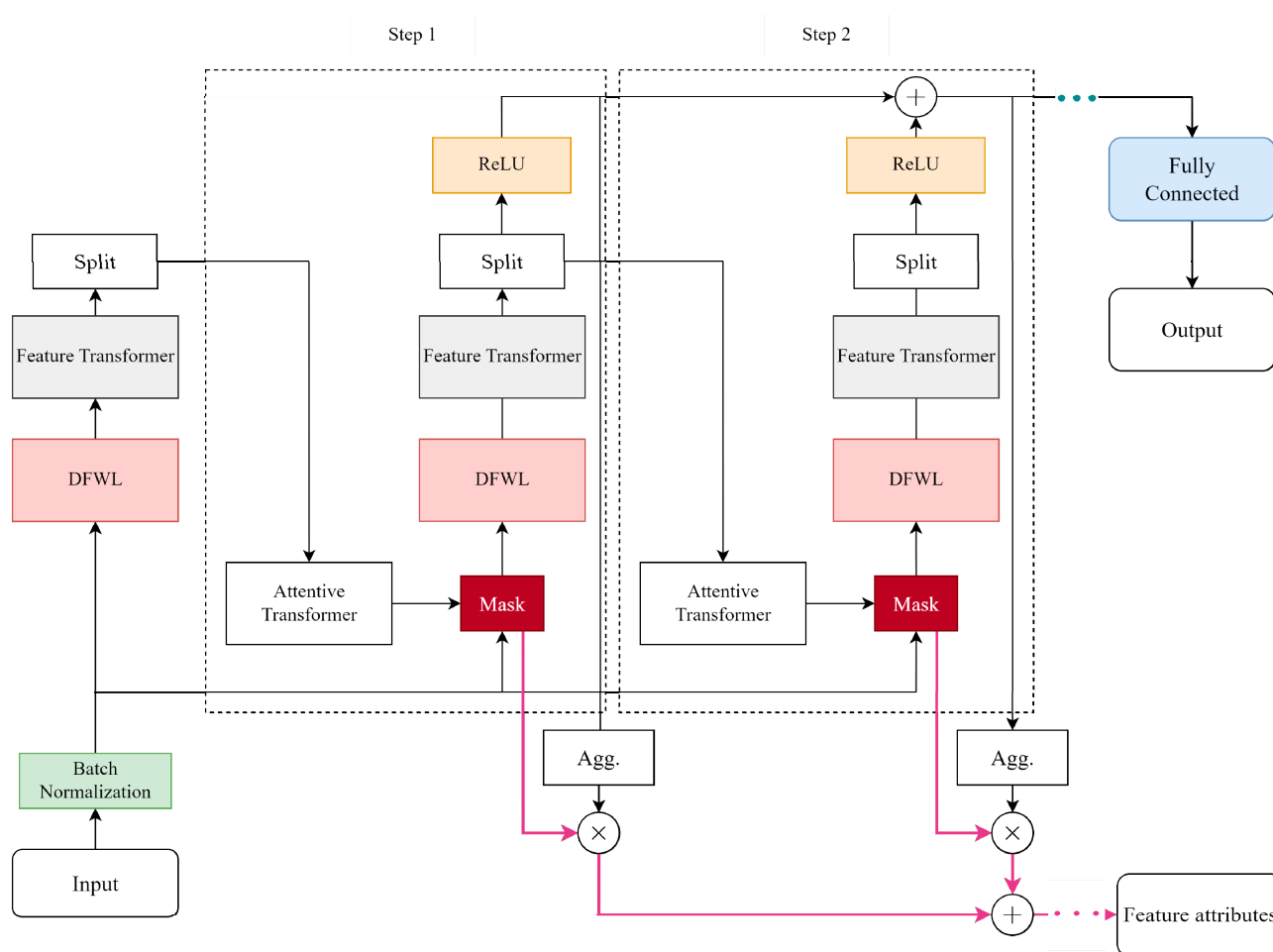


Figure 5. Network structure of TabNet-DFWL.

The DFWL proposed in this study is a self-learning mechanism that can dynamically adjust feature weights based on feature importance. Its core consists of a learnable parameter vector with the same dimensionality as the input features. This vector is constrained by the sigmoid function to lie within the range of (0, 1), representing the importance weight of each feature. For a given input feature matrix $X \in \mathbb{R}^{B \times D}$ (where B represents the batch size and D represents the feature dimension), the output of the DFWL is calculated by element-wise multiplication, the formula is as follows:

$$X_{weighted} = X \otimes \sigma(w) = X \otimes \sigma(\delta(W_1 \cdot \text{GAP}(X) + b_1)) \quad (15)$$

where $\otimes$ denotes element-wise multiplication, w represents the weight vector, σ is defined as the Sigmoid function $\sigma(x) = 1/(1 + e^{-x})$, W_1 and b_1 are defined as the weights and biases of the fully connected layer, $\text{GAP}(X)$ denotes global average pooling, and δ represents the ReLU activation function.

In the TabNet-DFWL network, fish spatial features are first processed through the DFWL, where initial weights are assigned to each feature based on its importance, and the initial weight vector is output. Next, the initial weight vector, along with the fish spatial features, is passed into the attentive transformer, and then the attentive transformer generates an attention mask based on feature importance, filtering out the most useful subset of fish spatial features for the current step. Then, the selected subset of fish spatial features is processed through the feature transformer, where deep abstract features of the fish spatial features are extracted to support decision-making. Finally, the output of the feature transformer is split by the split layer into two parts: one part is used to output the fish feeding intensity assessment result, and the other part serves as the input for feature selection in the next step. After multiple rounds of training, the assessment of fish feeding intensity is completed by aggregating the outputs of each step.

2.6. Experimental Setup

To ensure fair and reproducible comparisons between the proposed method and the baseline models, this subsection details the training methodology, hyperparameter settings, and implementation specifics for each comparative model. All baselines were evaluated under identical data splits, preprocessing steps, and evaluation metrics as the proposed method.

2.6.1. Hardware and Software Environment

All experiments were conducted on a workstation with an Intel Core i7-12650H CPU (16 cores, 2.3 GHz), 32 GB DDR5 RAM, and an NVIDIA GeForce RTX 4060 GPU (8 GB VRAM). The software environment included Windows 11, Python 3.9, PyTorch 2.6.0 with CUDA 12.4 and OpenCV 4.7, and scikit-learn 1.6.1.

2.6.2. Comparison Experiment Setup

The proposed method and all comparison models were implemented using the same software and hardware platform as described in Section 2.6.1.

To ensure a fair comparison, the proposed method and all comparison models were provided with exactly the same input features, namely the normalized fish spatial features, and underwent the same data cleaning preprocessing procedure to remove the missing values before being input to any model. The models compared include XGBoost [29], Light Gradient Boosting Machine (LGBM) [30], Random Forest (RF) [31], Multi-Layer Perceptron (MLP) [32], Decision Tree (DT) [33], and 1D Convolutional Neural Network (1D-CNN) [34]. The main parameter settings used in the experiment are listed in Table 3.

Table 3. Training parameter settings.

Model	Parameter Name	Parameter Value
XGBoost	n_estimators	500
	learning_rate	0.001
	max_depth	−1
	subsample	0.5
LGBM	num_leaves	31
	max_depth	−1
	e_estimators	500
	num_threads	0
	learning_rate	0.001
RF	n_estimators	500
	max_depth	None
	min_samples_split	2
	min_sample_leaf	1
MLP	input_size	7
	hidden_size	2
	output_size	4
	hidden_dim	128
DT	max_depth	None
	min_samples_split	2
	min_sample_leaf	1
	criterion	gini
1D-CNN	num_of_Convolution	2
	in_channels_1	1
	out_channels_1	16
	in_channels_2	16
	out_channels_2	32
	kernel_size	3
	max_epochs	500
	batch_size	1024
	learning_rate	0.001

The parameter settings of the comparison models are strictly executed according to Table 3, and default values are used for parameters not mentioned in the table. In addition, an early stopping strategy was adopted to prevent overfitting. The patience value is set to 10, which means that if no improvement is observed for 10 consecutive epochs, the training will stop. The system will automatically save the current optimal weight and terminate the training process when the fluctuation amplitude of the indicator is less than +0.15% for 10 consecutive epochs.

2.7. Evaluation Metrics

To evaluate the performance of the proposed TabNet-DFWL, accuracy, precision, recall, specificity and F1-Score are used to assess the model's performance. The definitions of these evaluation metrics are given as follows:

$$Accuracy = (TP + FN) / (TP + TN + FP + FN) \times 100\% \quad (16)$$

$$Precision = TP / (TP + FP) \times 100\% \quad (17)$$

$$Recall = TP / (TP + FN) \times 100\% \quad (18)$$

$$Specificity = TN / (FP + TN) \times 100\% \quad (19)$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (20)$$

where TP , TN , FP , and FN represent true positives, true negatives, false positives, and false negatives, respectively.

3. Results and Discussion

3.1. Image Pre-Processing Results

A series of image pre-processing methods was used in this study to extract fish bodies and draw their contours. The pre-processing results are shown in Figure 6. Figure 6a–e represent the fish images, image segmentation results, image enhancement results, image binarization results, and contour extraction results, respectively.

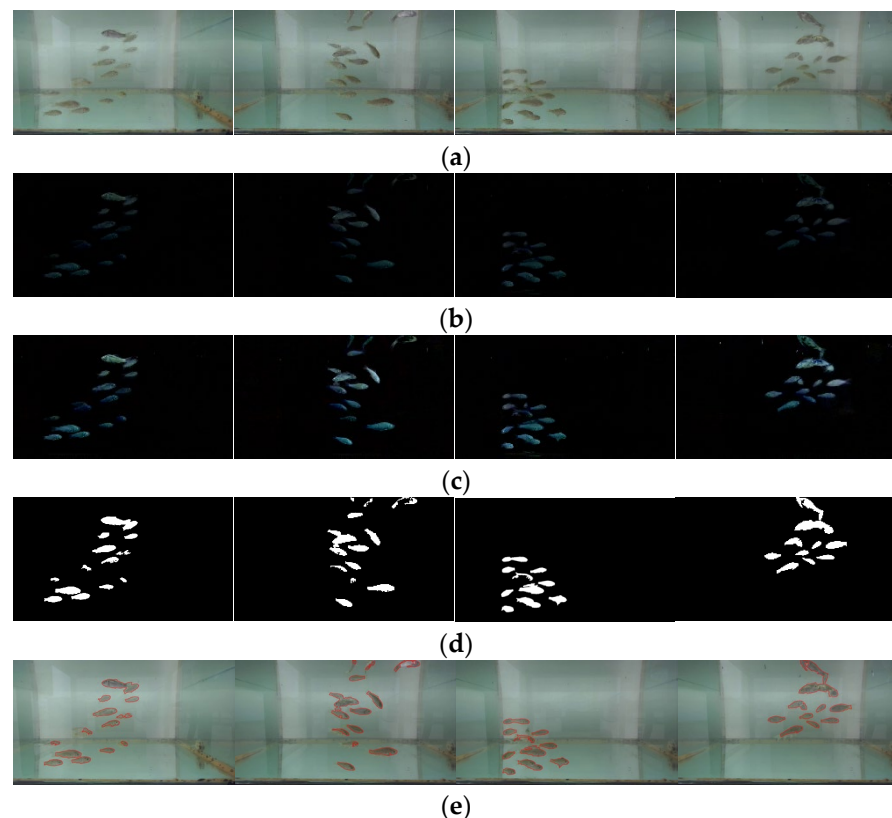


Figure 6. Image pre-processing results. (a) Fish images; (b) image segmentation results; (c) image enhancement results; (d) image binarization results; (e) contour extraction results.

As shown in Figure 6, the image pre-processing methods used in this study demonstrate good performance, with the fish targets and their contours being accurately extracted from the images. This is beneficial for the subsequent fish spatial feature extraction and feeding intensity assessment.

3.2. Fish Spatial Feature Extraction Results

Based on the precise segmentation of fish targets, a series of fish spatial features was extracted, and their results were converted into feature values. To intuitively display the numerical differences in features under different feeding intensities and highlight the effectiveness of the extracted features, boxplots of fish spatial features under various feeding intensities were plotted, as shown in Figure 7.

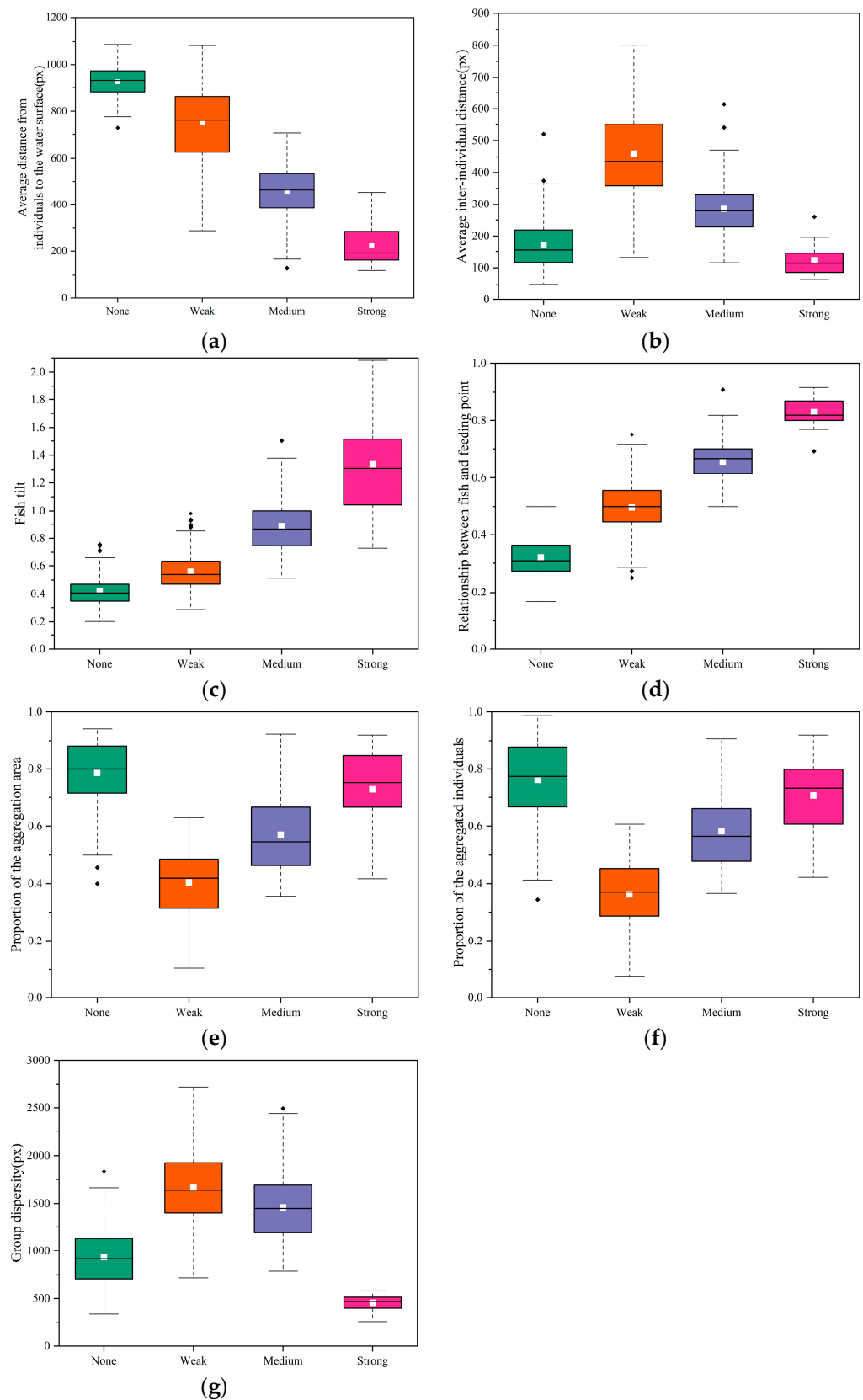


Figure 7. Boxplots of fish spatial features. (a) Average distance from individuals to the water surface (ADS); (b) average inter-individual distance (AID); (c) fish tilt (FT); (d) relationship between fish and feeding point (RFPR); (e) proportion of the aggregation area (PAA); (f) proportion of aggregated individuals (PAI); (g) group dispersity (GD).

As shown in Figure 7, the seven fish spatial features extracted in this study exhibit significant differences under different feeding intensities. Taking the average distance from individuals to the water surface as an example, it can be observed from the figure that as feeding intensity increases, the average distance from individuals to the water surface shows a gradual decreasing trend. Under the none state, fish tend to be concentrated in the deep and middle layers of the water, resulting in the maximum and most concentrated distribution of average distance from individuals to the water surface. In the weak state, some individuals rise to feed, while others remain submerged, causing the distribution of average distance from individuals to the water surface to be wider, and the overall value shows a decreasing trend compared to the none state. Under medium conditions, most individuals float to the water surface to feed, further reducing the overall value of the average distance from individuals to the water surface. In the strong state, individuals move to the water surface to compete for food, resulting in the smallest value of average distance from individuals to the water surface and a relatively concentrated overall distribution. This phenomenon aligns with the behavioral features of fish rising to feed under feeding demand, indicating that the average distance from individuals to the water surface can effectively distinguish between different levels of feeding intensity.

In addition, the other features extracted also exhibit significant differences under different feeding intensities, which proves that the extracted features can distinguish between different levels of feeding intensity too. Thus, the effectiveness of the feature extraction process is validated. These features provide good data input for the fish feeding intensity assessment model and are beneficial for the precise assessment of fish feeding intensity.

3.3. Fish Feeding Intensity Assessment Results

In this study, a method for fish feeding intensity assessment based on spatial features and the TabNet-DFWL is proposed. The DFWL introduced in this method improves the performance of the TabNet model and enables the evaluation of fish feeding intensity. The fish feeding intensity is classified into four categories: None, Weak, Medium, and Strong, for recognition and assessment. The fish feeding intensity assessment results are obtained, and the confusion matrix is shown in Figure 8.

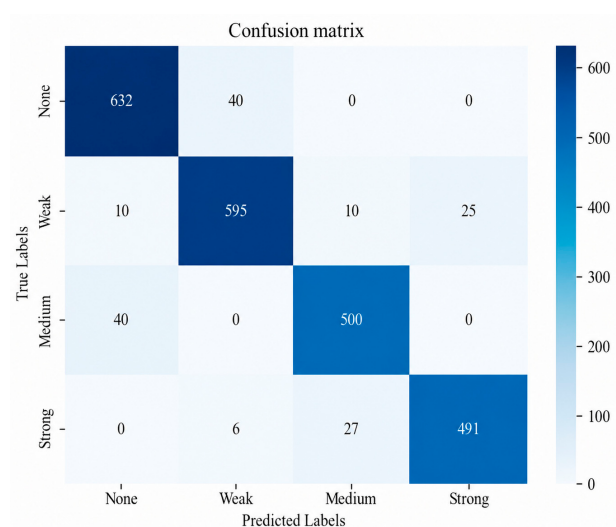


Figure 8. Confusion matrix of fish feeding intensity assessment.

The assessment accuracy based on TabNet-DFWL and spatial features reaches 95.96%. The precision, recall, specificity and F1-score for each category are shown in Table 4.

Table 4. Assessment results based on TabNet DFWL and spatial features.

Feeding Intensity	Precision (%) $\uparrow$ ¹	Recall (%) $\uparrow$	Specificity (%) $\uparrow$	F1-Score (%) $\uparrow$
None	92.67	94.05	97.31	93.35
Weak	92.82	92.97	96.15	92.89
Medium	93.11	92.59	99.60	92.85
Strong	95.16	93.70	99.56	94.42
Average	93.44	93.33	98.15	93.38

¹ The symbol “ $\uparrow$ ” in the table indicates that the higher the value of this metric is, the better the performance of the model will be.

As shown in Table 4, the method proposed in this study achieves good performance across the four metrics of precision, recall, specificity and F1-score. For the categories of none, weak, medium, and strong, all evaluation metrics exceed 92%, with the F1-score for the strong category being the highest at 94.42%. The recall for the none category is the highest, reaching 94.05%. In terms of overall averages, all metrics remain above 93.30%, indicating that the model has balanced recognition capabilities.

From the specific performance of each category, all metrics for identifying a strong state remain superior. This is due to the more distinct feature expression of fish behavior under strong intensity, such as the feature average distance from individuals to the water surface being significantly smaller than under other intensities. While the precision for identifying a none state is slightly lower, its recall reaches 94.05%, indicating that the model provides the best coverage for this category. The metrics for weak and medium behaviors are relatively close, reflecting the feature similarities between these two categories.

Overall, the method proposed in this study demonstrates stable assessment performance across all feeding intensity categories, with an average accuracy of 95.96%, precision of 93.44%, recall of 93.33%, specificity of 98.15% and F1-score of 93.38%. The fluctuation of metrics across categories is minimal, which indicates that the model has good generalization ability, thereby validating the effectiveness of the method.

Additionally, to assess the overall performance of the proposed TabNet-DFWL model, the Receiver Operating Characteristic (ROC) curve and Precision–Recall (PR) curve were generated, as shown in Figure 9.

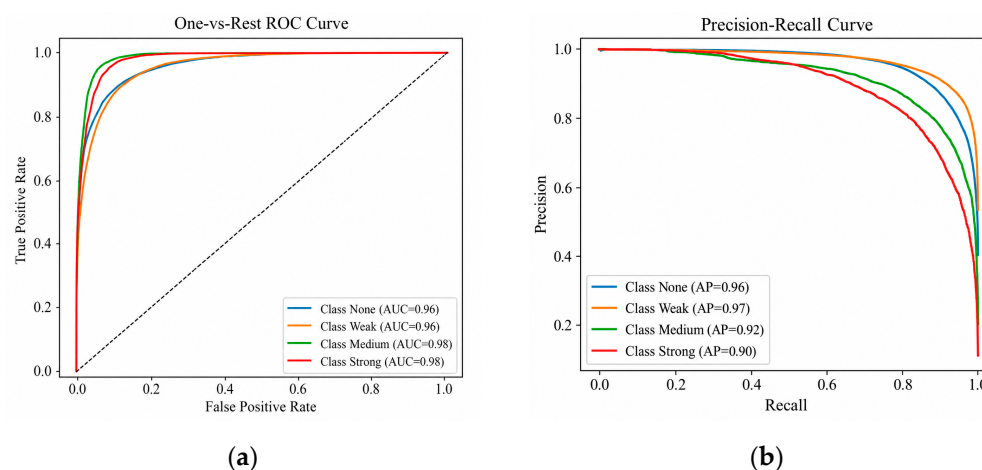


Figure 9. Performance analysis of fish feeding intensity assessment based on TabNet-DFWL. (a) Receiver operating characteristic of TabNet-DFWL; (b) precision–recall curve of TabNet-DFWL.

The area under the ROC curve is referred to as Area Under Curve (AUC), which is an important performance metric for evaluating the quality of a model. A higher AUC indicates better performance in classification tasks. The area under the PR curve is referred to as the Average Precision (AP) value, and typically, a higher AP value indicates

better classification performance. As shown in Figure 9a, the AUC scores for the none and weak categories are 0.96, while the AUC scores for the medium and strong categories are 0.98. Meanwhile, as shown in Figure 9b, the AP values for the none, weak, medium, and strong categories are 0.96, 0.97, 0.92, and 0.90, respectively. All these metrics demonstrate that the TabNet-DFWL model proposed has superior performance in classification assessment tasks.

3.4. Comparison Results

3.4.1. Comparison Results of Different Spatial Features

TabNet was used as the baseline model for fish feeding intensity assessment in this study. This model not only demonstrates superior classification performance but also offers advantages in feature selection and global feature visualization. Based on this model, the DFWL was introduced, and by aggregating the feature selection masks of the training samples, the global feature importance ranking of the model was obtained. The result of the feature importance ranking is shown in Figure 10. This visualization clearly reveals the key behavioral features that influence fish feeding intensity, enhancing the transparency of model decisions and providing reliable support for biological interpretation.

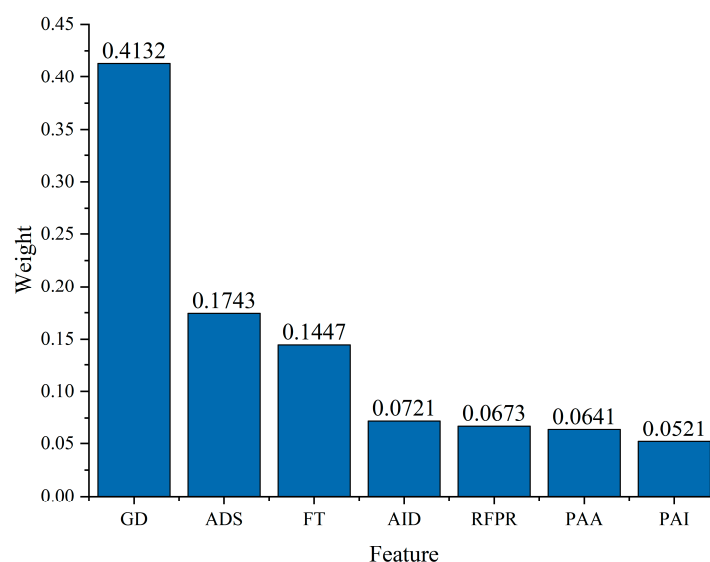


Figure 10. Result of the feature importance ranking.

As shown in Figure 10, the feature weights obtained using the DFWL algorithm are used to rank the features, with the importance ranking of the features being group dispersity (GD), average distance from individuals to the water surface (ADS), fish tilt (FT), average inter-individual distance (AID), relationship between fish and feeding point (RFPR), proportion of the aggregation area (PAA) and proportion of aggregated individuals (PAI). To validate the effectiveness of these features, a comparison of feature forward selection was conducted, sequentially adding the feature with the highest importance to the feature set, and recording the evaluation metrics after feature addition. The results of these experiments are shown in Table 5.

From Table 5, it can be seen that the comparison of feature forward selection has achieved the expected results. After adding each feature to the input model's feature set based on its importance, the evaluation metrics of the model steadily improved. This shows that all features make a positive contribution to the model. The monotonic increase in performance, therefore, confirms the effectiveness of feature selection, meaning that each feature carries unique information related to feeding intensity.

Table 5. Experiment results of ablation of feature subsets.

Feature Subset Used	Accuracy (%) ↑	Average Precision (%) ↑	Average Recall (%) ↑	Average Specificity (%) ↑	Average F1-Score (%) ↑
Proposed method	95.96	93.44	93.33	98.15	93.38
GD	69.77	68.74	65.78	89.86	67.22
GD, ADS	74.65	71.16	69.75	92.36	70.45
GD, ADS, FT	88.48	87.19	87.50	93.18	87.34
GD, ADS, FT, AID	91.46	90.71	89.68	95.80	90.19
GD, ADS, FT, AID, RFPR	93.40	91.98	91.46	96.89	91.72
GD, ADS, FT, AID, RFPR, PAA	94.62	92.51	91.98	97.72	92.25

3.4.2. Comparison Results of the Effectiveness of DFWL

In the proposed TabNet-DFWL method, the DFWL is introduced to address the issue of performance degradation caused by the implicit expression of satiety in fish feeding states. To validate the performance of the improved TabNet-DFWL, it is compared with the original TabNet network. The comparison results are shown in Figure 11.

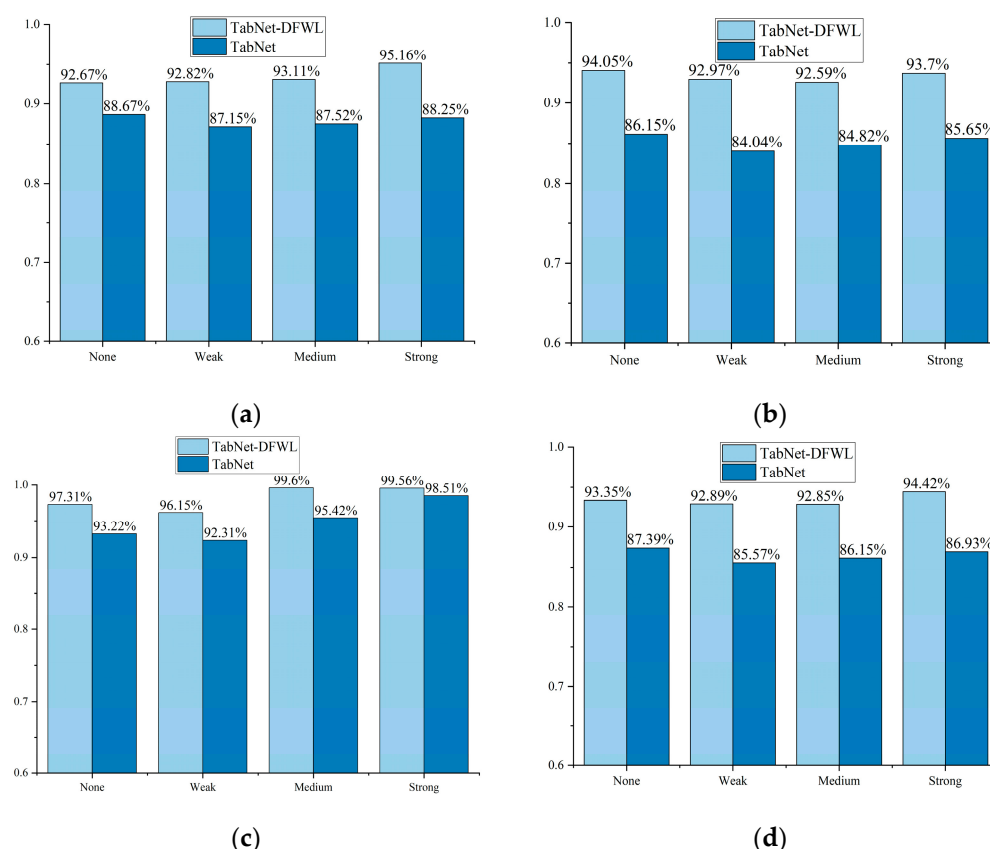


Figure 11. Comparison of the effectiveness of DFWL on various categories. (a) Precision comparison results; (b) recall comparison results; (c) specificity comparison results; (d) F1-Score comparison results.

The accuracy of the TabNet-DFWL model is 95.96%, while the accuracy of the original TabNet model is 91.70%, representing an improvement of 4.65%. Additionally, the comparison of evaluation metrics for each category in Figure 11 shows that the proposed method

demonstrates superior performance. This improvement is attributed to the DFWL, which can learn feature importance, thereby enhancing model performance.

The average values of the evaluation metrics for the four categories are also recorded in this study. The results are shown in Figure 12.

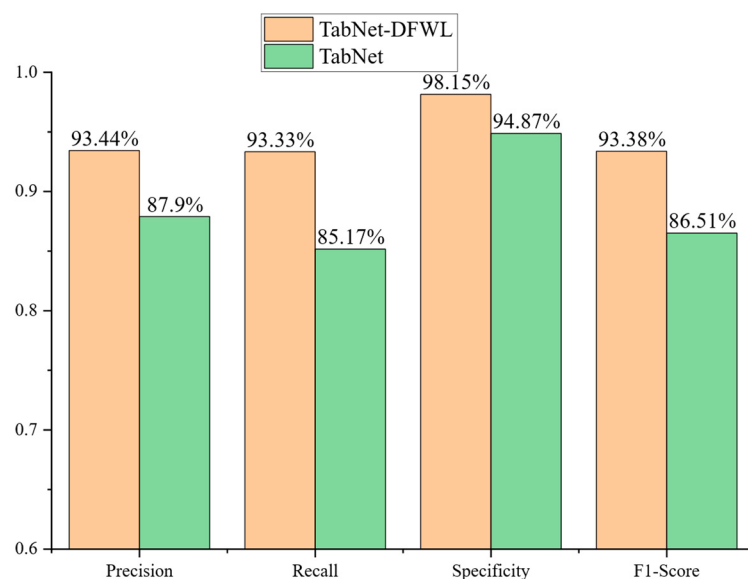


Figure 12. Evaluation metric average values of comparison DFWL.

As shown in Figure 12, in summary, the average values of all evaluation metrics for TabNet-DFWL are higher than those of the original TabNet. Therefore, the DFWL proposed in this study can address the issue caused by the implicit expression of satiety in fish feeding states, enhancing the model recognition performance.

Additionally, the loss and valid accuracy changes in TabNet and TabNet-DFWL on the validation set were recorded and analyzed in this study. The loss and valid accuracy curves are shown in Figure 13.

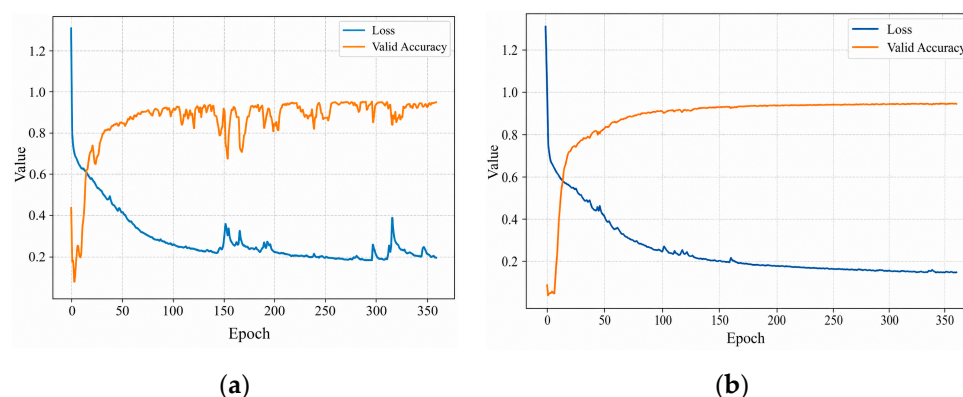


Figure 13. Loss and valid accuracy curves of TabNet and TabNet-DFWL on the validation set. (a) The loss and valid accuracy curves of TabNet; (b) the loss and valid accuracy curves of TabNet-DFWL.

As shown in Figure 13, the introduced DFWL significantly improves the model convergence stability and feeding intensity assessment performance. From Figure 13a, it can be observed that the Loss curve of the original TabNet model shows significant fluctuations, especially after 150 epochs, where instability and oscillations persist. At the same time, the Valid Accuracy frequently fluctuates within the range of 0.80 to 0.95. In contrast, after introducing the DFWL (Figure 13b), the Loss curve of the TabNet-DFWL model declines

more smoothly and monotonically, stabilizing around 200 epochs. The validation accuracy increases rapidly and remains consistently above 0.95, with a significant reduction in fluctuations. It indicates that the model's training process is more robust, and the convergence is improved.

The main reason for the performance improvement is that the DFWL can dynamically adjust the feature weights during the model training process based on feature importance and stability. By strengthening the discriminative key features, the model can focus on discriminative feature information, thereby improving the overall performance and robustness of the model.

In summary, the improved TabNet-DFWL model outperforms the original model in terms of convergence speed, training stability and validation performance. This result indicates that the DFWL effectively mitigates the issue caused by the implicit expression of satiety in fish feeding states, significantly enhancing the model learning stability and overall recognition performance.

3.4.3. Comparison Results of Different Classification Models

To validate the performance of the proposed TabNet-DFWL model, it is compared with several other classification models. The models compared include XGBoost [29], LGBM [30], RF [31], MLP [32], DT [33], and 1D-CNN [34]. The comparison results are shown in Table 6.

Table 6. Comparison results of different classification models.

Assessment Model	Accuracy (%) ↑	Average Precision (%) ↑	Average Recall (%) ↑	Average Specificity (%) ↑	Average F1-Score (%) ↑
Proposed method	95.96	93.44	93.33	98.15	93.38
XGBoost	80.39	81.07	74.66	91.45	77.73
LGBM	84.44	84.91	81.75	93.09	83.30
RF	86.40	86.08	84.32	94.22	85.19
MLP	88.93	88.99	87.56	94.95	88.27
DT	90.99	89.40	89.21	96.72	89.30
1D-CNN	73.56	71.54	65.36	88.59	68.31

As shown in Table 6, the TabNet-DFWL method proposed in this study outperforms other comparison methods while maintaining a model size of only 1.78MB. The accuracy, average precision, recall, specificity and F1-Score of the TabNet-DFWL were 95.56%, 93.44%, 93.33%, 98.15% and 93.38%, respectively. Compared with XGBoost, these evaluation metrics of TabNet-DFWL increase by 19.37%, 15.26%, 25.01%, 7.33% and 20.13%, respectively. Compared with LGBM, these evaluation metrics increase by 13.64%, 10.05%, 14.17%, 5.44% and 12.10%, respectively. Compared with RF, these evaluation metrics increase by 11.06%, 8.55%, 10.69%, 4.17% and 9.61%, respectively. Compared with MLP, these evaluation metrics increase by 7.91%, 5.00%, 6.59%, 3.37% and 5.79%, respectively. Compared with DT, these evaluation metrics increase by 5.46%, 4.52%, 4.62%, 1.48% and 4.57%, respectively. Compared with 1D-CNN, these evaluation metrics increase by 30.45%, 30.61%, 42.79%, 10.79% and 36.70%, respectively.

Additionally, to provide a comprehensive comparison of the overall metrics of different classification models, the model size, number of parameters, and test time metrics are also compared. The comparison results are shown in Figure 14.

As shown in Figure 14, the model size of the TabNet-DFWL proposed in this study is 1.78 MB, which is only slightly larger than XGBoost and much smaller than RF. This makes the model more suitable for deployment and use on devices. Furthermore, we tested the model on a dataset of 2376 entries from the test set, obtaining a total testing time of 0.30888 s, with an average assessment time of 0.00013 s per data point, enabling real-time feeding intensity assessment and guidance for feeding.

Therefore, taking model accuracy, model size, and speed into consideration, the TabNet-DFWL model proposed in this study emerges as the better choice, providing a more suitable solution to meet the needs of the aquaculture industry.

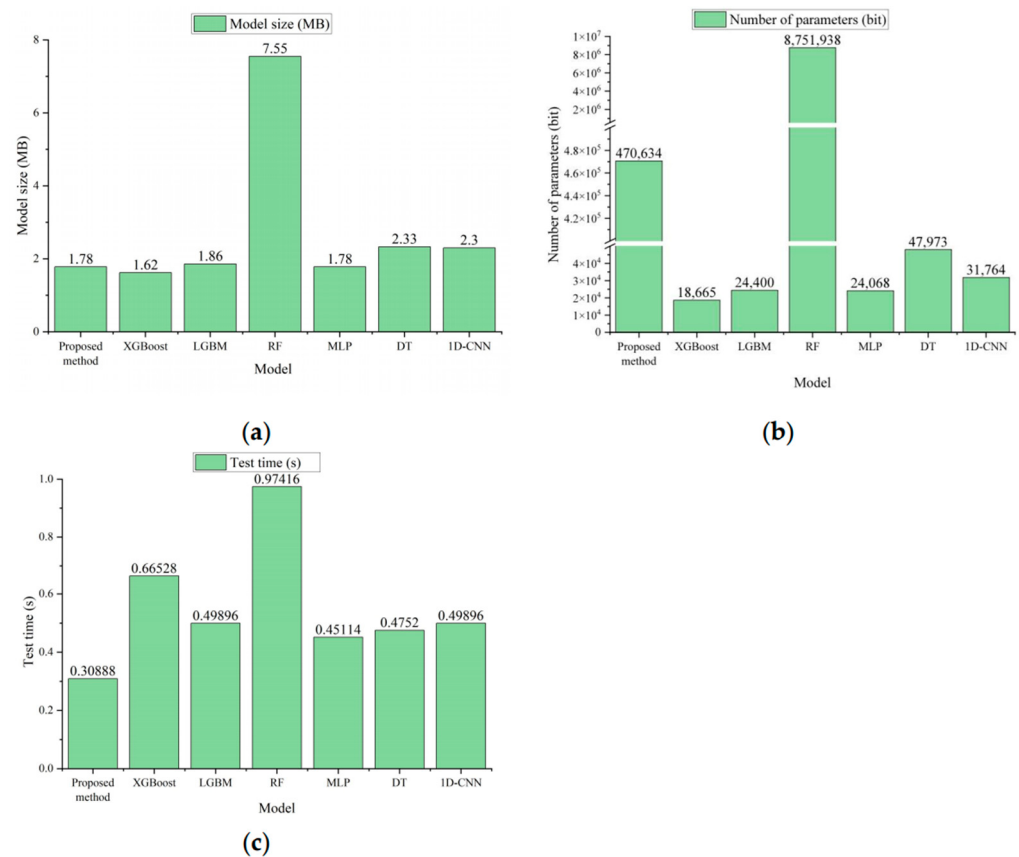


Figure 14. Comparison results of various parameters of the classification model. (a) Model size comparison results; (b) number of parameters comparison results; (c) test time comparison results.

3.4.4. Comparison Results with Other Thesis Methods

To further validate the effectiveness of the fish feeding intensity assessment method proposed in this study, it is compared with the methods from [9,35,36]. The comparison results are shown in Table 7.

Table 7. Comparison results with other thesis methods.

Method	Accuracy (%) ↑	Average Precision (%) ↑	Average Recall (%) ↑	Average Specificity (%) ↑	Average F1-Score (%) ↑
Proposed method	95.96	93.44	93.33	98.15	93.38
Huang et al. [35]	44.44	46.93	43.84	71.22	45.33
Yuan and Zhu [9]	51.61	53.34	53.80	79.70	53.57
Yu et al. [36]	83.27	81.14	81.52	93.47	81.33

As shown in Table 7, the accuracy of the method proposed in this study is 95.96%, which represents an improvement of 115.93%, 85.93%, and 15.23% compared with the 44.44%, 51.61%, and 83.27% of the comparison methods, respectively. The average precision is 93.44%, which is significantly higher than the 46.93%, 53.34%, and 81.14% of the comparison methods, representing improvements of 99.11%, 75.18%, and 15.16%, respectively. The average recall is 93.33%, which represents an increase of 112.89%, 73.48%, and 14.49% compared with the 43.84%, 53.80%, and 81.52% of the comparison methods, respectively. The specificity is 98.15%, which is an improvement of 37.81%, 23.15%, and 5.01%

compared with the 71.22%, 79.70%, and 93.47% of the comparison methods, respectively. For the F1-Score, the method proposed in this study achieves the highest value of 93.38%, which is an increase of 106.00%, 74.31%, and 14.82% compared with the 45.33%, 53.57%, and 81.33% of the comparison methods, respectively.

4. Conclusions and Future Work

To address the issue of unreliable and untrustworthy results in existing fish feeding intensity assessments, a method for fish feeding intensity assessment based on spatial features and the TabNet-DFWL is proposed in this study, enabling the accurate assessment of fish feeding intensity. This method has significant implications for achieving precise feeding in the aquaculture industry, promoting healthy fish growth, reducing farming costs, and improving productivity. A series of spatial features highly correlated with feeding behavior is proposed, which can enhance the effectiveness of feature representation and improve the reliability of feeding intensity assessment. The DFWL is proposed to automatically strengthen the weights of key features within the TabNet model, addressing the degradation in assessment performance caused by the implicit expression of fish satiety and improving the model performance when processing fish feeding data. A fish feeding intensity assessment method based on the TabNet-DFWL network is proposed to avoid the black-box risk of conventional deep learning models, enhance model interpretability and credibility, and provide a reliable basis for precision feeding in aquaculture.

The method proposed in this study was tested on a real fish feeding dataset and achieved promising experimental results. The assessment accuracy was 95.96%, with an average precision of 93.44%, average recall of 93.33%, average specificity of 98.15%, and average F1-score of 93.38%. Compared with the algorithms XGBoost, LGBM, RF, MLP, DT, and 1D-CNN, the assessment accuracy was increased by 19.37%, 13.64%, 11.06%, 7.91%, 5.46%, and 30.45%, respectively, demonstrating that the proposed method can achieve accurate fish feeding intensity assessment.

Although the proposed method has achieved promising results, it still has some limitations. For example, fish actually move in three-dimensional space, but current methods mainly extract two-dimensional information from side-view images. This may limit the completeness of the description of fish spatial behavior, especially in terms of vertical movement and depth-related information. Therefore, in future work, we plan to introduce three-dimensional spatial features to obtain more comprehensive spatial information of fish during feeding. By incorporating depth information, the spatial behavior of fish schools can be described more accurately, which is expected to further improve the accuracy and reliability of fish feeding intensity assessment.

Author Contributions: Conceptualization, L.Z.; methodology, L.Z., S.Z. and Z.L.; software, Z.L. and H.Y.; validation, S.Z. and W.N.; formal analysis, S.Z. and Z.L.; investigation, L.Z. and Y.L.; resources, L.Z.; data curation, Z.L. and W.N.; writing—original draft preparation, L.Z. and Z.L.; writing—review and editing, L.Z. and S.Z.; visualization, S.Z. and H.Y.; supervision, Y.L.; project administration, L.Z.; funding acquisition, L.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 32303070, the Jiangsu Provincial Agricultural Science and Technology Independent Innovation Fund, grant number CX(24)3064 and the China Postdoctoral Science Foundation, grant number 2023M732995.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: For data supporting the results of this study, please contact the corresponding author. Due to privacy, these data have not been made public.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhou, C.; Xu, D.; Lin, K.; Sun, C.; Yang, X. Intelligent feeding control methods in aquaculture with an emphasis on fish: A review. *Rev. Aquac.* **2018**, *10*, 975–993. [\[CrossRef\]](#)
2. Wu, M.; Wang, L.; Huang, T.; Pang, H.; Liu, S.; Cui, M.; Xu, L. STCA-MobileViTv3: A spatiotemporal collaborative attention network for fish feeding intensity recognition in underwater videos. *Aquac. Eng.* **2026**, *113*, 102695. [\[CrossRef\]](#)
3. Zhang, L.; Liu, Z.; Zheng, Y.; Li, B. Feeding intensity identification method for pond fish school using dual-label and MobileViT-SENet. *Biosyst. Eng.* **2024**, *241*, 113–128. [\[CrossRef\]](#)
4. Zhang, L.; Wang, J.; Li, B.; Liu, Y.; Zhang, H.; Duan, Q. A MobileNetV2-SENet-based method for identifying fish school feeding behavior. *Aquac. Eng.* **2022**, *99*, 102288. [\[CrossRef\]](#)
5. Atoum, Y.; Srivastava, S.; Liu, X. Automatic feeding control for dense aquaculture fish tanks. *IEEE Signal Process. Lett.* **2014**, *22*, 1089–1093. [\[CrossRef\]](#)
6. Oddsson, G.V. A definition of aquaculture intensity based on production functions—The aquaculture production intensity scale (APIS). *Water* **2020**, *12*, 765. [\[CrossRef\]](#)
7. Zhang, C.; Chen, M.; Feng, G.; Guo, Q.; Zhou, X.; Shi, G.; Chen, G. A fish feeding behavior detection method based on multi-feature fusion and machine learning. *J. Hunan Agric. Univ.* **2019**, *45*, 97–102.
8. Chen, M.; Zhang, C.; Feng, G.; Chen, X.; Chen, G.; Wang, D. Intensity assessment method of fish feeding activities based on feature weighted fusion. *Trans. Chin. Soc. Agric. Mach.* **2020**, *51*, 245–253. [\[CrossRef\]](#)
9. Yuan, C.; Zhu, R. Research on fish school feeding behavior detection based on KPCA multi-feature fusion support vector machine. *Aquaculture* **2020**, *41*, 17–21.
10. Wang, Y.; Yu, X.; Liu, J.; Zhao, R.; Zhang, L.; An, D.; Wei, Y. Research on quantitative method of fish feeding activity with semi-supervised based on appearance-motion representation. *Biosyst. Eng.* **2023**, *230*, 409–423. [\[CrossRef\]](#)
11. Dong, Y.; Zhao, S.; Wang, Y.; Cai, K.; Pang, H.; Liu, Y. An integrated three-stream network model for discriminating fish feeding intensity using multi-feature analysis and deep learning. *PLoS ONE* **2024**, *19*, e0310356. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Liu, S.; Liu, X.; Zou, H. Assessment method for feeding intensity of fish schools using MobileViT-CoordAtt. *Fishes* **2025**, *10*, 253. [\[CrossRef\]](#)
13. Wang, G.; Muhammad, A.; Liu, C.; Du, L.; Li, D. Automatic recognition of fish behavior with a fusion of RGB and optical flow data based on deep learning. *Animals* **2021**, *11*, 2774. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Zhang, S.; Yao, M.; Zhao, J.; Liu, X.; Wang, H. A multi-stage augmented multimodal interaction network for fish feeding intensity quantification. *arXiv* **2025**, arXiv:2506.14170. [\[CrossRef\]](#)
15. Iqbal, U.; Li, D.; Qureshi, M.F.; Mushtaq, Z.; Rehman, H.A.U. LightHybridNet-Transformer-FFIA: A hybrid Transformer based deep learning model for enhanced fish feeding intensity. *Aquac. Eng.* **2025**, *111*, 102604. [\[CrossRef\]](#)
16. Wang, Z.; Wen, S.; Yang, L.; Mei, Y.; Yang, Q.; Li, Y. A dual-stream spatiotemporal fusion method for fish school feeding intensity identification. *Aquac. Int.* **2026**, *34*, 33. [\[CrossRef\]](#)
17. Varshney, P.; Lucieri, A.; Balada, C.; Dengel, A.; Ahmed, S. Discovering concept directions from diffusion-based counterfactuals via latent clustering. *Pattern Recognit. Lett.* **2025**, *in press*. [\[CrossRef\]](#)
18. McDonnell, K.; Sheehan, B.; Murphy, F. Bridging transparency in insurance claims prediction: A comparative study of explainable AI and traditional linear models using vehicle telematics data. *Technol. Forecast. Soc. Change* **2026**, *223*, 124418. [\[CrossRef\]](#)
19. Takahashi, Y.; Komeyama, K. Development of a feeding simulation to evaluate how feeding distribution in aquaculture affects individual differences in growth based on the fish schooling behavioral model. *PLoS ONE* **2023**, *18*, e0280017. [\[CrossRef\]](#)
20. Kong, H.; Wu, J.; Liang, X.; Xie, Y.; Qu, B.; Yu, H. Conceptual validation of high-precision fish feeding behavior recognition using semantic segmentation and real-time temporal variance analysis for aquaculture. *Biomimetics* **2024**, *9*, 730. [\[CrossRef\]](#)
21. Feng, G.; Kan, X.; Chen, M. A multi-step image pre-enhancement strategy for fish feeding behavior analysis using EfficientNet. *Appl. Sci.* **2024**, *14*, 5099. [\[CrossRef\]](#)
22. Xiao, Y.; Huang, L.; Zhang, S.; Bi, C.; You, X.; He, S.; Guan, J. Feeding behavior quantification and recognition for aquaculture. *Appl. Anim. Behav. Sci.* **2025**, *285*, 106588. [\[CrossRef\]](#)
23. Eriksen, M.S.; Færevik, G.; Kittilsen, S.; McCormick, M.I.; Damsgård, B.; Braithwaite, V.A.; Braastad, B.O.; Bakken, M. Stressed mothers—Troubled offspring: A study of behavioural maternal effects in farmed *Salmo salar*. *J. Fish. Biol.* **2011**, *79*, 575–586. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Lai, C.L.; Tsai, S.T.; Chiu, Y.T. Analysis and comparison of fish posture by image processing. In *2010 International Conference on Machine Learning and Cybernetics*; IEEE: Piscataway, NJ, USA, 2010; pp. 2559–2564. [\[CrossRef\]](#)

25. Zheng, J.; Zhao, F.; Lin, Y.; Chen, Z.; Gan, Y.; Pang, B. Evaluation of fish feeding intensity in aquaculture based on near-infrared depth image. *J. Shanghai Ocean Univ.* **2021**, *30*, 1067–1078.
26. Irmak, E.; Ertas, A.H. A review of robust image enhancement algorithms and their applications. In *2016 IEEE Smart Energy Grid Engineering (SEGE)*; IEEE: Piscataway, NJ, USA, 2016; pp. 371–375. [\[CrossRef\]](#)
27. Zhou, C.; Xu, D.; Lin, K.; Chen, L.; Zhang, S.; Sun, C.; Yang, X. Evaluation of fish feeding intensity in aquaculture based on near-infrared machine vision. *Smart Agric.* **2019**, *1*, 76–84. [\[CrossRef\]](#)
28. Helforouh, Z.; Shojaie, M.; Arghamiri, S. Metaheuristic-optimized TabNet ensemble for accurate and interpretable obesity classification. *Swarm Evol. Comput.* **2025**, *98*, 102128. [\[CrossRef\]](#)
29. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794. [\[CrossRef\]](#)
30. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T. LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems 30*; NeurIPS: San Diego, CA, USA, 2017; pp. 3146–3154.
31. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
32. Suasono, Z.S.; Setiawardhana, S.; Gunawan, A.I.; Winarno, I. Performance evaluation of water quality for shrimp farming using deep learning classification. *Aquac. Eng.* **2026**, *112*, 102648. [\[CrossRef\]](#)
33. Quinlan, J.R. Induction of decision trees. *Mach. Learn.* **1986**, *1*, 81–106. [\[CrossRef\]](#)
34. Kiranyaz, S.; Avci, O.; Abdeljaber, O.; Ince, T.; Gabbouj, M.; Inman, D.J. 1D convolutional neural networks and applications: A survey. *Mech. Mech. Syst. Signal Process.* **2021**, *151*, 107398. [\[CrossRef\]](#)
35. Huang, Z.; He, J.; Song, X. Recognition and quantification of fish feeding behavior based on motion feature of fish body and image texture. *Period. Ocean Univ. China* **2022**, *52*, 32–41. [\[CrossRef\]](#)
36. Yu, X.; Wang, Y.; An, D.; Wei, Y. Identification methodology of special behaviors for fish school based on spatial behavior characteristics. *Comput. Electron. Agric.* **2021**, *185*, 106169. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.