Modified kinetic energy feature-based graph convolutional network for fish appetite grading using time-limited data in aquaculture

Dan Wei[1], Baimin Ji[1], Haijun Li[1], Songming Zhu[1,2], Zhangying Ye[1,2]* and Jian Zhao[1]*

[1]College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, China,
[2]Ocean Academy, Zhejiang University, Zhoushan, China

Feed has the greatest impact on the carbon footprint of the aquaculture, and also determines the water quality in aquaculture to a great extent. Making appropriate feeding control strategies is one of the most effective ways to promote cleaner production as well as fish welfare in aquaculture. Reliable and accurate fish appetite grading especially based on time-limited data is a prerequisite for achieving high-precision and reasonable feeding control in practical production. To date, however, few efforts have been done on this challenge. For these, regarding *Micropterus salmoides* as the experimental fish, a novel and practical method, based on a modified kinetic energy feature-based graph convolutional network (GCN), was developed in this study. First, graphs were constructed based on the extracted modified kinetic energy features and their temporal correlation. Then, with the help of a series of the convolution and global pooling operations, a GCN model was customized based on the constructed graphs. Following this, the customized GCN model was enriched by the self-attention pooling mechanism and customized network structure. Results show that the proposed GCN-based approach outperforms other typical state-of-the-art methods in fish appetite grading, and the grading accuracy obtained here could be 98.60% using only the first 4.2 seconds as well as the first 8.3 seconds of input data, which is not much different from that (98.89%) using full-length (25 second-long) input data. What's more, compared to the recurrent neural network (RNN)-based method which performance is closest to our method, the space complexity of the proposed approach here can better satisfy the requirements of real aquaculture, in which the quantity of the trainable parameters here is only 6.4% ~ 31.8% of the RNN-based method. In summary, the proposed modified kinetic energy feature-based GCN approach is favorable for the appetite grading of fish like *Micropterus salmoides* with time-limited data, which is a

promising approach in dealing with feeding control tasks and alleviating the water environmental burden in aquaculture.

Introduction

Aquaculture production has grown rapidly in the past few decades, thereinto, 52% consumption of the aquatic products worldwide in 2018 was provided by aquaculture (FAO, 2020). In the meantime, the concept of cleaner production and fish welfare is being emphasized due to its indispensable role in quality and yield of the aquatic products (Luna et al., 2019). Feeding is of great importance in managing aquaculture tasks, where the cost of feed is around 30%~70% of the total production costs (Føre et al., 2011; Atoum et al., 2015; Zhou et al., 2018). Underfeeding impedes fish growth, thus the strategy of overfeeding is commonly adopted in practical production to satisfy the nutritional needs of fish. However, overfeeding leads to left-over feed, which results in not only the extra cost, but also the poor water quality (Barraza-Guardado et al., 2014; Jescovitch et al., 2018; Zhao et al., 2019) and an extra load on water treatment equipment (Chang et al., 2005). Moreover, previous research has shown that feed has the greatest impact on the carbon footprint of the aquaculture (Luna et al., 2019). As a consequence, optimization of the feeding control is a prime consideration to realize cleaner production and promote fish welfare in aquaculture, especially in intensive modes.

Precise representation of fish appetite is the guarantee for the accurate feeding control. Relevant studies have shown that fish feeding behavior has significant advantages in fish appetite representation (Parra et al., 2018; Li et al., 2020; An et al., 2021), compared with many other mediums such as residual feed (Atoum et al., 2015; Wang et al., 2022) and water quality (Zhao et al., 2019; Zhao et al., 2020a). Until now, many works have been done to optimize the feeding based on fish school behavior. For example, the infrared photoelectric senor was used to capture the gathering behaviors of eels for the feeding control in indoor intensive aquaculture systems (Chang et al., 2005). Liu et al. (2014) proposed a computer vision-based feeding activity index for the automatic feeding of Atlantic salmon in recirculating aquaculture system (RAS) by analyzing differences in two consecutive frames. Ye et al. (2016) made use of the Lucas-Kanade optical flow and information entropy to assess and optimize the feeding of tilapia in RAS, and this method was then further improved by the quantification of fish spontaneous collective behaviors (Zhao et al., 2017). These methods, however, were based on human-made features (i.e., low-level features), which made them task-specific and weak in generalization capability. With the rapid development of deep learning, convolutional neural network (CNN) was gradually applied to fish appetite evaluation (Zhou et al., 2019). Profiting from the utilization of the high-level features of feeding behavior, fish appetite could be represented more precisely and robustly (Zhou et al., 2019; Ubina et al., 2021). From this, Wei et al. (2021) developed a method based on the modified kinetic energy model and customized recurrent neural network (RNN) to comprehensively utilize the spatial-temporal characteristics of fish feeding behavior, which therefore made fish appetite evaluation more accurate and practical. Similarly, by exploiting the spatial-temporal characteristics of fish feeding behavior, Feng et al. (2022) also realized the precise quantification of fish appetite resorted to a lightweight 3D ResNet-GloRe network, although a feeding strategy not commonly used in real production was adopted.

Methods mentioned above mainly rely on the characteristics of feeding behavior over time and have high requirement on the time-duration of data (i.e., data integrity). Generally speaking, the longer the time duration of data is, the better the data integrity and the better performance of the method would be. However, longer the time-duration of data normally means the more time taken in fish appetite assessment; what's more, the collection of data with long time-duration is a time-consuming process itself. In real production, in order to leave enough reaction time for the follow-up feeding control (including feeding strategy adjustment), the sooner of fish appetite assessment, the better. For this, increasing hardware investment seems to be a simplest and most direct solution, nevertheless, this will undoubtedly increase the production costs and affect the economic benefits (Feng et al., 2022). A promising alternative to the above solution is decreasing the time-duration of input data, namely, grading fish appetite using time-limited data particularly the beginning of time series data. But how to

construct an efficient fish appetite grading model with strong learning ability on this time-limited feeding behavior data is still a challenge. Few efforts on this challenge have been reported so far.

Given all that, regarding *Micropterus salmoides* as the experimental fish, a modified kinetic energy feature-based graph convolutional network (GCN), which could address the challenge mentioned above and grade fish appetite precisely with low space complexity, was developed in this study. In this network, each video frame was presented as a node. It's time information and the corresponding quantitative spatial information of fish feeding behavior were utilized as node features in the graph. In the meantime, the temporal connections between nodes were abstracted as edges in the graph. Benefiting by the specific graph constructed above and the customized network structure, the grading accuracy obtained by the proposed method here could be 98.60% using only the first 4.2 (one-sixth of the full-length data) seconds as well as the first 8.3 seconds (one-third of the full-length data) of input data, which is not much different from that on full-length (25 second-long) input data.

Materials and methods

Our experimental protocol was approved by the committee of the Care and Use of animals of the Zhejiang University. In addition, the experiments carried out on fish were conducted in strict accordance with the guidelines of the Association for the Study of Animal Behavior Use of Zhejiang University (ZJU20190074) in this study. Note that due to the indispensable role and rapid growth trend of industrial RAS in aquaculture, our experiment was conducted in RAS.

Fish

In this experiment, *Micropterus salmoides* were used. All experimental fish (quantity: 150) were first acclimated in the experimental RAS for one month. The average size of fish was 37.5 ± 5 g. During the entire experiment, the fish were placed under a 12h: 12h light-dark cycle (08:00-20:00 light, 20:00-08:00 dark) and fed 2 times a day (10:00 and 16:00) using commercial floating pellets. The feeding amount per day was set to 5% of the total mass of the fish.

Experimental system

The experimental aquaculture system (Figure 1) mainly consisted of a rearing tank (75 cm radius and 40 cm water depth), a feeding machine, and a computer vision system. During the entire acclimation and experiment, the following conditions were maintained: temperature at 26 ± 2°C, dissolved oxygen (DO) at (5.5 ± 0.5) mg/L, pH at 7.2 ± 0.5, nitrate ≤ 0.5 mg/L, and total ammonia nitrogen (TAN) ≤ 0.8 mg/L. The computer vision system possessed a Hikvision DS-2CD6233F-SDI camera, a Hikvision DS-7808NB-K2 Digital Video Recorder and a Server (GPU: NVIDIA 1080ti 11GB, CPU Intel Core i5-9400, 2.9 GHz, 8 GB memory). The camera was fixed 120 cm above the water surface of the rearing tank, with a 25fps frame rate and a 1080×1920 pixel.

To obtain sufficient fish feeding video data to verify the performance of the method proposed in this study, the overfed regime was adopted in this study. Food pellets were delivered with the same dose at intervals of ~25 s in each feeding event. Feeding wouldn't stop until fish showed no response to the delivered food. The residual pellets remaining on the water
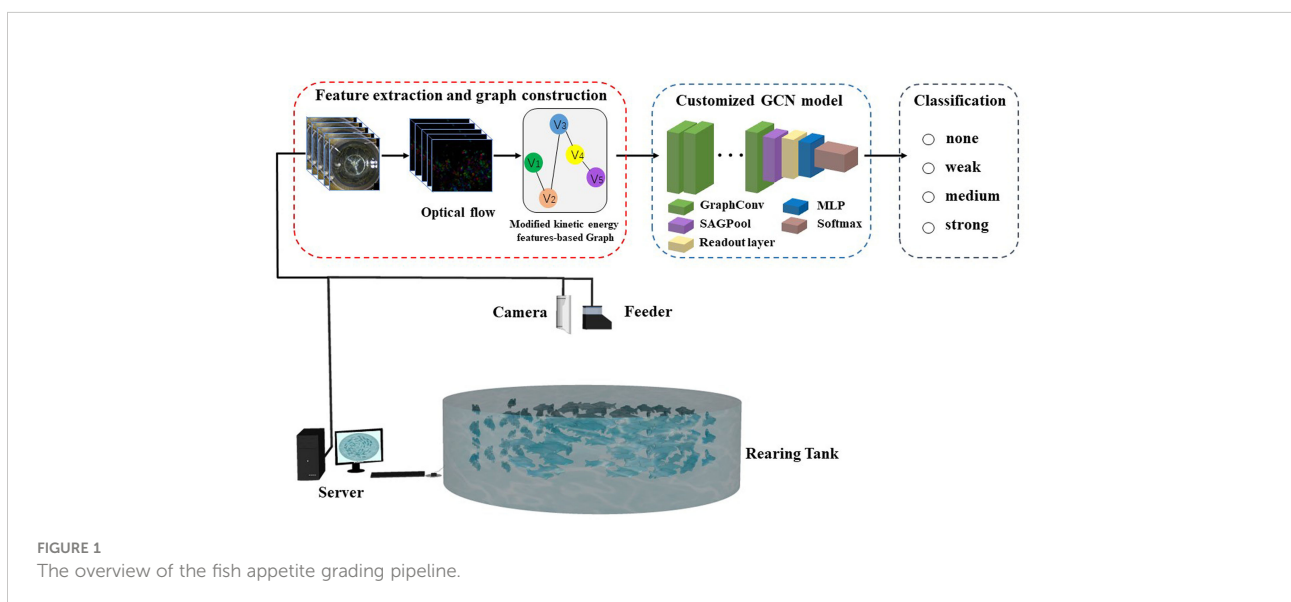

FIGURE 1
The overview of the fish appetite grading pipeline.

surface after feeding were then removed to prevent affecting the water quality.

Overview of the proposed approach

Accurate fish appetite grading is a prerequisite for intelligent feeding control and cleaner production in aquaculture. Therefore, we proposed a GCN-based fish appetite grading method, as shown in Figure 1. The method consists of two major steps: (1) feature extraction and graph construction: the improved kinetic energy model is used to extract the spatial characteristics of fish feeding behavior from feeding videos, and then a graph $G = (V, E)$ is constructed based on the modified kinetic energy features and their temporal correlation. (2) GCN-based classification model: the GCN-based model combines the graph structures and vertex features in the convolution, and propagated over the graph through multiple layers. Through the graph convolution layer by layer, the node features are extracted and updated. In addition, the GCN structure adopts the global graph pooling method based on self-attention mechanism, which fully considers the topology of nodes and graphs, and has significant advantages in graph classification tasks. In this study, the proposed method was developed using python3.7, and the customized neural networks grading model was trained on Pytorch1.10.0.

Feature extraction and graph construction

Graph, as a data structure that could simultaneously store target's feature information and its associated information, has been widely applied to efficient task classification (Lazer et al., 2009; Lee et al., 2019). This technique, however, is rarely used in aquaculture yet. For fish appetite grading, if it could be transformed into a simple graph classification task, the efficiency of this grading would be maximized. But how to extract efficient features as the graph features and construct the graph are the key to achieving this graph classification task. The spatial-temporal characteristics of fish feeding behavior shows great potential in fish appetite assessment (Wei et al., 2021; Feng et al., 2022), therefore, the feature extraction and graph construction in this study are carried out for the representation of these spatial-temporal characteristics.

First, the modified kinetic energy model (Eq. (1)) was used to extract spatial characteristics of fish behavior due to their strong motion feature extraction ability (Zhao et al., 2017).

$$E_K = C_E \times v_E^2 \tag{1}$$

Where C_E and v_E denote the disorder degree and velocity of the change in target areas, respectively.

The Gunner Farneback optical flow algorithm was used to calculate the v_E (Eq. (2)) of the changes in target areas in this study. Then, the scope of velocity was divided into a number of

sections. As shown in Eq. (3), v_E was classified into the corresponding sections. In addition, to avoid the influence of the fish body length change on the motion feature extraction during the experiment, the v_E was calculated by the normalization method.

$$v_E = \frac{\sum_{x,y}|F_n(x,y)|}{N} \tag{2}$$

$$p(j) = \{(k(j)/N) \times 100\%, \ 1 \le j \le m\} \tag{3}$$

Where F_n is the normalized optical flow between two consecutive frames of images, (x, y) represents the coordinates of the reflective region in the current frame. m is the number of sections of v_E, N is the number of motion vectors in the current frame. k and p are the set of statistical numbers and statistical probability in each section, respectively. In this section, v_E was counted at intervals of 0.04 bl (bl is the average body length of *Micropterus salmoides*), and m was set to 25.

Then, C_E was calculated as:

$$C_E = -\sum_{j=1}^{m} p(j) log_2(p(j)) \tag{4}$$

Combined with Eq. (4), the normalized kinetic energy of the whole areas was defined as:

$$E_k = -\left(\sum_{j=1}^{m} p(j) log_2(p(j))\right) \times \left(\frac{\sum_{x,y}|F_n(x,y)|}{N}\right)^2 \tag{5}$$

Figure 2 shows the modified kinetic energy over time within a single round feeding event. As shown below, fish appetite, graded following the criterion of Øverli et al. (2006) and Eriksen et al. (2011), can be described well by the modified kinetic energy here.

Finally, graph G was constructed by the following two elements: (1) adjacency matrix A ($A \in R^{N \times N}$). This element is used to represent the connection between video frames (i.e., time correlation). The adjacency matrix contains only elements of 0 and 1. The element is 0 if there is no link between two video frames and 1 denotes there is a link. (2) feature matrix X. We regard the spatial-temporal characteristics of fish behavior (i.e., the normalized kinetic energy features and their temporal correlation) extracted from video frames as the attribute features of the node in the networks, expressed as $X \in R^{N \times P}$, where P represents the number of node attribute features.

To express the above algorithm more intuitively, the feature extraction and graph construction process are outlined in Algorithm 1.

Input: video frames; modified kinetic energy model (MKEM); normalization strategy N; adjacency matrix A ($A \in R^{N \times N}$); feature matrix X ($X \in R^{N \times P}$);
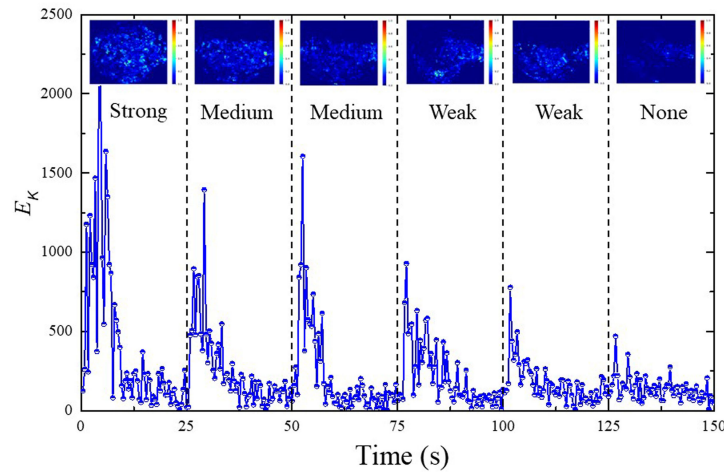
FIGURE 2
Diagram of the modified kinetic energy over time within a single round feeding event (None: fish do not respond to food; Weak: fish eat only pellets that fall directly in front of them but do not move to take food; Medium: fish move to take food, but return to their original positions; Strong: fish move freely between food items and consume all the available food. Noted that, the heat maps generated by the optical flow of the points with the maximum kinetic energy are used to better visualized fish appetite here).

```
Output: G = (V , E)
1: function feature extraction (video
frames, MKEM, N)
2: C_E ← video frames
3: F_n(x,y) ← video frames & N
4: v_E ← F_n(x , y)
5: for x, y in video frames do
6: E_K / C_E × v_E^2
7: return E_K
8: function graph construction (E_K, A, X)
9: A ← video frames
10: X ← video frames & E_K
11: G = (V , E) ← A & X
12: return G
```

ALGORITHM 1
Feature extraction and graph construction.

GCN-based fish appetite grading model

Advanced methods of applying deep learning to structured data such as graphs have been proposed in recent years. In particular, the method of generalizing the convolution operation to graphs has been proven to improve performance and has been widely used (Lee et al., 2019; Zhao et al., 2020b).

Given this, a customized GCN (Figure 3) following the constructed graph above was proposed in this study to achieve accurate fish appetite grading using time-limited data. This model consists of seven graph convolutional layers, and outputs of each layer are concatenated. Node and graph feature are updated and aggregated in the pooling layer with self-attention mechanism, and then transmitted to the MLP layer through the readout layer. Finally, the fish appetite level is determined with the help of softmax layer. The propagation rule of GCN can be summarized by the following expression:
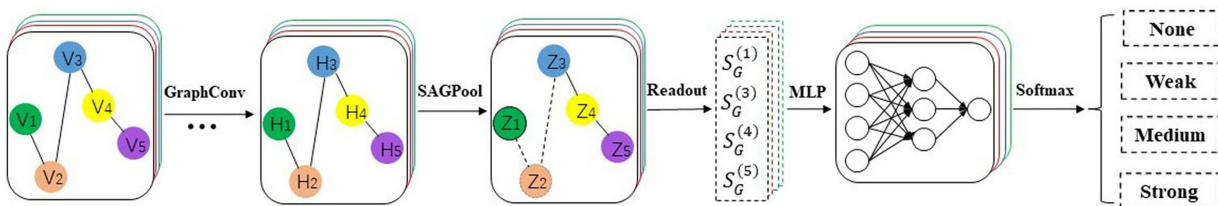

FIGURE 3
Schematic diagram of the proposed fish appetite grading model.

$$h^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}h^{(l)}\Theta\right) \qquad (6)$$

Where $h^{(l)}$ represents the node representation of l-th layer and $\Theta \in \mathbb{R}^{F \times F'}$ represents the convolution weight with input feature dimension F and output feature dimension F', $\tilde{A} = A + I_N$ represents the matrix with added self-attentions, I_N represents the identity matrix, $\tilde{D} = \sum \tilde{A}_{ij}$, and σ is the activation function. The Rectified Linear Unit (Relu) function was used as an activation function in this study.

The attention mechanism has been widely used in recent deep learning studies (Cheng et al., 2016; Lee et al., 2019). Such a mechanism enables the model to focus more on important features and less on noncritical features (Lee et al., 2019), especially the self-attention mechanism. Thus, the self-attention pooling mechanism was used in this study. The self-attention score $Z \in \mathbb{R}^{N \times 1}$ was calculated as follow.

$$Z = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}X\Theta_{att}\right) \qquad (7)$$

Where $X \in \mathbb{R}^{N \times F}$ is the input feature of the graph with N nodes and F-dimensional features, and $\Theta_{att} \in \mathbb{R}^{F \times 1}$ is the parameter of the self-attention pooling layer.

In the GCN-based classification model, the readout layer was used to aggregates node features to make a fixed size representation. The summarized output feature of the readout layer was as follows:

$$s = \frac{1}{N}\sum_{i=1}^{N}x_i \parallel \max_{i=1} x_i \qquad (8)$$

Where N is the number of nodes, x_i is the feature vector i-th node, and $\parallel$ denotes concatenation.

Data collection and training set production

In this study, the video frames of fish school behavior under four feeding intensities were intercepted at equal intervals (12 frames per second). After data augmentation (referring to Wei et al., 2021), the total number of samples in the training set increased to 24300 (see https://github.com/Doubleblindpeerreview/fish-appetite-grading for details of dataset and codes). Of those data, 80% were used as the training set, 10% were used as the validation set, and 10% were used as the test set.

Setting appropriate parameters is the key step to training a robust model. In this study, Adam optimizer, early stopping criterion and hyperparameter selection strategy were used as the model architecture. If the validation loss did not improve for 60 epochs in an epoch termination condition with a maximum of 100k epochs, the training would be stopped. After many trials, the initial parameters and training strategies of the GCN-based method were set to the values shown in Table 1.

TABLE 1 The main parameters of GCN model.

Parameter	Value
Learning rate	5×10^{-4}
Batch size	128
Dropout rate	0.5
Pooling rate	0.8
Hidden size	256
Weight decay	1×10^{-4}

Performance evaluation

The results of testing for all approaches were arranged in confusion matrices, including true positive (TP), true negative (TN), false positive (FP), and false negative (FN). In this context, TP and TN respectively denote the numbers of the same samples with the current feeding appetite pertaining to the other fish appetite recognition results and actual results; FP and FN are the numbers of different videos with the current fish appetite pertaining to the other fish appetite recognition results and actual results, respectively.

To evaluate model performance, five widely used measures were calculated: accuracy, precision, recall, specificity and $F1$ score. Accuracy is the ratio of the number of correctly graded samples to the total number of samples; precision is the ratio of the number of samples for a specific level of fish appetite in the test set to the number of samples for that fish appetite in the recognition results, which shows the ability of the model to accurately grade the fish appetite; recall is the proportion of correctly classified items among all items to be classified; specificity is the ratio of the number of samples with wrong recognition to the number of samples with other fish appetite in the test set; $F1$ score is a harmonic means of the precision and recall (Jiang et al., 2020). All the above five measures are ranged from 0 to 1, high value means the good predictive ability of the model, their definitions are as follows:

$$accuracy = \frac{TP + TN}{FP + FN + TP + TN} \qquad (9)$$

$$precision = \frac{TP}{TP + FP} \qquad (10)$$

$$recall = \frac{TP}{TP + FN} \qquad (11)$$

$$specificity = \frac{TN}{FP + FN} \qquad (12)$$

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \qquad (13)$$

TABLE 2　Grading results for each fish appetite level.

Fish appetite	Grading result				Precision (%)	Recall (%)	Specificity (%)	Accuracy (%)
	none	weak	medium	strong				
none	537	1	0	0	99.44	99.81	99.84	**98.89**
weak	3	669	0	0	98.82	99.55	99.54	
medium	0	7	611	3	97.92	98.39	92.28	
strong	0	0	13	586	99.49	97.83	99.84	

Results and discussion

Performance of the proposed method

As shown in Table 2, the average accuracy of the method under four fish appetite levels reached 98.89% (Precision: 98.92%, Recall: 98.90%, $F1$ score: 98.90%), which indicates the effectiveness of the proposed fish appetite grading method.

To verify the performance of the method, we compared it with the three most widely used GNNs including ChebNet, GraphSAGE, and GAT. In addition, to better reveal the performance of our method in fish appetite grading, the hierarchical pooling architecture and global pooling architecture were used in the same GNN-based method as a comparison. The hierarchical pooling architecture consists of three blocks, each consisting of a graph convolution layer and a graph pooling layer. The outputs of each block are summarized in the readout layer. The sum output of each readout layer is input to the linear layer for classification. For a fair comparison, we performed the same self-attention graph pooling strategy, training strategy, and hyperparameter optimization strategy for each method, and use the same dataset (full-length data).

Figure 4 and Figure 5 (see Table S1 in supplementary materials for details) show the grading results of fish feeding behavior datasets using different GNN-based models, where the suffix h indicates that the model adopts the hierarchical pooling mechanism, and g indicates that the model adopts the global pooling mechanism. Pleased node that based on the same global pooling or hierarchical pooling architecture, our GCN-based method dramatically outperforms other GNN-based methods. In addition, for the data structure used in this paper, the global pooling method is better than the hierarchical pooling method. The global pooling architecture minimizes the loss of information and outperforms hierarchical pooling on datasets with fewer nodes. Due to the limited number of nodes in the constructed graph dataset, the global pooling method shows better performance in this study.

Graph attention network (GAT), as a novel convolution-style neural networks that operate on graph-structured data, leveraging masked self-attentional layers. The network allows for assigning operations, and is parallelizable across all nodes within a neighborhood while dealing with different sized neighborhoods, and does not depend on knowing the entire graph structure upfront. As opposed to GCN, the GAT model
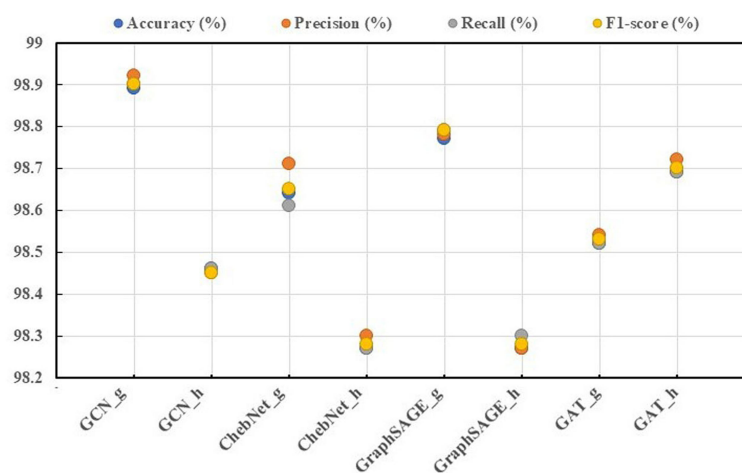

FIGURE 4
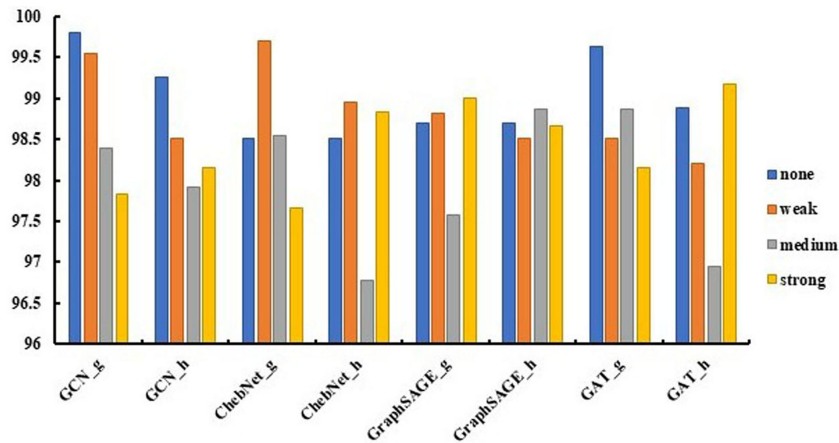Fish appetite grading results of different GNN-based methods.

FIGURE 5
The accuracy of different GNN-based methods under different fish appetite levels.

can dynamically learn neighbor weights, but ignores the relationship between nodes (Xiang et al., 2021), which makes it not as effective as GCN under some conditions. As shown in Figure 4, although the accuracy of the GAT-based method under the two pooling mechanisms is similar, their accuracy is slightly lower than that of the GCN-based method. The GraphSAGE-based method uses an inductive method to calculate the node representation (Hamilton et al., 2017). Specifically, the method first extracts a fixed number of nodes from the adjacent nodes of each node, and then integrates the information of these neighbor nodes. This method has achieved good results in many large-scale inclusive learning problems. However, compared with the GraphSAGE-based method, the GCN-based method can capture the global information of the graph so as to better represent the characteristics of nodes, which is suitable for small-scale graphs. The dataset graph used in this study has a simple structure and few nodes. For this type of dataset, the GCN-based method has more advantages. The ChebNet-based method has strong expression ability. Its K-order convolution operator can cover the K-order neighbor nodes of nodes, but its complexity and parameter quantity are higher than GCN (Kipf and Welling, 2016). By stacking multiple GCN layers or expanding the empirical domain of graph convolution, the expressivity of the GCN- based method can be greatly improved. Therefore, under the dataset used in this study, the training accuracy of the ChebNet-based method is lower than that of the GCN-based method.

Feasibility demonstration

For the feeding control in real production, the most important is how to accurately assess fish appetite as soon as possible, and then leave enough reaction time for the next feeding operation. Therefore, it is crucial to evaluate the feasibility of the method proposed in this study whether the fish appetite can be accurately and effectively evaluated with time-limited data. In view of this, we divided the dataset into four subsets (as illustrated in Figure 6) to meet the needs of feasibility verification, which includes 1) the first 4.2 seconds of data (Dataset 1, one-sixth of the full-length data), 2) the first 8.3 seconds of data (Dataset 2, one-third of the full-length data), 3) the first 12.5 seconds of data (Dataset 3, one-half of the full-length data), and the first 25 seconds of data (Dataset 4, full-length data).

To verify the performance of the proposed method here, we compared it with two typical and state-of-the-art fish appetite grading methods, namely, the RNN-based method (Wei et al., 2021) and the CNN-based method (Zhou et al., 2019). Note that to make comparison more comprehensive, not only the RNN-based method but its normalized version (i.e., RNN'-based method) was used for the comparison here, allowing for the fact that the normalized motion features was adopted in our method. In particular, for the CNN-based method above, its image dataset in this study was obtained by extracting video frames owning the most obvious feeding behavior characteristic from the original video. Specifically, five consecutive frames of images with the strongest fish appetite were extracted from each feeding videos as the original image samples. And then, the original image samples were augmented to 24300 samples using rotation, flip, and translation image expansion techniques. Following this, the dataset was divided into a training set, a validation set, and a test set in a ratio of 8:1:1. It should be noted that, to obtain the optimal performance of the adopted methods above, the corresponding optimum hyperparameters were adopted here (details in Table S2 and Table S3 in the supplementary material).
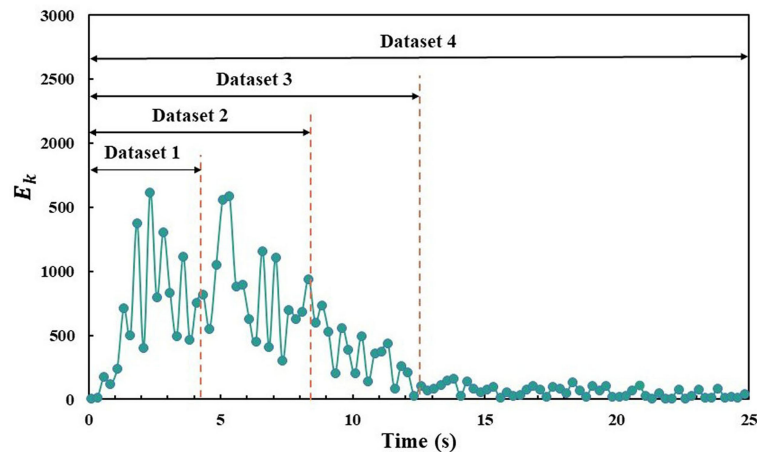
FIGURE 6
The illustration of dataset partition.

The grading accuracy of fish appetite based on our method (i.e., GCN-based method) and RNN-based method under different datasets were analyzed in this study, as well as the grading accuracy of CNN-based method (see Table S4 in supplementary materials for details). The average accuracy, precision, recall and $F1$-score of the CNN-based method were 83.54%, 83.90%, 84.15%, and 83.97%, respectively. Because the CNN-based method only uses the behavioral spatial characteristics of fish school to grade fish appetite, its performance is far from that of the other two fish feeding desire grading methods. Benefit from the spatial-temporal behavioral characteristics, the RNN-based method showed better performance than the CNN-based method. As shown in Figure 7 (since the results of the CNN-based fish appetite grading method are quite different from those of the GCN-based and RNN-based method, it is not shown in the figure to avoid affecting the expression), the RNN-based method achieved similar fish appetite grading results on normalized and non-normalized version. It should be noted that with the increase in the time duration of the dataset, the effect of the RNN-based fish appetite grading method shows an obvious upward trend. The RNN-based method achieved the best performance in Dataset 4, but the fish appetite grading accuracy is relatively low in time-limited datasets. Fish appetite representation is closely related to time (Wei et al., 2021). The RNN-based method is designed to counter the effect of diminishing gradients through layers and is suitable for time series data. However, the length of time series data also restricts the grading performance of RNN-based method on fish appetite. As presented in Figure 7, the GCN-based fish appetite grading method achieved the best performance in Dataset 4, but there was only a minor difference from the test results obtained from Dataset 1 and Dataset 2. Benefiting from the construction of the modified

kinetic energy-based graph and the customization of GCN structure, our method indicated stronger learning ability than the state-of-the-art fish appetite assessing methods especially on time-limited feeding behavior data.

The t-SNE technique, which visualizes high-dimensional data by giving each datapoint a location in a two or three-dimensional map (Van Der Maaten and Hinton, 2008), is becoming more and more popular in data analysis. Thus, the t-SNE based two-dimensional analysis was used in this study to visualize the grading effect of the GCN-based, RNN-based, and CNN-based method on fish feeding desire. As illustrated in Figure 8, the CNN-based appetite grading method failed to divide fish appetite into four significant clusters (different colors represent different appetite levels), which also showed that the CNN-based method achieved poor fish feeding desire grading accuracy. On the contrary, both GCN-based and RNN-based grading methods can divide fish appetite into four significant clusters, so both methods have great appetite grading performance. However, compared with the RNN-based method, there are only a few data points with different colors mixed together in the GNN-based method (i.e., fewer data sample points prone to misclassification). To further analyze the causes of RNN-based method error identification, some false recognition examples are shown in Figure 9. It can be seen that there are a small number of "medium" or "weak" samples that were incorrectly recognized as "weak" or "strong" in all the four datasets. The main reason is that the modified kinetic energy has similar variation characteristics when the appetite level is "medium" or "weak" (as shown in Figure 2). Hence, the samples of these two are sometimes mis-recognized. In addition, Dataset 1 and Dataset 2 cannot reflect the whole process of fish feeding, which is also the key to limiting the accuracy of RNN-based method in fish appetite classification on
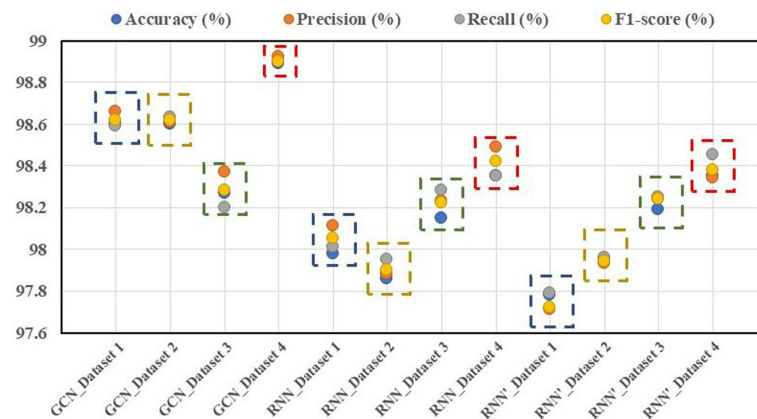
FIGURE 7
Fish appetite grading results under different datasets.

these two datasets. This reason also makes some "strong" samples incorrectly recognized as "medium".

In addition, Figure 10 shows the confusion matrixes of fish appetite grading results of the RNN-based and GCN-based methods on different datasets. It was obvious that in comparison with the RNN-based method, the classification accuracy of GCN-based method was equal or higher in each class to some extent, especially in Dataset 2 and Dataset 4. That means the proposed GCN-based approach learns new feature representation from the neighbor nodes through graph convolution, which improves the recognition ability under different datasets. Compared with the RNN-based method, the GCN-based method is more suitable to characterize the spatial and temporal topological information of fish feeding behavior. Therefore, the GCN-based method has achieved effective fish appetite grading results under different datasets, including the time-limit datasets.

It should be noted that, in order to achieve efficient grading of fish appetite in real production, the complexity especially the space complexity of the grading method itself is very important, as the valid training samples are limited in practical farming (Pan et al., 2019). We therefore calculated the quantity of the trainable parameters of RNN-based and GCN-based methods in this study, respectively (as illustrated in Figure 11). Combined with Figure 7 and Figure 9, the GCN-based method proposed in this study could not only obtain high accuracy in fish appetite grading, but take only 6.4% ~ 31.8% space complexity of that in RNN-based method, which can greatly improve the feasibility of fish appetite assessment in practical production.

The proposed modified kinetic energy feature-based GCN approach in this paper can effectively grade fish appetite with time-limited data, which is a promising approach in dealing with feeding control tasks and alleviating the water environment burden in aquaculture. Nonetheless, limitations still exist in
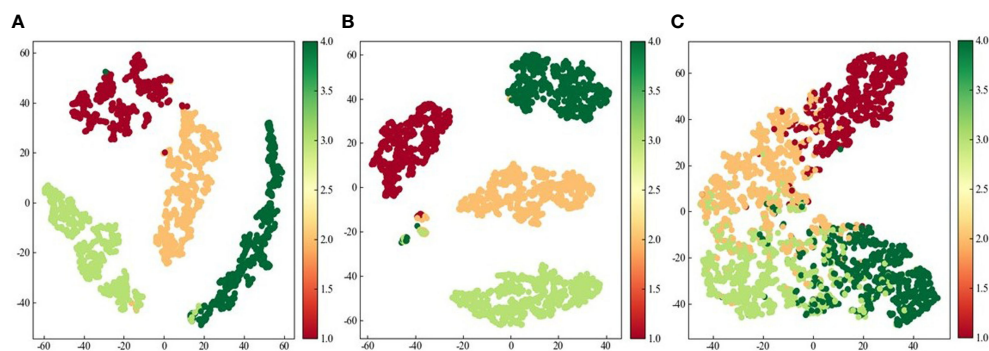

FIGURE 8
t-SNE based two-dimensional clustering analysis of different fish appetite grading methods (see Figure S1 in supplementary materials for more details). **(A)** GCN_Dataset 4. **(B)** RNN_Dataset 4. **(C)** CNN.
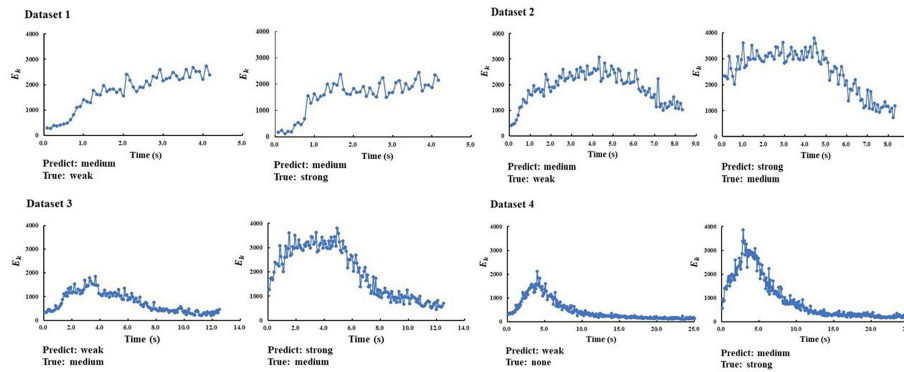
FIGURE 9
False recognition examples of the RNN-based method model in different datasets.

this method. First, our training on the model is based on experimental data under ideal conditions, which were derived from videos of specific growth periods of *Micropterus salmoides* in RAS, without monitoring the entire growth period of *Micropterus salmoides*. Therefore, when it comes to other scenarios, the practicability of our method may reduce. Besides, since the method proposed in this study is based on computer vision techniques, feed property also affects the performance of the model here to some extent, floating feed would be more beneficial to the performance maximization of the method here in contrast to sinking feed.

Conclusions

In order to leave enough reaction time for the follow-up feeding control and alleviate the water environment burden of

the aquaculture, a novel, practical and promising fish appetite grading method with low space complexity was proposed in this study. Benefiting from the construction of the modified kinetic energy feature-based graph and the customization of GCN structure, our method indicated stronger learning ability than the typical state-of-the-art fish appetite assessing methods especially on time-limited feeding behavior data. And the grading accuracy of fish appetite obtained by the proposed method could reach 98.60% using only the first 4.2 (Precision: 98.66%, Recall: 98.59%, $F1$ score: 98.62%) as well as the first 8.3 seconds (Precision: 98.61%, Recall: 98.63%, $F1$ score: 98.62%) of input data, which is not much different from that (98.89%) on full-length (25 second-long) (Precision: 98.92%, Recall: 98.90%, $F1$ score: 98.90%) input data. Although limitations (such as feed property) still exist in this study, the findings here could not only provide references for the accurate control of fish feeding, but is of significance for the realization of cleaner production in practical aquaculture.
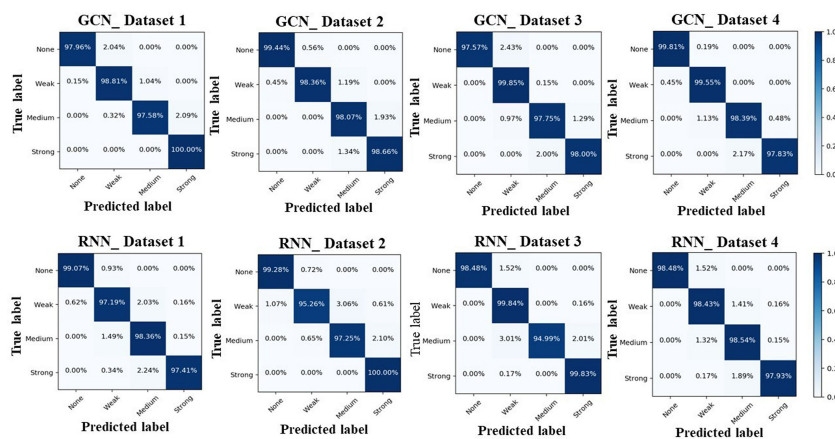

FIGURE 10
Confusion matrixes of fish appetite grading results of GNN-based and RNN-based methods on different datasets.
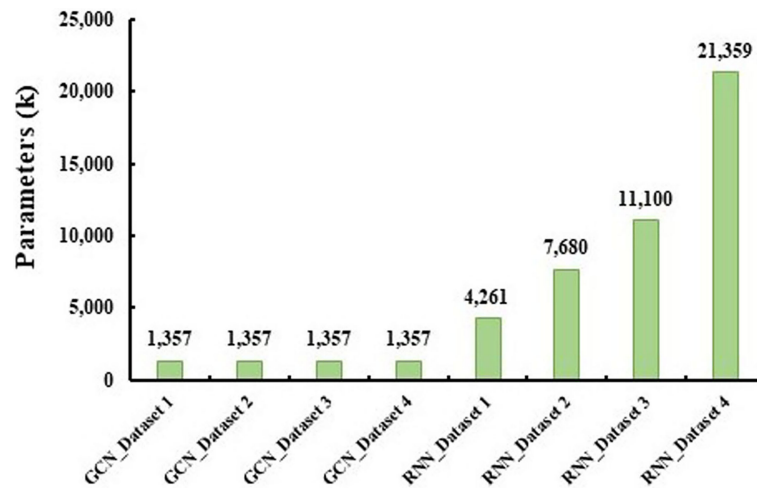
FIGURE 11
Space complexity of GCN-based and RNN-based methods on the different datasets.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://github.com/Doubleblindpeerreview/fish-appetite-grading.

Ethics statement

The animal study was reviewed and approved by the Association for the Study of Animal Behavior Use of Zhejiang University (ZJU20190074).

Author contributions

DW: Conceptualization, Methodology, Formal analysis, Writing-Original Draft, Visualization, Writing-Review & Editing. BJ: Formal analysis; Writing-Original Draft. HL: Methodology, Software; Writing-Original Draft. SZ: Supervision, Funding acquisition. ZY: Resources, Supervision, Funding acquisition. JZ: Conceptualization, Writing-Review & Editing, Project administration, Funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.1021688/full#supplementary-material

References

Øverli, Ø., Sørensen, C., and Nilsson, G. E. (2006). Behavioral indicators of stress-coping style in rainbow trout: Do males and females react differently to novelty? *Physiol. Behav.* 87, 506–512. doi: 10.1016/j.physbeh.2005.11.012

An, D., Huang, J., and Wei, Y. (2021). A survey of fish behaviour quantification indexes and methods in aquaculture. *Rev. Aquac.* 13, 2169–2189. doi: 10.1111/raq.12564

Atoum, Y., Srivastava, S., and Liu, X. (2015). Automatic feeding control for dense aquaculture fish tanks. *IEEE Signal Process Lett.* 22, 1089–1093. doi: 10.1109/LSP.2014.2385794

Barraza-Guardado, R. H., Martínez-Córdova, L. R., Enríquez-Ocaña, L. F., Martínez-Porchas, M., Miranda-Baeza, A., and Porchas-Cornejo, M. A. (2014). Effect of shrimp farm effluent on water and sediment quality parameters off the coast of Sonora, Mexico. *Cienc. Mar.* 40, 221–235. doi: 10.7773/cm.v40i4.2424

Chang, C. M., Fang, W., Jao, R. C., Shyu, C. Z., and Liao, I. C. (2005). Development of an intelligent feeding controller for indoor intensive culturing of eel. *Aquac. Eng.* 32, 343–353. doi: 10.1016/j.aquaeng.2004.07.004

Cheng, J., Dong, L., and Lapata, M. (2016). "Long short-term memory-networks for machine reading," in *In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing.* 551–561. (EMNLP Conference) doi: 10.18653/v1/d16-1053

Eriksen, M. S., Færevik, G., Kittilsen, S., McCormick, M. I., Damsgård, B., Braithwaite, V. A., et al. (2011). Stressed mothers-troubled offspring: a study of behavioural maternal effects in farmed salmo salar. *J. Fish Biol.* 79, 575–586. doi: 10.1111/j.1095-8649.2011.03036.x

Føre, M., Alfredsen, J. A., and Gronningsater, A. (2011). Development of two telemetry-based systems for monitoring the feeding behaviour of Atlantic salmon (Salmo salar l.) in aquaculture sea-cages. *Comput. Electron. Agric.* 76, 240–251. doi: 10.1016/j.compag.2011.02.003

FAO (2020). *The state of world fisheries and aquaculture 2020. Sustainability in action* (Rome: World series of the Food and Agriculture Organization of the United Nations).

Feng, S., Yang, X., Liu, Y., Zhao, Z., Liu, J., Yan, Y., et al. (2022). Fish feeding intensity quantification using machine vision and a lightweight 3D ResNet-GloRe network. *Aquac. Eng.* 98, 102244. doi: 10.1016/j.aquaeng.2022.102244

Hamilton, W. L., Ying, R., and Leskovec, J. (2017). Inductive representation learning on large graphs. *NIPS* 2017, 1025–1035. doi: 10.48550/arXiv.1706.02216

Jescovitch, L. N., Ullman, C., Rhodes, M., and Davis, D. A. (2018). Effects of different feed management treatments on water quality for pacific white shrimp litopenaeus vannamei. *Aquac. Res.* 49, 526–531. doi: 10.1111/are.13483

Jiang, H., Zhang, C., Qiao, Y., Zhang, Z., Zhang, W., and Song, C. (2020). CNN Feature based graph convolutional network for weed and crop recognition in smart farming. *Comput. Electron. Agric.* 174, 105450. doi: 10.1016/j.compag.2020.105450

Kipf, T. N., and Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *ICLR* 2017, 1–14. doi: 10.48550/arXiv.1609.02907

Lazer, D. D., Pentland, A. A., Adamic, L. L., Aral, S. S., Barabasi, A. L. A. L., and Brewer, D. D. (2009). Life in the network: The coming age of computational social science. *Science* 323, 721–723. doi: 10.1126/science.1167742

Lee, J., Lee, I., and Kang, J. (2019). Self-attention graph pooling. *ICML* 2019, 6661–6670. doi: 10.48550/arXiv.1904.08082

Liu, Z., Li, X., Fan, L., Lu, H., Liu, L., and Liu, Y. (2014). Measuring feeding activity of fish in RAS using computer vision. *Aquac. Eng.* 60, 20–27. doi: 10.1016/j.aquaeng.2014.03.005

Li, D., Wang, Z., Wu, S., Miao, Z., Du, L., and Duan, Y. (2020). Automatic recognition methods of fish feeding behavior in aquaculture: A review. *Aquaculture* 528, 735508. doi: 10.1016/j.aquaculture.2020.735508

Luna, M., Llorente, I., and Cobo, Á. (2019). Integration of environmental sustainability and product quality criteria in the decision-making process for feeding strategies in seabream aquaculture companies. *J. Clean Prod.* 217, 691–701. doi: 10.1016/j.jclepro.2019.01.248

Pan, J.-S., Lee, C.-Y., Sghaier, A., Zeghid, M., and Xie, J. (2019). Novel systolization of subquadratic space complexity multipliers based on toeplitz matrix-vector product approach. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* 27, 1614–1622. doi: 10.1109/TVLSI.2019.2903289

Parra, L., Sendra, S., García, L., and Lloret, J. (2018). Design and deployment of low-cost sensors for monitoring the water quality and fish behavior in aquaculture tanks during the feeding process. *Sensors (Basel Switzerland).* 18, 750. doi: 10.3390/s18030750

Ubina, N., Cheng, S. C., Chang, C., and C. Chen, H. Y. (2021). Evaluating fish feeding intensity in aquaculture with convolutional neural networks. *Aquac. Eng.* 94, 102178. doi: 10.1016/j.aquaeng.2021.102178

Van Der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2625.

Wang, Y., Yu, X., Liu, J., An, D., and Wei, Y. (2022). Dynamic feeding method for aquaculture fish using multi-task neural network. *Aquaculture* 551, 737913. doi: 10.1016/j.aquaculture.2022.737913

Wei, D., Bao, E., Wen, Y., Zhu, S., Ye, Z., and Zhao, J. (2021). Behavioral spatial-temporal characteristics-based appetite assessment for fish school in recirculating aquaculture systems. *Aquaculture* 545, 737215. doi: 10.1016/j.aquaculture.2021.737215

Xiang, Z., Gong, W., Li, Z., Yang, X., Wang, J., and Wang, H. (2021). Predicting protein-protein interactions *via* gated graph attention signed network. *Biomolecules* 11, 799. doi: 10.3390/biom11060799

Ye, Z., Zhao, J., Han, Z., Zhu, S., Li, J., and Lu, H. (2016). Behavioral characteristics and statistics-based imaging techniques in the assessment and optimization of tilapia feeding in a recirculating aquaculture system. *Trans. ASABE* 59, 345–355. doi: 10.13031/trans.59.11406

Zhao, J., Bao, W. J., Zhang, F. D., Ye, Z. Y., Liu, Y., and Shen, M. W. (2017). Assessing appetite of the swimming fish based on spontaneous collective behaviors in a recirculating aquaculture system. *Aquac. Eng.* 78, 196–204. doi: 10.1016/j.aquaeng.2017.07.008

Zhao, S., Ding, W., Zhao, S., and Gu, J. (2019). Adaptive neural fuzzy inference system for feeding decision-making of grass carp (Ctenopharyngodon idellus) in outdoor intensive culturing ponds. *Aquaculture* 498, 28–36. doi: 10.1016/j.aquaculture.2018.07.068

Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., and Lin, T. (2020b). T-GCN: A temporal graph convolutional network for traffic prediction. *IEEE Trans. Intell. Transp. Syst.* 21, 3848–3858. doi: 10.1109/TITS.2019.2935152

Zhao, S., Zhu, M., Ding, W., Zhao, S., and Gu, J. (2020a). Feed requirement determination of grass carp (Ctenopharyngodon idella) using a hybrid method of bioenergetics factorial model and fuzzy logic control technology under outdoor pond culturing systems. *Aquaculture* 521, 734970. doi: 10.1016/j.aquaculture.2020.734970

Zhou, C., Xu, D., Chen, L., Zhang, S., Sun, C., and Yang, X. (2019). Evaluation of fish feeding intensity in aquaculture using a convolutional neural network and machine vision. *Aquaculture* 507, 457–465. doi: 10.1016/j.aquaculture.2019.04.056

Zhou, C., Xu, D., Lin, K., Sun, C., and Yang, X. (2018). Intelligent feeding control methods in aquaculture with an emphasis on fish: A review. *Rev. Aquac.* 10, 975–993. doi: 10.1111/raq.12218