International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 5937-5946

https://doi.org/10.70517/ijhsa464431

Underwater Target Recognition Algorithm Based on Improved Deep Convolutional Neural Network

Yang Zhang[1], Luyao Wang[1,*], Hongping Xie[1], Kaixin Gu[1], Zijian Ye[1] and Cheng Yan[1]
[1] State Grid Jiangsu Electric Power Co., Ltd. Construction Branch, Nanjing, 210000, Jiangsu, China
Corresponding authors: (e-mail: wongsixrone@163.com).

Abstract This article aimed to address the challenges of underwater target recognition algorithms in the face of large differences in lighting conditions, image blurring, and distortion by improving deep convolutional neural networks. By introducing the attention mechanism, the purpose of strengthening the representation of target information in the convolutional feature map is achieved, thereby reducing the interference of irrelevant information. Firstly, the pre-trained and improved deep convolutional neural network model ResNet-50 was used to fine tune the underwater target dataset to capture local and global information of the target and construct an attention mechanism network; secondly, attention mechanism was used to calculate the weight of each pixel in the feature map, making the target area more prominent. The attention weighted features can be fused with the original features, and the channel attention mechanism can be used to weight the importance of features between channels; finally, this article designed a classifier to recognize underwater targets, based on the Softmax classifier for target classification, and outputs the probability distribution of target categories. The research results indicate that the deep convolutional neural network improved by applying attention mechanism performs the best compared to other models in terms of accuracy, recall, and F1 value, reaching 0.85, 0.82, and 0.83, respectively. This indicates that deep convolutional neural networks improved by applying attention mechanisms can play a greater role in underwater target recognition.

Index Terms Underwater Target Recognition Algorithm, Deep Convolutional Neural Network, Attention Mechanism, ResNet-50 Model, Softmax Classifier

I. Introduction

Nowadays, underwater target recognition plays an increasingly important role in marine resource development and environmental monitoring, but with it comes the challenges brought by the complexity of the underwater environment. Light attenuation, water turbidity, and illumination changes make underwater target recognition difficult [1], [2]. Therefore, in order to meet the current new needs for underwater target recognition, new technical methods need to be adopted to deal with the difficulties.

Previous studies have shown that researchers have proposed various methods for underwater target recognition. Among them, some studies use feature engineering [3], [4] and traditional machine learning methods [5]-[7], for example, technologies such as Support Vector Machine (SVM) and Random Forest (RF) are used to handle target recognition tasks in underwater images. Among them, scholars like LIN Xinghua [8] have successfully overcome the issue of autonomous underwater robots' inadequate perception of the flow field environment by developing a support vector machine model. Scholars such as LIU Xionghou [9] have effectively solved the problems faced by underwater slow small target classification and recognition using the joint classification method of support vector machines. Through the study of feature selection methods for random forests, Wang Jiawei et al. [10] effectively solved the problem of multi-base underwater small target classification and recognition. In response to the problem of interference in the performance of target detectors caused by underwater environments, random forests were employed by academics like Han Yongqiang [11] to increase the tracking accuracy of underwater target identification. Numerous academics have made significant advances in the study of underwater target recognition techniques in recent years [12]-[14], but these methods still cannot well overcome the challenges brought by the underwater environment.

In the realm of deep learning, scientists have attempted to incorporate deep learning techniques into the domain of underwater target identification. They have utilized deep learning models like the Deep Convolutional Neural Network [15]-[17] to extract features from underwater images [18]-[20]. Wang Chenyu and other researchers [21] have proposed a deep learning-based method for underwater image recognition to address the issue of limited target image data in marine environments. Na Wang and other scholars [22] applied the timbre feature model of multiple regression analysis to underwater noise recognition, which effectively solved the problem of difficulty in classifying and identifying underwater target radiation noise signals. In response to the problem of low recognition accuracy of moving objects in water caused by weak light and complex environment, Chen Yuliang et al. [23] have effectively solved the problem of imprecise underwater moving object recognition methods based on deep learning algorithms. By improving convolutional neural networks, Tian Bin et al. [24] effectively solved the problem of difficulty in feature extraction of power frequency underwater magnetic target signals. The research of Cao Jianrong et al. [25] has effectively solved the various factors that affect the target segmentation algorithm. Although many scholars have proposed their own methods for underwater target recognition [26]-[28], as well as research on different problems [29], [30]. However, existing deep learning methods still have some problems when dealing with underwater target recognition, such as strong sensitivity to lighting conditions and insufficient ability to handle image blurring.

This study enhances the performance of a deep convolutional neural network through the incorporation of an attention mechanism aimed at addressing challenges related to variations in lighting conditions, image blurring, and distortion encountered in underwater target recognition. The result indicates that the enhanced network exhibits superior performance in terms of accuracy, recall, and F1 score, achieving values of 0.85, 0.82, and 0.83, respectively.

II. Underwater Target Feature Extractioneature Extraction

In tasks involving the recognition of underwater targets, the process of extracting features is a vital first step. By effectively extracting features from underwater targets, key features of the targets can be more accurately captured, providing a solid foundation for subsequent recognition processes. An incomplete illustration following the process of feature extraction is depicted in Figure 1:

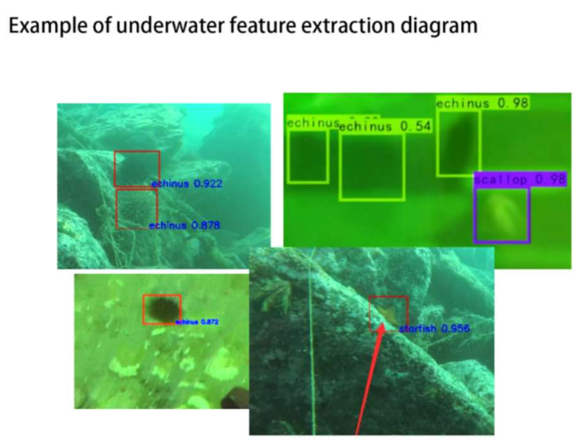

Figure 1: Example of underwater feature extraction map

From the example figure, it can be seen that underwater feature extraction relies on effectively capturing and representing key information in underwater images. Usually, underwater feature extraction utilizes image processing and computer vision techniques to extract texture, color, shape, and other features in the underwater environment through filtering, edge detection, color space conversion, and other methods.

In the feature extraction stage, this paper first uses the pre-trained improved deep convolutional neural network model ResNet-50 to fine tune the underwater target dataset, with the aim of capturing local and global information of the target. Deep learning models can be used to extract feature representations with rich semantic information from underwater images, laying the foundation for the subsequent introduction of attention mechanisms and feature fusion, and constructing feature representations suitable for underwater target recognition. The concept behind ResNet-50 is illustrated in Figure 2:

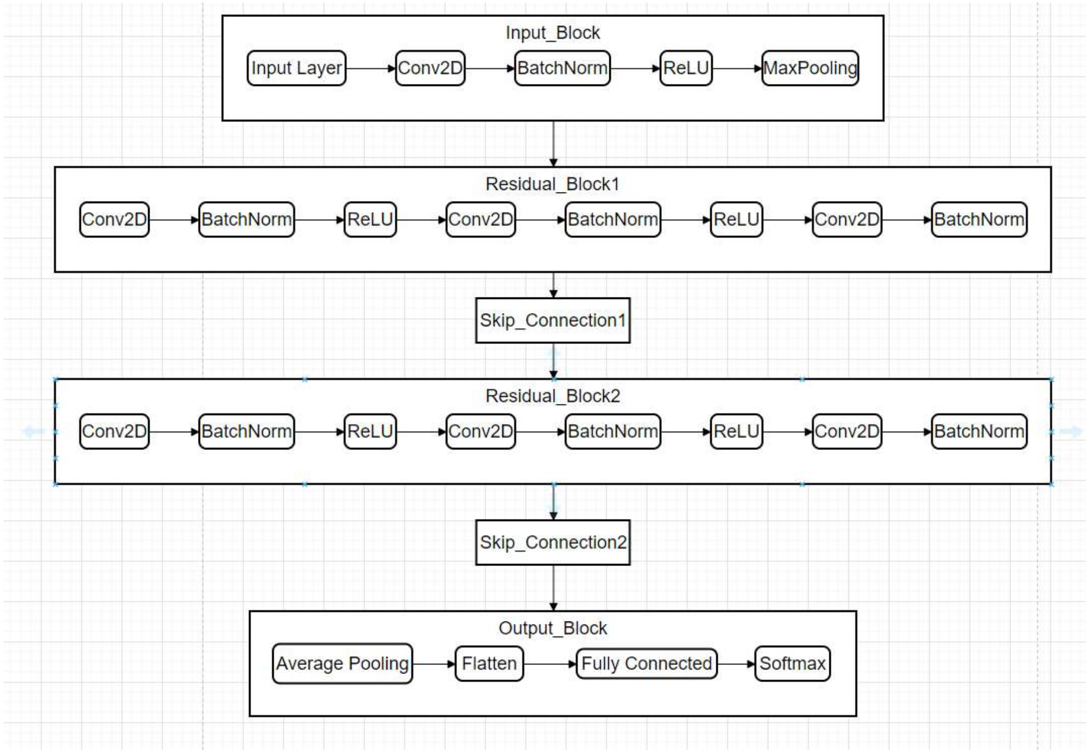

Figure 2: ResNet-50 schematic diagram

ResNet-50 is a complex neural network architecture that includes multiple layers of convolutional operations, where the first convolutional layer extracts features. The second convolutional layer can transform features and directly add input features to the output through skip connections. The main feature of ResNet-50 is the introduction of residual connections, which transmit information through direct connections across layers, solving the problems of gradient vanishing and exploding during deep network training, effectively improving the training effect and performance of the network. The network structure of ResNet-50 is as follows:

Input→Conv1_7x7→MaxPool→ResBlock1→···→ResBlock4→AvgPool→FC→Output

Among them, the Conv1_7x7 denotes a convolutional layer with a 7x7 filter size, MaxPool signifies the maximum pooling layer, ResBlock1 to ResBlock4 indicate four residual blocks, AvgPool denotes the global average pooling layer, FC represents the fully connected layer, and Output signifies the network's final output. By fine-tuning the ResNet-50 model, features with rich semantic information are extracted from underwater image data, laying a solid foundation for the subsequent introduction of attention mechanisms and feature fusion. The process of fine-tuning can be expressed as:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}\left(f_\theta(x_i), y_i\right) \tag{1}$$

Among them, $f_\theta(x_i)$ represents the output of the ResNet-50 model to the input underwater image x_i, y_i represents the true label corresponding to the underwater image x_i. The symbol $\mathcal{L}$ denotes the loss function, while θ stands for the parameters of the ResNet-50 model, and N signifies the quantity of training samples. Through modifying the ResNet-50 model's parameters θ, the loss function is reduced on the underwater target dataset, leading to a better feature representation for underwater environments.

III. Introduction of Attention Mechanism

In this article, a crucial aspect introduced is the attention mechanism. The self-attention mechanism enables the network to automatically learn the importance of each pixel based on the input feature map and highlight the target area by calculating the weight of each pixel in the feature map. The calculation process of the self-attention mechanism is:

$$\text{Attention}(x_i) = \text{softmax}(W_q x_i) \cdot W_v x_i \tag{2}$$

Among them, x_i represents the i-th pixel in the feature map, and W_q and W_v represent the weight matrix of the query and value in the self-attention mechanism respectively. The attention score of each pixel was obtained by multiplying the feature map x_i with the query weight matrix W_q. The attention mechanism mimics the human visual attention mechanism, allowing neural networks to focus on important parts and improving performance and efficiency. To better comprehend the concept of attention mechanism, Figure 3 illustrates the flowchart of how it works:

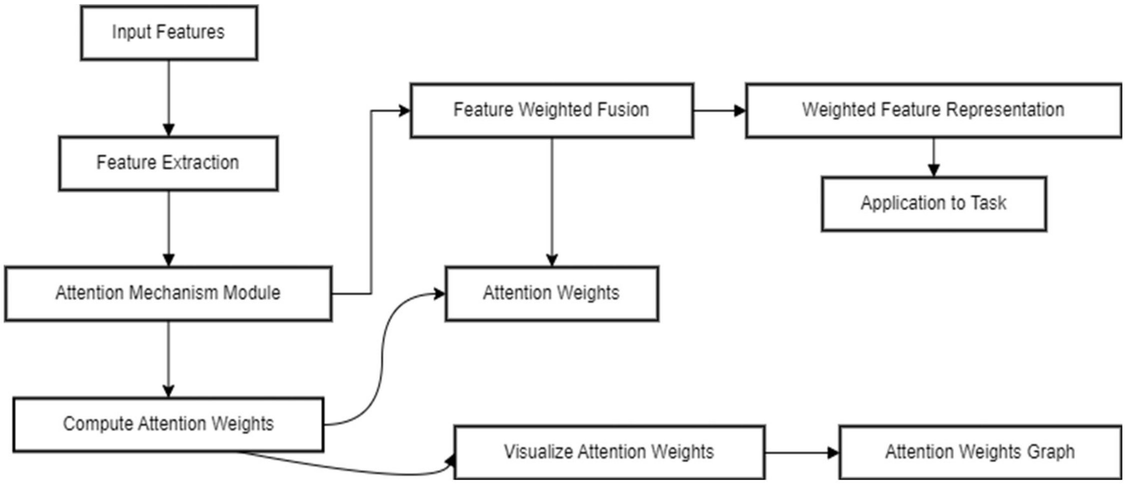

Figure 3: Attention mechanism flowchart

Firstly, the network extracts feature representations of input data, which can be pixels of images, words of text, or time steps of sequences; Then, the attention weight of each feature is calculated based on the context, indicating its importance to the current task, which is usually obtained through learning; next, important features can be highlighted and secondary features can be suppressed based on weighted fusion features, achieving the goal of enhancing network performance; finally, weighted features can be applied to tasks such as image classification and text generation to improve network performance and efficiency.

Subsequently, the attention score is adjusted using the softmax function to calculate the attention weights for each pixel. Ultimately, the feature map x_i is multiplied by the weight matrix W_v to create a weighted feature representation, enhancing the focus on the desired region. The Softmax function here is a commonly used activation function, typically used in multi-class classification tasks. It has the ability to transform a group of input values into a probability distribution where each value falls within the range of 0 to 1, and the total sum of all values equals 1. In order to demonstrate the role of the Softmax function more clearly, this article uses a simple rendering in Figure 4 for demonstration:

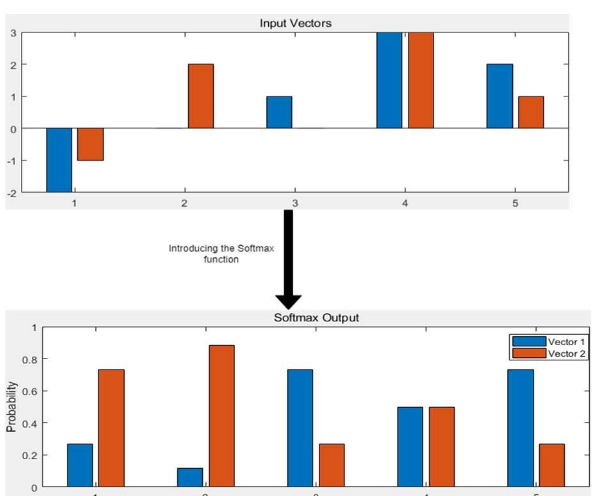

Figure 4: Display of softmax function effect

From the rendering, it can be seen that the Softmax function takes a real number vector as input, and by exponentiating and normalizing each element, it ultimately acquires a vector of probability distribution, with each element indicating the likelihood of the respective category. Specifically, the Softmax function normalizes the index value of each element by dividing it by the sum of the index values of all elements, ensuring that the total output probability is equal to 1. At the same time, the original score is converted into probability form for easy output in multi-classification tasks.

In deep learning, the Softmax function is frequently utilized as the ultimate activation function in Softmax classifiers to transform the neural networks' output into probabilities for different classes. The mathematical expression for the Softmax function is presented as follows:

$$\mathrm{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j})} \tag{3}$$

z_i represents the i-th element in the input vector, and C represents the total number of categories. The Softmax function indexes each element z_i in the input vector, then sums all the indexed values, and finally divides each indexed value by the summed result to get the probability value of each element.

IV. Feature Fusion

In the feature fusion stage, this paper adopts a channel attention mechanism to fuse the attention-weighted features and original features to enhance the representation ability of underwater targets. The channel attention mechanism discussed is an adapted form of the attention mechanism commonly utilized in deep learning architectures, tailored to assign weights to channels within the network to amplify or diminish the feature representation of particular channels. The channel attention mechanism is usually used to process feature maps with multiple channels. For example, in convolutional neural networks, each convolutional layer generates feature maps with multiple channels. The principle of channel attention is shown in Figure 5:

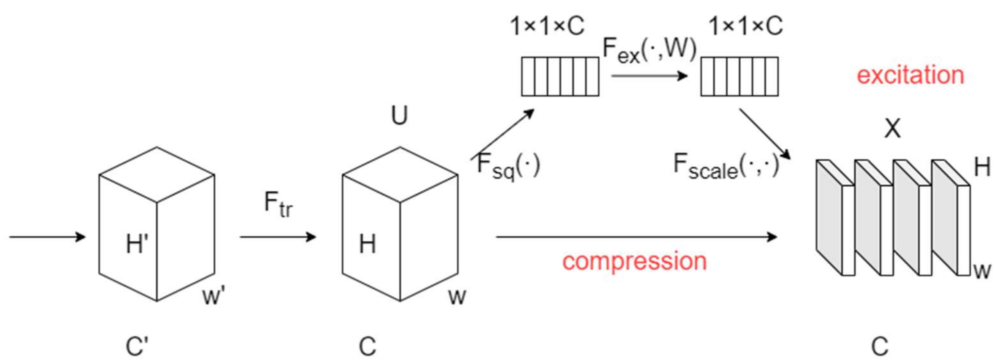

Figure 5: Channel Attention Principle Diagram

In the compression part, the size of the input element feature map is represented as H×W×C, with H, W, and C denoting height, width, and number of channels. The compression part's role is to reduce the dimensions from H×W×C to 1×1×C, essentially compressing the H×W dimensions to a 1 × 1 dimension. This is done using global average pooling. In the excitation stage, it is essential to integrate the 1x1xC dimension from the compression stage into the fully connected layer. This can forecast the significance of each channel and subsequently activate actions on the related channels within the preceding feature map.

The channel attention mechanism enables the network to dynamically adjust the weights of each channel by considering their relative importance, thereby enabling the network to better utilize the correlation information between features and improve underwater target recognition performance. The formula for feature fusion is:

$$\mathrm{Fused\_Feature} = \mathrm{ReLU}(W_f[\mathrm{Original\_Feature}; \mathrm{Weighted\_Feature}]) \tag{4}$$

Among them, ReLU represents the activation function, $[\mathrm{Original\_Feature}; \mathrm{Weighted\_Feature}]$ represents concatenating the original features and attention weighted features by channels, W_f represents the weight matrix in the fusion operation. By fusing the original features and attention weighted features, the information of both can be fully utilized to improve the recognition performance of underwater targets.

This feature fusion method can better utilize the information introduced by attention mechanisms, further improving the performance of underwater target recognition. By adaptively learning the weights of each channel, the network can more effectively capture the correlation between features, thereby enhancing the representation ability of underwater targets. The final fused features can fully leverage the advantages of different features, enabling the network to more accurately recognize underwater targets.

V. Classifier Design

In the classifier design stage, this article uses the Softmax classifier to recognize underwater targets. Softmax classifier is a commonly used multi-class classifier, commonly used in the output layer of deep learning. The Softmax function mentioned earlier is commonly used as the last activation function in Softmax classifiers.

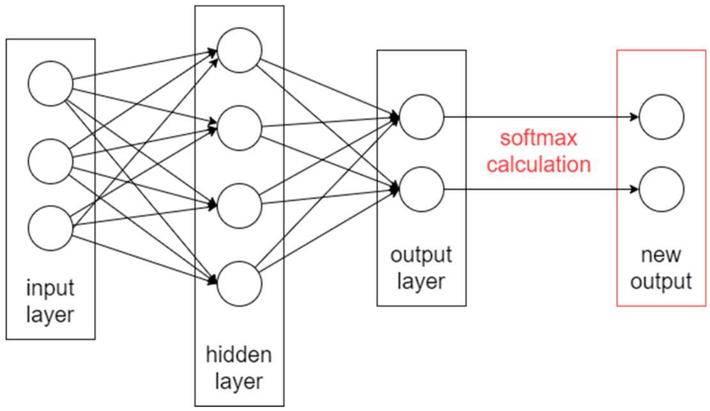

Figure 6: Example of Softmax classifier

Figure 6 summarizes the implementation process of the Softmax classifier. Firstly, the input features can be received and linearly combined to transform them into real vectors representing the original scores of each category; subsequently, the score vector can be input into the Softmax function for exponentiation and normalization to obtain the probability of each category. Finally, a probability distribution vector can be output, where each element represents the predicted probability of the corresponding category.

The Softmax classifier transforms the neural network's output into a distribution of probabilities for a group of categories, ensuring that the likelihood of each category is predicted to be within the range of 0 to 1. The sum of the predicted probabilities for all categories is 1, thus achieving accurate identification of underwater targets. The calculation process of Softmax classifier can be described as follows:

$$P(y=i|x)=\frac{e^{W_i^T x+b_i}}{\sum_{j=1}^{C} e^{W_j^T x+b_j}} \tag{5}$$

$P(y=i|x)$ represents the probability that the target belongs to category i given input x, W_i and b_i represent the weight and bias parameters of the Softmax classifier, and C represents the total number of categories. The Softmax classifier exponentiates the linear combination of input feature x and weight W_i with bias b_i, and normalizes the probability distribution of each category. By selecting the category with the highest probability as the prediction result, the classification and recognition of underwater targets can be completed.

VI. Effect Evaluation

To verify the real effectiveness of the deep convolutional neural network improved by applying the attention mechanism in underwater target recognition, a systematic effect evaluation must be conducted. By comprehensively evaluating the model's performance in different underwater environments and measuring key indicators such as accuracy, recall, and F1 value, it can objectively evaluate the model's performance in identifying underwater targets and the degree of performance improvement. The particular procedure of impact assessment is illustrated in Figure 7:

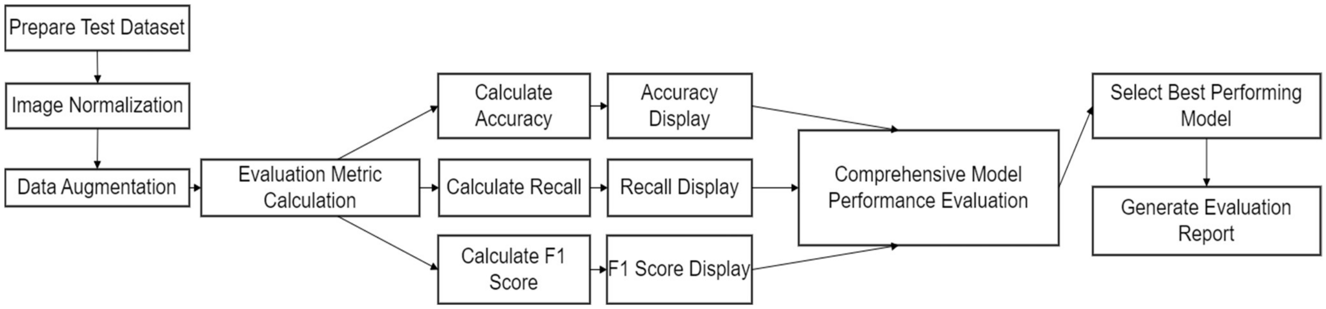

Figure 7: Effect evaluation flowchart

The flow chart of the effect evaluation is shown in Figure 7. First, it can prepare the test data set, input the test data set into the already trained model to predict and obtain the prediction results; secondly, the model on the test data set can be calculated for indicators such as accuracy, recall, and F1 value according to the selected evaluation indicators; then it can select the model with the best performance through comprehensive evaluation and finally generate an evaluation report based on the evaluation results.

VI. A. Selection of Evaluation Indicators

In order to comprehensively evaluate the performance of the model, this article selects accuracy, recall and F1 value as the criteria for evaluating the model.

Accuracy is the proportion of the number of samples correctly predicted by the model to the total number of samples. The precision can be assessed by employing the subsequent equation:

$$Accuracy = \frac{TP+}{TP+TN+FP+FN} \tag{6}$$

In this formula, TP represents the number of occurrences in which the model accurately identifies true data as true, TN signifies the number of occurrences in which the model correctly identifies false data as false, FP denotes the number of occurrences in which the model inaccurately identifies false data as true, and FN indicates the number of occurrences in which the model inaccurately identifies true data as false.

Recall rate refers to the proportion of the number of correct samples successfully predicted by the model to the number of all real samples. The recall rate can be determined using the following formula:

$$Recall = \frac{TP}{TP+FN} \tag{7}$$

The F1 value is the harmonic mean of precision and recall. The formula is as follows:

$$F1 = \frac{2 \times Precision \times Recall}{Precision+R} \tag{8}$$

Among them, Precision is one of the metrics used to evaluate models, representing the ratio of correctly predicted positive samples to all predicted positive samples. It is calculated using a specific formula:

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

VI. B. Preparation of Test Datasets

When creating a test data set, it should make sure to include a variety of scenarios and ensure that the data is representative to some degree. In addition, the test data set also needs to be annotated to ensure that the label information of each sample is true and reliable. The model’s generalization ability and performance can be objectively and thoroughly assessed only with a well-prepared test dataset. This article addresses the issues of poor sensitivity to lighting conditions and insufficient ability to handle image blurring. Based on actual situations, various image samples of underwater scenes were collected. Figure 8 shows some image sample images of underwater scenes:

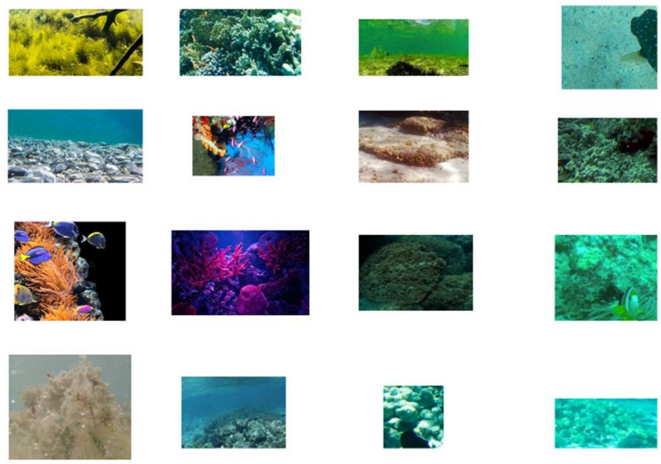

Figure 8: Partial underwater sample images

These image samples cover different underwater environments, different types of targets, and different lighting conditions, with the aim of ensuring the diversity and representativeness of the dataset. Table 1 shows the corresponding test set data table:

Table 1: Test set data

	Rock	Fish	Shellfish	Other
No Processing	100	50	60	30
Blur Processing	70	40	20	10
Lighting Processing	80	60	40	20
Blur + Lighting Processing	90	70	50	40

The data content in Table 1 is the number of test data. This article selected four common underwater images, namely fish, shellfish, rocks, and other types. To more effectively evaluate the model's performance in dealing with sensitivity to lighting variations and its capability to manage image blurriness, this paper processed the selected images with control variables and divided them into four groups: unprocessed group, blurred group, illuminated group, and both blurred and illuminated group. To reflect the recognition ability of the model in complex underwater situations, different processing methods are applied to the images.

VI. C. Model Evaluation

This article highlights five traditional machine learning models: deep convolutional neural network, convolutional neural network combined with long short-term memory network, support vector machine, random forest, and K-nearest neighbor. It has improved the attention mechanism for deep convolutional neural network models, forming a control group and an experimental group. Subsequently, the test dataset can be input into these models and the prediction results of each model can be obtained.

The reason for choosing these models as controls is because each of these classic models has its own characteristics in underwater target recognition. For instance, deep convolutional neural networks excel in tasks related to recognizing images by extracting image characteristics using multiple layers of convolution and pooling operations, and then categorizing them using fully connected layers. The merging of convolutional neural networks with long short-term memory networks combines the spatial feature extraction capability of CNNs with the time series data processing capability of LSTM. Support Vector Machine (SVM) is a classic supervised learning model that separates data of different categories by finding an optimal hyperplane. It is commonly used for underwater target classification. The random forest algorithm is a type of ensemble learning method that utilizes decision trees. It creates several decision trees and then merges them together to perform classification or regression tasks. It has good robustness and generalization ability in underwater target recognition. K nearest neighbor is an instance-based learning method. KNN does not make any assumptions about the distribution of data, and does not require complex parameter adjustment or training processes. Therefore, KNN can adapt to different types and different scales of targets, and can handle underwater target data when the distribution is unclear or irregular. The prediction results of the six models are compared and evaluated, and the final results can provide a reference for the best model.

VI. D. Evaluation Results

The test dataset can be input into different models to obtain the predicted results of the models (Table 2):

Table 2: Evaluation of the effectiveness of different models

Model Name	Accuracy	Recall	F1 Score
DCNN	0.82	0.78	0.80
CNN + LSTM	0.75	0.72	0.73
SVM	0.79	0.75	0.76
RF	0.81	0.77	0.79
KNN	0.73	0.68	0.69
Attention Mechanism Enhanced -DCNN	0.85	0.82	0.83

The DCNN model has achieved a high level of accuracy, recall, and F1 score. In terms of evaluation indicators, the best is the accuracy, which reaches 0.82, but the recall rate and F1 value are slightly lower, 0.78 and 0.8 respectively. The occurrence of this result indicates that the model has misclassification. The CNN+LSTM model performs poorly in terms of accuracy, recall, and F1 value, with values of 0.75, 0.72, and 0.73, respectively. Especially in terms of accuracy, the CNN+LSTM model's accuracy is just 0.02 points better than the lowest-performing KNN model. The LSTM model is usually used to process sequence data. Combining CNN can improve the utilization of time series information. Based on the evaluation findings, the impact of the CNN+LSTM model appears to be minimal, while the SVM and random forest models showed strong performance in the evaluations. The performance is equivalent and at the mid-range level. Their accuracy, recall rate and F1 value are 0.79, 0.75.0.76 and 0.81, 0.77, 0.79 respectively. The three indicator data of the random forest model appear to be slightly elevated compared to those of the SVN model, nevertheless, their performance is not as strong as the deep convolutional neural network model with enhanced attention mechanism due to constraints in feature representation capabilities. The KNN model has the lowest accuracy, recall rate, and F1 value compared to all other models, which are 0.73 respectively, 0.68 and 0.69. Because the KNN model is very simple and intuitive, its performance on complex data sets is not as good as other complex models, so more feature engineering or model adjustments are needed to improve performance. Ultimately, the deep convolutional neural network model with an enhanced attention mechanism has achieved the highest levels of accuracy, recall rate, and F1 value compared to all other models. They are 0.85, 0.82 and 0.83 respectively, which are comprehensively ahead of the better-performing DCNN model.

VI. E. Discussions

In terms of the selected traditional models, although the five selected traditional models have achieved a certain degree of accuracy, recall and F1 value in the underwater target recognition task, they all have certain limitations, such as errors in classification situations, poor performance in dealing with problems with large differences in lighting conditions, image blur and distortion. The deep convolutional neural network improved by introducing the self-attention mechanism shows obvious advantages in underwater target recognition tasks, leading other models with an accuracy of 0.85, a recall of 0.82, and an F1 value of 0.83. This shows that the attention mechanism has a good effect on improving the model for problems such as lighting conditions and image blur.

VII. Conclusions

A large number of experiments have shown that introducing an attention mechanism to improve the underwater target recognition algorithm of a deep convolutional neural network has significant advantages in dealing with illumination-sensitive and blurred images. However, the stability of this method under extreme lighting conditions is still limited. Future research can continue to improve the attention mechanism, explore more effective solutions to underwater environment challenges, and improve algorithm performance and robustness by combining advanced technologies such as reinforcement learning. It is believed that with the development of deep learning technology, underwater target recognition algorithms based on deep convolutional neural networks can provide strong support for marine resource development, environmental protection and other fields.

Acknowledgment

This research was supported by the 2024 Science and Technology R&D Project (Science and Technology Project of State Grid Jiangsu Electric Power Co., Ltd.) (JF2024005).

References

[1] Zhu Daqi, Pang Wenren and Kemeng. Research status and prospects of multi-AUV underwater collaborative search [J]. Journal of University of Shanghai for Science and Technology, 2022, 44(5):417-428.
[2] Gao Han, Liu Weidong, Gao Li'e. Research on underwater target detection algorithm based on forward-looking sonar [J]. Computer Measurement and Control, 2022, 30(11):17-24.
[3] Zhang Muxing, Shen Xiaohong, He Lei, et al. A maximum information coefficient feature selection method for underwater target recognition [J]. Journal of Northwestern Polytechnical University, 2020, 38(3): 471-477.
[4] Li Hui, Wang Xiaoyu, Liu Yun, et al. Underwater target detection integrating multi-scale features and multi-attention [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(20):129-139.
[5] Fu Tongqiang, Hu Qiao, Liu Yu, et al. Underwater target recognition method based on optimized two-dimensional variational mode decomposition and transfer learning [J]. Torpedo Technology, 2021, 029(002):153-163.
[6] Tang Zheng, Wang Zhigang, Zhao Lulu, et al. Research on fusion reasoning method for underwater target recognition based on DS evidence theory [J]. Modern Navigation, 2017, 008(002):113-117.
[7] SONG Kuiyong, ZHOU Lianke, WANG Hongbin. Feature-level Fusion Method for Underwater Multi-source Data[J]. Journal of Jilin University (Information Science Edition), 2021, 39(3): 331-338.
[8] LIN Xinghua, WU Jianguo, QIN Qing, et al. Research on Target Perception Method of Underwater Robot Based on Lateral Line Mechanism[J]. Journal of Ship Mechanics, 2020, 24(5): 559-569.
[9] LIU Xionghou, SHI Rongwei, YANG Yixin. Underwater Slow Small Target Classification Method Using Tracking Trajectory Features and SVDD-SVM Joint Classifier[J]. Acta Acoustica Sinica, 2023, 48(4): 646-655.
[10] Wang Jiawei, Xu Feng, Yang Juan. Multistatic Small Underwater Target Recognition Based on Joint Sparse Representation and Exponential Smoothing in Kernel Space[J]. Acta Electronica Sinica, 2024, 52(1): 217-231.
[11] Han Yongqiang, Zhang Lucheng, Li Lihua, Liu Yongqing. Underwater Anti-jamming Target Tracking Method Based on DFI Integrated Network[J]. Chinese Journal of Inertial Technology, 2022, 30(2): 240-247.
[12] Zhang Yin. Survey of Airborne LiDAR Small Target Detection Technology in Shallow Water[J]. Digital Ocean and Underwater Attack and Defense, 2023, 6(4): 450-457.
[13] Zhang Ming, Ma Xiaochen, LI Jianlong. Self-noise control and experimental verification of autonomous underwater vehicles [J]. Ship Science and Technology, 2020, 42(12): 146-149.
[14] Liu Cong, Han Dong, Zhang Xinyang, et al. Research on underwater acoustic target recognition based on original radiated noise [J]. Electroacoustic Technology, 2023(8):30-37.
[15] Su Changwei, Liang Ran, Wang Xueren, et al. Research progress on line spectrum vibration and noise of underwater vehicles [J]. Ship Science and Technology, 2023, 45(9): 1-8.
[16] Qiao Meiying, Zhao Yan, Shi Jianke, et al. Underwater target detection using high-frequency enhancement network and FPN fusion [J]. Electronic Measurement Technology, 2023, 46(13):146-154.
[17] Lu Jiamin, Song Sanming, Jing Yan, et al. Underwater moving target noise fundamental frequency detection based on DEMON spectrum and LSTM network [J]. Applied Acoustics, 2021, 40(5): 745-753.
[18] Du Libin, Wang Zhengkai, Lu Zhichao, et al. Review of ship target recognition and classification methods based on radiation noise characteristics [J]. Digital Ocean and Underwater Attack and Defense, 2023(5):613-621.
[19] Li Xuegang, Jiang Guoqing, Sun Guocang, et al. Vertical nested array measurement method of radiated noise from underwater low-noise targets [J]. Acta Acoustica Sinica, 2023, 48(5):971-977.
[20] Tian Yu, Tian Wenfeng, Su Yuchen, et al. Research on underwater target recognition technology based on deep neural network [J]. Science and Technology Information, 2023, 21(17):13-16.
[21] Wang Chenyu, Liu Liwen, Liu Gang, et al. Underwater echo generation method based on small sample conditions [J]. Journal of Underwater Unmanned Systems, 2022, 30(5): 612-620.
[22] Na Wang, Kean Chen, et al.Regression model of timbre attribute for underwater noise and its application to target recognitionRegression model of timbre attribute of underwater noise and its application to target recognition[J]. Acta Physica Sinica, 2010, 59(4):2873-2881.
[23] Chen Yuliang, Dong Shaojiang, Sun Shizheng, et al. Improved YOLOv5s low-light underwater biological target detection algorithm [J]. Journal of Beihang University, 2022, 50(2):499-507.
[24] Tian Bin, Li Jun, Yang Chao, et al. Power frequency underwater magnetic target detection based on improved convolutional neural network [J]. Journal of Detection and Control, 2023, 45(1):78-83.
[25] Cao Jianrong, Han Fatong, Wang Ming, et al. Multi-target semantic segmentation algorithm for underwater images based on YOLOv5 and FCN-DenseNet [J]. Computer System Applications, 2022, 31(12):309-315.
[26] Guo Ge, Wang Xingkai, Xu Huipu. Review of underwater target detection, recognition and tracking research based on sonar images [J]. Control and Decision, 2018, 33(5): 906-922.
[27] Zhu Huiqing. Method for extracting structural features of active sonar target echo highlight points based on deep learning [J]. Acoustics and Electronic Engineering, 2023(3):11-15.
[28] Li Mengfei, Liang Jing, Feng Chao. Research on underwater static target recognition method based on improved mean filtering [J]. Electronic Technology and Software Engineering, 2023(8):178-181.
[29] Chen Guojun, Chen Wei, Yu Hanqi. Research on monocular vision underwater robot target tracking method based on deep learning [J]. Machine Tools and Hydraulics, 2019, 47(23):79-82.
[30] Shi Yuchen. Wake detection based on remote sensing image underwater target review [J]. Modern defense technology, 2024, 52 (1) : 83-91.