JURNAL TEKNOLOGI DAN OPEN SOURCE
Vol. 7, No. 2,December 2024, pp. 195~202
e-ISSN: 2622-1659, accredited Four Grade by Kemenristekdikti, Decree No: 152/E/KPT/2023
DOI: 10.36378/jtos.v7i2.3968

Digital Fish Image Segmentation Using U-Net for Shape Feature Extraction

Fathorazi Nur Fajri[1], Mohammad Dzikrillah[2], Ahmad Khairi[3]
[1,2,3]Department of Informatics, Nurul Jadid University, Probolinggo Indonesia

Article Info

Article history:

Received 12 17, 2024
Revised 12 21, 2024
Accepted 12 29, 2024

Keywords:

Digital fish
Feature Extraction
Segmentation
U-Net

ABSTRACT

Segmentation of digital images of fish is an important challenge in image processing in the field of marine biology and aquaculture. Extraction of fish shape features through image segmentation can improve accuracy in species identification and fish population monitoring. The U-Net method, which is based on deep learning, has been proven effective in medical image segmentation and is beginning to be applied in fish image segmentation. This study aims to develop a fish digital image segmentation method using U-Net architecture for accurate and efficient fish shape feature extraction. The dataset used consists of 500 fish images of various shapes and sizes collected from various sources. The fish images were processed using a U-Net artificial neural network, which was trained and tested to obtain the best segmentation results, with evaluation using Intersection over Union (IoU). The segmentation results show that the U-Net method can produce precise segmentation, with a high degree of accuracy in extracting fish shape features. Evaluation of the segmentation metrics resulted in an IoU value of 0.88, indicating excellent performance in distinguishing the fish object from the background and accurately mapping the fish shape. The fish digital image segmentation method using U-Net is effective for fish shape feature extraction and can be applied in fish species identification and aquatic ecosystem monitoring.

Corresponding Author:

Fathorazi Nur Fajri
Department of Informatics
Nurul Jadid University
Probolinggo, Indonesia
Email: Fathorazi@unuja.ac.id

1. Introduction

Segmentation of digital images of fish is one of the major challenges in image processing that is important for marine biology and aquaculture. Fish species identification and population monitoring rely heavily on the ability to accurately extract fish shape features from digital images. In previous studies [1], image segmentation techniques have often faced problems such as fish shape variation, complex backgrounds, and limited accuracy in feature extraction [2]. For example, the study by [3] [4] [5] used a thresholding-based technique, but the results were inadequate in the case of fish with irregular shapes. Therefore, more sophisticated methods are needed to improve precision and efficiency in the process of fish image segmentation.

The U-Net method, which is based on artificial neural networks [6], has proven effective in medical image segmentation [7] [8] and is beginning to be applied in fish image segmentation. The advantage of U-Net lies in its ability to preserve the detailed shape of objects in the image by using deep convolution layers

[9] and decomposition layers for spatial information processing [10]. U-Net has been successfully used in various image segmentation applications, including in a [11] who introduced this architecture for medical image segmentation with highly accurate results. Therefore, the application of the U-Net method to fish images has great potential to significantly improve the quality of fish shape feature extraction.

Although deep learning-based segmentation methods have shown progress, there is still a gap in the specific application of these techniques for fish shape segmentation in aquatic images [12] [13]. Several studies related to fish image segmentation, such as the one conducted [1], rely on traditional segmentation methods or simpler deep learning-based techniques, which are less effective in dealing with the complexity of fish shapes. Recent research applying the U-Net method for fish image segmentation is limited, and the results have not been able to overcome the problem of irregular backgrounds and variations in fish shapes in large datasets [14]. Conclusion: This research aims to fill the gap by developing a more efficient and accurate digital image segmentation method of fish using U-Net.

The purpose of this research is to develop a fish image segmentation method using U-Net architecture that can extract fish shape features with more precision and efficiency. By using U-Net, it is expected to obtain better segmentation results, which can be applied in fish species identification and aquatic ecosystem monitoring. Evaluation of segmentation performance through metrics such as Intersection over Union (IoU) will give an idea of how effective this method is in distinguishing fish from the background and retaining their shape accurately [15]. The main contribution of this research is the application of U-Net to improve the quality of fish image segmentation, which may affect the field of marine biology and aquaculture in the long run.

2. Research Method

This research uses a U-Net architecture-based fish digital image segmentation method that has proven effective in various image segmentation applications [16] can be seen in the figure 1. U-Net is a deep learning method designed for high-fidelity image segmentation, with the ability to handle images that have complex backgrounds and objects that have irregular shapes. Research [17] showed that U-Net can produce highly accurate segmentation in medical applications, and the method has since been used in a variety of fields including aquatic biology image processing. In the context of fish segmentation, some studies [18] [19] also applied U-Net with very satisfactory results in improving the accuracy of fish shape segmentation. Therefore, the U-Net method was chosen for its ability to handle fish image segmentation challenges and is expected to provide optimal results in this study.

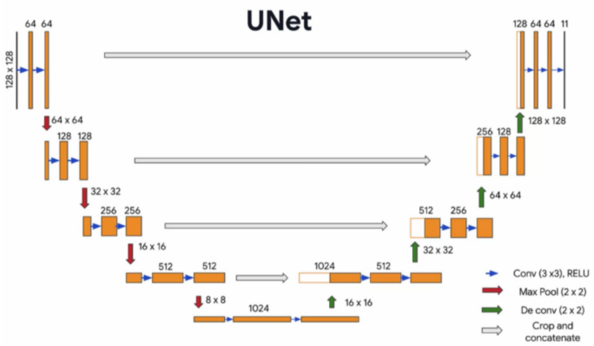

Figure 1. U-Net Architecture

The data in this study was obtained through the collection of digital images of fish caught by fishermen with the type of tuna. Collecting diverse datasets is important to ensure that the model can be trained to recognize different forms of fish present in aquatic ecosystems. The dataset used in this study consists of 500 fish images taken directly from fishermen. Each image was selected with respect to the variety of fish body shapes as well as different backgrounds to provide more challenges to the segmentation model. The use of this varied dataset will increase the generalization of the model in recognizing different shapes of fish in digital images.

The research procedure consists of several stages starting with data preparation, model training, evaluation, and analysis of results [20] can be seen in the figure 2. Each stage is performed chronologically to ensure that the model is trained effectively and the results are evaluated with appropriate metrics. The first step is image data preprocessing which includes normalization and image augmentation to increase data diversity. Once the data was prepared, the U-Net model was trained on the dataset using Adam optimizer-based optimization and binary cross-entropy loss function. The training process was carried out for 50 epochs with a batch size of 16. After the training was completed, the model was tested using a dataset that had never been used before to avoid overfitting. By following this systematic procedure, it is expected that the U-Net model can produce accurate and reliable segmentation.

Figure 2. Research Procedure

Segmentation results will be measured, tested, and evaluated using standard evaluation metrics in image segmentation such as Intersection over Union (IoU). Formula IoU can be seen in the formula 1. These metrics are widely used in image segmentation research as they provide a clear [21] and objective measure of segmentation accuracy [22]. IoU measures how much overlap there is between the modeled segmentation area and the ground truth area measures the similarity between the two. In this study, higher IoU values indicate better segmentation quality. For example, the study [15] used IoU to evaluate medical image segmentation with the results showing a strong correlation between the values of these metrics and segmentation accuracy. Evaluation using IoU will provide a clear picture of how well the model performs fish shape segmentation.

$$IoU = \frac{TP}{TP+FP+FN} \quad (1)$$

3. Result and Discussion

This section discusses research procedures such as data collection, data preparation, model training, evaluation and analysis of results..

3.1. Data Collect

The data in this study were collected from fishermen's catches of tuna. Diversity in the dataset is essential to train the model to recognize different forms of fish in different environmental conditions. The dataset used consists of 500 fish images collected from fishermen's catches can be seen in the figure 3. The images represent a variety of backgrounds as well as lighting. Some images were taken with higher quality, while others contained noise and more complex backgrounds. The collection of diverse and representative data ensures that the model can cope with variations in fish shape and condition in digital images.

Figure 3. Sample dataset Fish

3.2. Data Preparation

The data preparation process involves preprocessing and augmentation to improve data quality and diversity [23] [24]. Preprocessing and augmentation are necessary to ensure that the model can learn effectively from limited data and prevent overfitting. The collected images are processed with noise and orientaion normalization to ensure consistency can be seen in the figure 4. Image augmentation, such as rotation, flipping, and zooming, is performed to enrich the variety of data. This allows the model to learn from various image representations without requiring new data collection. Proper preprocessing and augmentation increases data diversity, allowing the model to generalize better to more complex images.

Figure 4. Remove Noise and Orientation image

3.3. Model Training

The U-Net training model for fish image segmentation starts with an Input Layer that receives a fish image of 256x256 pixels (or any other size adapted to the dataset). This image data has been previously processed with preprocessing techniques, such as size and color normalization, to ensure image consistency.

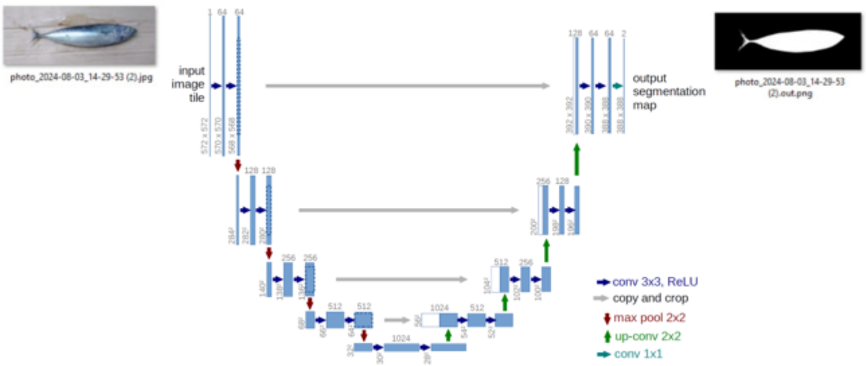

Figure 5. Architecture U-Net Fish Segmentation

In the Encoder or Contracting Path section, a series of 3x3 convolution layers are used to extract features from the input image. Each convolution layer is followed by a ReLU activation function to add nonlinearity and 2x2 max pooling to reduce image dimensionality, while retaining important spatial information. This process continues as the number of filters increases (e.g., 8 → 16 → 32 → 64 → 128), which helps the model recognize increasingly complex features. After that, Bottleneck serves to dig deeper into the feature representation. Here, more convolution layers are applied to increase the complexity level of the model in processing larger and finer image features.

In the Decoder or Expanding Path section, the image size will be enlarged back to its original size (256x256) using transposed convolution or upsampling techniques. Each upsampling layer is followed by a 3x3 convolution. One of the important elements in U-Net are skip connections, which connect the encoder and decoder sections to ensure that information lost during the pooling process can be recovered through concatenation between these two sections [25]. This process ensures that the model can recover spatial details that may have been lost in the encoder stage.

At the Output Layer, the model generates a binary segmentation (fish vs. background) using a 1x1 convolutional layer with sigmoid activation [26], which outputs pixel probabilities to determine whether the pixel belongs to a fish object or not. Training is done using binary cross-entropy as the loss function, as this segmentation is a binary problem, where pixels can only belong to two classes: fish or background. Adam Optimizer is used to speed up the convergence process, and evaluation is done using IoU (Intersection over Union) metrics, which are commonly used in image segmentation problems.

3.3.1. Evaluation

In the first experiment, with the standard model, an IoU value of 0.85 were obtained, indicating good segmentation results. Experiments with data augmentation showed a slight performance improvement, with an IoU value of 0.87, indicating that the additional variation in the dataset helped improve the model's ability to segment fish images. On the other hand, using L2 regularization and decreasing the learning rate slightly degrades the performance, although the model still provides good results. However, by using a larger dataset and larger batch size, the model showed the best performance with an IoU value of 0.88. This indicates that the larger and more diverse the dataset, the better the model is at segmenting, as it is able to handle variations in fish shape and a more complex background.

Table 1. Experiment and result training model

Eksperiment	IoU	Epoch	Learning Rate	Batch Size
1	0.85	50	0.001	16
2	0.87	50	0.001	16
3	0.84	50	0.001	16
4	0.83	50	0.0001	16
5	0.88	60	0.001	32

3.3.2. Analysis of Result

The results of fish image segmentation using the U-Net method show excellent performance in extracting the shape of the fish from the background. Fish objects were separated accurately can be seen in the figure 6 and 7, with clear and detailed contours, even on the complex and thin tail. Consistency of results is seen in various variations of fish orientation, both horizontal and diagonal positions, where the fish shape remains unbroken. However, there is a slight loss of detail in the smaller and thinner areas of the tail, which may be due to the input resolution or the model's sensitivity to very fine features. Nonetheless, the segmentation had minimal background noise, demonstrating the effectiveness of the model in focusing on key features. Overall, these results prove that the U-Net method is capable of performing binary segmentation with a high degree of accuracy, making it highly suitable for marine biology and aquatic ecosystem research. Further refinements, such as the use of higher image resolution or fine-tuning the training parameters, can be made to increase the sensitivity of the model to more complex features.

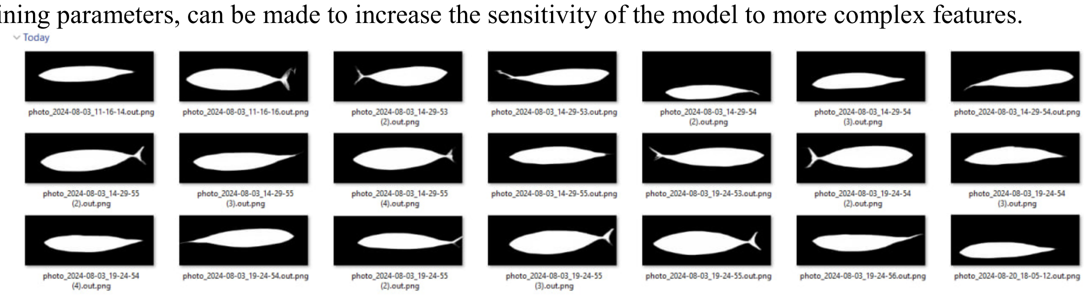

Figure 6. Result Fish Segmentation

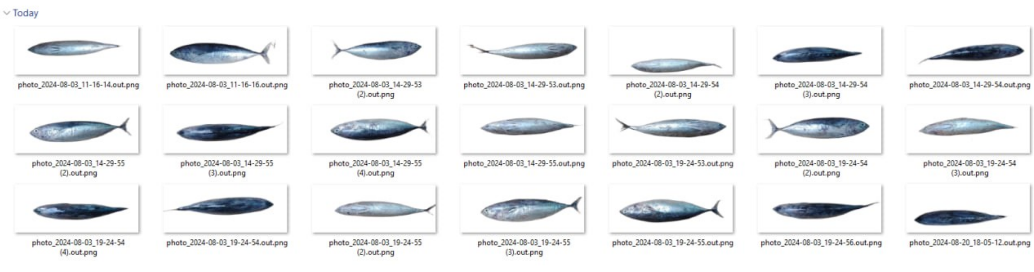

Figure 7. Result Remove Background

4. Conclusion

This research successfully developed a fish image segmentation method using U-Net architecture to extract fish shape features with high accuracy. Through research stages that include data collection, preprocessing, model training, evaluation, and result analysis, the model shows good and consistent performance. Evaluation results using the Intersection over Union (IoU) metric showed a value of 0.88 in the best experiment, indicating the model's ability to accurately distinguish fish objects from the background.

The data augmentation process and the use of a larger dataset proved to improve the model's performance in capturing the details of the fish shape, although some challenges were still found in the very small and thin parts of the features, such as the fish tail. This shows that the input image resolution and training parameters play an important role in precise segmentation. Good fish image segmentation also facilitates the fish shape feature extraction process. With accurate and consistent segmentation results, shape features such as contour, length, width, and proportion of the fish body can be obtained more quickly and precisely.

References

[1] M. R. Kumaseh, L. Latumakulita e N. Nainggolan, "Segmentasi citra digital ikan menggunakan metode thresholding," *Jurnal Ilmiah Sains,* pp. 74-79, 2013.

[2] P. Nabilla, M. F. Saputra e R. A. Saputra, "Perbandingan Ruang Warna RGB, HSV Dan YCbCr Untuk Segmentasi Citra Ikan Kembung Menggunakan K-Means Clustering.," *JATI (Jurnal Mahasiswa Teknik Informatika),* vol. 6, nº 2, pp. 476-481, 2022.

[3] A. Puspaningrum, N. Nur, O. S. Riza e A. Z. Arifin, "Image thresholding based on hierarchical clustering analysis and percentile method for tuna image segmentation," *Nusantara Journal of Computers and its Applications,* vol. 2, nº 1, pp. 1-8, 2018.

[4] A. B. Kaswar, A. Z. Arifin e A. Y. Wijaya, "Segmentasi Citra Ikan Tuna dengan Mahalanobis Histogram Thresholding dan Mahalanobis Fuzzy C-Means," *Jurnal Buana Informatika,* vol. 7, nº 3, 2016.

[5] R. Nuraini, "Implementasi Euclidean Distance dan Segmentasi K-Means Clustering Pada Identifikasi Citra Jenis Ikan Nila," *KLIK: Kajian Ilmiah Informatika dan Komputer,* vol. 3, nº 1, pp. 1-8, 2022.

[6] J. Jing, Z. Wang, M. Rätsch e H. Zhang, "Mobile-Unet: An efficient convolutional neural network for fabric defect detection," *Textile Research Journal,* vol. 92, nº 1-2, pp. 30-42, 2022.

[7] P. Agrawal, N. Katal e N. Hooda, "Segmentation and classification of brain tumor using 3D-UNet deep neural networks," *International Journal of Cognitive Computing in Engineering,* vol. 3, pp. 199-210, 2022.

[8] A. Kermi, I. Mahmoudi e M. T. Khadir, "Deep convolutional neural networks using U-Net for automatic brain tumor segmentation in multimodal MRI volumes," *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop,* pp. 37-48, 2018.

[9] F. Liu e L. Wang, "UNet-based model for crack detection integrating visual explanations," *Construction and Building Materials ,* vol. 322, p. 126265, 2022.

[10] B. S. Vittikop e S. R. Dhotre, "Automatic segmentation of MRI images for brain tumor using unet," *2019 1st International Conference on Advances in Information Technology (ICAIT),* pp. 507-511, 2019.

[11] R. Azad, E. K. Aghdam, A. Rauland, Y. Jia, A. H. Avval, A. Bozorgpour, S. Karimijafarbigloo, J. P. Cohen, E. Adeli e D. Merhof, "Medical image segmentation review: The success of u-net," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2024.

[12] O. Ulucan, D. Karakaya e M. Turkan, "A large-scale dataset for fish segmentation and classification," *2020 Innovations in Intelligent Systems and Applications Conference (ASYU),* pp. 1-5, 2020.

[13] N. F. F. Alshdaifat, A. Z. Talib e M. A. Osman, "Improved deep learning framework for fish segmentation in underwater videos," *Ecological Informatics,* vol. 50, p. 101121, 2020.

[14] N. A. Nezla, T. M. Haridas e M. H. Supriya, "Semantic segmentation of underwater images using unet architecture based deep convolutional encoder decoder model," *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS),* vol. 1, pp. 28-33, 2021.

[15] B. Cheng, R. Girshick, P. Dollár, A. C. Berg e A. Kirillov, "Boundary IoU: Improving object-centric image segmentation evaluation," *Proceedings of the IEEE/CVF conference on computer vision and*

pattern recognition, pp. 15334-15342, 2021.

[16] I. Bidari, S. Chickerur e S. Kadam, "Semantic Segmentation Using U-Net Architecture for Change Detection on Hyperspectral Imagery," *2023 International Conference on Sustainable Communication Networks and Application (ICSCNA),* pp. 932-937, 2023.

[17] G. Du, X. Cao, J. Liang, X. Chen e Y. Zhan, "Medical Image Segmentation based on U-Net: A Review.," *Journal of Imaging Science & Technology,* vol. 64, nº 2, 2020.

[18] Z. Zhang, W. Li e B. Seet, "A lightweight underwater fish image semantic segmentation model based on U-Net," *IET Image Processing,* vol. 18, nº 12, pp. 3143-3155, 2024.

[19] L. Zhang, Y. Qiu, J. Fan, S. Li, Q. Hu, B. Xing e J. Xu, "Underwater fish detection and counting using image segmentation," *Aquaculture International,* pp. 1-19, 2024.

[20] F. N. Fajri, S. Syaiful e W. G. Priambodo, "Fire and Smoke Object Detection Using Mask R-CNN," *Journal of Advanced Research in Informatics,* vol. 2, nº 2, pp. 1-7, 2024.

[21] M. A. Rahman e Y. Wang, "Optimizing intersection-over-union in deep neural networks for image segmentation," *International symposium on visual computing,* pp. 234-244, 2016.

[22] P. Luc, N. Neverova, C. Couprie, J. Verbeek e Y. LeCun, "Predicting deeper into the future of semantic segmentation," *Proceedings of the IEEE international conference on computer vision,* pp. 648-657, 2017.

[23] J. Brownlee, Data preparation for machine learning: data cleaning, feature selection, and data transforms in Python, Machine Learning Mastery, 2020.

[24] D. Klaoudatos, M. Vlachou e A. Theocharis, "From Data to Insight: Machine Learning Approaches for Fish Age Prediction in European Hake," *Journal of Marine Science and Engineering,* vol. 12, nº 9, p. 1466, 2024.

[25] K. Hao, S. Lin, J. Qiao e Y. Tu, "A generalized pooling for brain tumor segmentation," *IEEE Access,* vol. 9, pp. 159283-159290, 2021.

[26] X.-C. Tai, H. Liu, R. H. Chan e L. Li, "A mathematical explanation of UNet," *arXiv preprint arXiv:2410.04434,* 2024.