Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 6, No. 3, Jun 2025, Page. 1669-1682
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

Improving Semantic Segmentation of Flood Areas Using Rotation and Flipping-Based Feature Augmentation

Naili Suri Intizhami*[1], Eka Qadri Nuranti[2], Nur Inaya Bahar[3]

[1,2,3]Computer Science, Institut Teknologi Bacharuddin Jusuf Habibie, Indonesia

Email: [1]naili.suri@ith.ac.id

Received : Mar 24, 2025; Revised : Jul 3, 2025; Accepted : Jul 3, 2025; Published : Jul 9, 2025

Abstract

Semantic segmentation is one of the powerful methods for analyzing flood video or picture data captured by smartphones. However, achieving accurate semantic segmentation requires the application of several methods. In this work, we address the task of feature augmentation approach using rotation (90°, 180°, 270°) and flipping (horizontal, vertical) to improve semantic segmentation of flood areas in Parepare city using a Fully Convolutional Network (FCN). The experimental results demonstrate that the best augmentation scenario 270° rotation achieved an accuracy of 88% and 90° rotation achieved an mean Intersection over Union (mIoU) of 43%, significantly outperforming the baseline FCN model without augmentation, which achieved 86% accuracy and 35% mIoU.

Keywords : *Deep Learning, Feature Augmentation, Fully Convolutional Network (FCN), Semantic Segmentation.*

1. INTRODUCTION

In January 2023, Parepare was back in the spotlight due to the massive flood disaster. It has been reported that the flood occurred due to two simultaneous factors, such as heavy rainfall and tidal waves occurring at the same time [1], [2]. The flood submerged several areas in Parepare city with varying water levels, ranging from ankle-deep and knee-deep to the most severe, reaching waist-deep for an adult. In some locations, the flood was accompanied by strong currents capable of sweeping away objects in its path [3]. Figure 1. Shows the example of the flood with a strong current.

(a) (b)

Figure 1. Flood incident: (a) motorcycle being swept away, (b) water level reaching an adult's thigh

A flood monitoring system is an essential component that must be in place in areas with severe flooding. It is useful for helping related parties manage flood disasters effectively, reducing casualties,

Vol. 6, No. 3, Jun 2025, Page. 1669-1682
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

and minimizing material losses [4]. Flood monitoring can be conducted using visual data, such as flood videos recorded by citizens with smartphone and shared on social media [5]. These video data can be analyzed as they provide a close-up view of flood conditions [6].

Deep learning like convolutional neural network (CNN), is a method that capable of analyzing images or videos and has been widely applied in tasks such as semantic segmentation. Several studies have shown that deep learning for semantic segmentation continues to advance with the emergence of CNN architectures specifically designed for this task, such as Fully Convolutional Network (FCN).

Wang et al [7] introduces a Fully Convolutional Network (FCN) with an end-to-end object detection approach that eliminates the need for post-processing steps such as Non-Maximum Suppression (NMS). The proposed Convolutional Neural Network (CNN) is built upon an FCN architecture integrated with two key innovations: Prediction-aware One-to-One (POTO) label assignment and 3D Max Filtering (3DMF). The research successfully embeds both techniques into the FCN architecture and demonstrates superior performance compared to the baseline.

Ji et al [8] also introduced an advancement of the Fully Convolutional Network within their architecture. This study focuses on improving the accuracy of semantic segmentation, particularly in detecting object boundaries in images, by integrating FCN with edge detection techniques using Holistically-Nested Edge Detection (HED). The proposed approach demonstrates a significant performance improvement compared to baseline methods.

Huang et al [9] provide a comprehensive overview of the development of Fully Convolutional Networks in semantic segmentation for medical imaging. The article argues that although FCNs have demonstrated promising performance in studies involving medical images, further improvements are still needed in terms of network architecture, pre-processing, and post-processing to develop more robust FCN models in the future.

These studies about FCN also emphasize that semantic segmentation based on deep learning holds great potential and can be applied across various domains, including disaster monitoring using videos or images data. Such visual data can be processed using semantic segmentation method, which aims to differentiate all objects in the video or image [10]. This allows for more precise flood analysis. Semantic segmentation works by assigning a label to each pixel [11]. This capability makes semantic segmentation advantageous in identifying and analyzing the flood-affected areas, as well as distinguishing between flooded and non-flooded areas in an image.

Semantic segmentation has several challenges, such as labeling each pixel causes the number of class labels to typically follow the number of objects in an image [12]. Additionally, some objects have a high degree of similarity in their visual appearance, making segmentation less effective [13]. Another challenge is that semantic segmentation performs well on single images, its effectiveness tends to decrease when applied to a large number of images.

Images captured using smartphones also sometimes suffer from poor quality, leading to blurry images or rapid changes in perspective due to random movements of the smartphone user [14]. Other challenges faced by this method for flood images is the similarity of water and the surrounding environment or other objects in the image.

To overcome this challenge, this study employs feature augmentation techniques to improve the accuracy of pixel-level predictions in semantic segmentation. Feature augmentation involves modifying existing images in the dataset to generate additional training samples, thereby increasing data diversity and simulating various real-world conditions [11].

Several studies have shown that feature augmentation is useful for improving the performance and robustness of deep learning models. Alomar et al [15] discuss various data augmentation techniques applicable to image classification and segmentation tasks. The study categorizes augmentation methods into two main groups: traditional techniques such as rotation, flipping, scaling, and translation, and deep

Vol. 6, No. 3, Jun 2025, Page. 1669-1682
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

learning-based techniques such as Generative Adversarial Networks (GANs). The article emphasizes that combining conventional and deep learning-based augmentation methods can significantly enhance model performance when adapted to the characteristics of the dataset and the specific visual task.

Kumar et al [16] also discuss data augmentation techniques commonly used for image data in the context of deep learning. The article categorizes these techniques into two main groups: basic methods such as rotation, flipping, and color manipulation, and advanced methods based on deep learning, including GANs and autoencoders. These techniques are evaluated across various computer vision tasks such as image classification, object detection, and semantic segmentation, with the results demonstrating significant improvements in model performance.

In this study, feature augmentation techniques such as rotation and flipping are applied to address the challenge of varying viewpoints. Flipping alters the spatial orientation of objects through geometric transformations along a specific axis, while rotation changes the position of objects in an image by rotating them at predefined angles [13]. This work systematically evaluates the impact of rotation and flipping on FCN performance for flood segmentation, highlighting the importance of orientation diversity in training data for improving model robustness.

2. METHOD

This section describes several aspects related to this study, including the dataset used, feature augmentation techniques, and the Fully Convolutional Network (FCN) as the CNN model. However, before discussing the sub-sections, the authors need to explain the scenarios carried out in this study as illustrated in Figure 2.

Figure 2 shows an illustration of the proposed method, which is a combination of several feature augmentation techniques and FCN.

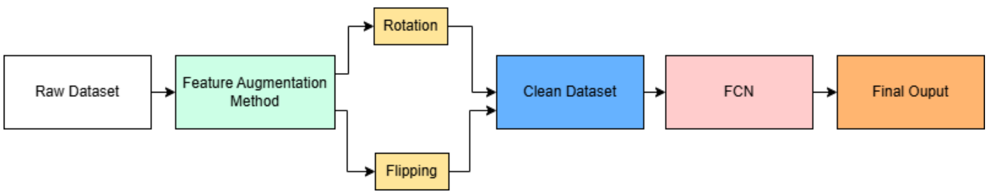


Figure 2. Illustration of the proposed method

Figure 2 illustrates the proposed method in this study, showing that the research begins with the collection of flood image datasets. The data collection process is explained in subsection 2.1 Following this, the study applies two types of feature augmentation techniques, namely rotation and flip, which are described in subsection 2.2. Once the dataset has undergone feature augmentation, it is then used to train the FCN model for semantic segmentation. The discussion about FCN is presented in subsection 2.3. After obtaining the model's testing results, the next step is to evaluate the model using several evaluation metrics, which are explained in the results section.

2.1. Dataset

This study uses a previously annotated flood image dataset. The dataset consists of approximately 3000 flood images from flood events that occurred in Parepare city in 2023. The dataset includes images showcasing different areas affected by the flood, ranging from minor to severe flooding. These images have several distinguishing criteria, including location, lighting conditions, and objects within the images [17].

This dataset is specifically created for semantic segmentation tasks, ensuring that all objects in the image have been classified into several classes, namely vegetation, buildings, vehicles, flood, sky,

P-ISSN: 2723-3863
https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

and people. These classes of objects will be used to develop a CNN model capable of performing semantic segmentation, not only to distinguish between flood and non-flood areas but also to differentiate each object present in an image. The number of datasets used in this study is approximately 800 images. We do not utilize all the images available in the dataset, as the total number of images will increase with the application of feature augmentation. Figure 3 shows an example of flood data images.

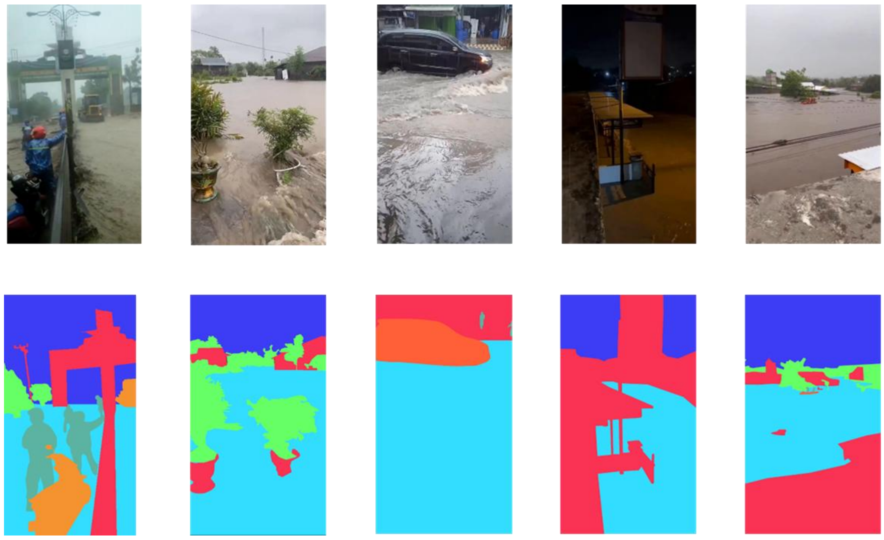


Figure 3. Example of images to be used: (a) original image, (b) annotated image [17]

2.2. Feature Augmentation

This study applies several feature augmentation techniques, including rotation [18] and flipping [19]. These techniques are used as part of the preprocessing stage before training with the CNN model, Fully Convolutional Network (FCN). The purpose of these feature augmentation techniques is to expand the diversity of training data [16], enabling the CNN model to have better generalization capabilities and accurately recognize objects. Although the techniques used may seem simple, their implementation can effectively enhance data variation and strengthen pixel prediction performance.

The rotation technique is a feature augmentation technique used to rotate images at a specific angle [20]. This technique modifies the position of objects based on the angle, forcing the CNN model to learn to recognize objects from different perspectives. Three types of rotation: 90 degrees, 180 degrees, and 270 degrees, which are applied separately to each image used. With the variation in perspectives generated by this technique, the CNN model is expected to recognize objects in various orientations and maintain segmentation accuracy, even when images have unusual angles and orientations. Figure 4 shows the examples of images that have been rotated: (a) 90 degrees, (b) 180 degrees, (c) 270 degrees.

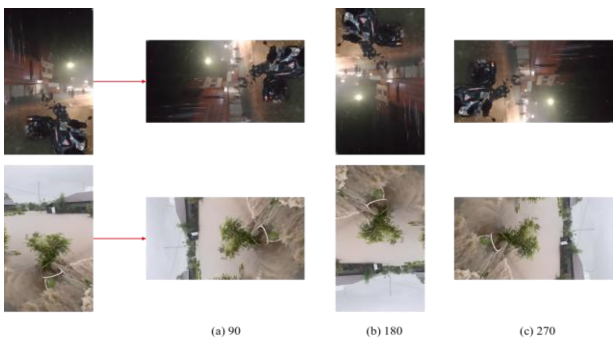
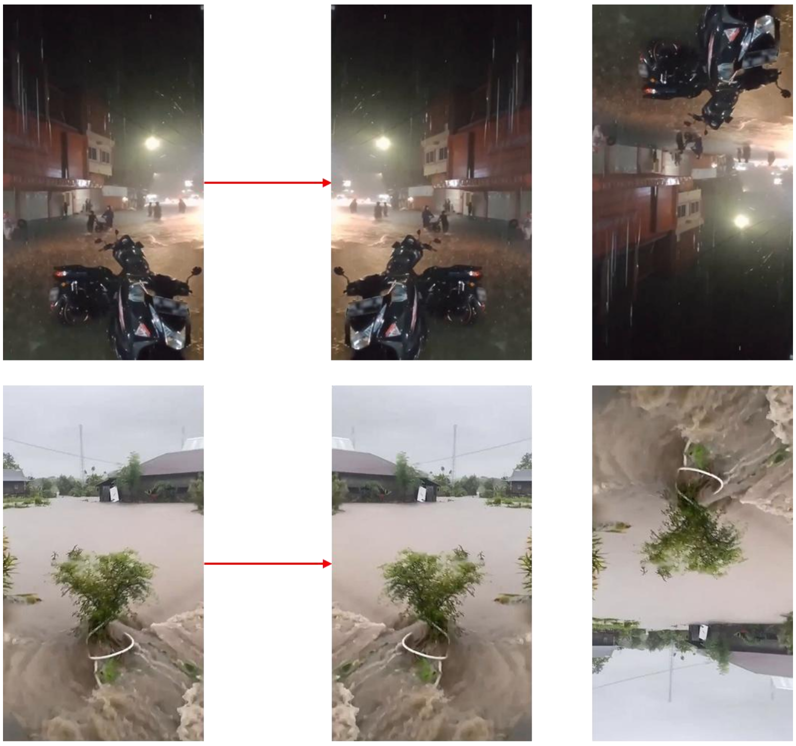


Figure 4. Examples of images that have been rotated: (a) 90 degrees, (b) 180 degrees, (c) 270 degrees

Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 3, Jun 2025, Page. 1669-1682
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

The flipping technique is a data augmentation technique that flips an image along a specific axis to create new image variations [21]. There are two types of flipping: horizontal flip and vertical flip. In this study, both flipping techniques are applied separately to all images. This technique forces the CNN model to recognize objects that may appear in the opposite direction from their original orientation. Figure 5 shows the examples of images that have been flipped: (a) horizontal, (b) vertical.



(a) Horizontal (b) Vertical

Figure 5. Examples of images that have been flipped: (a) horizontal, (b) vertical

As explained in the previous subsection, the number of original images used from the dataset is only around 370 since both augmentation techniques also generate new images, bringing the total to approximately 2,220 images, which will be used for training the CNN model.

2.3. Fully Convolutional Network

The Fully Convolutional Network (FCN) is a convolutional network architecture designed to process images pixel by pixel [22]. This approach aims to preserve the spatial information of the original image. FCN can accept input in the form of an image and produce an output of the same size, with objects clearly segmented according to their respective color classes. This study will utilize FCN as the CNN model for semantic segmentation on images.

FCN uses an encoder-decoder-based architecture [23], which consists of two main stages. In the encoder stage, FCN performs a downsampling process to reduce the image resolution, extracting features that contain essential information from the image while preserving them. This stage consists of three parts, where each part aims to extract deeper features from the image.

The encoder consists of six layers as follows:

- Conv2D 1: 64 filters, 3x3 kernel, stride 1, padding 1, followed by a ReLU activation function.
- Conv2D 2: 128 filters, 3x3 kernel, stride 1, padding 1, followed by a ReLU activation function dan 2x2 MaxPooling.

- Conv2D 3: 256 filters, 3×3 kernel, stride 1, padding 1, follwed by 2x2 MaxPooling.
- Conv2D 4: 512 filter, kernel 3×3, stride 1, padding 1, followed by 2x2 MaxPooling,

After the features are successfully extracted in the encoder stage, the process continues to the decoder stage. This stage functions to restore the resolution of the predicted image to match the original size. The decoder in this study consists of three transposed convolution layers responsible for the upsampling process, and one final convolution layer responsible for generating the output.

The decoder structure used in this study is as follows:

- ConvTranspose2D 1: 256 filters, 2x2 kernel, stride 2, followed by ReLU activation function.
- ConvTranspose2D 2: 128 filters, 2x2 kernel, stride 2, followed by ReLU activation function.
- ConvTranspose2D 3: 64 filters, 2x2 kernel, stride 2, followed by ReLU activation function.
- Conv2D Output: 7 filters (corresponding to the number of classes), 1x1 kernel, no activation function (since softmax is applied through CrossEntropyLoss).

The output of this FCN model is a segmentation prediction with a spatial dimension of 256×256 pixels, matching the image size used in this study. The depth consists of 7 channels representing the probability of each pixel belonging to one of the predefined classes. Each pixel is classified into one class based on the maximum prediction score.

Figure 6 illustrates the FCN architecture used in this study for the image-based semantic segmentation task.

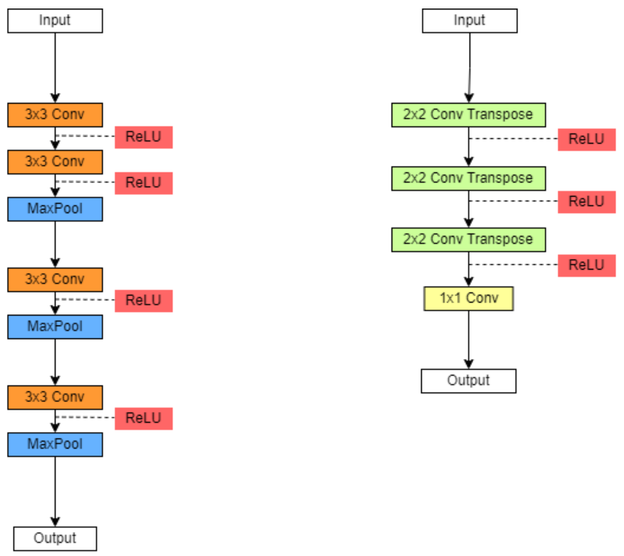


Figure 6. The FCN architecture will be used.

3. RESULT

In this section, we present the results of the proposed method. The evaluation is conducted in two ways: visually examining the semantic segmentation results and assessing performance through several evaluation metrics, including accuracy, precision, F1-score, and mean Intersection over Union (MIoU). This evaluation provides a comprehensive overview of the proposed method's ability to perform semantic segmentation.

Semantic segmentation uses an additional metric, MIoU [8], [24], which measures the degree of overlap between the predicted segmentation and the ground truth for each pixel in the image. A high MIoU value indicates that the model not only performs segmentation well but is also capable of accurately distinguishing objects within the image.

The experiments in this study are divided into several scenarios, as presented in Table 1 and Table 2. The main focus of these experiments is to enhance pixel prediction in semantic segmentation through

Vol. 6, No. 3, Jun 2025, Page. 1669-1682
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

feature augmentation. The purpose of these experiments is to evaluate the impact of feature augmentation on improving FCN prediction results.

Table 1. Scenario with rotation method

Num of Exp	FCN Settings			
	Batch Size	Feature Augmentation	Num of Images	Epoch
1	64	Rotation 90	740	50
2	64	Rotation 90	740	100
3	64	Rotation 180	740	50
4	64	Rotation 180	740	100
5	64	Rotation 270	740	50
6	64	Rotation 270	740	100
7	64	Baseline	370	50
8	64	Baseline	370	100
9	64	Baseline, Rotation 90, 180, 270	1478	50
10	64	Baseline, Rotation 90, 180, 270	1478	100

Table 2. Scenario with flip method

Num of Exp	FCN Settings			
	Batch Size	Feature Augmentation	Num of Images	Epoch
1	64	Baseline	370	50
2	64	Baseline	370	100
3	64	Flip Horizontal	740	50
4	64	Flip Horizontal	740	100
5	64	Flip Vertical	740	50
6	64	Flip Vertical	740	100
7	64	Baseline, Flip Horizontal, Flip Vertical	1110	50
8	64	Baseline, Flip Horizontal, Flip Vertical	1100	100

It has been previously explained that this study applies two feature augmentation techniques, namely rotation and flipping. Each of these techniques has its own scenarios with several adjusted parameters. However, before that, all images will be resized to a resolution of 256 x 256 pixels. This step is useful for ensuring that all images have a uniform size and for reducing the risk of segmentation errors caused by differences in image scale [25].

Additionally, the experiment is conducted by adjusting several key parameters that play a role in the training process, including batch size, number of epochs, augmentation techniques, and the amount of training data. Each parameter is selected to optimize the performance of the FCN model, ensuring accurate semantic segmentation results.

Batch size is closely related to the number of data samples processed by the model in a single iteration [26]. The chosen batch size affects computational efficiency. The number of epochs represents the total number of full cycles the model performs on the training data. The determination of the number of epochs influences how well the model learns data patterns but also carries the risk of overfitting. Next, the amount of training data refers to the number of images used to train the model, which directly affects the segmentation results produced by FCN. Lastly, feature augmentation techniques vary for each scenario, as this is the key parameter being evaluated. The goal is to determine whether applying specific augmentation techniques can enhance semantic segmentation accuracy.

Vol. 6, No. 3, Jun 2025, Page. 1669-1682
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

The rotation technique is divided into ten scenarios, which can be seen in Table 1. These scenarios apply three different rotation angles, namely 90, 180, and 270 degrees. Additionally, there is a scenario as baseline that used as a comparison. As an additional experiment, a scenario combining all images from each rotation scenario is included to evaluate the reliability of FCN.

The flipping technique is divided into eight scenarios, as shown in Table 2. The experiments involve two different flipping techniques, namely horizontal flip and vertical flip. Similar to the rotation technique scenarios, there is also a scenario as baseline that used as a comparison, along with a scenario that combines all images from each flipping scenario for further analysis.

Figure 7 presents the results of the scenario using the rotation technique. This image demonstrates the impact of rotation on the reliability of FCN in performing semantic segmentation. It also shows that FCN's prediction results are improved when using the rotation technique compared to those baseline.

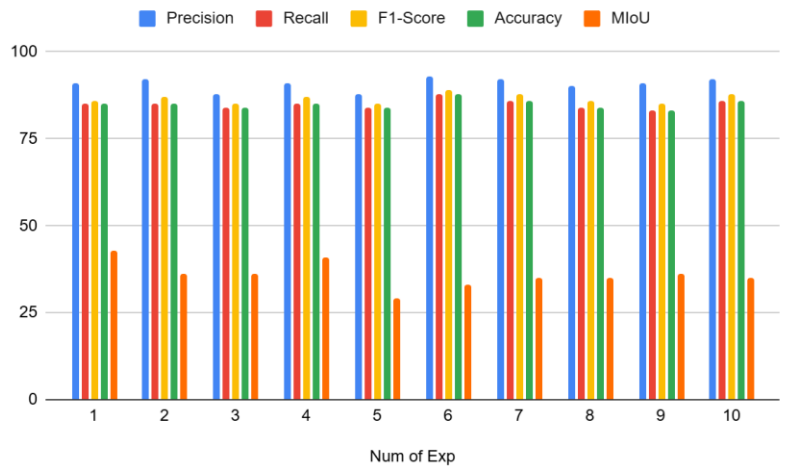


Figure 3. Evaluation Results for Experiments Using the Rotation Technique

In Figure 7, it can be seen that experiment number 6 provides better results across almost all evaluation metrics used. This scenario utilizes a batch size of 64, a 270-degree rotation, 100 epochs, and a total of 740 image data samples. However, in scenario 6, the evaluation for MIoU did not yield optimal results. Based on these findings, Figure 8 presents the visualization of semantic segmentation results for scenario 6.

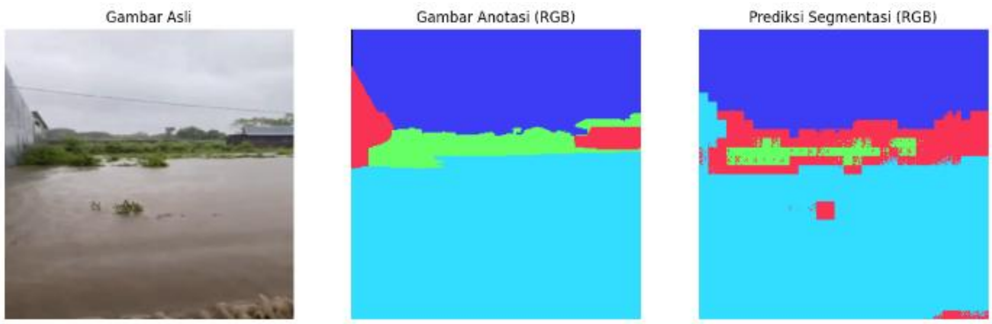


Figure 4. The Results of Scenario 6

Figure 8 displays three images: the original image (left), the ground truth (center), and the prediction result (right) from Scenario 6. From the prediction image, it can be observed that there are some inaccurate predictions in certain classes. For example, areas that should have been classified as buildings were predicted as flood regions. This misclassification contributes to the lower MIoU score

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 3, Jun 2025, Page. 1669-1682
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

compared to other scenarios. However, as shown in Figure 7, aside from MIoU, Scenario 6 performs quite well across all other evaluation metrics. This indicates that Scenario 6 is capable of making accurate predictions, especially for classes with a large number of pixels, such as flood and sky.

Then, referring back to Figure 7, it can be seen that Scenario 1 achieved the highest MIoU result among all other scenarios. Scenario 1 utilized a batch size of 64, a 90-degree rotation, 50 epochs, and a total of 740 image data samples. Figure 9 presents the visualization of semantic segmentation results for Scenario 1, also showing the original image (left), ground truth (center), and the prediction result (right) from Scenario 1.

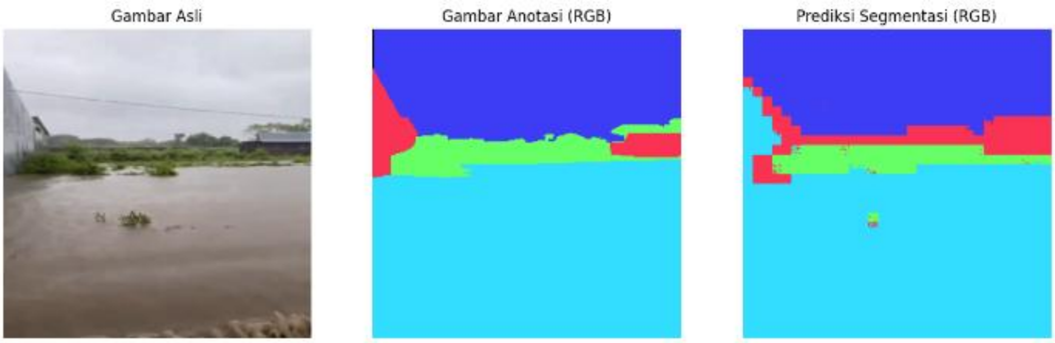


Figure 5. The Results of Scenario 1

In Figure 9, it can be observed that there are still some misclassifications in the building and vegetation classes. However, overall, the prediction results of Scenario 1 are better than those of Scenario 6. This indicates that the FCN model is capable of accurately predicting object areas. Therefore, it can be concluded that Scenario 1 and Scenario 6, which achieved MIoU scores of 43% and accuracy of 88% respectively using the rotation technique, were able to improve prediction results compared to the baseline. According to Figure 7, the baseline achieved an accuracy of 81% and an MIoU of 32%.

Meanwhile, the results of the flip scenario can be seen in Figure 10. This image illustrates the impact of the flip technique on the reliability of FCN in performing semantic segmentation. From the figure, it can be observed that FCN's prediction results using the flip technique are not optimal. In fact, the prediction results from baseline were able to match the performance of the scenario using the flip technique.

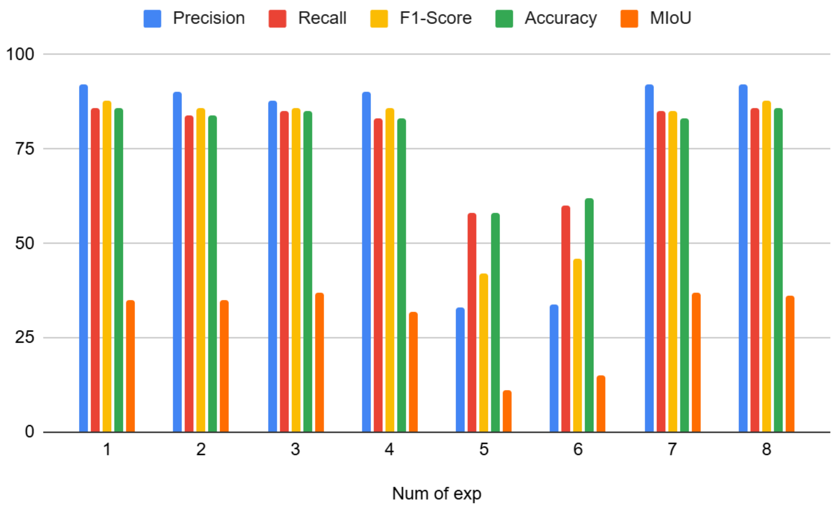


Figure 6. Evaluation Results for Experiments Using the Flipping Technique

Vol. 6, No. 3, Jun 2025, Page. 1669-1682
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

Figure 10 presents the results across all scenarios using the flip augmentation technique. It can be observed that the baseline achieves comparable results in terms of precision, recall, F1-score, and accuracy to those of the flip-based scenarios. It is also evident that some scenarios produced very low evaluation scores.

Although the baseline was able to match most of the evaluation results obtained from scenarios using the flip technique, there was one scenario that achieved a better MIoU score which is Scenario 3. This scenario applied the horizontal flip technique, with a batch size of 64, 50 epochs, and 740 image samples. Figure 11 shows the visualization of the results from Scenario 3 using the flip technique, where the left image is the original image, the center image is the ground truth, and the right image is the predicted output.

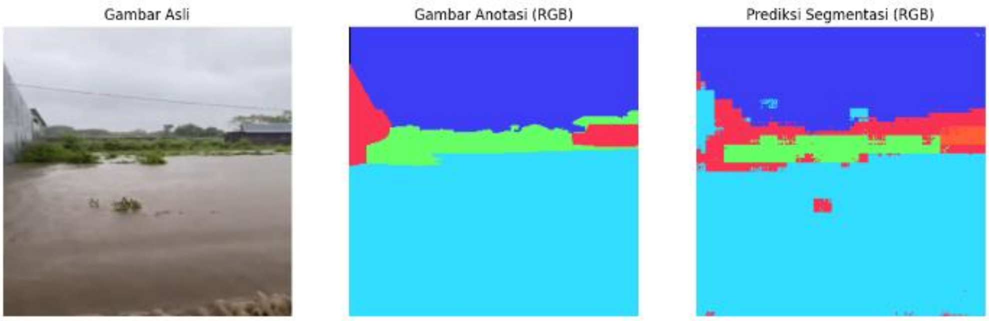


Figure 7. The Results of Scenario 3

Figure 11 shows that the prediction results are not very accurate, with many classes being misclassified. This contributes to the less-than-optimal performance in several evaluation metrics. However, this scenario was able to accurately predict classes with a large number of pixels, which resulted in a better MIoU score compared to other scenarios even slightly higher than the baseline MIoU.

Furthermore, referring back to Figure 10, it can be seen that Scenario 7 also produced relatively good results. This scenario used a batch size of 64 and combined all image variations, including horizontal flips, vertical flips, and original images, with 50 epochs and a total of 1,110 images. Figure 12 presents the visualization of the prediction results for Scenario 7 using the flip technique.

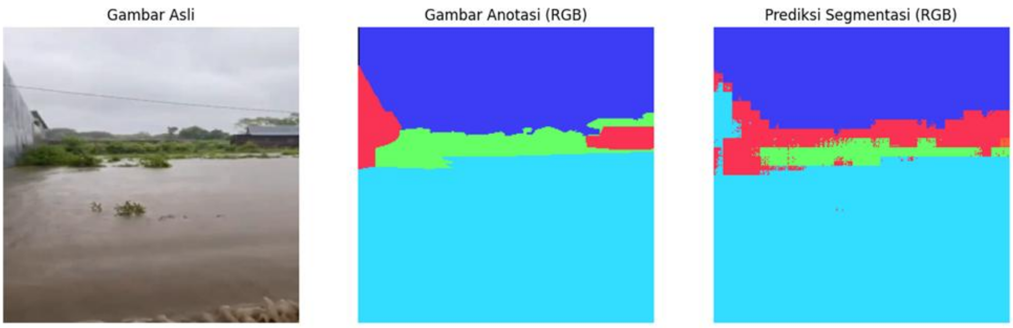


Figure 8. The Results of Scenario 7

In Figure 12, it can be observed that there are still some misclassifications in the prediction results, such as pixels that should belong to the vegetation class but are instead predicted as buildings. This scenario achieved an MIoU score of 37%, which is similar to the metric evaluation result of Scenario 3 in Figure 11. However, when comparing the visualizations of both scenarios, a clear difference can be seen. In Scenario 7, although some misclassifications still occur, they are fewer than in Scenario 3. This indicates that the FCN model in Scenario 7 is able to produce smoother segmentation results, as evidenced by improvements in minor classes such as buildings and vegetation.

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

All experiments ran successfully. Nevertheless, several scenarios produced less than optimal results. Not all augmentation techniques applied were effective; some worked well with the FCN model, while others did not. These results also show that rotation techniques can improve the FCN model's ability to handle various object orientations. Horizontal flipping, which introduces realistic variations, still allowed the FCN to perform semantic segmentation effectively. On the other hand, vertical flipping disrupted the visual context of the images, which led to lower segmentation performance. A more detailed explanation of these experimental findings will be discussed in the following section.

4. DISCUSSIONS

This section provides an analysis of all the experimental results described in the previous section. The analysis focuses on evaluating the performance of each scenario, selecting the most effective feature augmentation techniques, and examining how these techniques influence the accuracy and reliability of semantic segmentation using the FCN model.

Table 3 presents the scenarios from both augmentation techniques used in this study. In this figure, it can be seen that the best scenarios for the rotation technique are Scenario rotation 90 and Scenario rotation 270, while the best scenarios for the flip technique are Scenario flip horizontal and Scenario that combine all feature augmentation and without augmentation. Table 3 also displays the feature augmentation techniques applied along with their corresponding evaluation metric results.

Table 3. Evaluation Results of the Optimal Scenarios from Both Feature Augmentation Techniques and Baseline

No	Feature Augmentation	Results				
		Precision	Recall	F1-Score	Accuracy	MIoU
1	Flip Horizontal	88	85	86	85	37
2	Baseline, Flip Horizontal, Flip Vertical	92	85	85	83	37
3	Rotation 90	91	85	86	85	43
4	Rotation 270	93	88	89	88	33
5	Baseline	92	86	88	86	35

Table 3 presents a comparison of the results from all scenarios discussed in the previous section. This table also shows a comparison between the model results with and without augmentation. It demonstrates that the rotation technique yields better results than the baseline, while the flip technique also provides good results, although it still falls short of outperforming the rotation technique and tends to be similar to the baseline results.

The table also indicates that the use of augmentation techniques can improve the FCN's performance in segmenting flood areas. This is further supported by the visualization results, where the segmented flood regions are clearly delineated in most scenarios presented in the previous section. In fact, for the rotation technique, the visualization results are nearly identical to the ground truth images. This suggests that augmentation techniques play a crucial role in providing diverse orientation in the training data, thus enhancing the robustness of the FCN model.

Vertical flip, however, yielded poor results in this study. This is due to the unnatural perspective it creates, which confuses the FCN model during training. This emphasizes that the choice of augmentation technique should be aligned with the real-world orientation of objects.

Misclassifications in some minor classes reveal a lack of visual variation in the images, which prevents the FCN model from gaining enough information, resulting in poor or incorrect predictions. For example, in Figure 11, it is shown that objects belonging to the building class are incorrectly predicted as vehicles. This is caused by the limited number of pixels representing buildings or vehicles in the images, making it difficult for the model to recognize them.

Vol. 6, No. 3, Jun 2025, Page. 1669-1682
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

In addition to the lack of visual variation, misclassification in minor classes may also occur due to data imbalance among classes. As a result, the FCN model in this study tends to prioritize accurate predictions for majority classes. This explains why the visual prediction results for flood and sky appear very accurate, while predictions for plants, vehicles, humans, and buildings are less reliable. However, upon closer inspection, the FCN model is still able to correctly predict some pixels for the plant or building classes.

Nevertheless, although the FCN model delivers fairly good results, it still requires further development. The results presented in the table and figures show that the FCN model can perform predictions, but some classes are still misclassified. The relatively low MIoU scores also indicate that the FCN model used in this study needs to be further improved.

All scenarios conducted in this study ran smoothly. However, a clear difference was observed in the evaluation results between the two feature augmentation techniques. Differences in shape, axis, and object position significantly affect the FCN model's ability to perform segmentation. The rotation technique provides diverse visual perspectives of objects without significantly altering their recognizability. This is also true for the horizontal flip scenario.

5. CONCLUSION

This study proposes the use of feature augmentation techniques to improve pixel prediction for semantic segmentation. The feature augmentation techniques used are rotation and flipping, while the CNN model employed is the Fully Convolutional Network (FCN), applied to a flood image dataset. Experimental results show that the rotation technique provides good and still realistic visual variations, resulting in high accuracy and MIoU values. For the flip technique, horizontal flipping proved to be more effective than vertical flipping, achieving an MIoU of up to 37%. Vertical flipping presents a perspective that the FCN model was unable to recognize, thus lowering segmentation performance.

Overall, these findings confirm that choosing the appropriate feature augmentation technique can enhance the reliability and prediction capability of the FCN model for semantic segmentation. The results also emphasize the importance of selecting the right augmentation strategy, especially in real-world applications such as flood monitoring using images.

This study highlights the limitations of the prediction results produced by the FCN model used. Therefore, in future research, the FCN architecture will be further developed and integrated with augmentation techniques such as kernels/filters designed to enhance image sharpness and refine object edges, in order to improve performance.

CONFLICT OF INTEREST

The authors declares that there is no conflict of interest between the authors or with research object in this paper.

ACKNOWLEDGEMENT

Acknowledgement is only addressed to funders or donors and object of research. Acknowledgement can also be expressed to those who helped carry out the research.

The authors would like to thank the 2024 DRTPM (Grant No: 001/IT13.D/KH-DRTPM/2024) from the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia for the financial support. The authors would like to thank Institut Teknologi Bacharuddin Jusuf Habibie (ITH) and LPPM-PM ITH (Institute for Research, Community Service, and Quality Assurance) for their support throughout the research process.

Vol. 6, No. 3, Jun 2025, Page. 1669-1682
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

REFERENCES

[1] RENY SRI AYU ARMAN, "Banjir Melanda Parepare, Dua Warga Meninggal," Feb. 02, 2023. [Online]. Available: https://www.kompas.id/baca/nusantara/2023/02/02/banjir-melanda-parepare-dua-orang-warga-meninggal-dunia?open_from=Tagar_Page

[2] U. Sideng, S. Nyompa, and N. C. Rahayu, "PEMETAAN TINGKAT KEKRITISAN UNTUK DAERAH RESAPAN (WILAYAH STUDI KASUS KOTA PAREPARE)," *JES*, vol. 4, no. 1, Dec. 2021, doi: 10.35580/jes.v4i1.22393.

[3] Muh. N. Imam, M. Manaf, and K. Aksa, "Mitigasi dan Adaptasi Bencana Banjir Rob di Kelurahan Cappa Galung Kecamatan Bacukiki Barat Kota Parepare," *ursj*, vol. 7, no. 2, pp. 224–235, Jun. 2025, doi: 10.35965/ursj.v7i2.6051.

[4] L. Sabbatini, L. Palma, A. Belli, F. Sini, and P. Pierleoni, "A Computer Vision System for Staff Gauge in River Flood Monitoring," *Inventions*, vol. 6, no. 4, p. 79, Nov. 2021, doi: 10.3390/inventions6040079.

[5] R. J. Pally and S. Samadi, "Application of image processing and convolutional neural networks for flood image classification and semantic segmentation," *Environmental Modelling & Software*, vol. 148, p. 105285, Feb. 2022, doi: 10.1016/j.envsoft.2021.105285.

[6] U. K. Panchal, H. Ajmani, and S. Y. Sait, "Flooding Level Classification by Gait Analysis of Smartphone Sensor Data," *IEEE Access*, vol. 7, pp. 181678–181687, 2019, doi: 10.1109/ACCESS.2019.2959557.

[7] J. Wang, L. Song, Z. Li, H. Sun, J. Sun, and N. Zheng, "End-to-End Object Detection with Fully Convolutional Network," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA: IEEE, Jun. 2021, pp. 15844–15853. doi: 10.1109/CVPR46437.2021.01559.

[8] J. Ji, X. Lu, M. Luo, M. Yin, Q. Miao, and X. Liu, "Parallel Fully Convolutional Network for Semantic Segmentation," *IEEE Access*, vol. 9, pp. 673–682, 2021, doi: 10.1109/ACCESS.2020.3042254.

[9] S.-Y. Huang, W.-L. Hsu, R.-J. Hsu, and D.-W. Liu, "Fully Convolutional Network for the Semantic Segmentation of Medical Images: A Survey," *Diagnostics*, vol. 12, no. 11, p. 2765, Nov. 2022, doi: 10.3390/diagnostics12112765.

[10] Z. Luo, W. Yang, Y. Yuan, R. Gou, and X. Li, "Semantic segmentation of agricultural images: A survey," *Information Processing in Agriculture*, vol. 11, no. 2, pp. 172–186, Jun. 2024, doi: 10.1016/j.inpa.2023.02.001.

[11] N. S. Inthizami *et al.*, "Flood video segmentation on remotely sensed UAV using improved Efficient Neural Network," *ICT Express*, vol. 8, no. 3, pp. 347–351, Sep. 2022, doi: 10.1016/j.icte.2022.01.016.

[12] M. Rahnemoonfar, T. Chowdhury, A. Sarkar, D. Varshney, M. Yari, and R. R. Murphy, "FloodNet: A High Resolution Aerial Imagery Dataset for Post Flood Scene Understanding," *IEEE Access*, vol. 9, pp. 89644–89654, 2021, doi: 10.1109/ACCESS.2021.3090981.

[13] B. Neupane, T. Horanont, and J. Aryal, "Deep Learning-Based Semantic Segmentation of Urban Features in Satellite Images: A Review and Meta-Analysis," *Remote Sensing*, vol. 13, no. 4, p. 808, Feb. 2021, doi: 10.3390/rs13040808.

[14] S. Soares, G. M. Fernandes, and F. R. P. Rocha, "Smartphone-based digital images in analytical chemistry: Why, when, and how to use," *TrAC Trends in Analytical Chemistry*, vol. 168, p. 117284, Nov. 2023, doi: 10.1016/j.trac.2023.117284.

[15] K. Alomar, H. I. Aysel, and X. Cai, "Data Augmentation in Classification and Segmentation: A Survey and New Strategies," *J. Imaging*, vol. 9, no. 2, p. 46, Feb. 2023, doi: 10.3390/jimaging9020046.

[16] T. Kumar, R. Brennan, A. Mileo, and M. Bendechache, "Image Data Augmentation Approaches: A Comprehensive Survey and Future Directions," *IEEE Access*, vol. 12, pp. 187536–187571, 2024, doi: 10.1109/ACCESS.2024.3470122.

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4564

[17] N. S. Intizhami, E. Q. Nuranti, and N. I. Bahar, "Dataset for flood area recognition with semantic segmentation," *Data in Brief*, vol. 51, p. 109768, Dec. 2023, doi: 10.1016/j.dib.2023.109768.

[18] T. Agustin, E. Utami, and H. A. Fatta, "Implementation of Data Augmentation to Improve Performance CNN Method for Detecting Diabetic Retinopathy," in *2020 3rd International Conference on Information and Communications Technology (ICOIACT)*, Yogyakarta, Indonesia: IEEE, Nov. 2020, pp. 83–88. doi: 10.1109/ICOIACT50329.2020.9332019.

[19] R. Bravin, L. Nanni, A. Loreggia, S. Brahnam, and M. Paci, "Varied Image Data Augmentation Methods for Building Ensemble," *IEEE Access*, vol. 11, pp. 8810–8823, 2023, doi: 10.1109/ACCESS.2023.3239816.

[20] E. Goceri, "Medical image data augmentation: techniques, comparisons and interpretations," *Artif Intell Rev*, vol. 56, no. 11, pp. 12561–12605, Nov. 2023, doi: 10.1007/s10462-023-10453-z.

[21] G. Zhou, Y. Chen, and C. Chien, "On the analysis of data augmentation methods for spectral imaged based heart sound classification using convolutional neural networks," *BMC Med Inform Decis Mak*, vol. 22, no. 1, p. 226, Aug. 2022, doi: 10.1186/s12911-022-01942-2.

[22] H. Li, J. Fan, Q. Hua, X. Li, Z. Wen, and M. Yang, "Biomedical sensor image segmentation algorithm based on improved fully convolutional network," *Measurement*, vol. 197, p. 111307, Jun. 2022, doi: 10.1016/j.measurement.2022.111307.

[23] G. Chen *et al.*, "SDFCNv2: An Improved FCN Framework for Remote Sensing Images Semantic Segmentation," *Remote Sensing*, vol. 13, no. 23, p. 4902, Dec. 2021, doi: 10.3390/rs13234902.

[24] M. M. Rahman and R. Marculescu, "Medical Image Segmentation via Cascaded Attention Decoding," in *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, HI, USA: IEEE, Jan. 2023, pp. 6211–6220. doi: 10.1109/WACV56688.2023.00616.

[25] D. Tribuana, Hazriani, and A. L. Arda, "Image Preprocessing Approaches Toward Better Learning Performance with CNN," *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 8, no. 1, pp. 1–9, Jan. 2024, doi: 10.29207/resti.v8i1.5417.

[26] J.-S. Hwang, S.-S. Lee, J.-W. Gil, and C.-K. Lee, "Determination of Optimal Batch Size of Deep Learning Models with Time Series Data," *Sustainability*, vol. 16, no. 14, p. 5936, Jul. 2024, doi: 10.3390/su16145936.