

## Refining Semantic Segmentation of Flood Images Using Edge Sharpening and CNN

Naili Suri Intizhami<sup>1\*</sup>, Eka Qadri Nuranti<sup>2</sup>, Muhammad Anugrah<sup>3</sup>, Sri Sukma Tahir<sup>4</sup>

<sup>1,2,3,4</sup>Institut Teknologi Bacharuddin Jusuf Habibie, Indonesia

<sup>1</sup>[naili.suri@ith.ac.id](mailto:naili.suri@ith.ac.id), <sup>2</sup>[eka.qadri@ith.ac.id](mailto:eka.qadri@ith.ac.id), <sup>3</sup>[muhhammadanugrah.221011090@mahasiswa.ith.ac.id](mailto:muhhammadanugrah.221011090@mahasiswa.ith.ac.id),

<sup>4</sup>[srisukmatahir.231011137@mahasiswa.ith.ac.id](mailto:srisukmatahir.231011137@mahasiswa.ith.ac.id)



### \*Corresponding Author

#### Article History:

Submitted: 12-05-2026

Accepted: 26-05-2026

Published: 03-06-2026

#### Keywords:

Flood Image Segmentation;  
Semantic Segmentation; Unsharp  
Masking; Edge Sharpening;  
Efficient Neural Network.

**Brilliance: Research of  
Artificial Intelligence** is licensed  
under a Creative Commons  
Attribution-NonCommercial 4.0  
International (CC BY-NC 4.0).

### ABSTRACT

Post-disaster impact analysis is an important component in supporting mitigation planning, emergency response, and evidence-based decision-making after flood events. Visual data, such as flood images, can be used to identify affected areas and analyze environmental conditions through semantic segmentation. Semantic segmentation is a pixel-level classification process that assigns each pixel in an image to a specific object class. However, flood images collected from real-world conditions often have low visual quality, unclear object boundaries, and complex backgrounds, which may reduce the quality of segmentation results. This study proposes an edge-sharpening-based preprocessing approach combined with a Convolutional Neural Network (CNN) model to improve semantic segmentation performance on flood images. The proposed method applies unsharp masking to enhance edge and contour information before the images are processed by the CNN model. The experiments were conducted using flood and non-flood image datasets and compared with the ENet baseline and a modified ENet model. The evaluation was performed using visual comparison and quantitative metrics, including precision, recall, F1-score, accuracy, and mean Intersection over Union (mIoU). The results show that the proposed method achieved the best performance on the flood image dataset, with 98% precision, 98% recall, 98% F1-score, 97% accuracy, and 54% mIoU. These results outperform the comparative CNN models and indicate that edge sharpening can improve object boundary representation, particularly for flood images with blurred or low-quality visual characteristics.

### INTRODUCTION

Floods are natural disasters that cause a significant degree of damage (Intizhami, Nuranti, & Bahar, 2025). In Indonesia, floods remain the most frequently occurring disaster to this day. The main causes of flooding include high rainfall intensity, the reduction of water catchment areas, and the lack of preparedness in many regions to face flood events (Bukhari et al., 2025). In addition, the impacts of flooding are highly detrimental to both humans and infrastructure. Therefore, flood monitoring has become an essential effort to minimize the resulting impacts (Nair, Vallimeena, Gopalakrishnan, Rao, & Krishnamoorthy, 2024).

Flood monitoring is an activity carried out to observe, measure, and analyze flooding periodically in order to understand the situation and potential occurrence of floods in a particular area. This monitoring process is an integral part of flood mitigation, as it can influence the effectiveness of a region's mitigation efforts (Bagheri & Liu, 2025; Muhadi, Abdullah, Bejo, Mahadi, & Mijic, 2021). Flood monitoring activities include measuring rainfall, conducting direct observations and documentation by field officers, as well as analyzing drone or satellite imagery (A. Hammood et al., 2021). However, these activities still have several limitations and are prone to potential errors in both data collection and analysis.

Therefore, flood monitoring that utilizes technologies such as the Internet of Things (IoT) and artificial intelligence (AI) is highly needed. Technology offers convenience and automation, allowing repetitive tasks to be performed efficiently by machines. However, at present, the implementation of technology-based flood monitoring systems is still in its early stages and continues to be developed further (Bukhari et al., 2025; Mandal, Sharma, & Chutia, 2025).

In this study, to predict flood occurrences within an image, the authors employed a Convolutional Neural Network (CNN) model to perform semantic segmentation on flood images (Farhadi, Ebadi, Kiani, & Asgary, 2024). Semantic segmentation is a technique aimed at classifying each pixel in an image into its corresponding class (Cheng, Deng, Su, An, & Wang, 2024). This technique enables machines to understand and distinguish the various objects present within an image. Several studies that specifically discuss the development of CNN models for semantic segmentation in images have shown promising results.

The study by (Emek Soylu et al., 2023) conducted a comprehensive review of the development and current state of semantic segmentation methods, ranging from traditional CNN-based approaches to transformer architectures. This



research presented a complete overview of semantic segmentation methods by comparing their accuracy, efficiency, and model complexity.

The study by (Li, 2021) developed a real-time semantic segmentation method by combining a Generative Adversarial Network (GAN) with ENet (Efficient Neural Network). The results of this study demonstrated an improvement in the performance of the proposed hybrid model, achieving a segmentation accuracy of over 93% and a semantic segmentation simulation time of less than 0.171 seconds.

The study by (Suresh, Sundaralingam, Akilan, & Jahan, 2025) developed a semantic segmentation method by integrating edge context and an attention mechanism. The results of this study showed an improvement in the quality of the semantic segmentation output images compared to the baseline methods used in the research.

The existing studies illustrate that the development of CNN models for semantic segmentation holds great potential and can be applied to various tasks, including flood monitoring. However, research on CNN-based semantic segmentation models still faces several challenges that need further improvement, such as high computational requirements, limited model generalization, and segmentation outputs that do not closely resemble the original input images (Safavi & Rahneemoonfar, 2023).

Therefore, this study focuses on improving semantic segmentation accuracy by applying an image sharpening technique during the pre-processing stage and utilizing a CNN model, namely the Efficient Neural Network (ENet). The study also aims to evaluate the performance of the ENet model and analyze the effect of feature augmentation techniques on the model's performance and prediction results.

### METHOD

This research employs a real-time semantic segmentation CNN model, namely the Efficient Neural Network (ENet), which is further developed to improve its performance, particularly in detecting and segmenting objects more accurately. Therefore, an image sharpening technique is incorporated as part of the data augmentation process to enhance segmentation results.

Data collection is the first stage of this research, involving two types of data: flood images and non-flood images. The flood image data were obtained from the authors' previous study, which documented flood events in Parepare City, South Sulawesi, Indonesia, in 2023. This dataset was specifically developed for semantic segmentation tasks, consisting of original images and their corresponding annotated images (Intizhami, Nuranti, & Bahar, 2023). The non-flood image data used in this study were obtained from the CamVid dataset (Upadhyay, Malhotra, Gupta, & Mishra, 2024). These non-flood images were utilized as a comparative benchmark to evaluate the performance of the CNN model and to assess its capability in performing semantic segmentation. This study utilized an equal number of flood and non-flood images, with approximately 1,478 images in each category. In total, the dataset used in this research comprised around 2,956 images. Fig.1 show the example of data in this research.

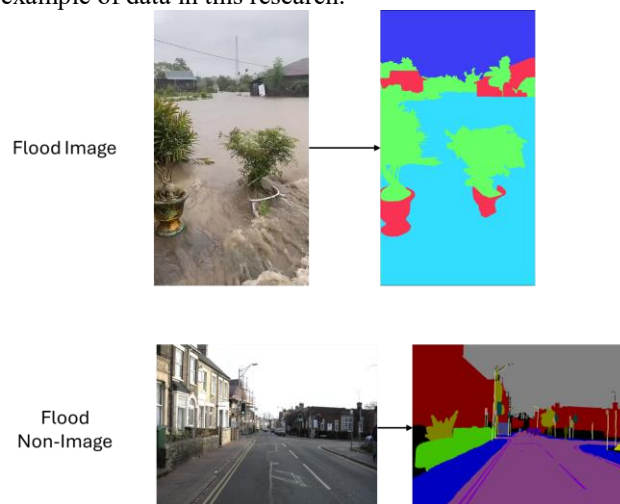


Figure 1. Example of the data used in this study: (top) flood images, (bottom) non-flood images

This study applied feature augmentation with the aim of producing greater data diversity. Feature augmentation also helps improve the performance of the CNN model by generating more varied data samples. In this research, feature augmentation was primarily used as part of the pre-processing stage, namely image sharpening. Image sharpening can enhance the quality of the data by improving the visual clarity of all objects within the image (Balaji, Song, Malekzadeh, Heidari, & Dutta, 2024). As a result, the CNN model can more easily recognize the shapes and textures of each object class in the image.

Image sharpening in this study is implemented using a convolution filter (kernel), specifically through the unsharp masking technique. Unsharp masking enhances contour and edge details in an image, resulting in clearer and

more distinguishable object boundaries (Mishra, Vishwakarma, & Kumar, 2025). This technique is deemed suitable for the flood image dataset used in this study, as the dataset was developed from social media images that are generally of low visual quality. The unclear boundaries between objects in such images often cause CNN models to produce abstract or imprecise object predictions. In addition, unsharp masking can sharpen image edges without altering the overall structure of the image, which differs significantly from techniques such as rotation or flipping that change the original point of view of the image (Lu, Chen, & Pan, 2024).

In this study, the unsharp masking technique is employed to clarify the boundaries between object classes, allowing the CNN model to perform pixel-level recognition and classification more accurately. Furthermore, this sharpening process introduces additional data variations, enriching the learning process of the CNN model. Fig. 2 below illustrates the pre-processing pipeline employed in this study.

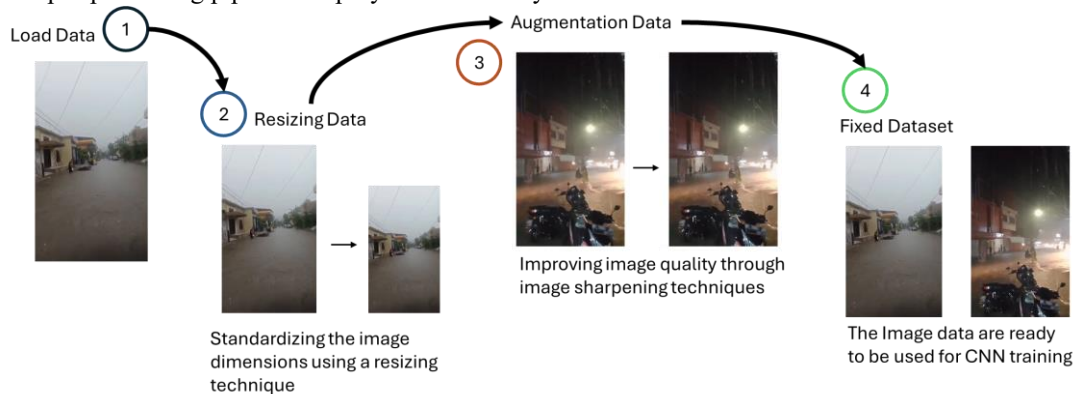


Figure 2. The pre-processing pipeline

The Efficient Neural Network (ENet) is a CNN-based model for semantic segmentation that features a relatively shallow architecture, enabling reduced computational complexity and lower power consumption. It is primarily designed to achieve real-time semantic segmentation performance (Wang, 2022). As a CNN-based model for semantic segmentation, ENet adopts an encoder–decoder architecture that extracts essential features from the input data by reducing spatial dimensions while increasing the number of feature channels. The data are then reconstructed to their original spatial size but in a transformed representation. In the case of semantic segmentation, the output is a segmentation map in which each pixel is assigned to a corresponding object class (Doğan & Ergen, 2024). This study utilizes ENet as the CNN model for semantic segmentation. The ENet architecture implemented in this work is adapted from the version proposed by (Inthizami et al., 2022), which improves the original ENet structure through the integration of depthwise (DSC) and atrous separable convolutions (ASC) on Encoder (EN) and Decoder (DC) steps.

## RESULT

In this study, the authors aim to improve the semantic segmentation results of the CNN model, ENet, on flood images by applying an image sharpening technique. Table 1 presents the experimental scheme conducted in this research.

Table 1. The Experiment Scenarios

No	Method	Dataset	Batch Size	Epochs
1	ENet Baseline	Flood	8	30
2	ENet Baseline	Non-Flood	8	30
3	Enet with EN: ASC DC: DSC	Flood	8	30
4	Enet with EN: ASC DC: DSC	Non-Flood	8	30
5	Proposed Method	Flood	8	30
6	Proposed Method	Non-Flood	8	30

Table 1 includes several experimental scenarios designed to evaluate the performance of the proposed method in comparison with existing approaches. As previously explained in the dataset section, two types of image data are used, flood and non-flood images. All experimental scenarios are carried out using the same hyperparameter configuration. The hyperparameters defined in Table 1 are batch size and epochs. Standardizing the hyperparameters ensures that the model performance can be evaluated fairly and validly, with a batch size of eight and a total of thirty epochs.

The model evaluation is conducted through two approaches: visual analysis of the semantic segmentation outputs and quantitative comparison using evaluation metrics such as accuracy, precision, recall, F1-score, and mean Intersection over Union (mIoU). The proposed method was also compared with the original ENet and the modified ENet.

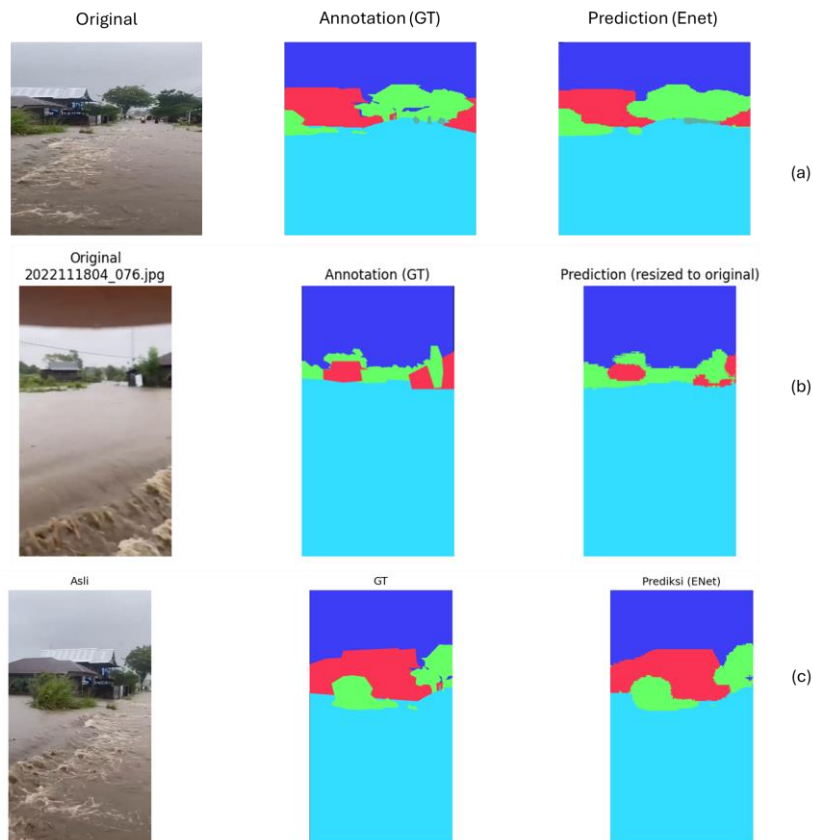


Figure 3. Prediction results of each method on the flood image dataset: (a) ENet Baseline, (b) ENet model (EN: ASC, DC: DSC), and (c) the proposed method

Figure 3 presents the prediction results on the flood image dataset. The visual comparison shows that the proposed method produced segmentation outputs that were more consistent with the ground truth annotation. Compared with the baseline models, the proposed method generated object regions with clearer boundaries and more recognizable shapes. The flood area and surrounding objects were segmented with better spatial consistency, indicating that the use of image sharpening contributed to improving the visual quality of the segmentation output.

Figure 4 presents the prediction results on the non-flood image dataset. In this dataset, all models produced segmentation outputs with several misclassified object regions. Although the proposed method was able to detect some dominant object areas, its segmentation output did not show a substantial visual improvement over the baseline models. This result indicates that the effect of image sharpening was less prominent on non-flood images, particularly when the original images already contained relatively clear visual structures and object boundaries.

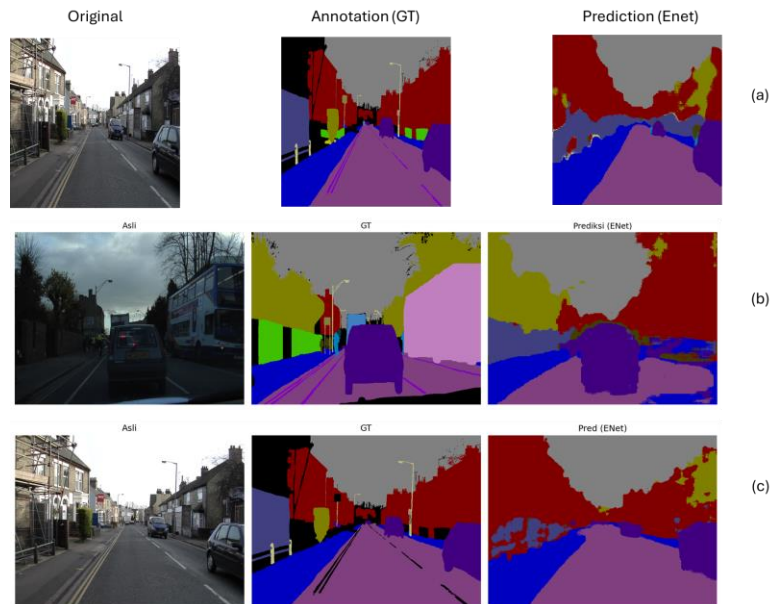


Figure 4. Prediction results of each method on the non-flood image dataset: (a) ENet Baseline, (b) ENet model (EN: ASC, DC: DSC), and (c) the proposed method

The quantitative results are presented in Table 2. On the flood image dataset, the proposed method achieved the highest overall performance, with a precision of 98%, recall of 98%, F1-score of 98%, accuracy of 97%, and mIoU of 54%. These results were higher than those of the original ENet model, which obtained 97% precision, 96% recall, 96% F1-score, 96% accuracy, and 52% mIoU. The proposed method also outperformed the modified ENet model, which achieved 97% precision, 97% recall, 94% F1-score, 97% accuracy, and 51% mIoU.

Table 2. The Experimental Scenarios Performance Comparison

No	Method	Dataset	Batch Size	Epochs	Precision	Recall	F1-Score	Accuracy	mIoU
1	ENet Baseline	Flood	8	30	97	96	96	96	52
2	ENet Baseline	Non-Flood	8	30	86	86	86	86	46
3	Enet with EN: ASC DC: DSC	Flood	8	30	97	97	94	97	51
4	Enet with EN: ASC DC: DSC	Non-Flood	8	30	86	86	86	86	43
5	Proposed Method	Flood	8	30	98	98	98	97	54
6	Proposed Method	Non-Flood	8	30	90	93	89	79	41

Table 2 shown that in contrast, the performance of the proposed method on the non-flood image dataset was lower than that of the baseline models in terms of accuracy and mIoU. The proposed method achieved 90% precision, 93% recall, 89% F1-score, 79% accuracy, and 41% mIoU. Meanwhile, the original ENet model achieved 86% precision, 86% recall, 86% F1-score, 86% accuracy, and 46% mIoU. The modified ENet model obtained 86% precision, 86% recall, 86% F1-score, 86% accuracy, and 43% mIoU. These results show that the proposed method was more effective on the flood image dataset than on the non-flood image dataset.

Overall, the experimental results indicate that the proposed method improved semantic segmentation performance on flood images, particularly in terms of precision, recall, F1-score, and mIoU. However, its performance on non-flood images was less consistent, suggesting that the benefit of image sharpening depends on the visual characteristics of the input image.

## DISCUSSION

The results demonstrate that the proposed method is effective for improving semantic segmentation performance on flood images. This improvement can be explained by the characteristics of the flood image dataset, which was collected from real-world flood documentation and social media sources. Such images often contain low visual quality, blurred regions, unclear object boundaries, and complex environmental conditions. Under these conditions, image sharpening helps enhance edge information and object contours before the images are processed by the CNN model.

The improvement in the flood image dataset is reflected not only in the quantitative metrics but also in the visual segmentation results. The proposed method achieved the highest mIoU value of 54%, compared with 52% for the original ENet and 51% for the modified ENet. Although the improvement in mIoU appears moderate, it is meaningful because mIoU is a stricter metric than accuracy. Accuracy can remain high when dominant classes occupy large portions of the image, while mIoU provides a more detailed evaluation of the overlap between predicted regions and ground truth annotations. Therefore, the increase in mIoU suggests that the proposed method improved the model's ability to localize object regions more accurately.

The higher precision, recall, and F1-score achieved by the proposed method on flood images also indicate that the model became more balanced in identifying relevant object classes. The sharpening process likely strengthened the visual distinction between floodwater, road surfaces, vegetation, buildings, and other surrounding objects. As a result, the CNN model could extract more discriminative features during training and produce segmentation maps that better preserved the shape of objects.

However, the results on the non-flood image dataset show a different pattern. Although the proposed method achieved higher precision and recall than the baseline models, its accuracy and mIoU were lower. This suggests that image sharpening does not always improve segmentation performance across all image domains. In non-flood images, object boundaries are generally clearer and the visual quality is more stable. Applying sharpening to images that already have clear edges may introduce excessive edge contrast, amplify noise, or distort subtle texture information. This condition may cause the model to produce less accurate pixel-level segmentation, particularly for small or visually complex object classes.

These findings indicate that image sharpening is more suitable for images with poor visual quality, blurred boundaries, or low contrast between objects. In the context of flood monitoring, this is an important result because disaster images are often captured under uncontrolled conditions, such as during heavy rain, low lighting, unstable camera movement, or emergency situations. Therefore, the proposed method has practical relevance for supporting post-disaster image analysis, especially when the available image data are visually inconsistent or degraded. Nevertheless, the relatively low mIoU values show that the proposed method still has limitations. Although the accuracy reached 97%, the mIoU of 54% indicates that pixel-level overlap between predictions and annotations can still be improved. This gap suggests that future work should not rely only on image sharpening but should also explore adaptive preprocessing, class-balancing strategies, improved annotation quality, and more advanced segmentation architectures. Future studies may also investigate whether sharpening should be applied selectively based on image quality, rather than uniformly to all input images.

In summary, the proposed method contributes to improving semantic segmentation of flood images by enhancing object boundaries and supporting better feature extraction in CNN-based segmentation. However, its effectiveness is domain-dependent. The method performs better on low-quality flood images than on clearer non-flood images, indicating that preprocessing techniques should be selected according to the visual characteristics of the dataset.

## CONCLUSION

This study aimed to assess the effectiveness of image sharpening in enhancing the performance of CNN-based semantic segmentation models. The findings indicate that edge sharpening is particularly effective for low-quality or blurry images, such as those obtained from social media. The proposed method successfully improved segmentation accuracy by enhancing object boundaries. However, for high-quality images with clear edges, the sharpening technique provided limited benefits and, in some cases, slightly reduced performance. Future work will focus on exploring adaptive pre-processing strategies and integrating more advanced feature enhancement methods to enable the proposed approach to generalize effectively across both blurry and clear image datasets, thereby achieving more consistent and robust segmentation performance.

## ACKNOWLEDGMENT

The authors would like to thank the 2025 DRTPM (Grant No: 014/C3/DT.05.00/PL/2025) from the Ministry of Higher Education, Science, and Technology of Indonesia for the financial support. The authors would like to thank Institut Teknologi Bacharuddin Jusuf Habibie (ITH) and LPPM-PM ITH for their support throughout the research process.

## REFERENCES

- A. Hammood, W., Abdullah Arshah, R., Mohamad Asmara, S., Al Halbusi, H., A. Hammood, O., & Al Abri, S. (2021). A Systematic Review on Flood Early Warning and Response System (FEWRS): A Deep Review and Analysis. *Sustainability*, 13(1), 440. <https://doi.org/10.3390/su13010440>
- Bagheri, A., & Liu, G.-J. (2025). Climate change and urban flooding: Assessing remote sensing data and flood modeling techniques: a comprehensive review. *Environmental Reviews*, 33, 1–14. <https://doi.org/10.1139/er-2024-0065>



- Balaji, V., Song, T.-A., Malekzadeh, M., Heidari, P., & Dutta, J. (2024). Artificial Intelligence for PET and SPECT Image Enhancement. *Journal of Nuclear Medicine*, 65(1), 4–12. <https://doi.org/10.2967/jnumed.122.265000>
- Bukhari, S. A. S., Shafi, I., Ahmad, J., Butt, H. T., Khurshaid, T., & Ashraf, I. (2025). Enhancing flood monitoring and prevention using machine learning and IoT integration. *Natural Hazards*, 121(4), 4837–4864. <https://doi.org/10.1007/s11069-024-06986-3>
- Cheng, J., Deng, C., Su, Y., An, Z., & Wang, Q. (2024). Methods and datasets on semantic segmentation for Unmanned Aerial Vehicle remote sensing images: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 211, 1–34. <https://doi.org/10.1016/j.isprsjprs.2024.03.012>
- Doğan, G., & Ergen, B. (2024). A new CNN-based semantic object segmentation for autonomous vehicles in urban traffic scenes. *International Journal of Multimedia Information Retrieval*, 13(1), 11. <https://doi.org/10.1007/s13735-023-00313-5>
- Emek Soylu, B., Guzel, M. S., Bostanci, G. E., Ekinci, F., Asuroglu, T., & Acici, K. (2023). Deep-Learning-Based Approaches for Semantic Segmentation of Natural Scene Images: A Review. *Electronics*, 12(12), 2730. <https://doi.org/10.3390/electronics12122730>
- Farhadi, H., Ebadi, H., Kiani, A., & Asgary, A. (2024). Near Real-Time Flood Monitoring Using Multi-Sensor Optical Imagery and Machine Learning by GEE: An Automatic Feature-Based Multi-Class Classification Approach. *Remote Sensing*, 16(23), 4454. <https://doi.org/10.3390/rs16234454>
- Inthizami, N. S., Ma'sum, M. A., Alhamidi, M. R., Gamal, A., Ardhiyanto, R., Kurnianingsih, & Jatmiko, W. (2022). Flood video segmentation on remotely sensed UAV using improved Efficient Neural Network. *ICT Express*, 8(3), 347–351. <https://doi.org/10.1016/j.icte.2022.01.016>
- Intizhami, N. S., Nuranti, E. Q., & Bahar, N. I. (2023). Dataset for flood area recognition with semantic segmentation. *Data in Brief*, 51, 109768. <https://doi.org/10.1016/j.dib.2023.109768>
- Intizhami, N. S., Nuranti, E. Q., & Bahar, N. I. (2025). Improving Semantic Segmentation of Flood Areas Using Rotation and Flipping-Based Feature Augmentation. *Jurnal Teknik Informatika (Jutif)*, 6(3), 1669–1682. <https://doi.org/10.52436/1.jutif.2025.6.3.4564>
- Li, H. (2021). Image semantic segmentation method based on GAN network and ENet model. *The Journal of Engineering*, 2021(10), 594–604. <https://doi.org/10.1049/tje2.12067>
- Lu, R.-H., Chen, T.-J., & Pan, J. (2024). Unsharp Mask Sharpening Detection via Global Analysis. *2024 35th Irish Signals and Systems Conference (ISSC)*, 1–4. Belfast, United Kingdom: IEEE. <https://doi.org/10.1109/ISSC61953.2024.10602852>
- Mandal, R., Sharma, B., & Chutia, D. (2025). Edge-Integrated IoT and Computer Vision Framework for Real-Time Urban Flood Monitoring and Prediction. *International Journal of Advanced Computer Science and Applications*, 16(9). <https://doi.org/10.14569/IJACSA.2025.0160946>
- Mishra, S., Vishwakarma, A., & Kumar, A. (2025). Multi-headed U-Net: An automated nuclei segmentation technique using Tikhonov filter-based unsharp masking. *Smart Science*, 13(2), 237–248. <https://doi.org/10.1080/23080477.2024.2373551>
- Muhadi, N. A., Abdullah, A. F., Bejo, S. K., Mahadi, M. R., & Mijic, A. (2021). Deep Learning Semantic Segmentation for Water Level Estimation Using Surveillance Camera. *Applied Sciences*, 11(20), 9691. <https://doi.org/10.3390/app11209691>
- Nair, B. B., Vallimeena, P., Gopalakrishnan, U., Rao, S. N., & Krishnamoorthy, S. (2024). Enhanced Urban Flood Monitoring: Integrating Advanced Semantic Segmentation and Human Facial Feature and Posture Analysis. *IEEE Access*, 1–1. <https://doi.org/10.1109/ACCESS.2024.3513239>
- Safavi, F., & Rahnemoonfar, M. (2023). Comparative Study of Real-Time Semantic Segmentation Networks in Aerial Images During Flooding Events. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16, 4–20. <https://doi.org/10.1109/JSTARS.2022.3219724>
- Suresh, T., Sundaralingam, H., Akilan, T., & Jahan, N. (2025). ECASeg: Enhancing Semantic Segmentation with Edge Context and Attention Strategy. *Procedia Computer Science*, 262, 274–282. <https://doi.org/10.1016/j.procs.2025.03.202>
- Upadhyay, D., Malhotra, S., Gupta, M., & Mishra, S. (2024). Implementation of Pruned and Quantized Semantic Segmentation Neural Network Using Cambridge-Driving Labeled Video Database (CamVid) Dataset. *2024 2nd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)*, 1–6. Dehradun, India: IEEE. <https://doi.org/10.1109/DICCT61038.2024.10532844>
- Wang, Y. (2022). Remote sensing image semantic segmentation network based on ENet. *The Journal of Engineering*, 2022(12), 1219–1227. <https://doi.org/10.1049/tje2.12200>