LAqua: Laplacian Pyramids for Aquatic Segmentation (Student Abstract)

Laven Srivastava*, **Ishaan Gakhar***

Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal - 576104, India
lavensri@gmail.com, ishaangakhar04@gmail.com

Abstract

Semantic segmentation of marine environments is essential for autonomous navigation of unmanned surface vessels (USVs) as well as the detection of environmental hazards such as oil spills. To tackle the challenges of accurate environmental perception, we propose a lightweight semantic segmentation network, LAqua (Laplacians for Aquatic Segmentation), which leverages Laplacian pyramids to enhance edge detection in marine imagery. Our method drastically reduces computational requirements while maintaining high accuracy in generating semantic masks for marine environments. We evaluate LAqua on two distinct datasets: one focused on detecting oil spills in port environments and another on environmental segmentation for USVs. Results show that LAqua not only performs well across varied marine settings but also achieves comparable or superior segmentation accuracy with far fewer parameters than other models. This efficiency highlights LAqua's potential for applications in real-time detection for marine environments.

Introduction and Previous Work

The advent of drone technology has brought significant advancements in autonomous detection and segmentation of environments, particularly for unmanned surface vehicles (USVs). USVs are commonly used in environmental monitoring, search and rescue operations, inspections, and scientific exploration. This work focuses on two key use cases for autonomous drones in marine environments: detecting and segmenting surrounding environmental features and identifying critical ecological hazards like oil spills. Both use cases rely on data captured by unmanned vehicles, which necessitates fast, accurate, and near-real-time segmentation to ensure timely and effective decision-making.

This paper introduces LAqua, a novel framework able to achieve near state-of-the-art metrics on its datasets while having up to **28x** fewer parameters than other methods. The use of traditional computer vision techniques such as the Laplacian Pyramid, along with modern deep learning methods of non-linear activations and multi-level processing, allows for efficient feature extraction by leveraging the vital edge information, thus leading to highly accurate segmentation.

Previous work on marine segmentation, such as WaSRNet (Bovcon and Kristan 2022), uses a deep encoder-decoder with fusion modules to handle water variability and reduce false positives. BEMSNet (Zhang et al. 2024) builds upon WASRNet by focusing on boundary segmentation, using boundary-enhancing modules and a boundary-aware loss, achieving SOTA results with 71 million parameters.

Proposed Methodology

We propose **LAqua**, a hierarchical network architecture for multi-class image segmentation. The model leverages multi-scale feature extraction and residual learning to efficiently process high-resolution images through a three-stage Laplacian Pyramid decomposition.

The **Laplacian Pyramid module** decomposes the input image into multiple scales, each half the resolution of the previous, enabling the network to extract both coarse and fine-grained details. At each level, a Gaussian kernel smooths the image, followed by downsampling, capturing low- and high-level features.

For each decomposed level, the **Low Branch** generates the primary segmentation mask with a series of 3x3 convolutional layers, Instance Norm, and Leaky ReLU activations. The network then applies 7 residual blocks to further refine feature representations. The final segmentation mask captures semantic information across multiple classes, while Conv layers upsample the output.

The residual blocks used throughout the model consist of a Leaky ReLU activation between 2 conv layers with a skip connection.

In parallel, high-frequency details are captured by residual branches using Conv layers for upsampling. The **Middle Branch** applies 7 residual blocks to reconstruct high-resolution residuals, while the **Top Branch** applies 2 residual blocks. Both branches utilize convolutions, Leaky ReLU, and residual blocks to preserve finer details throughout the pyramid levels.

At each level, upsampled segmentation masks and residual components are concatenated with higher-resolution feature maps. Additional convolutional layers and residual blocks process the combined representations to refine segmentation at each scale. The final segmentation output is

*These authors contributed equally.

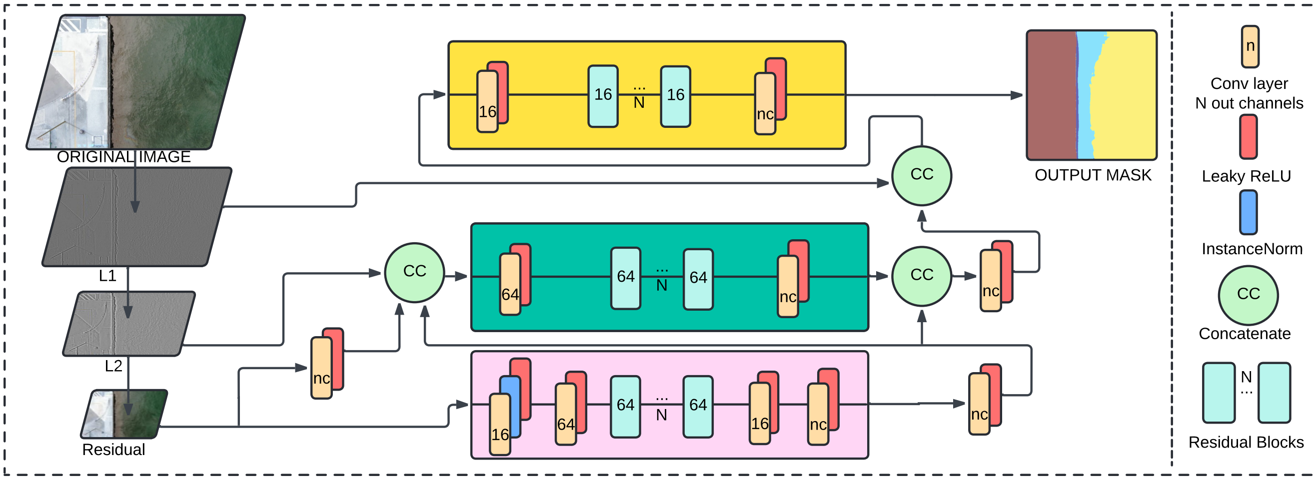

Figure 1: Our proposed model architecture for L-Aqua. We demonstrate the integration of multi-scale Laplacian Pyramid decomposition, residual blocks, and non-parametric upsampling as a lightweight, effective learning technique. Here, 'nc' refers to the number of output channels (number of classes in the mask for that dataset).

Model	Mastr1325	Oil spills	#Params (M)
PSPNet	93.46	97.03	56.0
DeepLabv3	97.67	97.14	48.0
HRNet	97.87	-	63.0
UNet	99.36	97.51	31.0
WaSRnet	99.80	-	71.4
BEMSNet	99.91	-	71.4
LAQUA	**98.97**	**94.8**	**1.1**

Table 1: Comparison of various models against ours on mIOU across two datasets - Mastr1325 (Bovcon et al. 2019) and Oil Spill Drone (Sels, Vanlanduit, and Kerf 2024). The results of UNet, PSPNet, and DeepLabv3 are on the val set.

generated through hierarchical upsampling, where Conv layers increase resolution and merge features across all levels. The **Top Branch** module produces the highest-resolution segmentation map, capturing both global structure and fine details.

LAqua employs 3x3 convolutions, Leaky ReLU activations, and residual blocks to extract and integrate multi-scale features. Conv layers handle upsampling, while multiple branches perform segmentation and residual reconstruction in parallel. This architecture enables accurate multi-class segmentation by progressively refining both high- and low-level features across the image.

Experimental Setup and Results

We conducted experiments with our model and realized the best results with a number of residual blocks as 7, 7, and 2 in the lowest, middle and top branches, respectively. We utilised a combination loss function of cross entropy and focal loss.

We benchmarked both datasets using SOTA models such as UNet, PSPNet, and DeepLabv3. All models mentioned in 1 are trained for 300 epochs at a learning rate of 1e-4. Our findings indicate that our model produces SOTA results with upto 28x lesser trainable parameters.

Conclusion

In this paper, we introduced LAqua, a lightweight segmentation network designed for marine environments, leveraging Laplacian pyramids for efficient edge detection and multi-scale feature extraction. By evaluating two distinct datasets, LAqua demonstrated competitive performance while drastically reducing the number of parameters compared to state-of-the-art models. This efficiency makes LAqua well-suited for real-time applications, such as USV navigation and environmental hazard detection, where fast and accurate segmentation is crucial.

Ackowledgments

We would like to thank Mars Rover Manipal, an interdisciplinary student team of MAHE, for providing the resources needed for this project. We also extend our gratitude to Dr. Ujjwal Verma for his guidance as well as Mohammed Sulaiman for his support in our work.

References

Bovcon, B.; and Kristan, M. 2022. WaSR—A Water Segmentation and Refinement Maritime Obstacle Detection Network. *IEEE Transactions on Cybernetics*, 52(12): 12661–12674.

Bovcon, B.; Muhovič, J.; Perš, J.; and Kristan, M. 2019. The MaSTr1325 dataset for training deep USV obstacle detection models. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 3431–3438.

Sels, S.; Vanlanduit, S.; and Kerf, T. D. 2024. Annotated RGB images of Oil Spills in a Port Environment.

Zhang, L.; Sun, X.; Li, Z.; Kong, D.; Liu, J.; and Ni, P. 2024. Boundary Enhancement-Driven Accurate Semantic Segmentation Networks for Unmanned Surface Vessels in Complex Marine Environments. *IEEE Sensors Journal*, 24(15): 24972–24987.