LaFINet: Laplacian-Based Frequency Injection Network for Camouflage Object Detection (Student Abstract)

Aravinthakshan A S[*], Aditya Prashant Naidu[*], Aadiv Rath[*]

Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal - 576104, India
aravinthakshanmain@gmail.com, adityanaidu2004@gmail.com, aadivrath04@gmail.com

Abstract

Camouflaged object detection is critical for military, defense, and security operations, where targets evade conventional surveillance by mimicking the background or exhibiting low-contrast differences. It also supports non-invasive monitoring of elusive wildlife and endangered species, improving population estimates, habitat management, and biodiversity assessments by recovering objects that are visually indistinguishable from their surroundings. Existing solutions are computationally heavy, with large model parameters and high computational demands, which hinder deployment in real-world applications. Lightweight models have been explored, but they often compromise fine boundary fidelity. This paper introduces a lightweight Laplacian pyramid–based feature extractor that progressively aggregates multiscale Laplacian features with frequency information. The proposed architecture emphasizes object edge boundaries, enabling precise localization under subtle target–background differences while maintaining realtime efficiency. The design achieves performance comparable to the state of the art (SOTA) convolution based methods on CHAMELEON and NC4K datasets.

Introduction and Previous Work

Camouflaged object detection (COD) aims to segment targets that are deliberately concealed within their surroundings. Unlike traditional object detection, COD is substantially more challenging because target objects blend seamlessly with the background through highly similar color, texture, and illumination patterns, making COD more difficult than conventional saliency detection or generic segmentation tasks. Heavyweight models such as (Fan et al. 2022) achieve strong detection performance; however, their substantial computational requirements and high FLOPs hinder deployment in real-time or resource-constrained environments. At the other end, TinyCOD (Xing et al. 2023) provides a lightweight alternative, but often suffers from imprecise localization and blurred boundaries. The multi-branch strategy in (Liang et al. 2024), complemented by the introduction of a Feature Fusion Module (FFM), represents a promising attempt to balance efficiency with accuracy, yet the design remains overly lightweight, limiting its capacity to fully exploit frequency feature information. Skeleton-guided frameworks such as (Ren, Bai, and Sun 2025) are shown to improve structural consistency and boundaries, but the reliance on explicit skeleton extraction introduces considerable overhead.

The proposed architecture addresses these challenges by employing Laplacian pyramid-based feature extraction, which captures high-frequency and structural information across multiple scales. This process yields skeletal-like structures crucial for camouflaged object detection without introducing unnecessary overhead, thereby enhancing boundary fidelity while maintaining a low parameter count. The main contributions of this work are as follows:

- **Laplacian feature extraction and injection** facilitates multiscale feature extraction by decomposing images with a three-stage Laplacian pyramid and integrating the details using Laplacian Injection Blocks (LIB).
- **Frequency Split-Modulation and Injection Module (FSIM)** that integrates multi-scale frequency cues with foreground–background modulation to enhance boundary fidelity and target–background separation efficiently inspired by (Khan et al. 2024), (Liang et al. 2024)
- **Channel and Position Attention for Camouflage Detection:** The Channel and Position Attention Mechanism (CPAM) (Kang et al. 2024), originally developed for Cell Instance Segmentation, is adapted to COD to adaptively focus on informative channels and spatial locations.

Proposed Methodolgy

As illustrated in Figure 1, LaFINet begins by decomposing the input image into a three-level Laplacian pyramid. The resulting frequency components, concentrated in the first three levels, are then prepared for injection into the encoder.

In the encoder stage, the raw input image is processed by an EfficientNet-B0 backbone to produce multistage feature maps (x1,x2,x3,x4). At the first three relevant stages, the corresponding Laplacian levels are introduced via Laplacian Injection Blocks (LIB). Here, the Laplacian features are fused via concatenation and convolution. This allows each encoder stage to retain frequency-sensitive details aligned with its scale. These reduced features are then processed by Frequency Split-Modulation and Injection Module (FSIM) modules at each encoder stage. These operate in parallel

[*]These authors contributed equally.

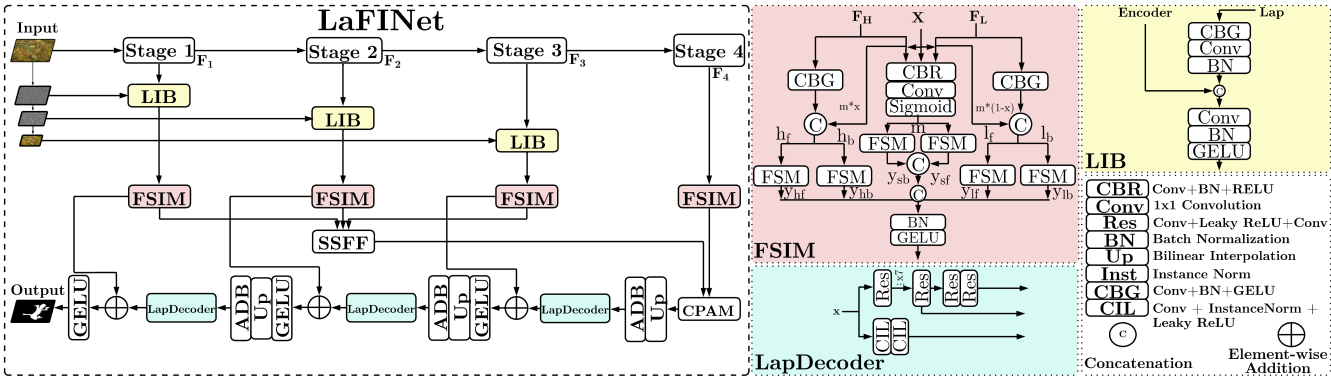

Figure 1: Proposed LaFINet architecture with Frequency Split-Modulation and Injection Module (FSIM), LapDecoders, Laplacian Injection Blocks (LIBs), Channel and Position Attention Mechanism (CPAM), Scaled Sequence Feature Fusion (SSFF).

with the semantic pathways, handling the fused signals from the injected Laplacian cues and ensuring a balanced propagation of frequency-rich detail alongside semantic structure.

Asymmetric Decoder Block (ADB) introduced in (Liang et al. 2024) is modified to a LapDecoder with 3 branches, which are fused into a unified output. The outputs of three decoders are fused using Scaled Sequence Feature Fusion (SSFF) (Kang et al. 2024), after which Channel and Position Attention Mechanism (CPAM) emphasizes key channels and spatial locations to improve contextual awareness and localization for segmentation of small structures.

Experimental Setup and Results

We employ the ADAM optimizer and train our model for 200 epochs with a cosine decay schedule, starting from an initial learning rate of 0.001, using a NVIDIA Tesla P100 GPU. The performance. Table 1 summarizes the evaluation of our method compared against existing baselines on the metrics, including $S_\alpha \uparrow$ (Structure Measure), $E_\phi^d \uparrow$ (Adaptive Measure), and $F_\beta^w \uparrow$ (Weighted F-measure).

Method	CHAMELEON			NC4K		
	$S_\alpha \uparrow$	$E_\phi^d \uparrow$	$F_\beta^w \uparrow$	$S_\alpha \uparrow$	$E_\phi^d \uparrow$	$F_\beta^w \uparrow$
SinetV2	0.888	0.930	0.816	0.847	0.901	0.770
TinyCOD	0.887	0.931	0.814	0.843	0.903	0.766
FINet	0.883	0.928	0.808	0.847	0.904	0.771
LaFINet	**0.892**	**0.940**	**0.829**	**0.845**	**0.906**	**0.772**

Table 1: Comparison of LaFINet with existing baselines on two benchmark datasets. Results are highlighted in **bold**.

Notably, LaFINet requires only 4.48M parameters, slightly more than FINet (3.74M), and significantly smaller than SINetV2 (24.9M) and TinyCOD (4.72M), and outperforms all three baselines on CHAMELEON with only a marginal increase in parameters. Figure 2 illustrates the intuition behind using Laplacian pyramid decomposition for camouflaged object detection. At Level 1, the decomposition highlights sharp boundaries and contours, which are crucial for distinguishing object edges even in cluttered environments, attributing to the high Weighted F-measure (F_β^w).

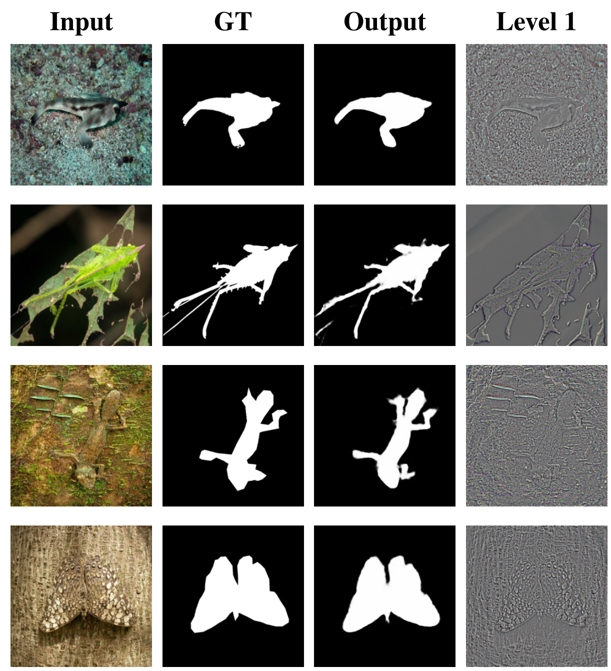

Figure 2: Visualization showing input image, ground truth, model output, and first-level laplacian pyramid output.

Conclusion

This work presents an architecture for camouflaged object detection that integrates Laplacian features, CPAM attention, SSFF module, and the novel FSIM block to capture fine boundary details under strict latency and parameter constraints. We achieve performance comparable to the state of the art (SOTA) on CHAMELEON and NC4K datasets. Future work will focus on extending this approach to both segmentation and detection tasks, as well as reducing FLOPs.

References

Fan, D.-P.; Ji, G.-P.; Cheng, M.-M.; and Shao, L. 2022. Concealed Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, $44(10_{Part_1}):6024$–6042.

Kang, M.; Ting, C.-M.; Ting, F. F.; and Phan, R. C.-W. 2024. ASF-YOLO: A novel YOLO model with attentional scale sequence fusion for cell instance segmentation. *Image and Vision Computing*, 147: 105057.

Khan, A.; Khan, M.; Gueaieb, W.; El Saddik, A.; De Masi, G.; and Karray, F. 2024. CamoFocus: Enhancing Camouflage Object Detection With Split-Feature Focal Modulation and Context Refinement. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 1434–1443.

Liang, W.; Wu, J.; Wu, Y.; Mu, X.; and Xu, J. 2024. FINet: Frequency Injection Network for Lightweight Camouflaged Object Detection. *IEEE Signal Processing Letters*, 31: 526–530.

Ren, P.; Bai, T.; and Sun, F. 2025. ESNet: An Efficient Skeleton-guided Network for camouflaged object detection. *Knowledge-Based Systems*, 311: 113056.

Xing, H.; Gao, S.; Tang, H.; Mok, T. Q.; Kang, Y.; and Zhang, W. 2023. Tinycod: Tiny and effective model for camouflaged object detection. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5. IEEE.