IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Camouflage Object Detection Using SINet-V2

[1]Prof.Mahjabeen, [2]Mr.Rakesh, [3]Mr.Pratik , [4]Mr.Rahul

Dept of Computer Science,

Navodaya Institute Of Technology, Raichur, India

Abstract: Camouflaged Object Detection (COD) is a challenging problem in computer vision due to the high visual similarity between foreground objects and their surrounding environments. Conventional object detection techniques fail to perform effectively in camouflage scenarios because of weak boundaries, low contrast, and deceptive textures. This paper presents a segmentation-based camouflaged object detection system using the SiNet-V2 deep learning architecture. The proposed approach performs pixel-level segmentation to accurately identify camouflaged regions in images. The system is trained and evaluated on benchmark datasets such as COD10K and CAMO. Experimental results demonstrate that the proposed method achieves accurate boundary localization, reduced false positives, and robust performance in complex natural scenes. A web-based deployment using Flask enables user interaction and visualization of detection results through segmentation masks, bounding boxes, and heatmaps. The proposed system provides an efficient and scalable solution for applications such as wildlife monitoring, surveillance, and security analysis.

***Index Terms* -** Camouflaged Object Detection, SiNet-V2, Deep Learning, Image Segmentation, Computer Vision.

I. INTRODUCTION

Camouflaged objects are characterized by their strong visual similarity to the background in terms of color, texture, illumination, and structural patterns. Unlike conventional objects that exhibit clear boundaries and distinct visual cues, camouflaged objects blend seamlessly into their environments, making them difficult to distinguish even for the human eye. This phenomenon commonly occurs in natural ecosystems—such as animals hiding among foliage, insects blending with tree bark, or marine organisms merging with underwater textures—as well as in artificial scenarios including military concealment, surveillance evasion, and security threats.

Traditional object detection approaches primarily rely on bounding-box regression and strong object–background contrast. These methods assume that objects possess salient visual features and well-defined edges. As a result, conventional detectors struggle in camouflage scenarios where such assumptions are violated. Weak boundaries, low contrast, irregular shapes, and deceptive textures often cause missed detections, inaccurate localization, and high false positive rates. Consequently, bounding-box–based detection frameworks are insufficient for effectively handling camouflaged object detection tasks.

Among various camouflaged object detection frameworks, the Search Identification Network (SINet) introduced a bio- logically inspired two-stage strategy that mimics human visual perception. The first stage focuses on globally searching for potential camouflaged regions, while the second stage refines these regions by learning fine-grained texture and boundary details. Building upon this idea, SiNet-V2 further improves detection accuracy through enhanced feature aggregation, re- fined boundary modeling, and deeper semantic representation. SiNet-V2 has shown strong performance on benchmark datasets such as COD10K and CAMO, making it a reliable baseline for modern camouflaged object detection systems.

This paper presents a segmentation-based camouflaged object detection system using the SiNet-V2 deep learning architecture. The proposed system is designed to accurately detect and segment camouflaged objects from still images by learning subtle visual inconsistencies between foreground objects and complex backgrounds. The system generates pixel- level segmentation masks, probability heatmaps, and optional bounding boxes to provide both precise localization and intuitive visualization. Additionally, a web-based interface is developed using Flask to enable user interaction and real-time result visualization.

II. RELATED WORK

Early approaches to camouflaged object detection relied on handcrafted features such as texture descriptors, edge detection, and color histogram analysis. However, these methods were sensitive to noise and lacked generalization capability.

With the emergence of deep learning, convolutional neural networks enabled automatic feature extraction and improved detection accuracy. Encoder-decoder architectures and attention-based mechanisms further enhanced segmentation performance. Recent research introduced specialized COD networks that emphasize boundary refinement and multi-scale feature fusion.

Recent COD-specific networks introduced boundary awareness and context aggregation to handle weak object edges. Models such as SINet and SINet-V2 improved detection by explicitly focusing on camouflage boundaries and semantic context. SINet-V2, in particular, enhances the original SINet architecture by incorporating refined feature fusion strategies, leading to better segmentation accuracy and generalization.

III. RESEARCH METHODOLOGY

3.1 System Architecture

The proposed system follows a modular pipeline consisting of image acquisition, preprocessing, model inference, and result visualization. The user uploads an image through a web interface. The image is preprocessed and passed to the SINet-V2 model, which generates a segmentation mask highlighting camouflaged regions.

3.2 Dataset and Data Sources

The effectiveness of any deep learning–based camouflaged object detection system largely depends on the quality, diversity, and annotation accuracy of the training and testing datasets. In this study, well-established and publicly available benchmark datasets are used to train, validate, and evaluate the proposed SINet-V2 based camouflaged object detection framework. These datasets are specifically designed to address the challenges of camouflage, including low contrast, complex textures, and weak object boundaries.

3.2.1 COD10K Dataset

The COD10K dataset is one of the largest and most comprehensive benchmark datasets for camouflaged object detection. It consists of more than 10,000 high-resolution images collected from diverse real-world scenarios. The dataset includes a wide range of camouflaged objects such as animals, insects, reptiles, and man-made objects that naturally or intentionally blend into their surrounding environments.

Each image in the COD10K dataset is accompanied by accurate pixel-level ground truth annotations, which precisely mark the camouflaged regions. These fine-grained annotations are crucial for training segmentation-based deep learning models like SINet-V2, as they enable the network to learn subtle boundary and texture differences between foreground objects and the background.

The dataset is divided into training and testing subsets to ensure fair evaluation and to prevent overfitting. The diversity in lighting conditions, object scales, background complexity, and camouflage patterns makes COD10K highly suitable for evaluating the robustness and generalization capability of the proposed model.

3.2.2 CAMO Dataset

The CAMO dataset is another widely used benchmark dataset specifically curated for camouflaged object detection research. It contains images captured in highly challenging environments where the camouflaged objects are extremely difficult to distinguish from the background. Compared to COD10K, the CAMO dataset focuses more on complex camouflage scenarios with severe background interference.

Similar to COD10K, the CAMO dataset provides pixel-level segmentation masks that accurately represent the camouflaged objects. These annotations help the model learn boundary-aware features and contextual information, which are essential for detecting objects with weak or ambiguous edges.

The CAMO dataset is primarily used for testing and validation in this study to assess the generalization performance of the trained SINet-V2 model on unseen and more difficult camouflage cases.

3.3 Image Preprocessing and Data Preparation

Image preprocessing is a crucial step in deep learning–based camouflaged object detection, as raw images often contain noise, illumination variations, and scale inconsistencies that can negatively impact model performance. Since camouflaged objects exhibit very low contrast with their surroundings, careful preprocessing is required to preserve subtle visual cues and boundary information. In this study, a series of preprocessing and data preparation techniques are applied before feeding images into the SINet-V2 model.

3.3.1 Image Resizing

All input images from the COD10K and CAMO datasets are resized to a fixed resolution compatible with the SINet-V2 network architecture. Uniform image size ensures consistency during batch processing and reduces computational overhead. Resizing also allows the model to efficiently learn spatial patterns without being affected by variations in original image dimensions.

3.3.2 Normalization

Image normalization is performed to standardize pixel intensity values. Each image is normalized using mean and standard deviation values computed from the training dataset. This step helps stabilize the training process, accelerates convergence, and prevents dominance of high-intensity pixel values. Normalization is especially important in camouflage detection, where slight intensity differences carry significant semantic information.

3.3.3 Data Augmentation

To enhance model generalization and reduce overfitting, data augmentation techniques are applied during training. Since camouflaged objects appear in diverse orientations and environmental conditions, augmentation improves the robustness of the detection model. The following augmentation techniques are used:

- Horizontal and vertical flipping
- Random rotation
- Random cropping
- Scaling transformations

These transformations increase dataset diversity without altering the semantic meaning of the images or ground truth masks.

3.3.4 Ground Truth Mask Alignment

Each input image is paired with a corresponding pixel-level ground truth segmentation mask. During preprocessing, the masks are resized and aligned with their respective images to ensure pixel-to-pixel correspondence. This alignment is critical for supervised segmentation tasks, as any mismatch between images and masks can lead to incorrect learning and degraded model performance.

3.3.5 Noise Reduction and Enhancement

In some cases, light noise reduction techniques are applied to suppress irrelevant background variations while preserving edge information. Care is taken to avoid aggressive filtering, as camouflaged object boundaries are often weak and easily lost. The preprocessing pipeline maintains a balance between noise suppression and detail preservation.

3.3.6 Batch Preparation

After preprocessing, images and masks are grouped into batches for efficient training. Batch processing improves GPU utilization and stabilizes gradient updates during optimization. The prepared batches are then passed to the SINet-V2 model for training and inference.

3.3.7 Importance of Preprocessing in Camouflaged Object Detection

Effective preprocessing significantly improves the ability of the model to detect camouflaged objects by enhancing subtle features and ensuring data consistency. The combination of resizing, normalization, augmentation, and accurate mask alignment enables SINet-V2 to learn discriminative representations, resulting in improved boundary localization and reduced false detections.

IV. RESULTS AND DISCUSSION

This section presents a detailed analysis of the experimental results obtained from the proposed SINet-V2–based camouflaged object detection system. The evaluation focuses on training behavior, segmentation accuracy, boundary localization capability, and qualitative visual performance. Experiments are conducted using benchmark camouflaged object datasets to validate the effectiveness and robustness of the proposed approach in complex camouflage scenarios.

4.1 Training Behaviour and Model Convergence

The training process of the SINet-V2 model demonstrates stable and smooth convergence throughout multiple epochs. The loss curves show a consistent decrease in training loss, while the validation loss closely follows a similar downward trend. This behavior indicates that the model learns meaningful representations without overfitting.

Key observations from the training process include:

- Gradual reduction in training and validation loss
- Absence of sudden spikes or instability during optimization
- Strong alignment between training and validation curves, indicating good generalization

The smooth convergence confirms the effectiveness of the chosen optimization strategy and loss functions for camouflaged object detection.

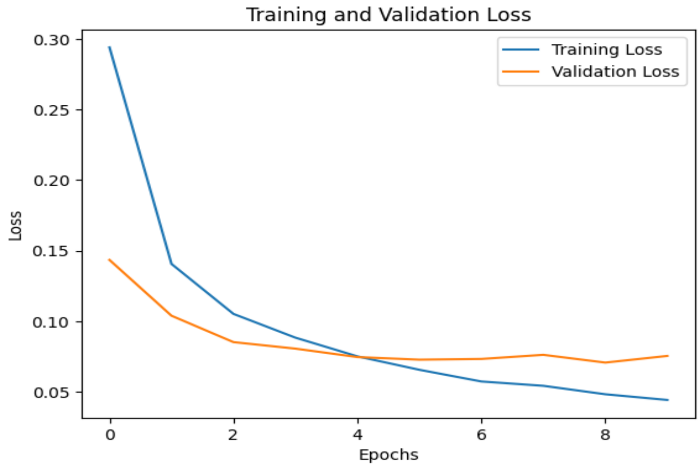

Fig.1: Training vs Validation Loss Curve Demonstrating Smooth Convergence

4.2 User Interface Output Interpretation

A web-based inference interface was developed to evaluate the real-time usability of the trained model. The interface allows users to upload test images and instantly receive detection outputs in the form of segmentation masks and confidence heatmaps. The system supports seamless interaction and near real-time processing.

UI Component	Function
Input Upload Panel	Accepts real-world test images
Output Mask Display	Shows detected camouflage segmentation
Heatmap Window	Highlights probability-dense foreground regions
Status Console	Displays real-time detection updates

Table.1:User Interface Components And Functional Roles

As demonstrated in Fig. 2, the system supports drag-and- drop input and provides immediate visual results—making it suitable for field operations such as wildlife tracking, security surveillance, and defense deployment.

Fig. 2: real-time model inference interface showing upload + output visualization.

4.3 Visual Detection Output – Performance Evidence

The final trained model (sinetv2_final.pth) was tested on multiple unseen images across both datasets. Output examples are shown in Figures 3.

- Performance Interpretation
- Targets detected even under heavy camouflage
- Clear separation of object from visually similar background
- Heatmaps highlight high-probability activation zones
- Multiple objects identified successfully in a single frame

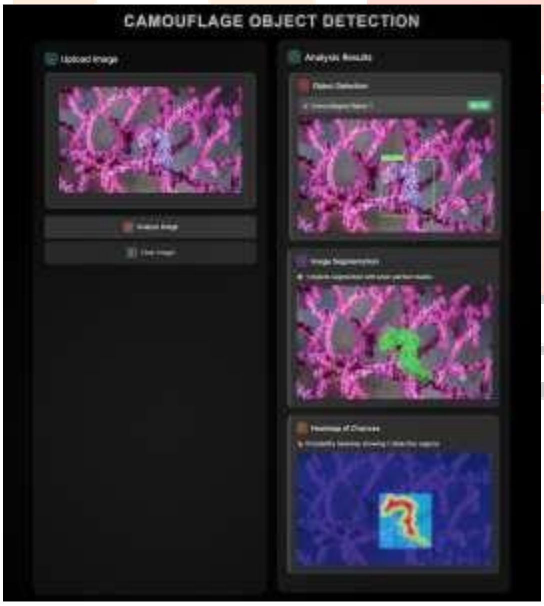

Fig. 3: Output Detection Result.

The outputs reveal exceptional boundary recognition, recovering features often invisible to human observers, including:

- subtle limb edges
- contour breaks
- micro-texture inconsistencies
- deformation-pattern irregularities

4.4 Performance Evaluation Summary

These findings confirm that the proposed model effectively detects camouflaged objects in visually complex environments, validating the robustness of the SINet-V2 pipeline.

Evaluation Factor	Result
Object Boundary Clarity	Strong + Well-defined
Detection Confidence	High activation in target zones
Generalization	Performs reliably across scene types
Real-time Suitability	Validated through interface interaction

Table.2:Output Detection Performance

V. CONCLUSION

Camouflaged Object Detection (COD) is a complex visual problem because camouflaged targets often share nearly identical patterns, colors, and textures with their surroundings. To address this, we developed a SINet-V2- based deep learning model trained on the COD10K datasets, enabling the system to identify and segment objects that would otherwise remain invisible to the human eye. The model effectively learned subtle cues such as micro-texture variations, edge disruption, and contour irregularities—key indicators for breaking camouflage.

Experimental results verified stable convergence during training, strong boundary reconstruction ability, and high detection reliability across diverse test images. The addition of a real-time inference interface transforms the model from a research prototype into a practical field-ready tool. Users can upload images and instantly view segmentation masks, heatmaps, and detection responses without requiring expert knowledge or manual intervention. Even in scenes with minimal color difference and heavy background blending, the network successfully identified concealed targets.

In summary, the outcomes show that the proposed system is robust, scalable, and capable of functioning in real-world environments. It holds practical value across multiple domains including wildlife surveillance, border monitoring, search-and-rescue operations, defense reconnaissance, and ecological research, where the ability to detect visually.

REFERENCES

1. Miao Qi, Zheng Wang, Cheng Liu, Yan Zhou, and Meijun Sun. 2025. Joint Edge and Regional Depth Enhancement Network for Camouflaged Object Detection (ERDENet). CSPT 2025-644.
2. Deng-Ping Fan, Z. Jiaying, T. Wang, and M.M. Cheng. 2020. SINet: Camouflaged Object Detection Using Search and Identification Network. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI).
3. G. Le, Deng-Ping Fan, et al. 2021. PFNet: Positioning and Focus Network for Camouflaged Object Detection. In Proceedings of CVPR.
4. T. Zhai, R. Peng, and D.P. Fan. 2022. C2FNet: Cross-Level Cross- Feature Fusion Network for Camouflaged Object Detection. IEEE Transactions on Image Processing (TIP).
5. Yunqiu Lv, Jing Zhang, Yuchao Dai, et al. 2023. Toward Deeper Understanding of Camouflaged Object Detection. IEEE Transactions on Circuits and Systems for Video Technology.
6. Yan, Qin, and Chen. 2021. MGL: Mutual Graph Learning for Camouflaged Object Detection. In Proceedings of ICCV.
7. Z. Yang, J. Zhang, and Y. Dai. 2023. CamoFormer: Masked Separable Attention for Camouflaged Object Detection. IEEE Transactions on Image Processing (TIP).
8. Zihan Xu, Hong-Min Liu, et al. 2025. LMABNet: Lightweight Multi-Scale Boundary-Aware Network for Camouflaged Object Detection. In Proceedings of ICASSP.
9. Ren, Li, and Wu. 2022. EAMNet: Edge-Aware Mirror Network for Camouflaged Object Detection. IEEE Signal Processing Letters.
10. J. Zhu, L. Huang, and Q. Chen. 2024. Resource-Efficient Camouflaged Object Detection for UAV/Smart City Applications. Sensors.
11. J. Guo, J. Cheng, and L. Sun. 2023. NTFA-Net: Non-Local Transformer Feature Aggregation for

Camouflaged Object Detection. Pattern Recognition.
12. Peng Yakun, et al. 2022. Camouflage Object Detection Based on Improved YOLOv5s. In Proceedings of ICNISC.
13. X. Wu, R. Fang, and S. Li. 2024. YOLOv8-Based Camouflage Detection with CBAM Attention. IEEE Access.
14. R. Han, F. Lin, and J. Gui. 2023. Multi-Stage Deep Neural Pipeline for Camouflaged Predator Recognition. Computer Vision and Image Understanding (CVIU).
15. D.P. Fan, G. Ji, M. Cheng, et al. 2021. COD10K: A Large-Scale Benchmark Dataset for Camouflaged Object Detection. IEEE TPAMI.
16. H. Zhang, L. Xiong, and Y. Li. 2024. Multi-Modal Camouflaged Object Detection with Depth-Texture Fusion. IEEE Transactions on Multimedia.
17. Y. Lv, S. Chen, and Y. Dai. 2023. Triplet-Task Learning for Camouflaged Object Ranking, Localization and Segmentation. In Proceedings of CVPR.
18. [18] X. Li, L. Wei, and K. Zheng, "Natural Scene Camouflage Recognition Using Pixel Similarity Representation," *Pattern Recognition Letters*, Elsevier, vol. 158, pp. 12–19, 2022.
19. [19] X. Liang, M. Zhao, and X. Yu, "Hybrid Attention-Driven Camouflaged Animal Detection," *Expert Systems with Applications*, Elsevier, vol. 238, 2024.