Exploring Camouflaged Object Detection Techniques for Invasive Vegetation Monitoring

Henry O. Velesaca[1,2], Hector Villegas[1], and Angel D. Sappa[1,3],
[1]*Escuela Superior Politécnica del Litoral, ESPOL, Campus Gustavo Galindo Km. 30.5 Vía Perimetral, Guayaquil, 090902, Ecuador*
[2] *Software Engineering Department, Research Center for Information and Communication Technologies (CITIC-UGR), University of Granada, 18071, Granada, Spain*
[3] *Computer Vision Center, 08193-Bellaterra, Barcelona, Spain*
{*hvelesac, hvillega, asappa*}*@espol.edu.ec, hvelesaca@correo.ugr.es*

Keywords: Camouflaged object detection, Weed detection, Pest detection, Deep learning, Computer vision, UAV

Abstract: This paper presents a novel approach to weed detection by leveraging state-of-the-art camouflaged object detection techniques. The work evaluates six camouflaged object detection architectures on agricultural datasets to identify weeds naturally blending with crops through similar physical characteristics. The proposed approach shows excellent results in detecting weeds using Unmanned Aerial Vehicle images. This work establishes a new framework for the challenging task of weed detection in agricultural settings using camouflaged object detection approaches, contributing to more efficient and sustainable farming practices.

1 Introduction

In recent years, the agricultural sector has faced increasing challenges in effective weed management, particularly when dealing with weeds that naturally blend with crops through similar physical characteristics (e.g., (Chauhan et al., 2017), (Westwood et al., 2018)). This phenomenon, known as biological mimicry or natural camouflage, presents a significant obstacle in traditional weed detection methods. The advancement of precision agriculture and computer vision technologies has opened new possibilities for addressing this complex challenge through innovative camouflage-based detection approaches (e.g., (Moldvai et al., 2024), (Wu et al., 2021)).

Traditional approaches to weed detection have primarily relied on conventional image processing techniques such as color analysis, texture features, and shape-based recognition. These methods typically employ techniques like RGB color space transformation, edge detection algorithms, and morphological operations to distinguish between crops and weeds (e.g., (Parra et al., 2020), (Agarwal et al., 2021)). Historically, researchers have utilized techniques such as the Normalized Difference Vegetation Index (NDVI), color indices, and spectral reflectance measurements to identify vegetation patterns. However, these conventional methods often fall short when confronted with sophisticated camouflage scenarios, where weeds closely mimic the visual characteristics of the desired crops (López-Granados et al., 2006).

Weed detection systems that can identify and differentiate camouflaged weeds from crops are becoming increasingly crucial for sustainable agriculture (Singh et al., 2024). These systems not only promise to reduce herbicide usage but also offer more precise and environmentally friendly weed control solutions. By leveraging advanced image processing techniques, machine learning algorithms, and sophisticated pattern recognition methods, camouflage-based weed detection represents a promising frontier in agricultural technology (Coleman et al., 2023).

This research explores various methodologies and techniques for detecting weeds that exhibit camouflage characteristics, focusing on overcoming the challenges posed by visual similarities between weeds and crops. The study examines both classical computer vision approaches and cutting-edge deep learning methods, aiming to develop more accurate and reliable detection systems for practical agricultural applications.

To address this work in detail, the manuscript is organized as follows. Section 2 introduces the background on using weed detection techniques based on classical and deep learning, and addresses the problem as a camouflaged object detection approach. Section 3 presents the proposed approach to identify weeds using camouflage detection techniques. Then, Section 4 shows the experimental results taking as reference a dataset with aerial images with the presence of weeds and a case study of banana crops evaluating COD techniques. Finally, conclusions are presented in Section 5.

2 BACKGROUND

The field of weed detection has evolved significantly over the past decades, transitioning from manual inspection methods to sophisticated automated systems. This evolution has been particularly marked by the integration of artificial intelligence and specialized detection techniques to address complex scenarios such as camouflaged weeds in agricultural settings.

Deep learning has revolutionized the field of computer vision and, by extension, weed detection systems. A recent work by Rehman et al. (Rehman et al., 2024) demonstrates the effectiveness of drone-based weed detection using feature-enriched deep learning approaches, achieving significant improvements in detection accuracy across various agricultural scenarios. This builds upon foundational work by Tang et al. (Tang et al., 2017), who pioneered the combination of K-means feature learning with convolutional neural networks for weed identification, establishing early benchmarks for automated detection systems. Further advances are made by Balabantaray et al. (Balabantaray et al., 2024), who have developed targeted weed management systems using robotics and YOLOv7 architecture, specifically focusing on Palmer amaranth detection and demonstrating the practical application of deep learning in real-world agricultural settings.

A particularly challenging aspect of weed detection involves scenarios where weeds exhibit camouflage characteristics within their environment. In (Singh et al., 2024), the authors address this complex issue through an innovative approach for identifying small and multiple weed patches using drone imagery. Their work specifically tackles the challenge of detecting weeds in scenarios where they naturally blend with crops, introducing novel techniques for distinguishing camouflaged weeds in complex agricultural environments. Their methodology demonstrated remarkable success in detecting small-scale infestations that are typically difficult to identify due to their visual similarity with surrounding vegetation. This breakthrough in handling camouflaged scenarios has opened new possibilities for addressing one of the most challenging aspects of automated weed detection systems.

3 PROPOSED STUDY

This section details the different stages followed to carry out the proposed study.

3.1 Dataset Description

Weed detection and classification datasets represent crucial resources in agricultural technology, playing a vital role in developing automated systems for sustainable farming practices. These specialized collections of images are specifically designed to address the challenges in identifying and managing unwanted vegetation in various agricultural settings. The datasets used for this work (e.g., (Future, 2024a), (Future, 2024b)) encompass different weed species captured under different lighting conditions, growth stages, and field scenarios, making them particularly valuable for developing robust detection systems. The images have been captured from aerial shots. These datasets serve as fundamental tools in developing more sophisticated and efficient weed management solutions, ultimately contributing to more sustainable and productive agricultural practices, while also highlighting the growing importance of data-driven approaches in modern agriculture.

3.2 COD Approaches

Camouflage Object Detection (COD) represents one of the most fascinating and complex challenges in computer vision. This problem, inspired by natural phenomena where certain organisms have evolved to blend with their surroundings, has motivated the development of various innovative techniques in recent years. The ability to identify objects that deliberately try to confuse themselves with their environment not only has applications in security and surveillance but also in biology, ecology, and agricultural applications, which is the focus of this investigation. Below are the most significant contributions of the state-of-the-art in the COD task that will be used to carry out the different experiments. It is worth mentioning that this approach to weed detection using COD techniques has not been widely addressed; there is only one precedent (e.g., (Singh et al., 2024)).

In the current work, off-the-shelf COD approaches are evaluated in the datasets mentioned above (e.g., (Future, 2024a), (Future, 2024b)). Table 1 shows the input size of the image, the backbone used as well as the number of parameters of each technique. These approaches are briefly described next.

The first chosen technique has been proposed by Fan et al. (Fan et al., 2021) (SINet-v2) it establishes a

Table 1: Comparison between different characteristics of the camouflage techniques used.

Technique	Year	Input size (px)	Backbone	#Param. (M)
SINet-v2 (Fan et al., 2021)	2021	352×352	Res2Net50 (Gao et al., 2019)	24.93
BGNet (Chen et al., 2022b)	2022	416×416	Res2Net50 (Gao et al., 2019)	77.80
C^{2}F-Net (Chen et al., 2022a)	2022	352×352	Res2Net50 (Gao et al., 2019)	26.36
DGNet (Ji et al., 2023)	2023	352×352	EfficientNet (Tan and Le, 2019)	8.30
HitNet (Hu et al., 2023)	2023	352×352	PVTv2 (Wang et al., 2022)	25.73
PCNet (Yang et al., 2024)	2024	352×352	PVTv2 (Wang et al., 2022)	27.66

Figure 1: Example of images from the datasets presented in (Future, 2024a) and (Future, 2024b).

fundamental benchmark in detecting camouflaged objects. Their pioneering research introduces a convolutional neural network-based approach that revolutionizes how this challenge is addressed. The technique is distinguished by its ability to extract distinctive features that allow differentiating camouflaged objects from their background, thus establishing a solid foundation for future research in the field. Building upon this foundational work and expanding its capabilities, Ji et al. (Ji et al., 2023) (DGNet) present an innovative approach based on deep gradient learning. Their method stands out for meticulously analyzing gradual changes in visual patterns, using gradient information to identify subtle differences between camouflaged objects and their surroundings. This technique excels not only in its computational efficiency but also in its ability to detect objects in highly complex camouflage situations, complementing and enhancing the features established by SINet-v2.

Taking this progress in a more specialized direction while incorporating these established principles, Yang et al. developed PlantCamo (Yang et al., 2024) (PCNet), a technique specifically designed for detecting camouflage in plants. This method represents a significant advance in understanding natural camouflage in the plant kingdom, incorporating specific biological knowledge and unique plant camouflage patterns to improve detection accuracy while building upon the gradient analysis concepts introduced by DGNet. Further advancing these developments and integrating previous insights, the contribution of Hu et al. (Hu et al., 2023) (HitNet) introduces a high-resolution iterative feedback network that marks an important milestone. Their innovative method uses an iterative process that progressively refines detection results, leveraging high-resolution information to capture the most subtle details in images. This feedback approach allows for continuous improvement in detection accuracy while incorporating elements from both SINet-v2's feature extraction and DGNet's gradient analysis.

Contrary to the previous approaches, Chen et al. (Chen et al., 2022b) (BGNet) propose a unique perspective with their boundary-guided network. This technique is distinguished by its focus on the precise identification of camouflaged object boundaries, using sophisticated edge and contour information to improve segmentation. The method proves particularly effective in cases where the boundaries between the camouflaged object and the background are especially diffuse, complementing the high-resolution analysis of HitNet. Building upon all these advances and synthesizing their strengths, Chen et al. (Chen et al., 2022a) (C^{2}F-Net) present an innovative method based on context-aware cross-level fusion. This technique represents a significant advance by integrating information from multiple feature levels, considering both spatial and semantic contexts. The intelligent fusion of different levels of information allows for a more holistic understanding of the scene, significantly improving detection robustness across various camouflage scenarios while incorporating the boundary awareness of BGNet and the iterative refinement of HitNet.

Recent advances in camouflaged object detection have enhanced our detection capabilities, creating opportunities across biodiversity conservation, security, and scientific research. While various techniques offer complementary approaches that contribute to the field's development, and despite rapid progress in computer vision and machine learning, challenges persist in achieving accurate detection and segmentation of camouflaged objects in real-world scenarios.

3.3 Metric Evaluation

To evaluate results for COD approaches different evaluation metrics have been proposed in the literature in the current work, five widely used metrics for COD tasks are adopted to evaluate the detection results of each model, namely, the S-measure (S_α) (Fan et al., 2017), weighted F-measure (F_β^w) (Margolin et al., 2014), Mean Absolute Error (M) (Perazzi et al., 2012), E-measure (E_ϕ) (Fan et al., 2018), and F-measure (F_β) (Achanta et al., 2009). S_α computes the structural similarity between prediction and ground truth. F_β^w is an enhanced evaluation metric that extends the traditional F_β by incorporating spatial weights to better assess segmentation quality, particularly emphasizing boundary accuracy and location-based importance of detected pixels in object detection tasks. M focuses on evaluating the error at the pixel level between the normalized prediction and the ground truth. E_ϕ simultaneously assesses the overall and local accuracy of COD based on the human visual perception mechanism. F_β is an overall measure that synthetically considers both precision and recall.

Since different scores of F-measure can be obtained according to different precision-recall pairs, there are the mean F-measure (F_β^{mean}) and the maximum F-measure (F_β^{max}). Similar to the F-measure, maximum, and mean denoted as E_ϕ^{mean} and E_ϕ^{max} are also used as evaluation metrics.

4 RESULTS AND DISCUSSION

This section presents the results obtained with the proposed study. For the performance evaluation, the metrics described in Sec. 3.3 are used.

Table 2: Distribution of dataset

Task	**# of images with class weed**		
	Dataset 1	**Dataset 2**	**Dataset**
Training	543	300	843
Validation	91	10	101
Testing	38	8	46
Total	672	318	990

4.1 Dataset

The dataset used for the comparisons is obtained from (Future, 2024a) and (Future, 2024b), each dataset consists of 672 and 318 images, respectively, and both datasets are merged to get a unique dataset of 990 images. Table 2 shows the dataset distribution for each subset. Additionally, Figure 1 shows examples of the images that are part of the dataset.

4.2 COD Techniques

To carry out the study, six different COD techniques are used (i.e., SINet-v2 (Fan et al., 2021), BGNet (Chen et al., 2022b), C^2F-Net(Chen et al., 2022a), DGNet (Ji et al., 2023), HitNet (Hu et al., 2023), PCNet (Yang et al., 2024)). Table 1 shows the comparison between different characteristics of each of the camouflage techniques. With each architecture, a fine-tuning of 100 epochs is carried out. Most techniques operate with a standard input size of 352×352 pixels, except BGNet, which uses a slightly larger input size of 416×416 pixels. This consistency in input size among most models suggests a standardized approach to image processing in this field.

Regarding the backbone architectures, there is a notable variety in the choices made by different researchers. Three of the techniques (SINet-v2, BGNet, and C^2F-Net) utilize Res2Net50 as their backbone network. PCNet and HitNet opt for PVT-V2 (Pyramid Vision Transformer V2), while DGNet employs EfficientNet, showing the diversity in architectural approaches to the problem.

The number of parameters varies significantly across these models, ranging from 8.30 million in DGNet to 77.80 million in BGNet. DGNet stands out as the most parameter-efficient model with just 8.30M parameters, while BGNet represents the other extreme with 77.80M parameters. The remaining models (SINet-v2, PCNet, HitNet, and C^2F-Net) fall within a relatively similar range, between 24-28 million parameters, suggesting a common sweet spot for model complexity in this domain.

Table 3: Metric evaluation results for each COD techniques using the state-of-the-art datasets (e.g., (Future, 2024a), (Future, 2024b))—notation as presented in Sec. 3.3. Top 3 results are shown in red, blue, and green

Technique	$S_\alpha \uparrow$	$F^w_\beta \uparrow$	$M \downarrow$	$E^{mean}_\phi \uparrow$	$E^{max}_\phi \uparrow$	$F^{mean}_\beta \uparrow$	$F^{max}_\beta \uparrow$
SINet-v2 (Fan et al., 2021)	0.7251	0.4738	0.0611	0.7943	0.8302	0.5110	0.5358
BGNet (Chen et al., 2022b)	0.7075	0.4468	0.0945	0.8203	0.8922	0.5799	0.6201
C^{2}F-Net (Chen et al., 2022a)	0.6297	0.2547	0.1357	0.6557	0.7386	0.3627	0.3962
DGNet (Ji et al., 2023)	0.6931	0.4019	0.0779	0.7481	0.8013	0.4379	0.4621
HitNet (Hu et al., 2023)	0.7631	0.5594	0.0458	0.8617	0.8828	0.5816	0.5972
PCNet (Yang et al., 2024)	0.7612	0.5516	0.0450	0.8517	0.8877	0.5780	0.5938

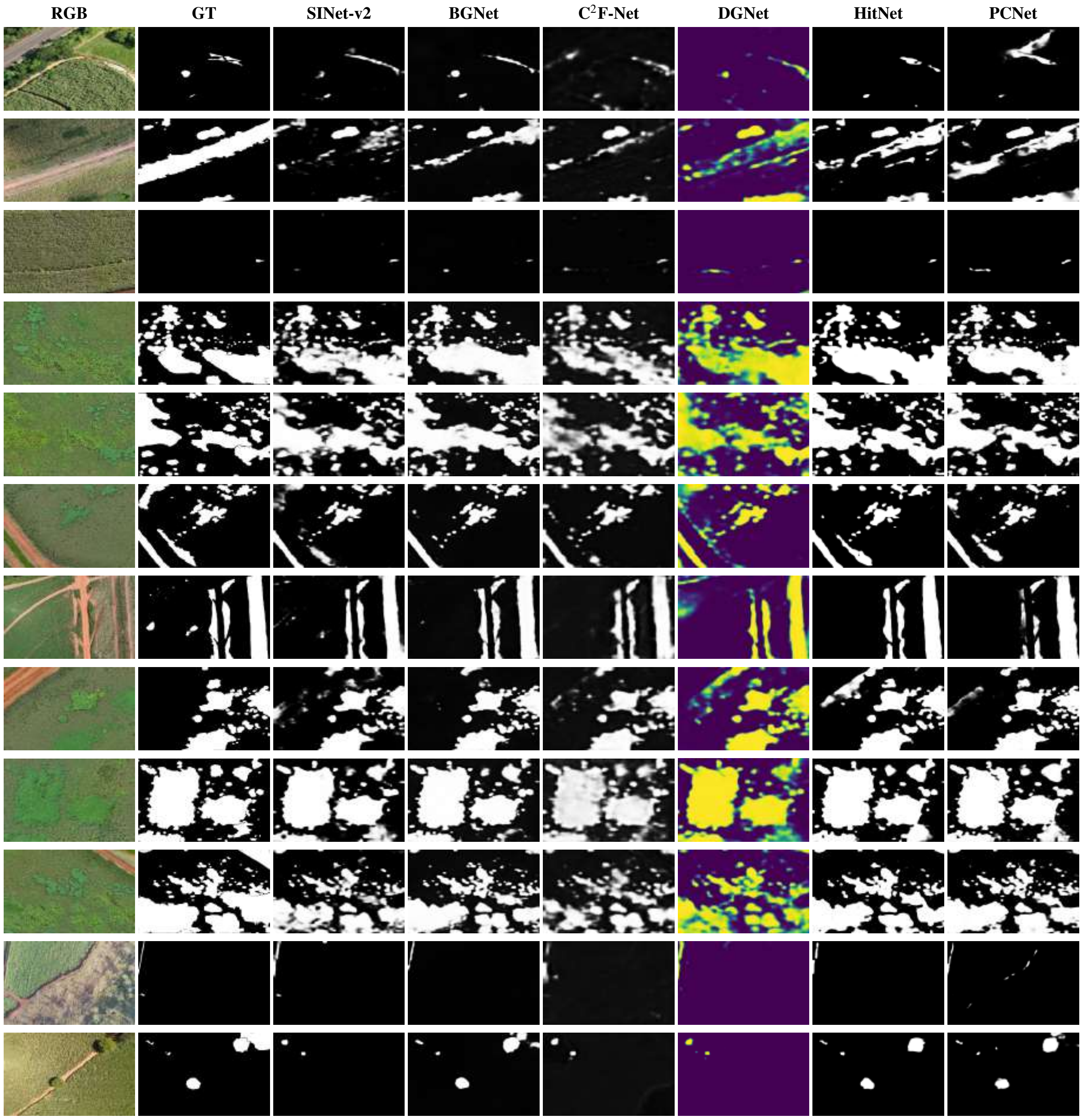

Figure 2: Prediction results using different state-of-the-art camouflage techniques. These example UAV images are part of the testing set (e.g., (Future, 2024a), (Future, 2024b)).

Table 4: Metric evaluation results on banana testing images with the presence of weeds. This is our dataset and consists of three images. Top 3 results are shown in red, blue, and green

Technique	$S_\alpha \uparrow$	$F_\beta^w \uparrow$	$M \downarrow$	$E_\phi^{mean} \uparrow$	$E_\phi^{max} \uparrow$	$F_\beta^{mean} \uparrow$	$F_\beta^{max} \uparrow$
SINet-v2 (Fan et al., 2021)	0.7254	0.5228	0.0348	0.8383	0.9631	0.5833	0.6353
BGNet (Chen et al., 2022b)	0.3301	0.0262	0.5185	0.2579	0.8176	0.0311	0.0582
C^{2}F-Net (Chen et al., 2022a)	0.4521	0.0528	0.1867	0.5711	0.8206	0.0710	0.0857
DGNet (Ji et al., 2023)	0.7048	0.4341	0.0498	0.8034	0.9489	0.5283	0.6262
HitNet (Hu et al., 2023)	0.8013	0.6280	0.0267	0.9394	0.9692	0.6849	0.6976
PCNet (Yang et al., 2024)	0.6655	0.3199	0.1033	0.7060	0.7848	0.3846	0.4697

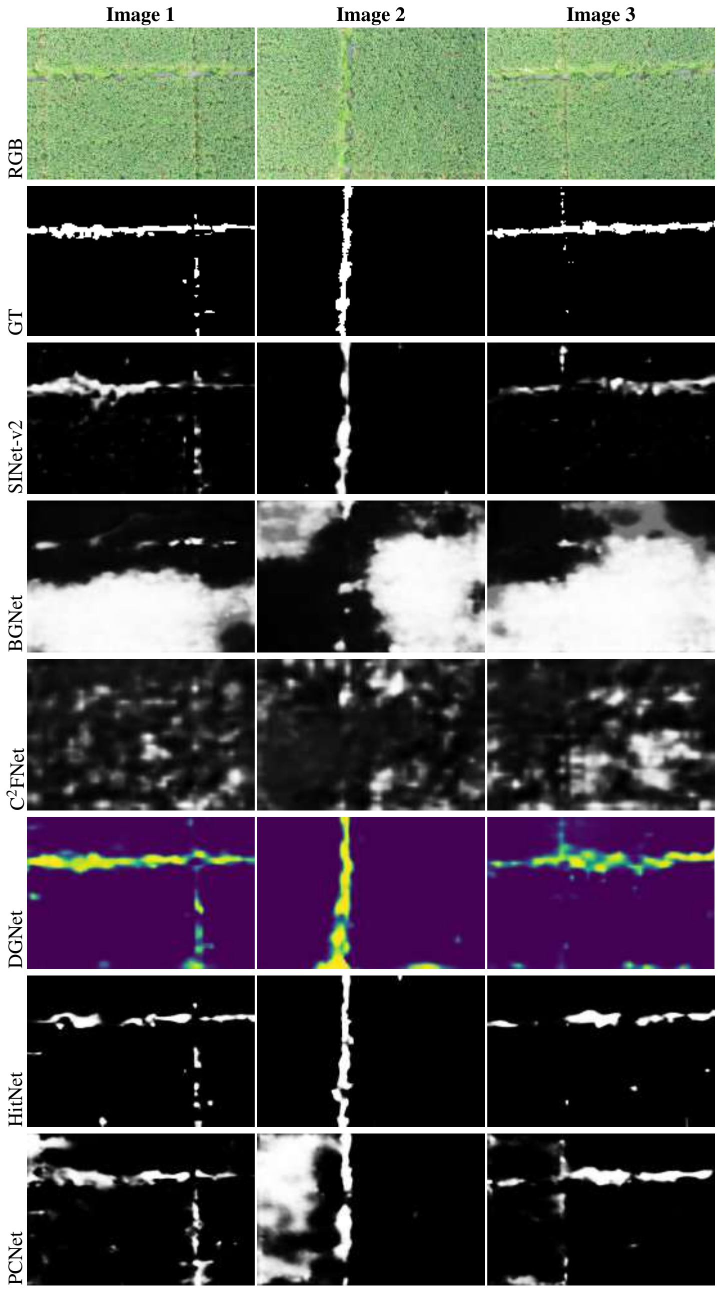

Figure 3: Prediction results of test UAV images belonging to banana crops with the presence of weeds. This is our dataset and consists of three images.

4.3 Comparisons

The comprehensive evaluation of six different COD architectures reveals significant insights into their performance in weed detection scenarios.

The quantitative analysis for the dataset described in Section 4.1 demonstrates that HitNet and PCNet consistently emerged as the leading performers across multiple evaluation metrics. HitNet achieved particularly impressive results with a S_α of 0.7631 and F_β^w of 0.5594, while PCNet closely follows with comparable performance metrics. In contrast, C^2F-Net shows the lowest performance across most metrics, indicating potential limitations in its application to weed detection tasks. The error analysis provides further validation of these findings, with PCNet and HitNet achieving the lowest M scores of 0.0450 and 0.0458, respectively. This represents a significant improvement over C^2F-Net, which shows the highest error rate with an M of 0.1357. The substantial difference in M between the best and worst-performing models (0.0907) shows the importance of architectural choices in achieving reliable weed detection results. In terms of boundary detection, BGNet demonstrates strong performance with an E_ϕ^{max} of 0.8922, while both HitNet and PCNet maintain consistent performance across both mean and maximum metrics. Table 3 shows metric evaluation results for each COD technique.

The qualitative analysis through visual results for the dataset described in Section 4.1 reveals important practical implications. HitNet and PCNet produce notably clearer boundaries between weed and non-weed regions, while BGNet shows good boundary detection but exhibits some over-segmentation tendencies. SINet-v2 and DGNet demonstrate moderate levels of false positives, whereas HitNet achieves a better balance between detection accuracy and false positives. C^2F-Net's tendency to under-detect camouflaged regions suggests limitations in its ability to handle subtle vegetation differences. On the other hand, it is also important to mention that the first four techniques in Fig. 2 (i.e., SINet-v2 (Fan et al., 2021), DGNet (Ji et al., 2023), PCNet (Yang et al., 2024), and HitNet (Hu et al., 2023)) can identify areas with very low weed density (see Fig. 2 *row 9 to 14*) identifying images where the presence of weeds is below 2% of the entire image, which indicates a good generalization capacity of these models for this type of datasets. Figure 2 shows some illustrations of the results obtained with the approaches evaluated in the current work.

To test the generalization of the different COD techniques, a dataset is captured and labeled that includes three images of banana crops in the presence of weeds. This dataset is not part of the training stage, only the pre-trained weights are used to test the generalization capacity of the architectures. Among the results obtained in banana crop scenarios, it can be highlighted the that HitNet significantly outperforms other models, obtaining the best result in all metrics. On the other hand, the SINet-v2 and DGNet metrics, in second and third place respectively, show regular performance, although not comparable with HitNet. Table 4 shows metric evaluation results on banana testing images in the presence of weeds. This specialized performance indicates HitNet's robust capability in handling complex agricultural scenarios. Although BGNet and C^2F-Net struggle significantly with banana crop images, SINet-v2 and DGNet maintain consistent performance, although PCNet shows somewhat reduced effectiveness compared to general testing scenarios. Figure 3 shows the prediction results of test images belonging to banana crops in the presence of weeds.

5 CONCLUSIONS

The comprehensive evaluation proposed in this work of six state-of-the-art COD techniques for weed detection has yielded several significant findings. HitNet and PCNet demonstrate superior performance in quantitative evaluation metrics for the dataset described in Section 4.1, with HitNet achieving the best overall results. These results can be contrasted with the qualitative analysis, showing that SINet-v2, PCNet, and HitNet present visual results where the weeds are correctly segmented. On the other hand, for the analysis carried out with images of banana crops with the presence of weeds, HitNet demonstrated exceptional performance in this type of scenario, indicating its robust capability of the model in specialized agricultural contexts despite not having been trained with images of banana crops and weeds. The significant variation in performance across different architectures shows the importance of model selection for specific agricultural applications. In summary, the success of the application of these COD techniques in the weed detection context demonstrates the viability of treating weed identification as a camouflage detection problem, opening new avenues for the field of precision agriculture.

ACKNOWLEDGEMENTS

This work was supported in part by the Air Force Office of Scientific Research Under Award FA9550-24-1-0206; in part by the ESPOL project "Advancing

Camouflaged Object Detection with a cost-effective Cross-Spectral vision system (ACODCS)" (CIDIS-003-2024); and in part by the University of Granada.

REFERENCES

Achanta, R., Hemami, S., Estrada, F., and Susstrunk, S. (2009). Frequency-tuned salient region detection. In *Conf. on Computer Vision and Pattern Recognition*, pages 1597–1604. IEEE.

Agarwal, R., Hariharan, S., Rao, M. N., and Agarwal, A. (2021). Weed identification using k-means clustering with color spaces features in multi-spectral images taken by uav. In *Int. Geoscience and Remote Sensing Symposium*, pages 7047–7050. IEEE.

Balabantaray, A., Behera, S., Liew, C., Chamara, N., Singh, M., Jhala, A. J., and Pitla, S. (2024). Targeted weed management of palmer amaranth using robotics and deep learning (yolov7). *Frontiers in Robotics and AI*, 11:1441371.

Chauhan, B. S., Matloob, A., Mahajan, G., Aslam, F., Florentine, S. K., and Jha, P. (2017). Emerging challenges and opportunities for education and research in weed science. *Frontiers in plant science*, 8:1537.

Chen, G., Liu, S.-J., Sun, Y.-J., Ji, G.-P., Wu, Y.-F., and Zhou, T. (2022a). Camouflaged object detection via context-aware cross-level fusion. *Transactions on Circuits and Systems for Video Technology*, 32(10):6981–6993.

Chen, T., Xiao, J., Hu, X., Zhang, G., and Wang, S. (2022b). Boundary-guided network for camouflaged object detection. *Knowledge-based systems*, 248:108901.

Coleman, G. R., Bender, A., Walsh, M. J., and Neve, P. (2023). Image-based weed recognition and control: Can it select for crop mimicry? *Weed Research*, 63(2):77–82.

Fan, D.-P., Cheng, M.-M., Liu, Y., Li, T., and Borji, A. (2017). Structure-measure: A new way to evaluate foreground maps. In *Int. Conf. on Computer Vision*, pages 4548–4557.

Fan, D.-P., Gong, C., Cao, Y., Ren, B., Cheng, M.-M., and Borji, A. (2018). Enhanced-alignment measure for binary foreground map evaluation. *arXiv*.

Fan, D.-P., Ji, G.-P., Cheng, M.-M., and Shao, L. (2021). Concealed object detection. *Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6024–6042.

Future (2024a). Machine segmentation dataset. https://universe.roboflow.com/future-cn6m4/machine-segmentation. visited on 2024-12-27.

Future (2024b). Weed segmentation dataset. https://universe.roboflow.com/future-cn6m4/weed-segmentation-mjudm. visited on 2024-12-27.

Gao, S.-H., Cheng, M.-M., Zhao, K., Zhang, X.-Y., Yang, M.-H., and Torr, P. (2019). Res2net: A new multi-scale backbone architecture. *Transactions on Pattern Analysis and Machine Intelligence*, 43(2):652–662.

Hu, X., Wang, S., Qin, X., Dai, H., Ren, W., Luo, D., Tai, Y., and Shao, L. (2023). High-resolution iterative feedback network for camouflaged object detection. In *AAAI Conf. on Artificial Intelligence*, volume 37, pages 881–889.

Ji, G.-P., Fan, D.-P., Chou, Y.-C., Dai, D., Liniger, A., and Van Gool, L. (2023). Deep gradient learning for efficient camouflaged object detection. *Machine Intelligence Research*, 20(1):92–108.

López-Granados, F., Jurado-Expósito, M., Peña-Barragán, J. M., and García-Torres, L. (2006). Using remote sensing for identification of late-season grass weed patches in wheat. *Weed Science*, 54(2):346–353.

Margolin, R., Zelnik-Manor, L., and Tal, A. (2014). How to evaluate foreground maps? In *Conference on Computer Vision and Pattern Recognition*, pages 248–255.

Moldvai, L., Ambrus, B., Teschner, G., and Nyéki, A. (2024). Weed detection in agricultural fields using machine vision. In *BIO Web of Conferences*, volume 125. EDP Sciences.

Parra, L., Marin, J., Yousfi, S., Rincón, G., Mauri, P. V., and Lloret, J. (2020). Edge detection for weed recognition in lawns. *Computers and Electronics in Agriculture*, 176:105684.

Perazzi, F., Krähenbühl, P., Pritch, Y., and Hornung, A. (2012). Saliency filters: Contrast based filtering for salient region detection. In *Conf. on Computer Vision and Pattern Recognition*, pages 733–740. IEEE.

Rehman, M. U., Eesaar, H., Abbas, Z., Seneviratne, L., Hussain, I., and Chong, K. T. (2024). Advanced drone-based weed detection using feature-enriched deep learning approach. *Knowledge-Based Systems*, 305:112655.

Singh, V., Singh, D., and Kumar, H. (2024). Efficient application of deep neural networks for identifying small and multiple weed patches using drone images. *Access*.

Tan, M. and Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning*, pages 6105–6114. PMLR.

Tang, J., Wang, D., Zhang, Z., He, L., Xin, J., and Xu, Y. (2017). Weed identification based on k-means feature learning combined with convolutional neural network. *Computers and electronics in agriculture*, 135:63–70.

Wang, W., Xie, E., Li, X., Fan, D.-P., Song, K., Liang, D., Lu, T., Luo, P., and Shao, L. (2022). Pvt v2: Improved baselines with pyramid vision transformer. *Computational Visual Media*, 8(3):415–424.

Westwood, J. H., Charudattan, R., Duke, S. O., Fennimore, S. A., Marrone, P., Slaughter, D. C., Swanton, C., and Zollinger, R. (2018). Weed management in 2050: Perspectives on the future of weed science. *Weed science*, 66(3):275–285.

Wu, Z., Chen, Y., Zhao, B., Kang, X., and Ding, Y. (2021). Review of weed detection methods based on computer vision. *Sensors*, 21(11):3647.

Yang, J., Wang, Q., Zheng, F., Chen, P., Leonardis, A., and Fan, D.-P. (2024). Plantcamo: Plant camouflage detection. *arXiv*.