Leveraging on foundation deep neural models for individual apple tree segmentation in dense orchards via prompt engineering in RGB images

Herearii Metuarea[1], Julien Garnier[2], Kévin Guerif[2], Frédérique Didelot[2], François Laurens[2], Ludovic Bervas[2], Pejman Rasti[1,2], Helin Dutagaci[3], and David Rousseau[1,2]

[1] Université d'Angers, LARIS, 62 Avenue Notre Dame du Lac, Angers, France
[2] INRAe-Institut Agro-Université d'Angers, 42 rue Georges Morel, Angers, France
[3] Eskisehir Osmangazi University, 26040, Eskisehir, Turkey
david.rousseau@univ-angers.fr

Abstract. We propose a strategy to prompt a vision foundation model in order to address, in a few-shot learning mode, the segmentation of individual trees in dense orchards from simple RGB images. The method produces similar segmentation results to those of a classical supervised segmentation method.

Keywords: Promptable vision foundational model · Prompt engineering · Instance segmentation · Tree detection · Trunk detection · Phenotyping

1 Introduction

We focus on the segmentation of individual apple tree in dense orchards. Due to the high density of such orchards, adjacent branches may be inter-wined which makes instance tree segmentation a very challenging task. Most of the current literature addresses this challenge by processing point clouds generated from LIDAR data or point clouds generated from sets of RGB images [9, 18]. Very recently, authors have tested the possibility to perform tree segmentation with a single RGB image [6] based on classical supervised deep learning. In this article, we propose to investigate this specific task with a prompt engineering approach as depicted in the visual abstract of Fig. 1.

2 Related work

Prompt engineering arises as a new paradigm with various strategies depending on the type of prompts and how they are used to boost few-shot learning [4, 10, 14, 15, 17]. Some specific prompt engineering strategies have been adapted to various application domains, as recently seen in medical imaging [1]. Similar approach is also deployed for plant imaging [2, 3, 8, 11–13, 16] but, to the best of our knowledge, this has never been applied to tree segmentation.

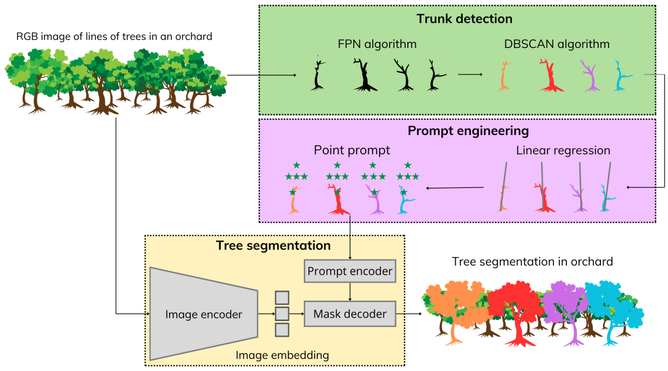

Fig. 1: Proposed workflow of individual tree segmentation. The trunk segmentation algorithm is the Feature Pyramid Network (FPN) developed by [7] which is trained on 3611 segmented trunks. The tree segmentation algorithm is the SAM v1 by [5] which is not fine-tuned but just used on a zero-shot learning mode based on prompts (5 points with diamond shape) positioned just above each individual detected trunks.

3 Results

Instance segmentation by the workflow of Fig. 1 of the tested trees (861 trees distinct from the 3611 used in the training of the trunk detection step) is assessed using the Dice and the average precision (AP) metric plus variants as shown in Table 1.

Table 1: Performance of the pipeline of Fig. 1 when using segment anything model (SAM v1) with optimized prompt located just above the detected trunks.

SAM	AP	AP_{50}	AP_{75}	$AP_{50:95}$	Dice
Unseen during training of the trunk detection	0.90	0.91	0.90	0.91	0.70 ± 0.16
Test Data from [6]	0.83	0.83	0.82	0.83	0.84 ± 0.06
Results from the supervised method of [6]	–	0.99	0.99	0.84	–

4 Conclusion

We have proposed a prompt engineering pipeline leveraging on a foundation model to tackle the difficult problem of individual apple tree segmentation in dense orchards from sole RGB images. The performance, although below the state-of-the-art for supervised learning, is competitive considering the zero-shot nature of the segmentation tree step and demonstrates that LIDAR data may not be necessary to segment the forefront rows of trees from the background.

5 Acknowledgements

This work was supported by the PHENET project funded from the EU Horizon research and innovation program under grant agreement N°101094587.

References

1. Ali, H., Bulbul, M.F., Shah, Z.: Prompt engineering in medical image segmentation: An overview of the paradigm shift. In: 2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings). pp. 1–4. IEEE (2023)
2. Carraro, A., Sozzi, M., Marinello, F.: The segment anything model (sam) for accelerating the smart farming revolution. Smart Agricultural Technology **6**, 100367 (2023)
3. Chen, Y., Yang, Z., Bian, W., Serikawa, S., Zhang, L.: Extraction study of leaf area and plant height of radish seedlings based on sam. In: Networking and Parallel/Distributed Computing Systems: Volume 18, pp. 69–83. Springer (2024)
4. Gu, J., Han, Z., Chen, S., Beirami, A., He, B., Zhang, G., Liao, R., Qin, Y., Tresp, V., Torr, P.: A systematic survey of prompt engineering on vision-language foundation models. arXiv preprint arXiv:2307.12980 (2023)
5. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Dollár, P., Girshick, R.: Segment anything. arXiv:2304.02643 (2023)
6. La, Y.J., Seo, D., Kang, J., Kim, M., Yoo, T.W., Oh, I.S.: Deep learning-based segmentation of intertwined fruit trees for agricultural tasks. Agriculture **13**(11), 2097 (2023)
7. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2117–2125 (2017)
8. Osco, L.P., Wu, Q., de Lemos, E.L., Gonçalves, W.N., Ramos, A.P.M., Li, J., Junior, J.M.: The segment anything model (sam) for remote sensing applications: From zero to one shot. International Journal of Applied Earth Observation and Geoinformation **124**, 103540 (2023)
9. Ozdarici, A., Ozgun, A.: Using remote sensing to identify individual tree species in orchards: A review. Scientia Horticulturae **321**, 112333 (2023)
10. Shtedritski, A., Rupprecht, C., Vedaldi, A.: What does clip know about a red circle? visual prompt engineering for vlms. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 11987–11997 (2023)
11. Sun, J., Yan, S., Alexandridis, T., Yao, X., Zhou, H., Gao, B., Huang, J., Yang, J., Li, Y.: Enhancing crop mapping through automated sample generation based on segment anything model with medium-resolution satellite imagery. Remote Sensing **16**(9), 1505 (2024)
12. Swartz, L.G., Liu, S., Cozatl, D.M., Palaniappan, K.: Segmentation of arabidopsis thaliana using segment-anything. In: 2023 IEEE Applied Imagery Pattern Recognition Workshop (AIPR). pp. 1–5. IEEE (2023)
13. Torres-Lomas, E., Lado-Jimena, J., Garcia-Zamora, G., Diaz-Garcia, L.: Segment anything for comprehensive analysis of grapevine cluster architecture and berry properties. arXiv preprint arXiv:2403.12935 (2024)
14. Wang, J., Liu, Z., Zhao, L., Wu, Z., Ma, C., Yu, S., Dai, H., Yang, Q., Liu, Y., Zhang, S., et al.: Review of large vision models and visual prompt engineering. Meta-Radiology p. 100047 (2023)
15. Zhang, C., Puspitasari, F.D., Zheng, S., Li, C., Qiao, Y., Kang, T., Shan, X., Zhang, C., Qin, C., Rameau, F., et al.: A survey on segment anything model (sam): Vision foundation model meets prompt engineering. arXiv preprint arXiv:2306.06211 (2023)

16. Zhang, W., Dang, L.M., Nguyen, L.Q., Alam, N., Bui, N.D., Park, H.Y., Moon, H.: Adapting the segment anything model for plant recognition and automated phenotypic parameter measurement. Horticulturae **10**(4), 398 (2024)
17. Zhou, K., Yang, J., Loy, C.C., Liu, Z.: Learning to prompt for vision-language models. International Journal of Computer Vision **130**(9), 2337–2348 (2022)
18. Zine-El-Abidine, M., Dutagaci, H., Galopin, G., Rousseau, D.: Assigning apples to individual trees in dense orchards using 3d colour point clouds. Biosystems Engineering **209**, 30–52 (2021)