Segmentation Quality Analysis Using Segment Anything Model in Very High-Resolution Imagery

Atik S. O.[1*], Ipbuker C.[1], Seker D. Z.[1]

[1] Department of Geomatics Engineering, Civil Engineering Faculty, Istanbul Technical University, Türkiye

*donmezsaz@itu.edu.tr

Abstract *With the widespread use of artificial intelligence (AI) in many fields, automatic information extraction from satellite images is among the trending topics. Segment Anything Model (SAM) is one of the critical developments that has become very popular in this field. In this study, it is aimed to apply the SAM method to high-resolution images with heterogeneous classes and to investigate the quality of the object boundaries obtained. As the application area of the study, the Pléiades satellite image of the selected area in Bergama district of Izmir was used. The region has coniferous forests, buildings, roads, fruit trees, and bare land. The boundaries of the obtained objects were analyzed with Over-segmentation (OS), Under-segmentation (US), Area fit index (AFI), and Quality rate (QR), which are parameters that measure segmentation quality. In the numerical analysis, the results obtained with SAM were examined comparatively with the results of the MRS algorithm, which is one of the methods on which the OBIA algorithm is based. In this context, the SAM algorithm showed superiority in several parameters used to measure segmentation quality compared to the building class's multi-resolution segmentation (MRS) algorithm. However, for all parameters, the MRS algorithm was superior to SAM for the tree class of the selected test samples. In this way, a comparative study is presented using SAM and the widespread use of AI in different band combinations in high-resolution satellite images.*

Keywords: Deep Learning, Segment Anything Model (SAM), Segmentation Quality, Multi-resolution Segmentation,

Introduction

As with many computer vision implementation domains, fewer human operators and faster methods are preferred. Deep Learning (DL) meets these needs admirably (Atik et al., 2022). DL methods extract the information content from EO data effectively and automatically. Convolutional neural networks (CNNs) can be used in these tasks, such as producing land cover, crop type maps, and determining cadastral or field boundaries. Segmentation, classification and object detection are main topics in computer vision tasks. Some of them have succeeded at the same time in the last years. Breaking the limits of segmentation, the recently presented segment anything model (SAM) has substantially advanced the growth of fundamental models for computer vision (Zhang et al., 2023).

In the Meta AI research team, Kirillov et al. (2023) introduced Segment, which is any project that involves tasks, models, and datasets. A pre-training algorithm is suggested by the promptable segmentation task, which produces a series of prompts (e.g., points, boxes, masks) for every training data and tests the model's masked predictions with the actual data (Kirillov, 2023). The dataset of SA model SA-1B is collected fully automatically.

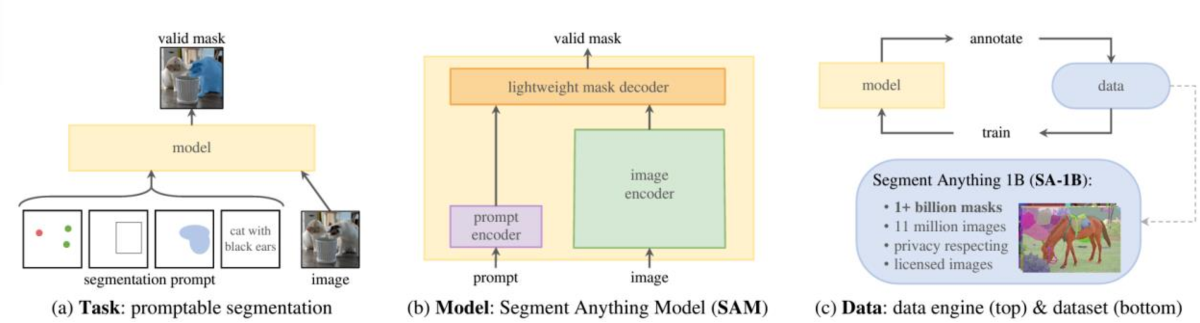

Figure 1: Schematic overflow of the SA model. It is borrowed from the original paper (Kirillov, 2023).

SAM has various usage areas, such as real-world scenes of crack detection (Ahmadi et al., 2023). Another example of usage SAM is moving object segmentation. There are improved frameworks such as Segment Any Anomaly + (SAA+) (Cao et al., 2024) and WS-SAM framework using weakly-supervised object detection (He et al., 2024). Another very crucial area is medical imaging using SAM. Computerized Tomography (CT) scan examples can help the segmentation of multi-phase live tumours (Hu & Li, 2023), organ (Roy et al., 2023) and multi-organ (Zhang & Liu, 2023). MRI imagery includes brain extraction (Mohapatra, 2023) and tumour segmentation (Putz et al., 2023). Colonoscopy (Zhou et al, 2023 and Li et al., 2023). Another SAM application is based on skin cancer segmentation (Hu et al., 2023). Besides, SAM-Track is utilised for multi-object detection from terrestrial cameras (Cheng et al., 2023). SAM is not suitable for implementing 3D scenes. However, the proposed 3A3D framework enables the segmentation of 3D scenes by leveraging NERFs (Cen et al., 2023).

Applying SAM models to very high-resolution images is still a new era. Several pioneer studies work on unmanned aerial vehicles (UAV) and satellite imagery. However, some gaps remain, especially in the classification phase after segmentation. There are some examples of merging CNNs with SAM, such as U-Net and SAM usage (Xiong et al, 2023). This study investigates different band combinations, showing the difference between the classical multi-resolution segmentation method and SAM results. Different band combinations cause different segment-quality performances.

Data

Satellite imagery regularly tracks the constant development and transformation of many man-made and natural regions on Earth. Spatial resolution directly affects classification accuracy. In the study, Pléiades o satellite imagery used Red-Green-Blue (R-G-B) and Near-Infrared (NIR) four bands dataset including trees, buildings, bare land, and road classes. The satellite image is pan-sharpened with 0.5 spatial resolution. The study is located in Bergama district in İzmir city, Türkiye.

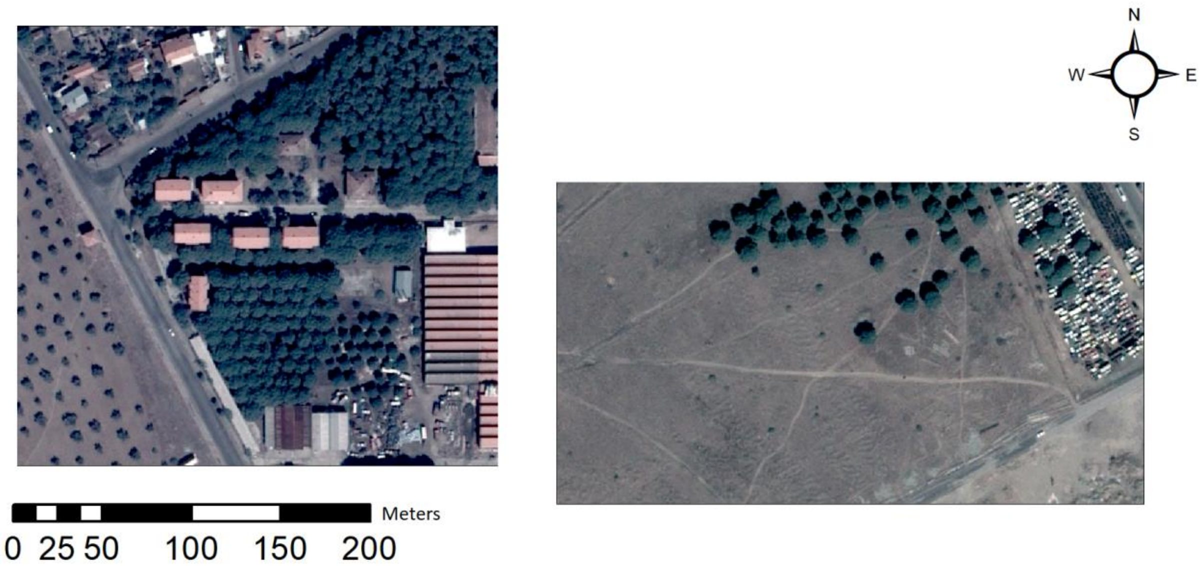

Figure 1: Study area in Bergama district: Region I is on the Left side, and Region II is on the Right.

In Figure 1, the study area is shown as two regions. In the first region, the building classification and the Region II tree segmentation are aimed at, and the results are compared.

Methodology

In the study, an algorithm of object-based image analysis (OBIA), multi-resolution segmentation, and a deep learning-based model, SAM, is used to classify a very high-resolution image with 0.5m spatial resolution in two band combinations. The first combination is RGB, and the second combines R-G-NIR bands. In the second combination, the NIR band is preferred over the blue band because it classifies trees more accurately.

a. Multi-resolution segmentation

Pixel-based image analysis and object-based image analysis (OBIA) are the two primary categories in remote sensing applications into which classification techniques can be divided. In this sense, the object-based approach can be seen as more sophisticated (İsiler et al., 2023)

Segmentation, also known as object-based image analysis, divides images into pieces consisting of related pixels (Atik, 2023). In the study for Region 1, the MRS parameters are chosen as the scale is 80, and the shape is equal to 0.1 for both band combinations. In Region II, the scale parameters are selected as 60, and the shape is implemented as 0.2.

b. Segment Anything Model (SAM)

SAM model gives different mask colours to the objects when processing the images. Figure 2 shows a SAM result for the study area of Region 1. In such cases, segmentation quality is better analysed for each object. Because the model shows various performances for different classes. The scale of the image is significant for running the model. When the sizes of the images change, even for the same imagery, the SAM results vary. When the image's borders are more extensive, the degree of generalization increases. The most miniature objects in the images lack segmentation because they are not masking. The algorithm gives different colours for each detected object by the model, like panoptic segmentation.

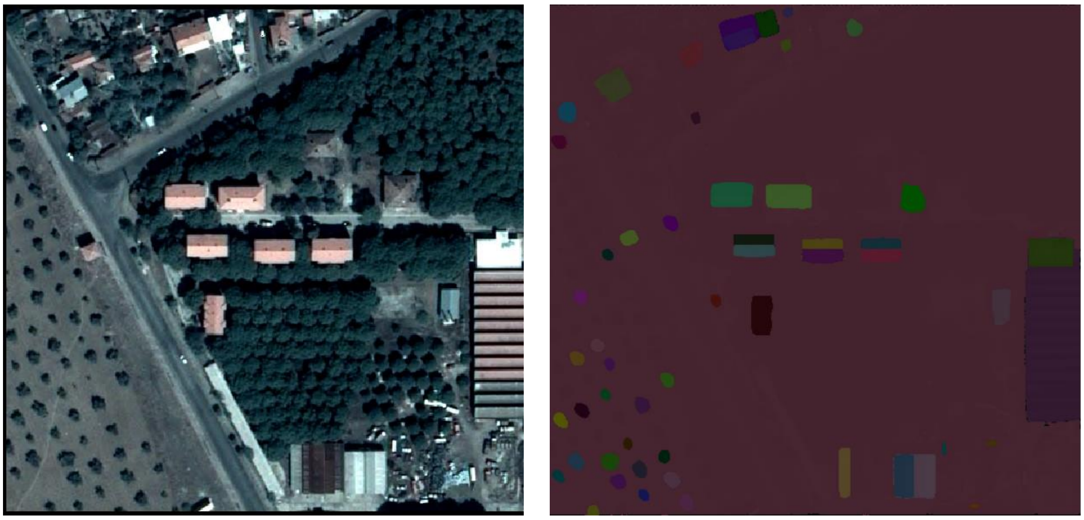

Figure 2: Region I is on the Left side, Region I RGB band combination on the Left side and SAM result for buildings on the right side.

c. Evaluation Metrics

The study used Equations 1-4 after segmentation as evaluation metrics.

$$Over - segmentation\ (OS) = 1 - \frac{A_{r(i)} \cap A_{s(i)}}{A_{r(i)}} \quad 1$$

$$Under - segmentation\ (US) = 1 - \frac{A_{r(i)} \cap A_{s(i)}}{A_{s(i)}} \quad 2$$

$$Area\ fit\ index\ (AFI) = \frac{A_{r(i)} - A_{s(i)}}{A_{r(i)}} \quad 3$$

$$Quality\ rate(QR) = \frac{A_{r(i)} \cap A_{s(i)}}{A_{r(i)} \cup A_{s(i)}} \quad 4$$

Under-segmentation (US) occurs when one segment corresponds to multiple objects (i.e., the referring segments are too large), and over-segmentation (OS) occurs when multiple segments define an individual object (i.e., the referring segments are too small). These are the two types of segmentation errors (Troya-Galvis, 2015). These figures are computed based on the reference object area (Ar) and the segmentation-obtained object area (As). The segmentation quality metrics used in this study were US, OS, area fit index (AFI) (Lucieer & Stein, 2022), and quality rate (QR) (Winter, 2000) (Equations (1)–(4)). US, OS, and AFI should all be near zero in a well-segmented system, while the QR value should be near one.

Results and Discussion

Which algorithm gives better segmentation results varies depending on the object class and the band combination in the dataset used. In the first test in Region I, the RGB band combination is used, and according to the performance metrics in three evaluation metrics, SAM gave better results, and for five metrics, MRS has superior values. In the second test, different metrics give similar results. According to the Region II test, segmenting the tree class gives higher values when using the MRS algorithm in both band combinations.

Figure 3: The RGB band combination on left is beloved to MRS, and the right is beloved to SAM in Region I.

Figure 4: The RGNIR band combination left is beloved to MRS, and the right is beloved to SAM in Region I. (The yellow line is Ground Truth, the Green is MRS, and the Red line is SAM).

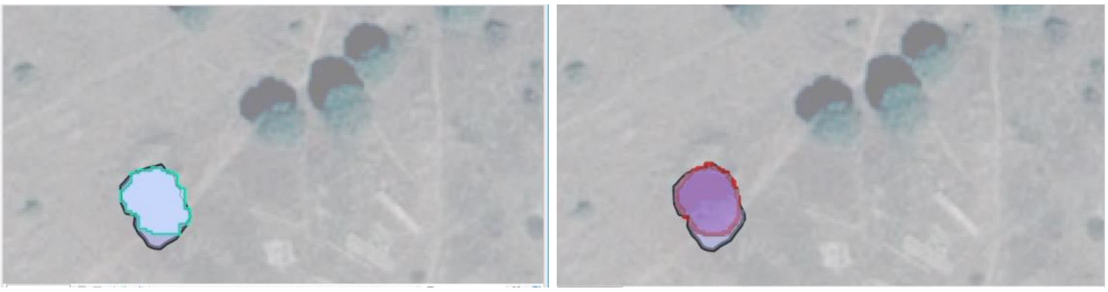

Figure 5: Tree segmentation in RGB band combination in Region II (black line GT, Left side is MRS, right side is SAM)

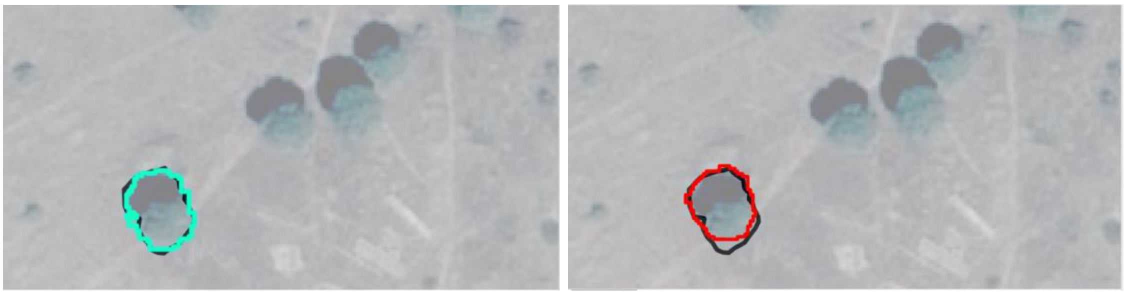

Figure 6: Tree segmentation in RGNIR band combination in Region II (black line GT, Left side is MRS, right side is SAM)

Figure 3,5 shows samples of segmentation from the RGB dataset. Segmentations from the RGNIR dataset are also shown in Figures 4 and 6. The red line refers to the SAM segments, the green line to the MRS segments, and the black line to the reference object in Figures 5 and 6. In the tables, OS, US, and AFI values are better near zero, and QR values near 1. Although the MRS algorithm mainly provides higher-quality results in segmentation, SAM comes to the fore in some metrics, especially in building segmentation.

Table 1: Region I, Building 1 Results of the Segmentation

Method	Dataset	OS	US	AFI	QR
MRS	RGB	**0.0018**	**0.0386**	**0.0383**	**0.9598**
	RGNIR	0.1245	**0.0215**	0.1052	0.8590
SAM	RGB	0.0794	0.1351	0.0645	0.8048
	RGNIR	**0.0189**	0.0362	**0.0180**	**0.9462**

Table 2: Region I, Building 1 Results of the Segmentation

Method	Dataset	OS	US	AFI	QR
MRS	RGB	**0.0540**	0.1406	0.1007	**0.8193**
	RGNIR	**0.0271**	0.0828	**0.0608**	**0.8943**
SAM	RGB	0.0969	**0.1358**	**0.0451**	0.7909
	RGNIR	0.2887	**0.0023**	0.2871	0.7101

Table 3: Region II results of the segmentation

Method	Dataset	OS	US	AFI	QR
MRS	RGB	**0.1579**	**0.0319**	**0.1301**	**0.8194**
	RGNIR	**0.0645**	**0.0651**	**0.0007**	**0.8783**
SAM	RGB	0.2004	0.0440	0.1635	0.7712
	RGNIR	0.1412	0.0784	0.0682	0.8004

Better segmentation quality values are shown as bold in the Tables. In the table 1, The highest OS value is 0.0018 is beloved MRS algorithm using RGB. The better US value is obtained using SAM as 0.0023 in the Region 1 using RGBNIR bands. The superior AFI values are yielded as 0.0007 in MRS algorithm test of Region II using RGNIR. Better QR results is obtained as 0.9598 using MRS algorithm in the test of Region I.

Conclusion

Although the SAM model produces very reliable and high quality results in many areas, in applications on high altitude images such as aerial photographs or satellite images, more experiments are needed to reach the results of classical methods. Some difficulties encountered in using SAM in this study can be overcome using the SAM model in conjunction with different CNN algorithms. Alternatively, the segmentation phase can be used together with classification to obtain results that serve many different purposes. These studies can be replicated under many different headings for the United Nations SDGs.

Acknowledgments

We would like to thank Istanbul Technical University's Implementation and Research Center for Satellite Communications and Remote Sensing (ITU - CSCRS) for the high-resolution satellite image of Pléiades that we used in our study.

References

Ahmadi, M., Lonbar, A. G., Sharifi, A., Beris, A. T., Nouri, M., & Javidi, A. S. (2023). Application of segment anything model for civil infrastructure defect assessment. *arXiv preprint arXiv:2304.12600*.

Atik, Ş.Ö. (2023). Object-Based Integration Using Deep Learning and Multi-Resolution Segmentation in Building Extraction from Very High Resolution Satellite Imagery. *Turkish Journal of Remote Sensing*, 5 (2), 67-77.

Atik, S. O., Atik, M. E., & Ipbuker, C. (2022). Comparative research on different backbone architectures of DeepLabV3+ for building segmentation. *Journal of Applied Remote Sensing*, *16*(2), 024510-024510.

Atik, S. O., & Ipbuker, C. (2021). Integrating convolutional neural network and multiresolution segmentation for land cover and land use mapping using satellite imagery. *Applied Sciences*, *11*(12), 5551.

Cao, Y., Xu, X., Sun, C., Cheng, Y., Du, Z., Gao, L., & Shen, W. (2023). Segment any anomaly without training via hybrid prompt regularization. *arXiv preprint arXiv:2305.10724*.

Cheng, Y., Li, L., Xu, Y., Li, X., Yang, Z., Wang, W., & Yang, Y. (2023). Segment and track anything. *arXiv preprint arXiv:2305.06558*.

He, C., Li, K., Zhang, Y., Xu, G., Tang, L., Zhang, Y., ... & Li, X. (2024). Weakly-supervised concealed object segmentation with sam-based pseudo labeling and multi-scale feature grouping. *Advances in Neural Information Processing Systems*, *36*.

Hu, C., Xia, T., Ju, S., & Li, X. (2023). When sam meets medical images: An investigation of segment anything model (sam) on multi-phase liver tumor segmentation. *arXiv preprint arXiv:2304.08506*.

Hu, M., Li, Y., & Yang, X. (2023). Skinsam: Empowering skin cancer segmentation with segment anything model. *arXiv preprint arXiv:2304.13973*.

Isiler, M., Yanalak, M., Atik, M. E., Atik, S. O., & Duran, Z. (2023). A Semi-Automated Two-Step Building Stock Monitoring Methodology for Supporting Immediate Solutions in Urban Issues. *Sustainability*, *15*(11), 8979.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R. (2023). Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 4015-4026).

Mohapatra, S., Gosai, A., & Schlaug, G. (2023). Sam vs bet: A comparative study for brain extraction and segmentation of magnetic resonance images using deep learning. *arXiv preprint arXiv:2304.04738*.

Li, Y., Hu, M., & Yang, X. (2024, April). Polyp-sam: Transfer sam for polyp segmentation. In *Medical Imaging 2024: Computer-Aided Diagnosis* (Vol. 12927, pp. 759-765). SPIE.

Lucieer, A.; Stein, A. Existential uncertainty of spatial objects segmented from satellite sensor imagery. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2518–2521.

Putz, F., Grigo, J., Weissmann, T., Schubert, P., Hoefler, D., Gomaa, A., ... & Huang, Y. (2023). The Segment Anything foundation model achieves favorable brain tumor autosegmentation accuracy on MRI to support radiotherapy treatment planning. *arXiv preprint arXiv:2304.07875*.

Roy, S., Wald, T., Koehler, G., Rokuss, M. R., Disch, N., Holzschuh, J., ... & Maier-Hein, K. H. (2023). Sam. md: Zero-shot medical image segmentation capabilities of the segment anything model. *arXiv preprint arXiv:2304.05396*.

Troya-Galvis, A., Gançarski, P., Passat, N., & Berti-Equille, L. (2015). Unsupervised quantification of under-and over-segmentation for object-based remote sensing image analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(5), 1936-1945.

Xiong, X., Wu, Z., Tan, S., Li, W., Tang, F., Chen, Y., ... & Li, G. (2024). SAM2-UNet: Segment Anything 2 Makes Strong Encoder for Natural and Medical Image Segmentation. *arXiv preprint arXiv:2408.08870*.

Winter, S. Location similarity of regions. *ISPRS J. Photogramm. Remote Sens.* **2000**, *55*, 189–200.

Zhang, C., Liu, L., Cui, Y., Huang, G., Lin, W., Yang, Y., & Hu, Y. A Comprehensive Survey on Segment Anything Model for Vision and Beyond. arXiv 2023. *arXiv preprint arXiv:2305.08196*.

Zhang, K., & Liu, D. (2023). Customized segment anything model for medical image segmentation. *arXiv preprint arXiv:2304.13785*.

Zhou, T., Zhang, Y., Zhou, Y., Wu, Y., & Gong, C. (2023). Can sam segment polyps?. *arXiv preprint arXiv:2304.07583*.