QUALITY ASSESSMENT FOR MULTI-RESOLUTION SEGMENTATION AND SEGMENT-ANYTHING MODEL USING WORLDVIEW-3 IMAGERY

Elif Ozlem Yilmaz[1*], Taskin Kavzoglu[1]

[1] Department of Geomatic Engineering, Gebze Technical University, Kocaeli - (eoyilmaz, kavzoglu)@gtu.edu.tr

KEY WORDS: Image Segmentation, Multiresolution Segmentation, Segment Anything Model, Segmentation Quality.

ABSTRACT:

Image segmentation, is of utmost importance in the disciplines of digital image processing, particularly remote sensing and computer vision, has seen an increasing demand for precise and efficient algorithms. This study focuses to conduct a comparative exploration of the segmentation capabilities of two sophisticated techniques namely, Multiresolution Segmentation (MRS) and Segment Anything Model (SAM), leveraging the high-resolution WorldView-3 (WV-3) satellite image. MRS adopts a hierarchical methodology, segmenting an image into various scales while retaining a profound understanding of its structure. Conversely, SAM employs a deep learning algorithm, prioritizing segment creation based on conceptual pixel similarity, irrespective of spatial adjacency. The WV-3 image, featuring diverse land cover elements like agricultural parcels, industrial structures, roads, red roofs, single trees, and water bodies, serves as the basis for assessing segmentation quality. Both methods are applied to the image, and their outcomes are individually evaluated against manually generated polygonal land use/cover objects. Segmentation quality metrics are employed for assessment. Results reveal MRS effectively preserves fine details and entity delineation, while SAM excels in capturing contextually similar regions. MRS outperforms SAM with a negligible discrepancy, yet SAM demonstrates superiority in the red roof object, achieving an Intersection over Union (IoU) value of 0.70 compared to MRS's 0.49. MRS tends to generate numerous segments for an item, while SAM produces only one segment. Nevertheless, it is important to recognise that both algorithms have specific constraints in particular scenarios, such as excessive segmentation in areas with abundant texture or inadequate segmentation in areas with slight changes.

1. INTRODUCTION

The utilisation of remote sensing data is a powerful way to analyse the Earth's surface. The latest studies on the acquisition of extremely detailed images and progress in digital image processing methods have yielded a wide range of uses for remote sensing (Kavzoglu, 2017). These applications include surveillance of the environment, management of natural resources, and the long-term preservation of environmental services (Colkesen and Kavzoglu, 2019; Kavzoglu, 2008; Tonbul et al., 2020). Moreover, due to current progresses in the development of sensor technologies, remote sensing imagery are being collected with greater precision, getting them useful for acquiring the surface information (Kavzoğlu and Yılmaz, 2022; Fan et al., 2024). There has been a substantial shift in the proportion between the dimensions of identified objects and the dimensions of pixels, thanks to the progress of satellite sensor technologies with a spatial resolution below 1 m. The pixels have become much smaller relative to the average dimensions of identified objects. Given the extensive level of detail in very high-resolution images, the traditional pixel-based approach is not suitable for such a kind of imagery (El-naggar, 2018). For instance, WorldView-3 (WV-3), which was launched in August 2014, possesses a remarkable spatial resolution of 0.31 m in panchromatic mode, 1.24 m in eight-band multispectral mode, and 3.70 m in eight-band SWIR mode. The improvement in digital image processing approaches has contributed to advances in numerous applications, including image classification, extraction of features and identification of changes. The complex nature of remotely sensed data requires the development of specialised image processing methodologies to extract meaningful insights and support informed decision making.

Pixel-based techniques employed for low and medium spatial resolution imagery are insufficient in leveraging the spatial variability of distinct land use and land cover (LULC) in high resolution images, as they neglect to consider adjacent pixels belonging to the identical LULC (Campbell and Wynne, 2011). As a result, object-based image analysis (OBIA) has been recognised as a highly effective technique for evaluating remotely sensed imagery with a high level of spatial detail (Hossain and Chen, 2019). This technique is a widely employed method for examining digital images (Blaschke, 2010). Unlike pixel-based classification, OBIA assesses real-world objects in an image by considering multiple criteria such as texture, colour, shape features, and their surrounding areas. Pixel-based classification focuses on individual spectral values, whereas OBIA considers information from a group of comparable nearby pixels, referred to as surface objects (Kavzoglu and Yildiz, 2014; Kavzoglu et al., 2016).

The implementation of OBIA typically involves two processing steps: image segmentation and classification. Image segmentation is of great importance in a variety of domains, consisting remote sensing and computer vision. The aim of segmentation is to divide an image into distinct sections that vary in specific attributes such as grayscale level, colour, form, dimensions, and texture (Lucchese and Mitray, 2001). The increasing demand for accurate and efficient image segmentation algorithms is driven by the advances in remote sensing technology (Kavzoglu et al., 2017). The segmentation of image objects significantly affects the accuracy of the classification process as it is a preliminary step prior to image classification. Hence, many approaches have been proposed for achieving remarkable accuracy in image classification applications. Numerous segmentation techniques are employed for image segmentation (i.e., Mean Shift, Multiresolution (MRS) and K-Means) in current works of literature (Kavzoglu and Tonbul, 2018). Recently, a novel Segment Anything Model (SAM) utilising deep learning is proposed, with the potential for producing image segmentation (Kirillov et al., 2023).

The selection of the segmentation technique is related to several issues, including the contents of the image, the tools utilised, the study goals, and the resolution of specific issues related to the study (Ez-zahouani et al., 2023). One of the most crucial aspects is the production of segments with remarkable quality. Evaluation of segmentation is critical for qualitatively and quantitatively determining the optimal segmentation and image analysis strategy (Maxwell et al., 2021a; 2021b). Approaches for assessing the quality of segmentation involve quantitative and quantitative (supervised and unsupervised) techniques. Visual assessment is a simple task, nonetheless, it lacks the ability to generate quantitative findings and is unavoidably influenced by subjective factors. The evaluation of segment quality is based on classification accuracy, where more accurate classification indicates a higher quality of segmentation application. The effectiveness of OBIA is directly influenced by the accuracy of segmentation in this technique. Nevertheless, it fails to encompass the numerous inherent characteristics of a segmentation method that are unrelated to specific applications (Zhang et al., 2015).

The aim of this research is to carry on a thorough comparative analysis, carefully assessing the segmentation capabilities demonstrated by two advanced algorithms: MRS and SAM. This comprehensive study primarily focuses on the extensive adoption of the high-resolution WV-3 satellite image. The primary objective is to assess and contrast the effectiveness of MRS and SAM in the process of dividing and defining different LULC components within the WV-3 image. The study surpasses a simple quantitative evaluation by integrating qualitative factors to provide an in-depth understanding of the segmentation findings produced by these advanced algorithms.

2. METHODOLOGY

2.1 Study Area and Dataset

The study region was situated in the Akyazı district of Sakarya province in the Marmara region of Turkey (Figure 1). Agriculture is the primary means of survival in this area, with an average annual precipitation of 800 mm. Typically, the cultivated crops include hazelnut, wheat, beets, potato, and a variety of vegetables, with corn being particularly prominent. Also, it encompasses objects related to various LULC categories, such as roads, agricultural parcels, rivers, woodland, residential and industrial areas. The research area covers approximately 422 km^2 of agricultural land and 402 km^2 of wooded area (Colkesen et al., 2023).

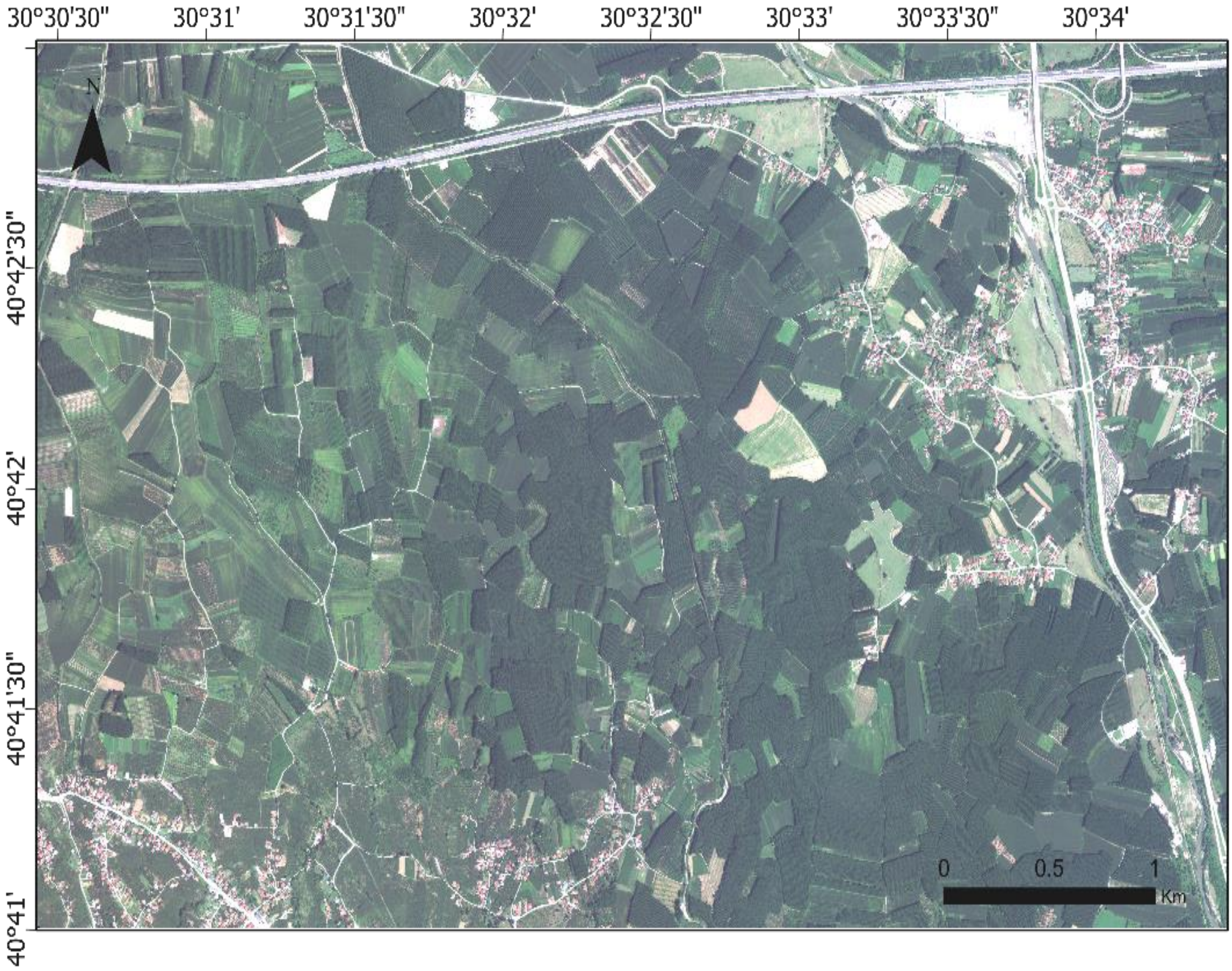

Figure 1. The study area is in Akyazı district of Sakarya in Türkiye.

The data source was a WV-3 image acquired on August 15, 2021. The WV-3 satellite captures imagery using 8 multi-spectral bands with a spatial resolution of 1.2 m. Furthermore, the satellite is equipped with an 8-band SWIR sensor that has a spatial resolution of 3.7 m (Table 1). Prior to the segmentation analysis, the SWIR bands were resampled to a resolution of 1.2 m using the closest neighbour approach. The resulting images were then combined with the multi-spectral bands. Moreover, WV-3 image, which contains a variety of LULC features, namely agricultural parcels, industrial buildings, roads, red roofs, single trees, and water bodies, was analysed to perform the segmentation quality analysis.

Band Name	Wavelength (nm)	Spatial Resolution
PAN	450-800	Nadir: 0.31 m, 20º off-nadir: 0.34 m
Coastal Blue Green Yellow Red Red Edge NIR-1 NIR-2	400 - 450 450 - 510 510 - 580 585 - 625 630 - 690 705 - 745 770 - 895 860 - 1040	Nadir: 1.24 m, 20º off-nadir: 1.38 m
SWIR-1 SWIR-2 SWIR-3 SWIR-4 SWIR-5 SWIR-6 SWIR-7 SWIR-8	1195 - 1225 1550 - 1590 1640 - 1680 1710 - 1750 2145 - 2185 2185 - 2225 2235 - 2285 2295 - 2365	Nadir: 3.70 m, 20º off-nadir: 4.10 m

Table 1. The characteristics of WV-3 imagery bands

2.2 Segmentation Algorithms

The term "segmentation" is used to describe all the processes involved in constructing, developing, merging, reducing, or dividing objects (Blaschke, 2010). Segmentation algorithms are essential in OBIA as they are employed to partition the image into significant components, enabling the identification of objects and features. It is a crucial initial step in the OBIA studies. The effectiveness of the image classification method is mostly determined by the quality of the procedure of segmentation, which in turn relies on the appropriate selection of parameter values for segmentation. An effective segmentation can be described as one that accurately identifies objects in an image, matching them precisely to the reference objects. It should avoid under-segmentation and may have a slight tendency towards over-segmentation. The effectiveness of a mathematical model for segmentation algorithms can be evaluated based on its ability to enable users to achieve precise segmentation without the need for exact choice of segmentation parameters (El-naggar, 2018). The MRS and SAM algorithms employed in the present study are thoroughly analysed from the perspectives of implementation. The MRS technique involves partitioning the image into a hierarchical structure at various sizes, whereas the SAM model utilises deep learning to segment the image by considering the conceptual similarity of pixels.

2.2.1 Multiresolution Segmentation

The MRS algorithm employs a hierarchical approach to partition an image into many segments based on various scales, so providing a comprehensive examination of the image's structure (Kavzoglu and Tonbul, 2017). In other words, it is a region-based method that computes the heterogeneity of spectral and form characteristics and also generates a collection of image objects by assigning weights to image layers (Emmanue et al., 2023). This algorithm depends on the user and the satellite image, making it relatively complex (Marpu et al., 2010). The primary variables of the object-oriented MRS algorithm intended for utilisation in this investigation are scale, shape, and compactness (Chen et al., 2023). The shape parameter has a value ranging from 0.1 to 0.9 to diminish the influence of the spectral diversity of pixels. Also, the compactness parameter plays a crucial role in distinguishing the boundaries of objects as either smooth or distinct during the segmentation procedure. The scale parameter is of utmost importance as it governs the relative dimension of objects and clearly influences the succeeding classification processes (Witharana and Civco, 2014).

2.2.2 Segment Anything Model

The SAM, based on a deep learning algorithm, prioritizes the creation of segments based on the conceptual similarity of pixels, regardless of their spatial adjacency to capture a greater amount of contextual data and provide a comprehensive representation of the image. The SAM algorithm comprises a Vision Encoder that functions as an image encoder, a Prompt Encoder that imposes constraints on the objects in the image, and Mask Decoder models that establish the relationship between the constraints and the image. By leveraging the advanced functionalities of SAM in segmentation and object recognition, it can effectively separate objects in remotely sensed images prior to the classification. Furthermore, SAM is a model trained with more than 11 million images and 1.1 billion segmentation masks and is used as an object detection and image enhancement algorithm in addition to its segmentation capability process in image processing applications (Kirillov et al., 2023).

2.3 Segment Quality Metrics

Segmentation is a crucial stage in OBIA applications, as it serves as the foundation for classification (Kavzoglu and Tonbul, 2017). Hence, it is vital to assess the quality of the generated segments. To begin, it is necessary to generate a dataset that includes reference segments to assess the segment quality (Jozdani and Chen, 2020). Therefore, since reference datasets are generally based on geometrically identifiable objects with smooth object boundaries (i.e., buildings, roads, vehicles, tree, parcel, etc.) can be used on the image. It is intended that the reference objects and the image segments created by the reference datasets have a one-to-one association. These models are created by analysing the geometric and/or mathematical link between the image objects being studied and the associated reference objects. Therefore, the evaluation of segmentation quality involved the use of mean Intersection over Union (IoU), Area-Fit Index (AFI), Precision, Recall, F1-score and Under-segmentation (US) and Over-segmentation (OS) measures (Zhang et al., 2015; Kavzoglu and Tonbul, 2017; Goodwin et al., 2022). To explain the quality metrics used in more detail, the IoU value, which is shown in equation 1, is computed by dividing the intersection of the reference object and the created segments by their union.

$$IoU = \frac{1}{k}\sum_{j=1}^{k} \frac{n_{jj}}{(n_{ij}+n_{ji}+n_{jj})}, \quad (1)$$

The AFI metric, shown in equation 2, quantifies the extent of similarity between reference objects and their related image

https://doi.org/10.5194/isprs-archives-XLVIII-4-W9-2024-383-2024 | © Author(s) 2024. CC BY 4.0 License.

segments. A value of zero indicates perfect overlap between the reference object and the relevant segment.

$$AFI = \frac{A_{r(i)} - A_{s(j)}}{A_{r(i)}}, \tag{2}$$

Precision evaluates the correctness of positive segment predictions by determining the proportion of accurately segment predicted positive cases out of all instances segment predicted as positive (equation 3). Recall measures a model's capacity to accurately identify all true positive segment cases, by calculating the ratio of properly predicted segment positives to the total number of true positive segment instances (equation 4). Besides, the F1-score metric is the harmonic average of precision and recall (equation 5).

$$Precision = \frac{\sum_{i=1}^{n}|s_i \cap R_{imax}|}{\sum_{i=1}^{n}|s_i|}, \tag{3}$$

$$Recall = \frac{\sum_{i=1}^{n}|R_i \cap s_{imax}|}{\sum_{i=1}^{n}|R_i|}, \tag{4}$$

$$F1 - score = 2\,x\,\frac{Precion\ x\ Recall}{Precion + Recall}, \tag{5}$$

OS refers to the situation where the obtained segments are smaller than they must be (equation 6). It is a metric calculated by dividing the intersection of the reference objects and the corresponding segments by the reference object area. On the other hand, US means to the situation where the obtained segments are larger than they should be (equation 7). The OS and US metrics range from zero to one, with a value of zero indicating optimum segmentation.

$$OS = 1 - \frac{A_{r(i)} \cap A_{s(j)}}{A_{r(i)}}, \tag{6}$$

$$\mathrm{US} = 1 - \frac{A_{r(i)} \cap A_{s(j)}}{A_{r(j)}}, \tag{7}$$

3. RESULTS

To extract the objects that belong to the WV-3 image, segmentation methods were applied to the image. The optimization technique was implemented to determine the parameters of MRS algorithm. Following the optimization process, the scale parameter was determined to be 40, the shape parameter to be 0.3, and the compactness parameter to be 0.7. MRS was employed to finish the segmentation process utilising eCognition software. On the other hand, the weights of the sam_vit_h_4b8939 model were utilised for SAM. An erosion kernel of size 3x3 was employed. The model had a prediction IoU threshold of 0.85, and also a minimum mask region area of 10 pixels was chosen. The segmentation analysis utilising the SAM algorithm was executed on a Jupyter notebook supplemented with a Python programming language.

Figure 2 displays the segments generated by the two segmentation algorithms for the study area. The red colour signifies the MRS results, whereas the green lines depict the result generated by SAM. A visual examination of the produced outcomes reveals that the MRS algorithm generates many segments for an image object. However, while analysing the outputs of the SAM algorithm, it is apparent the actual reference data matches more closely with the borders of objects, particularly for red roof and agricultural parcel objects.

Both algorithms were applied to the image and the results obtained from the segmentation algorithms were separately evaluated and compared with each of the manually drawn polygonal LULC objects. The study includes six categories of LULC objects with distinct boundaries: agricultural parcel, industrial structures, red roof, road, restricted water bodies, and single trees (Figure 3). The MRS is represented by the red line, the SAM is represented by the green line, and the manually created reference objects are represented by the yellow line.

Quantitative quality indicators, namely IoU, AFI, Precision, Recall, F-score, OS and US were used to evaluate their segmentation capabilities. Moreover, the average of the metrics was derived because multiple objects were employed for the segment quality analysis. The computed metrics are presented in Table 2 for MRS and Table 3 for SAM algorithms.

Objects / Metrics	Parcel	Industry	Roof	Road	Water	Single Tree
IoU	0.89	0.94	0.49	0.72	0.78	0.46
AFI	0.02	0.01	0.03	0.08	0.06	0.10
Precision	0.91	0.95	0.50	0.77	0.82	0.49
Recall	0.98	0.99	0.98	0.92	0.94	0.91
F1-score	0.94	0.97	0.66	0.84	0.88	0.63
OS	0.90	0.85	0.38	0.48	0.21	0.58
US	0.00	0.01	0.23	0.08	0.06	0.10

Table 2. Quantitative assessment of segment quality derived from MRS.

Objects / Metrics	Parcel	Industry	Roof	Road	Water	Single Tree
IoU	0.87	0.60	0.70	0.67	0.63	0.46
AFI	0.08	0.01	0.19	0.26	0.30	0.33
Precision	0.94	0.61	0.82	0.87	0.86	0.59
Recall	0.92	0.99	0.82	0.74	0.70	0.68
F1-score	0.93	0.75	0.82	0.80	0.77	0.63
OS	0.05	0.04	0.18	0.10	0.11	0.08
US	0.08	0.33	0.09	0.26	0.20	0.23

Table 3. Quantitative assessment of segment quality derived from SAM.

The results show that the MRS algorithm performs efficiently in preserving the intricate details and delineation of entities. However, the SAM exhibits exceptional efficiency in capturing contextually similar regions. Furthermore, it can be inferred that MRS outperforms SAM with a minimal difference. The MRS algorithm yielded high IoU values for many object categories. The categories "Parcel" and "Industry" have notably high IoU values (specifically 0.89 and 0.94, respectively). This demonstrates that MRS effectively divided various types of objects into segments. On the other hand, the SAM algorithm achieves IoU values over 0.60 for certain object categories, although it exhibits notably lower values, particularly in the "Industry" and "Single Tree" categories.

In the context of the red roof object, the IoU value for SAM is 0.70 whereas for MRS it is estimated as 0.49, which indicates that SAM outperforms MRS. In addition, AFI values were comparatively lower in segments generated using MRS as opposed to segments generated using SAM, suggesting a higher level of proximity to the reference object. Higher precision and recall values suggest that the algorithm tends to accurately detect both actual positive and true objects for MRS and SAM. In other words, both algorithms often exhibited good precision, recall, and F1-score values. Nevertheless, MRS exhibits inferior precision and recall metrics in the "Roof" and "Water" categories.

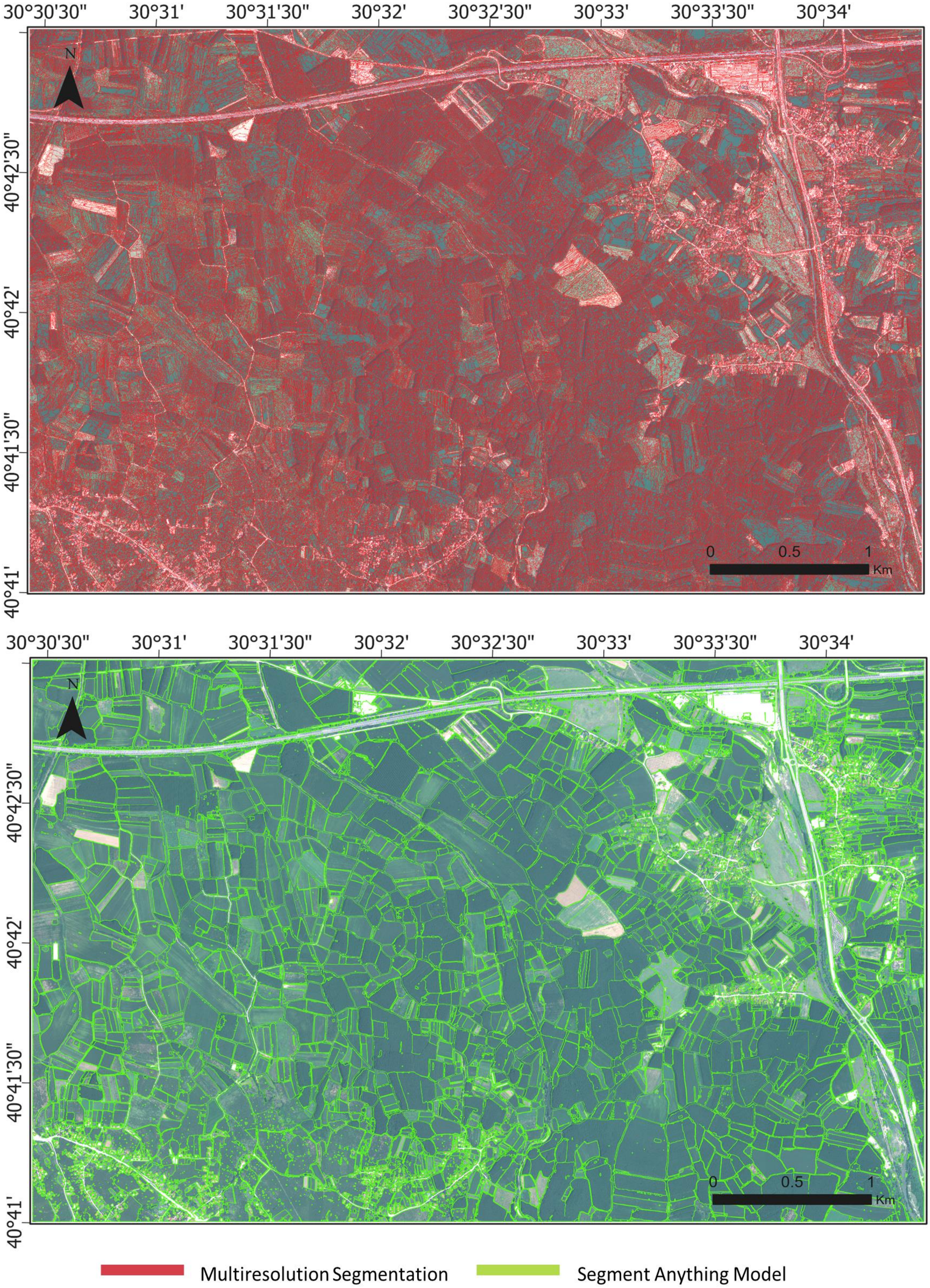

Figure 2. Segmentation results obtained with MRS (red line) and SAM (green line) algorithms.

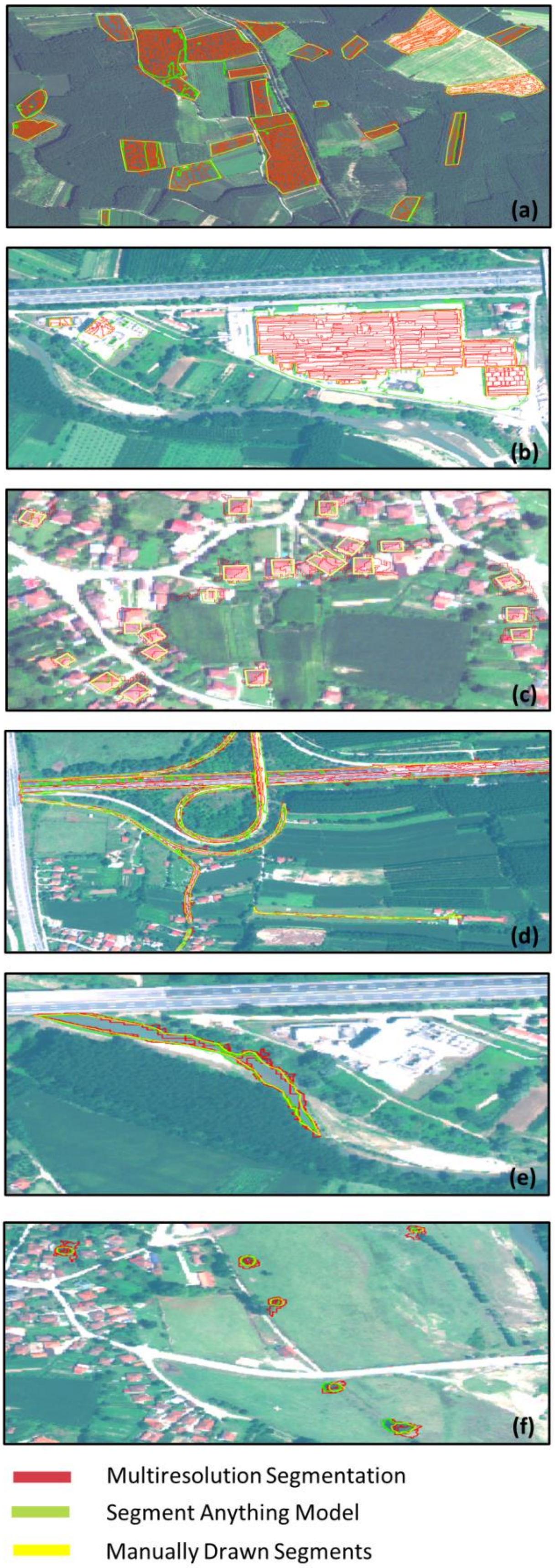

Figure 3. Examples of LULC objects: (a) parcels, (b) industrail buildings, (c) roof, (d) roads, (e) water bodies and (f) single tree, showing manually digitized object boundaries, and those of MRS and SAM algorithms.

The results of MRS showed low US values (lowest 0.00) and OS values (lowest 0.21). This suggests that MRS tended to break objects into excessively small parts and does not effectively evaluate objects. Conversely, the SAM method typically yielded low OS values (minimum of 0.04), but the US values exhibited greater variability (ranging from a minimum of 0.01 to a maximum of 0.33). From the produced results, one could claim that SAM tends to divide items into larger portions and, in parallel, struggles to accurately differentiate objects that belong to the industrial objects. As a result, an additional finding is that the MRS approach produces multiple segments for an object, whereas the SAM approach produces a single segment. However, it should be noted that both algorithms have certain limitations in certain situations, such as excessive segmentation in regions with high texture or insufficient segmentation in regions with subtle variations.

4. DISCUSSION & CONCLUSION

The study conducts a thorough analysis by utilising comprehensive segment quality criteria (IoU, AFI, Precision, Recall, F1-score, US, and OS measures) to objectively assess the performance of MRS and SAM algorithms when applied to WV-3 satellite imagery. MRS showcases its expertise in asset recognition by exhibiting an outstanding aptitude for preserving complex information. Conversely, SAM, an approach based on deep learning, effectively maintains total contextual consistency by precisely capturing contextual similarities. MRS outperformed SAM in specific areas, particularly in identifying objects, although SAM excels in attaining contextual coherence. Both algorithms have restrictions in specific situations such that MRS tends to generate many segments, while SAM has the capacity to ignore small distinctions. This study highlights the crucial significance of considering different measures of segment quality and emphasises the need for experts to carefully weigh the advantages and disadvantages between accuracy and contextual representation when using segmentation algorithms. The research results provide practical implications for continuing decision-making processes in LULC planning, environmental monitoring, and urban development. Additionally, they contribute to these fields by promoting a deeper understanding of the performance of segmentation algorithms in different geographical contexts, thereby laying the foundation for future advancements. Consequently, they enhance the efficacy of geospatial applications. To summarise, the study suggests that forthcoming investigations ought to explore hybrid methodologies that capitalise on the respective merits of SAM and MRS, thereby facilitating the development of more comprehensive and effective segmentation solutions. In addition, there are intentions to conduct a more comprehensive analysis of the results produced by the segmentation algorithms when the spatial resolution of the imagery is altered, whether it be reduced or increased. The purpose of this further inquiry is to enhance our comprehension of the performance of algorithms across different conditions and resolutions, thereby making a contribution to the ongoing advancement of segmentation techniques utilised in remote sensing applications.

ACKNOWLEDGEMENTS

The WorldView-3 satellite image utilised in this investigation was acquired through an initiative funded by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under the number 119O630. Our sincere appreciation goes out to the TUBITAK for their significant contributions to the conference paper.

REFERENCES

Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), 2-16.

Campbell, J. B., Wynne, R. H., 2011: *Introduction to Remote Sensing*. Guilford press.

Chen, X., Shen, X., Cao, L., 2023: Tree species classification in subtropical natural forests using high-resolution UAV RGB and Superview-1 multispectral imageries based on Deep Learning Network approaches: A case study within the Baima Snow Mountain National Nature Reserve, China. *Remote Sensing*, 15(10), 2697. doi.org/10.3390/rs15102697.

Colkesen, I., Kavzoglu, T., 2019: Comparative evaluation of decision-forest algorithms in object-based land use and land cover mapping. *In Spatial Modeling in GIS and R for Earth and Environmental Sciences*, 499-517. doi.org/10.1016/B978-0-12-815226-3.00023-5.

Colkesen, I., Kavzoglu, T., Atesoglu, A., Tonbul, H., Ozturk, M.Y., 2023: Multi-seasonal evaluation of hybrid poplar (P. Deltoides) plantations using Worldview-3 imagery and state-of-the-art ensemble learning algorithms. *Advances in Space Research*, 71(7), 3022-3044. doi.org/10.1016/j.asr.2022.10.044.

El-naggar, A.M., 2018: Determination of optimum segmentation parameter values for extracting building from remote sensing images. *Alexandria Engineering Journal*, 57(4), 3089–3097. doi.org/10.1016/j.aej.2018.10.001.

Emmanuel, H., Yu, J., Wang, L., Choi, S. H., Rwatangabo, D. E., 2023: Object-oriented remote sensing approaches for the detection of terrestrial impact craters as a reconnaissance survey. *Remote Sensing*, 15(15), 3807. doi.org/10.3390/rs15153807.

Ez-zahouani, B., Teodoro, A., El Kharki, O., Jianhua, L., Kotaridis, I., Yuan, X., Ma, L., 2023: Remote sensing imagery segmentation in object-based analysis: A review of methods, optimization, and quality evaluation over the past 20 Years. *Remote Sensing Applications: Society and Environment*, 32, 101031. doi.org/10.1016/j.rsase.2023.101031.

Fan, J., Li, J., Liu, Y., Zhang, F., 2024: Frequency-aware robust multidimensional information fusion framework for Remote Sensing Image Segmentation. *Engineering Applications of Artificial Intelligence*, 129, 107638, doi.org/10.1016/j.engappai.2023.107638.

Goodwin, M., Halvorsen, K.T., Jiao, L., Knausgård, K.M., Martin, A.H., Moyano, M., Oomen, R.A., Rasmussen, J.H., Sørdalen, T.K., Thorbjørnsen, S.H., 2022: Unlocking the potential of deep learning for marine ecology: Overview, applications, and outlook. *ICES Journal of Marine Science*, 79(2), 319–336. doi.org/10.1093/icesjms/fsab255.

Hossain, M. D., Chen, D., 2019: Segmentation for object-based image analysis (OBIA): A review of algorithms and challenges from Remote Sensing Perspective. *ISPRS Journal of Photogrammetry and Remote Sensing*, 150, 115–134. doi.org/10.1016/j.isprsjprs.2019.02.009.

Jozdani, S., Chen, D., 2020: On the versatility of popular and recently proposed supervised evaluation metrics for segmentation quality of remotely sensed images: An experimental case study of building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 160, 275–290. doi.org/10.1016/j.isprsjprs.2020.01.002.

Kavzoglu, T., 2008. Determination of environmental degradation due to urbanization and industrialization in Gebze, Turkey. *Environmental Engineering Science*, 25(3), 429-438. doi.org/10.1089/ees.2006.0271.

Kavzoglu, T., Yildiz Erdemir, M., Tonbul, H., 2016: A region-based multi-scale approach for object-based image analysis. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41, 241-247. doi.org/10.5194/isprs-archives-XLI-B7-241-2016

Kavzoglu, T., 2017. Object-oriented random forest for high resolution land cover mapping using Quickbird-2 imagery. *Handbook of Neural Computation*, 607–619. doi.org/10.1016/b978-0-12-811318-9.00033-8.

Kavzoglu, T., Erdemir, M.Y., Tonbul H., 2017: Classification of semiurban landscapes from very high-resolution satellite images using a regionalized multiscale segmentation approach, *Journal of Applied Remote Sensing*, 11(03), 035016. doi.org/10.1117/1.jrs.11.035016.

Kavzoglu, T., Tonbul, H., 2017: A comparative study of segmentation quality for multi-resolution segmentation and watershed transform, Proceedings of 8th International Conference on Recent Advances in Space Technologies, 113–117, Istanbul, Turkey, 19-22 August. doi.org/10.1109/RAST.2017.8002984.

Kavzoglu, T., Tonbul, H., 2018: An experimental comparison of multi-resolution segmentation, SLIC and K-Means clustering for object-based classification of VHR imagery. *International Journal of Remote Sensing*, 39, 6020–6036. doi.org/10.1080/01431161.2018.1506592.

Kavzoglu, T., Yildiz, M., 2014: Parameter-based performance analysis of object-based image analysis using aerial and Quikbird-2 Images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II–7, 31–37. doi.org/10.5194/isprsannals-ii-7-31-2014.

Kavzoğlu, T., Yılmaz, E.Ö., 2022: Analysis of patch and sample size effects for 2D-3D CNN models using multiplatform dataset: Hyperspectral Image Classification of ROSIS and Jilin-1 GP01 imagery. *Turkish Journal of Electrical Engineering and Computer Sciences*, 30(6), 2124–2144. doi.org/10.55730/1300-0632.3929

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo W., Dollár, P., Girshick, R., 2023: Segment anything. *arXiv preprint*, arXiv:2304.02643.

Lucchese, L., Mitra, S.K., 2001: Colour image segmentation: a state-of-the-art survey. *Proceedings-Indian National Science Academy Part A*, 67(2), 207-222. doi.org/10.1.1.84.4896.

Marpu, P.R., Neubert, M., Herold, H., Niemeyer, I. 2010: Enhanced evaluation of image segmentation results. *Journal of Spatial Science*, 55(1), 55–68. doi.org/10.1080/14498596.2010.487850.

Maxwell, A.E., Warner, T.A., Guillén, L.A., 2021a: Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—Part 1: Literature review, *Remote Sensing*, 13(13), 2450. doi.org/10.3390/rs13132450.

Maxwell, A.E., Warner, T.A., Guillén, L.A., 2021b: Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—part 2: Recommendations and best practices, *Remote Sensing*, 13(13), 2591. doi.org/10.3390/rs13132591.

Tonbul, H., Colkesen, I., Kavzoglu, T., 2020: Classification of poplar trees with object-based ensemble learning algorithms using Ssentinel-2A imagery. *Journal of Geodetic Science*, 10(1), 14–22. doi.org/10.1515/jogs-2020-0003.

Witharana, C., Civco, D.L., 2014: Optimizing multi-resolution segmentation scale using empirical methods: Exploring the sensitivity of the supervised discrepancy measure Euclidean distance 2 (ED2). ISPRS Journal of Photogrammetry and Remote Sensing, 87, 108–121. doi.org/10.1016/j.isprsjprs.2013.11.006.

Zhang, X., Feng, X., Xiao, P., He, G., Zhu, L., 2015: Segmentation quality evaluation using region-based precision and recall measures for remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 102, 73–84. doi.org/10.1016/j.isprsjprs.2015.01.009.