Bitlis Eren Üniversitesi Fen Bilimleri Dergisi

BİTLİS EREN UNIVERSITY JOURNAL OF SCIENCE
ISSN: 2147-3129/e-ISSN: 2147-3188
VOLUME: 13 NO: 3 PAGE: 681-691 YEAR: 2024
DOI:10.17798/bitlisfen.1473041

Evaluating the Effectiveness of Panoptic Segmentation Through Comparative Analysis

Cahide SARA[1*], İlhan DAŞDEMİR[2] , Sara ALTUN GÜVEN[1]

[1]*Tarsus University, Faculty of Engineering, Department of Computer Engineering, Tarsus/TURKEY*

[2] *Tarsus University, Faculty of Engineering, Department of Electrical Electronics Engineering, Tarsus/TURKEY*

(ORCID: 0009-0003-5432-3913) (ORCID: 0009-0004-4035-4425) (ORCID: 0000-0003-2877-7105)

Keywords: Image processing, Image segmentation, Panoptic segmentation, Instance segmentation.

Abstract
Image segmentation method is extensively used in the fields of computer vision, machine learning, and artificial intelligence. The task of segmentation is to distinguish objects in images either by their boundaries or as entire objects from the entire image. Image segmentation methods are implemented as instance, semantic, and panoptic segmentation. In this article, the panoptic segmentation method, seen as an advanced stage of instance and semantic segmentation, has been applied to three datasets and compared with the instance segmentation method. Experimental results are presented visually. Numerical results have been analyzed with the Panoptic Quality (PQ) and Semantic Quality (SQ) metrics. It has been observed that the segmentation outcome was best for the CityScapes dataset for panoptic segmentation.

1. Introduction

Computer Vision [1] is a subfield of computer science that enables computers to see and recognize objects or entities similarly to humans. In other words, it possesses the ability to perceive and interpret objects. Image segmentation techniques are frequently utilized in image and video analysis. These segmentation techniques are applied in various fields such as earth sciences [2],[3], smart cities [4]-[6], and healthcare services [7]-[9], and their applications are becoming increasingly widespread. There are three main approaches commonly used in image segmentation: semantic segmentation, instance segmentation, and panoptic segmentation.

Semantic segmentation is a computer vision task that divides an image into multiple segments corresponding to objects or areas of interest and assigns a class label to each pixel [10]. It is used in various fields, from self-driving cars to analyzing medical imagery, this capability enables the comprehension of image contents down to each individual pixel [11]. It is typically implemented with deep learning models, such as Convolutional Neural Networks (CNN) [12], and requires a large labeled dataset for training. Popular architectures include Fully Convolutional Networks [13], U-Net [14], and DeepLab [15], [16].

Instance Segmentation is a complex image segmentation task aimed at accurately identifying and delineating different objects within an image, making a distinction between multiple instances of the same object class. Unlike semantic segmentation, which groups pixels into categories, it enables the detailed pixel-level understanding of each distinct object instance. This means that each object instance is not only classified but also precisely outlined. Due to the necessity of distinguishing between overlapping or closely located objects of the same class, instance segmentation is often considered a more challenging

*Corresponding author: *cahidesr.04@gmail.com*

Received: 24.04.2024, Accepted: 23.06.2024

task than semantic segmentation. It finds applications in various fields such as autonomous driving, robotics, medical imaging, and object detection in complex scenes. Notable instance segmentation models include Mask R-CNN, Panoptic FPN, and YOLACT [17]-[19].

Panoptic Segmentation is an image segmentation task aimed at comprehensively understanding the image by assigning a category label to each pixel and facilitating the distinction between individual instances of objects. This task merges the benefits of semantic and instance segmentation, identifying the semantic category of every pixel in the image while also distinguishing between different instances of objects [20]. In panoptic segmentation, it is determined to which semantic category each pixel belongs. Objects belonging to the "object" category, such as people or cars, are identified individually, while areas belonging to non-object categories, such as the sky or road, are semantically labeled without specifying individual instances [20]. The goal of this task is to understand the entire scene and provide information at both the object and category levels. Panoptic segmentation is utilized in various fields including autonomous driving, robotics, scene understanding, and image editing [21].

Panoptic Segmentation models typically utilize a combination of CNNs [22] and specialized architectures that concurrently perform both Semantic Segmentation and Instance Segmentation tasks [23]. Leading Panoptic Segmentation frameworks include Panoptic FPN, UPSNet, and SOLO [24].

Many of the suggested panoptic segmentation techniques are applied to RGB images. For instance, in the study [25], A panoptic segmentation model called Panoptic-Fusion is introduced, serving as an online volumetric semantic mapping system designed to identify the class labels of background areas (items) and the individual segments of desired objects (things). For incoming RGB frames, it first predicts pixel-wise panoptic labels by combining traditional semantic and instance segmentation outputs. Similarly, in the study [26], Faraz et al. aimed to enhance the ability of networks to determine depth on a pixel-by-pixel basis from single RGB input images.. Various additional panoptic techniques have also been developed for the segmentation of RGB images[27]-[31]. For image segmentation using a panoptic approach, numerous frameworks have been initially suggested that employ both instance and semantic segmentation, subsequently merging the outcomes of each segment to produce the final panoptic segmentation results. To achieve this, Kirillov et al. [27] conduct instance and semantic segmentation independently before combining them.

Due to some limitations of instance segmentation, a panoptic architecture like Cell R-CNN has been proposed. In this case, the encoder part of the instance segmentation model is often used in learning global semantic features, in conjunction with a jointly trained semantic segmentation model [32]. Many other panoptic segmentation tools developed for segmenting medical images are widely used. For example, areas where these tools are frequently utilized include pathology image analysis [33], detection of prostate cancer [34], and segmentation of teeth in panoramic X-ray images [35].

The study [36] encompasses labeled LiDAR scans in various environments and car scenes. In [37], two fundamental approaches are employed for panoptic segmentation, combining semantic segmentation and 3D object detectors. Similarly, in [38], "The PointPillars object detection system is employed to determine bounding boxes and classify each object, while instance segmentation is executed for each category. The two principal networks undergo separate training and testing before their outcomes are merged in the final phase to achieve panoptic segmentation. An undisclosed test set is then utilized for the online assessment of LiDAR-driven panoptic segmentation.

In this article, the Mask R-CNN [17] method is used for instance segmentation. Mask R-CNN (Mask Region-Based Convolutional Neural Networks) is primarily based on Faster R-CNN, but it adds a segmentation mask for each detected object. It consists of three main components: a region proposal network (RPN), classification and bounding box regression, and segmentation mask prediction. Panoptic segmentation, on the other hand, uses the Panoptic-DeepLabmethod, which is based on the DeepLabv3+ model. It combines semantic segmentation and instance segmentation.

In this article, we focus on panoptic segmentation among image segmentation techniques (instance, semantic and panoptic segmentation), which are capable of distinguishing both the object and its boundaries within the image. The results of panoptic segmentation are visually compared and interpreted with those of instance segmentation. For image analyses, datasets such as COCO [39], CityScape [40], and an Industrial dataset have been utilized. Metric analysis employed includes Panoptic Quality (PQ) [27] and Semantic Quality (SQ) [27], which are used in panoptic segmentation methods. The results have been presented both in tabular and visual form.

The rest of the article is organized as follows. Section 2 covers the Materials and Methods. This section discusses the datasets used, the segmentation methods employed, and the implementation settings. Section 3 contains the experimental results and discussion. The final section includes the conclusions of the article.

2. Material and Method

In this study, computer vision and image segmentation methods will be examined, with a focus on the most common image segmentation algorithms, particularly emphasizing the panoptic segmentation method. The processing of RGB images outputs as data types in panoptic segmentation will be explained. Results will be evaluated by processing three popular datasets with the panoptic segmentation algorithm.

2.1. Datasets

The datasets used for the panoptic segmentation task include the COCO dataset, the CityScapes dataset, and an Industrial dataset created in the structure of the COCO dataset with images captured in industrial settings.

The COCO dataset [39] is a large-scale data set for object detection, where objects are labeled according to their types. Based on this dataset, objects are categorized by type, and each category is assigned a color code. Accordingly, objects in the image are detected, and their category is determined based on the label. The primary uses of the COCO dataset include object detection, human detection in security systems, and background removal in object detection, among many others. Being an extensive dataset, it contains images of all kinds. Our preference, however, is for dense images where objects are more intertwined.

The CityScapes dataset [40] is another large-scale dataset created for the detection and labeling of objects. Again, objects are categorized, and these categories are assigned color codes. This dataset consists of images taken in various cities at different times of the year, in different seasons, and at different times of the day. Our preferred images are those from different cities that contain denser objects.

The Industrial dataset is organized following the logic of the COCO dataset, encompassing sequential images of environments populated with industrial objects. The objectives for utilizing the COCO dataset are similarly relevant here. The images selected for our study are prioritized for their clarity, as consecutive shots might compromise image quality, and those that feature a higher density of objects. Figure 1 in the manuscript showcases actual images alongside their corresponding panoptic masks for all discussed datasets.

Figure 1. Examples from the COCO, CityScapes, and Industrial datasets (1st Column: original images; 2nd Column: panoptic segmentation masks)

2.2. Methods

In this section, panoptic segmentation and instance segmentation are discussed in detail.

2.2.1. Panoptic segmentation

Panoptic segmentation is a machine learning method developed to understand and interpret images comprehensively. Beyond simple image segmentation, it determines which object each pixel belongs to, enabling a holistic understanding of objects and scenes. This method aims to classify and segment all objects in images both individually (instance segmentation) and as part of the background (semantic segmentation), determining not only to which object each pixel belongs but also what that object is. By merging object detection and semantic segmentation, this approach allows for richer and more detailed image analyses. Figure 2 presents various network structures used in the creation of panoptic segmentation.

In panoptic segmentation, there are two categories: items and objects. Items refers to innumerable areas such as the sky, sidewalks, and floors, while objects encompass all countable entities, including trees, cars and peoples. Items and objects are differentiated in the panoptic approach by assigning them distinct colors, which distinguishes them from one another, unlike object segmentation and semantic approaches where overlap between objects of the same type might occur. Panoptic segmentation provides a comprehensive view by recognizing each object and background separately, encompassing all these components. This method identifies and classifies each object individually while also considering the relationship of these objects with the background. By enabling a detailed understanding of objects and areas in images, it allows for more complex image analysis and interpretations. Moreover, panoptic segmentation enables effective visualization of diverse components within a scene and can be described as a comprehensive approach that encompasses detection, localization, and classification of different scene elements. This outcome contributes to a clear and functional understanding of the scene.

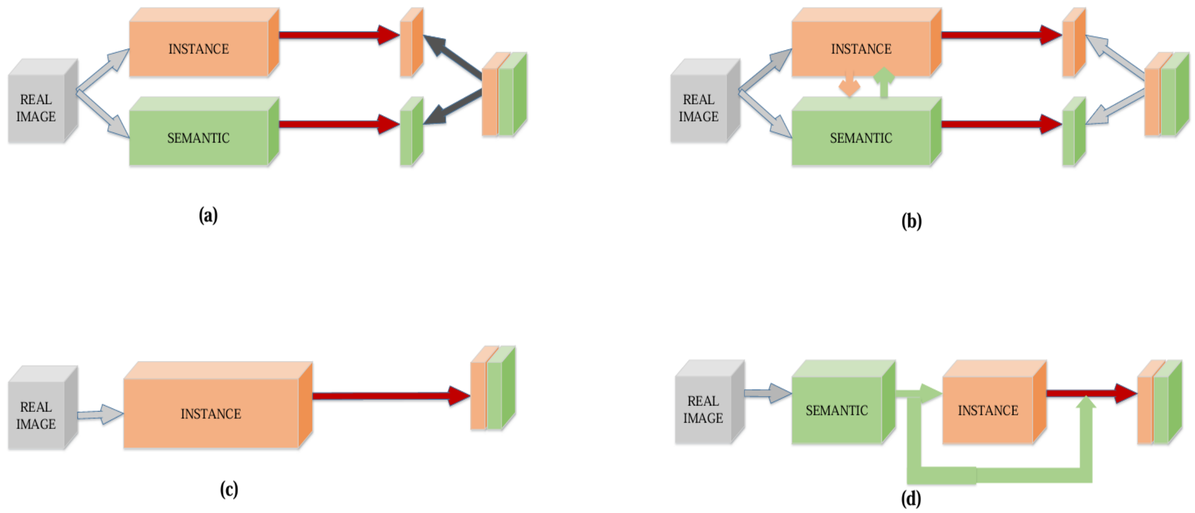

Figure 2. Network architectures for panoptic segmentation methods (a) Sharing image (b) Explicit coonection (c) One-shot Model (d) Cascade model

Panoptic segmentation, while merging instance and semantic segmentation, primarily uses architectures based on convolutional neural networks (CNN) [41]. This combination is shown in Equation 1.

$$P(x) = Combine(S(x), I(x)) \tag{1}$$

Here, $P(x)$ represents the panoptic segmentation output, $S(x)$ represents the semantic segmentation output, and $I(x)$ represents the instance segmentation output. Panoptic segmentation combines these two segmentations to produce the final output.

These networks are designed to perform both object detection and pixel-based segmentation tasks. In the first phase, the classes and locations of objects are detected. In the second phase, semantic segmentation is applied to classify each pixel. During model training, special loss functions are usually used for both object detection and semantic segmentation. Model parameters are optimized using stochastic gradient descent (SGD) or similar algorithms.

Panoptic segmentation has a wide range of applications, such as detailed identification and understanding of vehicles, pedestrians, and other significant elements on roads. By identifying structures, roads, and other urban elements in cities, it provides valuable insights for urban planning and management processes. In distinguishing diseased tissues from normal ones, it contributes to diagnosis and treatment processes by enabling more detailed analysis of medical images. It is used in identifying and analyzing natural elements like vegetation and water resources, as well as man-made structures in satellite imagery.

2.1.2. Instance Segmentation

Instance segmentation is an advanced image analysis method developed through the combination of deep learning and computer vision techniques. It identifies the boundaries of each object in images at the pixel level, while also classifying objects into different categories. This method allows for detailed analysis on images by identifying each object individually and delineating their boundaries. This type of segmentation plays a crucial role in understanding the count, locations, shapes, and sizes of objects. In short, instance segmentation enables the determination of what the objects are (their class) and where they are (their location and shape at the pixel level).

Each pixel in the image is classified as belonging to a specific object. Even if they belong to the same class, different objects are identified separately, and a unique mask is generated for each. This method can precisely define the boundaries and shapes of objects, allowing even the subtle differences between them to be detected. Training the model requires a rich dataset labeled with the classes of objects and pixel-wise masks for each object.

Figure 3 features a block diagram of a one-stage instance segmentation process. It illustrates how an image can yield masks, classes, and bounding boxes outputs through instance segmentation.

Instance segmentation fundamentally employs CNN-based architectures. Architectures such as Mask R-CNN, YOLO, and Faster R-CNN are popular choices for instance segmentation. Among the architectures commonly used for object segmentation are Mask R-CNN [42], R-CNN [43, 44], Path Aggregation Network (PANet) [45], and YOLACT [46,47]. Mask R-CNN [42] is a method specifically developed for instance segmentation. It is based on Faster R-CNN and provides both classification and pixel-level mask output for each object. Different loss functions are used for object classification, localization (bounding box), and mask generation. To maximize the model's performance, optimization algorithms like Stochastic Gradient Descent (SGD) or similar are preferred. This method offers valuable insights in various application areas, especially where the subtle differences between objects are significant. Instance segmentation is gaining increasing interest thanks to advancements in machine learning and artificial intelligence technologies, becoming a valuable tool across various industries.

2.3. Image Quality Metrics for Panoptic and Instance Segmentations

Panoptic segmentation merges object-based segmentation tasks (e.g., pedestrian, car) with scene segmentation tasks (e.g., road, sky). In this context, Panoptic Quality (PQ) [27] and Segmentation Quality (SQ) [27] are the most commonly used metrics.

Panoptic Quality (PQ) [27] evaluates both classification accuracy and segmentation quality simultaneously. PQ measures the agreement between predicted and true segments for each object, offering a single metric that combines both classification and segmentation performance. The formula for PQ is given in Equation 2. Segmentation Quality (SQ) [27] assesses only the accuracy of the segmentation, measuring how well the shapes and contours of the predicted segments match the true segments. SQ is typically calculated using the Intersection over Union (IoU) metric and is expressed as the average IoU for individual segments. The formula for SQ is given in Equation 3.

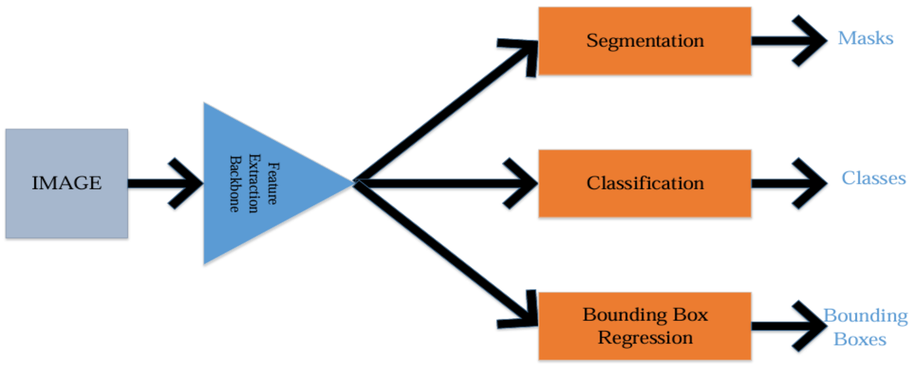

Figure 3: One-Stage instance segmentation process

PQ and SQ metrics are crucial for evaluating panoptic segmentation models. PQ comprehensively assesses a model's ability to correctly classify objects and accurately segment them, while SQ focuses on the quality of the segmentation. Understanding and utilizing these metrics play a significant role in the development and evaluation of panoptic segmentation models, leading to advancements in various application areas such as object recognition and scene understanding.

$$PQ = \frac{\sum_{(p,g)\in TP} IoU(p,g)}{|TP| + 1/2|FP| + 1/2|FN|} \qquad (2)$$

$$SQ = \frac{1}{|TP|}\sum_{(p,g)\in TP} IoU(p,g) \tag{3}$$

For the PQ equation; TP (True Positive) represents the number of correctly matched segment pairs, FP (False Positive) indicates the number of false positive predictions, and FN (False Negative) signifies the number of false negative predictions. $IoU(p,g)$ (Intersection over Union) expresses the ratio of the intersection to the union between the predicted segment p and the true segment g. For the SQ equation, TP again represents the correctly matched segment pairs, and $IoU(p,g)$ measures the ratio of the intersection to the union between the predicted and true segments.

Panoptic Quality (PQ) has been proposed for panoptic segmentation, but it is also used in instance segmentation. This metric measures both segmentation accuracy and detection accuracy. Segmentation Quality (SQ), on the other hand, measures the quality of segmentation masks. They are used as comparison metrics in instance segmentation methods.

3. Results and Discussion

This article aims to understand and interpret the results of instance segmentation and panoptic segmentation. Analyses were conducted on the COCO, CityScapes, and industrial datasets, with results presented visually. The analyses were based on the PQ and SQ metrics used for segmentation, from which conclusions were drawn and interpreted.

The figures show how the real images appear after instance segmentation and depict the boundaries of objects in images resulting from panoptic segmentation. Figure 4 presents the visual results of instance segmentation and panoptic segmentation for the COCO dataset.

Upon examining Figure 4, it's observed that instance segmentation focuses on objects, with the background not being given much importance. Objects are categorized into specific groups and labeled accordingly. In contrast, when looking at the panoptic segmentation results, both objects and the background are distinctly separated, with object labels being more clearly defined.

Figure 5 presents the visual outcomes of instance and panoptic segmentation for the CityScapes dataset.

Figure 4. Instance segmentation and panoptic segmentation results for the COCO dataset

When examining the instance segmentation results in Figure 5, it's noted that the objects are smaller and overlaid, with only the objects themselves being displayed. However, the types of objects are not clearly discernible. In the results of the panoptic segmentation, the entire environment appears to be color-coded with labels, illustrating a comprehensive labeling of both objects and the background.

Figure 6 features the visual outcomes of instance segmentation and panoptic segmentation for the Industrial dataset.

Figure 5. Instance segmentation and panoptic segmentation results for the CityScapes dataset

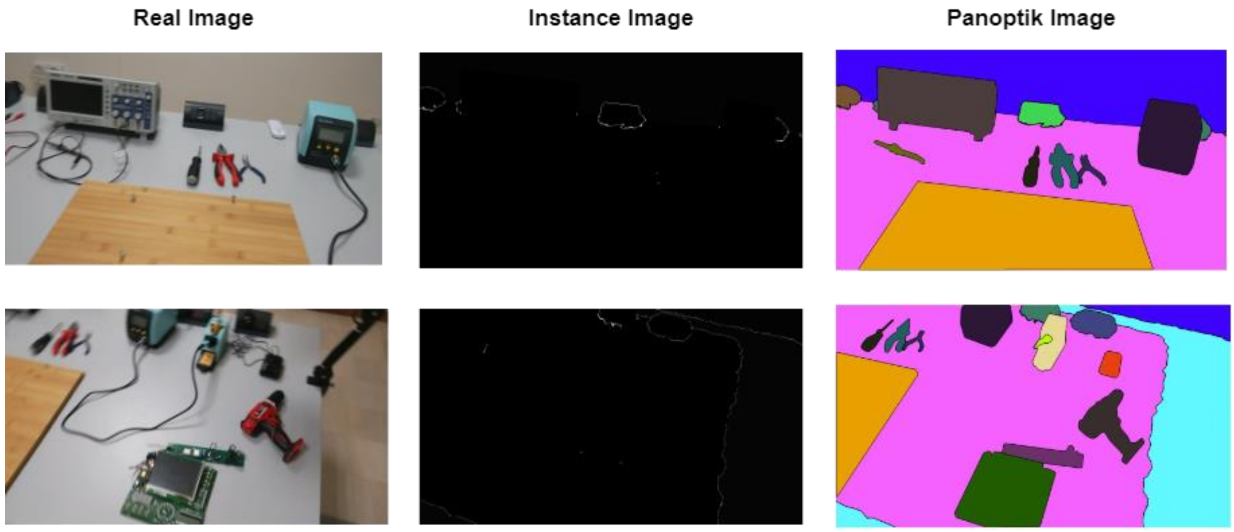

Figure 6. Instance segmentation and panoptic segmentation results for the Industrial dataset

In the real images depicted in Figure 6, due to the labels of objects not being distinctly clear, the outcome of instance segmentation shows that objects in the image are perceived similarly to the background and are not prominently displayed. However, in the results of panoptic segmentation, objects are displayed more distinctly.

Table 1 provides the metric results of PQ and SQ according to the datasets.

Table 1. PQ and SQ metric results for datasets used in panoptic segmentation

Datasets	**PQ**	**SQ**
COCO [39]	67.4	68.3
CityScapes [40]	**69.7**	**84.2**
Industrial	67.3	68.2

In Table 1, the PQ metric measures the quality of panoptic segmentation. It accounts for both correctly and incorrectly measured pixels, functioning by considering their ratio. The closer the ratio is to one hundred percent, the more accurate the results are. SQ, on the other hand, measures the quality of segmentation. Similarly, the closer the ratio is to one hundred percent, the more accurate the segmentation is deemed to be.

In the measurement of the PQ metric, a value of 69.7 and in the SQ metric, a value of 84.2 were observed, indicating that the Cityscape dataset showed the best performance for panoptic segmentation. The reason for the good results of the Cityscape dataset in panoptic segmentation is its ability to clearly distinguish objects and the background compared to other datasets. For example, when examining the panoptic segmentation visual results of the COCO dataset, it is understood that objects are not fully distinguished. It can be observed that there might be confusion in distinguishing between a horse and a human. In tabular results, considering these reasons, the COCO dataset was observed to be less successful than the Cityscape dataset. The industrial dataset showed similar results to the COCO dataset.

Upon examining the values in the table, we observe that the CityScapes dataset provides better results in terms of ratios compared to the other datasets.

Table 2 provides the metric results of PQ and SQ according to the datasets for instance segmentation.

Table 2. PQ and SQ metric results for datasets used in instance segmentation

Datasets	PQ	SQ
COCO [39]	**0.0009**	**0.1764**
CityScapes [40]	0.0002	0.0317
Industrial	0.0008	0.0476

Table 2 presents the instance segmentation results according to the PQ and SQ metrics. With a value of 0.0009 for PQ and 0.1764 for SQ, the COCO dataset has shown better results. Since instance segmentation is a technique that individually identifies and masks each object instance in an image, the COCO dataset has been observed to be more successful in instance segmentation compared to other datasets.

In essence, the article illustrates the distinct outcomes of instance and panoptic segmentation processes on various datasets, using PQ and SQ metrics to evaluate the effectiveness of these segmentation methods. The visual presentation of segmentation results allows for a direct comparison between the original images and their segmented counterparts, highlighting the precision and capabilities of both instance and panoptic segmentation techniques in delineating object boundaries and classifying scene elements. This comparative analysis provides valuable insights into the strengths and limitations of each segmentation approach, contributing to the ongoing development and refinement of computer vision and image analysis technologies.

4. Conclusion and Suggestions

In conclusion, the analysis delineates a clear advantage of panoptic segmentation over instance segmentation in achieving higher precision and distinctiveness in object delineation and scene element classification. The deployment of PQ and SQ metrics furnishes a comprehensive method for evaluating the segmentation quality, where the CityScapes dataset emerges as a benchmark for superior segmentation performance. This distinction underscores the capability of panoptic segmentation to more accurately discern and categorize objects from the background, enhancing the overall clarity and effectiveness of image analysis. The juxtaposition of visual results from instance and panoptic segmentation not only underscores the enhanced accuracy and detail provided by panoptic segmentation but also illustrates its critical role in advancing computer vision technologies. Panoptic segmentation has been observed to work in conjunction with the ability to understand relationships between objects. This is an important feature for understanding how objects interact together, such as how a car progresses on a road or how a person holds an object. Instance segmentation has been observed to be able to identify and classify each object separately in images, ensuring that different objects are fully separated from each other. When compared, it is observed that instance segmentation exhibits a deficiency in distinguishing between objects compared to panoptic segmentation. Through meticulous comparison and evaluation, this study contributes significantly to our understanding of segmentation technologies, offering insights that propel the field towards more nuanced and effective image analysis methods. The findings advocate for further exploration and refinement of panoptic segmentation, promising improvements in various applications from autonomous driving to urban planning and beyond, thus marking a pivotal step in the evolution of image analysis and computer vision capabilities.

Contributions of the authors

Conception: Cahide SARA
Design:Cahide SARA
Supervision: Sara ALTUN GÜVEN
Materials: Cahide SARA
Literature review: İlhan DAŞDEMİR
Writer: Sara ALTUN GÜVEN
Critical review: İlhan DAŞDEMİR
Other: Sara ALTUN GÜVEN

Conflict of Interest Statement

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

There were no ethical objections to the publication

References

[1] K. Ikeuchi, Computer Vision: *A Reference Guide. Cham, Switzerland:* Springer International Publishing, 2021.

[2] T. Hoeser and C. Kuenzer, “Object detection and image segmentation with deep learning on Earth observation data: A review-part I: Evolution and recent trends,” *Remote Sens. (Basel)*, vol. 12, no. 10, p. 1667, 2020.

[3] D. Galea, H.-Y. Ma, W.-Y. Wu, and D. Kobayashi, “Deep learning image segmentation for atmospheric rivers,” *Artificial Intelligence for the Earth Systems*, 2023.

[4] X. Chen et al., “Efficient Decoder and Intermediate Domain for Semantic Segmentation in Adverse Conditions,” *Smart Cities*, vol. 7, no. 1, pp. 254–276, 2024.

[5] J. Yuan, Z. Shi, and S. Chen, “Feature Fusion in Deep-Learning Semantic Image Segmentation: A Survey,” *in International Summit Smart City 360°, Cham:* Springer International Publishing, 2021, pp. 284–292.

[6] P. Garg, A. S. Chakravarthy, M. Mandal, P. Narang, V. Chamola, and M. Guizani, “ISDNet: AI-enabled Instance Segmentation of aerial scenes for smart cities*,” ACM Trans. Internet Technol*., vol. 21, no. 3, pp. 1–18, 2021.

[7] S. A. Güven and M. F. Talu, “Brain MRI high resolution image creation and segmentation with the new GAN method,” *Biomedical Signal Processing and Control*, vol. 80, 2023.

[8] Y. Xu, S. Hou, X. Wang, D. Li, ve L. Lu, "A medical image segmentation method based on improved UNet 3+ network," *Diagnostics*, vol. 13, no. 3, 576, 2023.

[9] K. Huang, Y. Zhang, H.-D. Cheng, and P. Xing, “Trustworthy breast ultrasound image semantic segmentation based on fuzzy uncertainty reduction,” *Healthcare (Basel)*, vol. 10, no. 12, p. 2480, 2022.

[10] B. Li, Y. Shi, Z. Qi, and Z. Chen, "A survey on semantic segmentation," *in 2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 1233-1240, Nov. 2018.

[11] C. Kaymak and A. Ucar, “Semantic image segmentation for autonomous driving using fully convolutional networks,” *in 2019 International Artificial Intelligence and Data Processing Symposium (IDAP)*, 2019.

[12] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, no. 4, pp. 193-202, 1980.

[13] Y. Lecun, D. Touresky, G. Hinton, and T. Sejnowski, “A theoretical framework for back-propagation,” *Proceedings of the 1988 connectionist models summer school*, vol. 1, pp. 21–28, 1988.

[14] O. Ronneberger, P. Fischer, ve T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *in Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference*, Munich, Germany, October 5-9, 2015, Proceedings, Part III, vol. 18, pp. 234-241, Springer International Publishing, 2015.

[15] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, ve A. L. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected crfs*," arXiv preprint arXiv:1412.7062*, 2014.

[16] İ. Kayadibi, U. Köse ve G. E. Güraksın, "Görüntü İşleme Teknikleri ve Evrişimsel Sinir Ağı Kullanılarak Bilgisayar Destekli Diş Segmentasyonu*," Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, cilt 1000, no. 1000, ss. 0-0.

[17] W. He, X. Wang, L. Wang, Y. Huang, Z. Yang, X. Yao, ... Z. Ge, "Incremental learning for exudate and hemorrhage segmentation on fundus images," *in Information Fusion*, vol. 73, pp. 157-164, 2021.

[18] Y. Zhang, X. Sun, J. Dong, C. Chen, ve Q. Lv, "GPNet: gated pyramid network for semantic segmentation," *in Pattern Recognition*, vol. 115, 107940, 2021.

[19] Q. Sun, Z. Zhang, and P. Li, "Second-order encoding networks for semantic segmentation," *Neurocomputing*, 2021.

[20] W. Mao, J. Zhang, K. Yang, ve R. Stiefelhagen, "Panoptic lintention network: Towards efficient navigational perception for the visually impaired," *2021 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, s. 857-862, July 2021.

[21] Liu, D., Zhang, D., Song, Y., Zhang, F., O'Donnell, L., Huang, H., ... & Cai, W. "Pdam: A panoptic-level feature alignment framework for unsupervised domain adaptive instance segmentation in microscopy images," *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, pp.154-165, 2020.

[22] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," *in 2017 International Conference on Engineering and Technology (ICET)*, pp. 1-6, Aug. 2017.

[23] J. Huang, D. Guan, A. Xiao, ve S. Lu, "Cross-view regularization for domain adaptive panoptic segmentation," *in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10133-10144, 2021.

[24] Y. Chen, W. Chu, F. Wang, Y. Tai, R. Yi, Z. Gan, ... X. Li, "CFNet: Learning correlation functions for one-stage panoptic segmentation," *arXiv preprint arXiv:2201.04796*, 2022.

[25] G. Narita, T. Seno, T. Ishikawa, ve Y. Kaji, "PanopticFusion: Online volumetric semantic mapping at the level of stuff and things," *in 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4205-4212, 2019.

[26] F. Saeedan and S. Roth, "Boosting monocular depth with panoptic segmentation maps," *in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 3853–3862, 2021.

[27] A. Kirillov, K. He, R. Girshick, C. Rother, ve P. Dollár, "Panoptic segmentation," *in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9404-9413, 2019.

[28] S. Liu, L. Qi, H. Qin, J. Shi, ve J. Jia, "An End-to-End Network for Panoptic Segmentation," *in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6172-6181, 2019.

[29] A. Nivaggioli, J.-F. Hullo, ve G. Thibault, "Using 3D models to generate labels for panoptic segmentation of industrial scenes," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, pp. 61-68, 2019.

[30] W. Mao, J. Zhang, K. Yang, ve R. Stiefelhagen, "Can we cover navigational perception needs of the visually impaired by panoptic segmentation?," *arXiv preprint arXiv:2007.10202*, 2020.

[31] L. Shao, Y. Tian, ve J. Bohg, "ClusterNet: 3D instance segmentation in RGB-D images," *arXiv preprint arXiv:1807.08894*, 2018.

[32] D. Liu, D. Zhang, Y. Song, H. Huang, ve W. Cai, "Cell R-CNN v3: A novel panoptic paradigm for instance segmentation in biomedical images," *arXiv preprint arXiv:2002.06345*, 2020.

[33] D. Zhang, Y. Song, D. Liu, H. Jia, S. Liu, Y. Xia, ... W. Cai, "Panoptic segmentation with an end-to-end Cell R-CNN for pathology image analysis," *in Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain*, September 16-20, 2018, Proceedings, Part II, vol. 11, pp. 237-244, Springer International Publishing, 2018.

[34] X. Yu, B. Lou, D. Zhang, D. Winkel, N. Arrahmane, M. Diallo, ... A. Kamen, "Deep attentive panoptic model for prostate cancer detection using biparametric MRI scans," *in Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru*, October 4–8, 2020, Proceedings, Part IV, vol. 23, pp. 594-604, Springer International Publishing, 2020.

[35] G. Jader, J. Fontineli, M. Ruiz, K. Abdalla, M. Pithon, ve L. Oliveira, "Deep instance segmentation of teeth in panoramic X-ray images," *in 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pp. 400-407, October 2018.

[36] J. Behley, M. Garbade, A. Milioto, J. Quenzel, S. Behnke, C. Stachniss, ve J. Gall, "SemanticKITTI: A dataset for semantic scene understanding of lidar sequences," *in Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9297-9307, 2019.

[37] J. Behley, A. Milioto, ve C. Stachniss, "A benchmark for LiDAR-based panoptic segmentation based on KITTI," *arXiv preprint arXiv:2003.02371*, 2020.

[38] A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, ve O. Beijbom, "PointPillars: Fast encoders for object detection from point clouds," *in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, s. 12697-12705, 2019.

[39] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, ... C. L. Zitnick, "Microsoft COCO: Common Objects in Context," *in Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V*, vol. 13, pp. 740-755, Springer International Publishing, 2014.

[40] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, ... B. Schiele, "The CityScapes dataset for semantic urban scene understanding," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3213-3223, 2016.

[41] J. Huang, D. Guan, A. Xiao, ve S. Lu, "Cross-view regularization for domain adaptive panoptic segmentation," *arXiv preprint arXiv:2103.02584*, 2021.

[42] X. Liu, D. Zhao, W. Jia, W. Ji, C. Ruan, Y. Sun, "Cucumber fruits detection in greenhouses based on instance segmentation," *IEEE Access*, vol. 7, pp. 139635-139642, 2019.

[43] M.-C. Roh ve J.-y. Lee, "Refining faster-rcnn for accurate object detection," *in 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA)*, IEEE, pp. 514-517, 2017.

[44] Y. Ren, C. Zhu, S. Xiao, "Object detection based on fast/faster rcnn employing fully convolutional architectures," *Mathematical Problems in Engineering*, vol. 2018, 2018.

[45] S. Liu, L. Qi, H. Qin, J. Shi, ve J. Jia, "Path aggregation network for instance segmentation," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8759-8768, 2018.

[46] D. Bolya, C. Zhou, F. Xiao, ve Y. J. Lee, "YOLACT: Real-time instance segmentation," *in Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9157-9166, 2019.

[47] D. Bolya, C. Zhou, F. Xiao, ve Y. J. Lee, "YOLACT++: Better real-time instance segmentation," *in Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019.