Semantic Segmentation based Approach for Waste Object Segmentation - An Analysis

R. Nivedhana
Phd Researcher Department of Computer Science Avinshilingam Institute for home Science and Higher Education for Women Coimbatore, Tamilnadu, India

Dr. V. Radha
Professor and Dean Department of Computer Science, School of Physical Sciences And Computational Sciences Avinshilingam Institute for Home Science and Higher Education for Women Coimbatore, Tamilnadu, India

Abstract—The planet Earth has been desecrated by humans because of trashing. Artificial intelligence has to play an important role in sorting out this problem with computer vision. Recent advancements in computer vision and deep learning algorithms will provide a better way to Waste object segmentation is required to find and separate the waste objects from the image data that was acquired. The evaluation of a State of Art model is done with trash object detection and segmentation. Sample Taco dataset images are analyzed and got better results. Trash Image data taken from the mobile phone with portrait mode acts as a salient image data and shows excellent results in semantic binary segmentation

Keywords—Trash detection, Waste detection, Semantic Segmentation, U2Net, Deep Learning

I. Introduction

As more processing capacity (GPUs) became accessible for training and research, the application of deep learning increased. We are now able to explore difficult problems with regards to significant environmental and ecological implications. Detection and classification of waste is one such area of concern. Trash that can be identified can be recycled effectively, which will ultimately benefit the environment and combat climate change. Domestic trash is not separated at the point of generation. An enormous amount of waste is deposited in landfills and mixed with other forms of waste, resulting in environmental contamination. All rubbish is gathered in one waste bin and stored in a neighboring public waste bin. The cost of the segregation procedure used by municipal solid waste management will decrease with source segregation.

II. Deeplearning segmentation techniques

Image segmentation is the process of identifying the regions in an image based on the collection of similar pixels together. Basic segmentation techniques are of two types, Manual Segmentation and Automatic Segmentation. The evolution of deep learning algorithms made easy processing of automated image segmentation. Instead of processing the entire image, the informative regions are segmented and proceeded towards further processing will give excellent results. The deep segmentation techniques are Semantic segmentation, Instance Segmentation and Panoptic Segmentation.

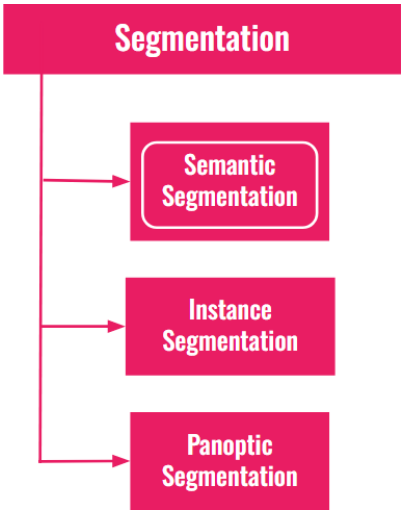

Fig. 1.Types of Deep Segmentation Techniques

DOI: https://doi.org/10.5281/zenodo.15428792

A. Semantic segmentation

The process of semantic segmentation entails categorizing each pixel in a given image. The image's objects are arranged into predetermined groupings or categories. Here for example, Trash/waste is considered as a group of similar pixels and the remaining is considered as background pixels. In this evaluation, semantic segmentation approach is used for segmenting the trash objects from its background.

B. Instance segmentation

In Instance Segmentation as the name says, class specific entities are identified. For example, different types of trash/waste like plastic, paper, metal, glass, etc., were segmented based on the instances.

C. Panoptic segmentation

Panoptic segmentation produces a single, unified result that integrates predictions from instance and semantic segmentation. It gives a detailed vision and understanding on the given image/scene by having the scene level and subject level knowledge.

Literature Study

The Literature study shows that, the challenging dataset for waste detection, use of knowledge graphs with CNN, deformable object detection and real time integrations for waste detection

A. Cluttered object dataset

The largest public waste detection dataset, ZeroWaste, is introduced[5]. ZeroWaste is intended to serve as a benchmark for the evaluation and training of fully, weakly, and semi-supervised detection and segmentation techniques. Baseline results were presented for the most common completely, weakly, semi-supervised, and transfer learning approaches.

B. Knowledge graph utilization

To store and model objects in the scene that can take the form of photos, videos, texts, and other multimodal formats, [9] makes use of knowledge graphs. The YOLOv5-Attention-KG model is created by combining the ESA attention mechanism with the built-in multimodal knowledge graph in order to improve the network's feature extraction capabilities. The model is then deployed to the service robot so that it can perform real-time perception on the objects in the scene. Finally, in order to reason about and analyze the data in real time, collaborative training is deployed from the cloud server side to the edge device side.

C. Intelligent road sweeper

An automatic road garbage identification system [8] is presented, capable of locating most types of road debris with high precision and accurately instructing a road sweeper concerning the way to handle them. The system will just require a camera and a PC board with the ability to read camera output and communicate via CanBus. The system's software will primarily consist of a deep neural network (Mini-UNet) for semantic image segmentation and a real-time software programme to time-synchronize the sweeper actuators.

D. Cluttered transparent object detection

The GlassLoc algorithm is presented [7] for employing plenoptic sensing to determine the grab stance of transparent items within transparent clutter. GlassLoc categorises places that can be grasped in space using a Depth Likelihood Volume (DLV) descriptor. We enhance the DLV so that it may be determined from various plenoptic viewpoints whether transparent objects are occupied inside a particular space.

E. Garbage pile detection

With the aid of a cutting-edge methodology, [4] suggest SWMACM-CA, a real-time Smart Waste Management and Classification Mechanism. It classifies and separates waste materials in a landfill area using the Internet of Things (IoT), deep learning (DL), and cutting-edge methodologies. Additionally, a waste grid segmentation method is suggested that divides the

DOI: https://doi.org/10.5281/zenodo.15428792

waste yard's pile into grid-like parts. To generate a waste grid, a camera takes a picture of the garbage dump and sends it to an edge node. The grid cell picture segments serve as a test image for deep learning that has been taught to anticipate a specific waste item. VGG16, a deep learning method with 16 layers, was used for this particular project. To reduce overall latency, the model is trained on a cloud server installed at the edge node.

F. Small Object Detection

DLSODC-GWM, a new deep learning-based small object detection and classification technique [2], is developed in the study. The major objective of DLSODC-GWM technique is to support intelligent waste management systems by identifying and classifying small rubbish waste objects. The object detection and classification algorithms make up the two main steps of the DLSODC-GWM approach. An improved RefineDet (IRD) model is used in conjunction with an arithmetic optimization algorithm (AOA) for object detection, where the AOA selects the hyper parameters of the IRD model in the best possible way. Second, multiple classes of waste objects were classified using the functional link neural network (FLNN) technique.

Methodology

The Trash segmentation methodology starts from inputting the trash images to the encoder decoder based semantic segmentation, U2Net model and then getting the binary mask as a segmented output. The binary mask is placed over the original image and foreground white (0) pixels were set over to get an RGB mask and remaining black (1) pixels were eliminated. Finally, the RGB mask with no background is obtained. By finding the contours of the extracted masks the border of the segmented masks are formed and by using region proposals four coordinates are generated and with that the bounding boxes were created. With contouring the extracted masks were overlaid on the original images.

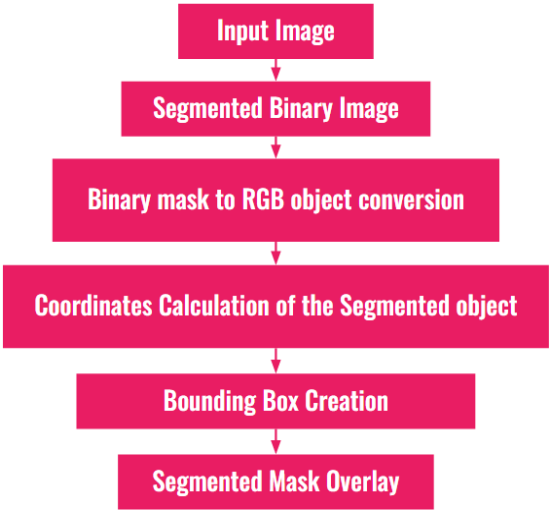

Fig. 2. Trash object Binary Segmentation Methodology

Fig. 3.

A. Dataset Description

For experimenting and evaluating the Waste object segmentation model, sample images from TACO Dataset [12] are taken. For real time image processing, sample images with different lighting conditions were taken from the Samsung M30 mobile in the portrait mode which is best suited for salient object capturing. These images were captured based on the type of plastics, glass, Medicines, toiletries, bottles, etc. Saliency denotes the important areas to be concentrated in the entire image.

B. Binary Semantic Segmentation Model

The binary segmentation deals with segmenting the foreground object from the background. So, here foreground is the trash object that will hold a pixel value of 0 and background is of various environments that holds the pixel value of 1. For deep binary segmentation the State of Art model U2Net [1] is used and the binary masks were obtained. U2Net is a Nested U-like structured architecture that is actually designed for salient object detection (SOD). The advantage of this model is that it segments with high accuracy at low computational costs.

DOI: https://doi.org/10.5281/zenodo.15428792

U2Net is similar to UNET, consists of encoder and decoder architecture with skip connections. The encoder down samples the original image while the decoder up samples back to its original resolution. In the bottom level of U structure, a novel ReSidual U-block module is used which extracts intra scale multi features without affecting the resolution of the feature map. On the top levels, each stage is filled with the ReSidual U-blocks (RSU). Local and global features are extracted and effective binary segmentation

Fig. 4. Trash object Binary Segmentation

Fig. 3 shows the binary segmented objects of the trash images using semantic segmentation model U2Net. Threshold is set to 0.9 and the Semantic salient object is done with sample TACO dataset images and real time images. As the real images were taken in a portrait mode, the salient object detection model performs well.

IoU scores shows that the performance of the model is good on the images. On the other hand, Taco dataset is a High resolution dataset and not based on the saliency. So comparatively the real time images were shown good results over the taco data. The results for the semantic segmentation process are shown in fig.4. As the image data is taken under different lighting conditions, some of the images undergoes over segmentation or under segmentation results.

results are obtained.

Performance Metrics

To analyse the performance of the segmentation model, some of the metrics are used. They are, Pixel Accuracy, Intersection over Union (Jaccard Index) and Dice Coefficient (F1 Score). Among these the common metric used to analyze the performance of the segmented model is Intersection over Union (IoU). The IoU metric is defined as the ratio of Intersection of actual ground truth mask and the predicted mask to the Union of the two masks. Therefore, the IoU metric measures the area of overlap between the actual ground truth mask and the predicted mask and the formula is given below in (1).

$$IoU = TP / (TP + FP + FN) \quad (1)$$

Where, TP, FP and FN denote True Positive, False Positive and False Negative respectively. True Positive is when the model correctly predicts the pixel of an object when it is actually a part of an object. If a model predicts the pixel as not on an object when it is actually a part of an object, then it is False Positive.

Conclusion

It is concluded that the State of Art model U2Net is performing good in the segmentation process and further it can be preceded with trash classification process. As given in the literature, this segmentation results can be further taken into consideration for the prototype implementation after doing the classification process.

Future direction can be considered as drone based trash detection, trash robotic picking and trash spilling areas under surveillance.

DOI: https://doi.org/10.5281/zenodo.15428792

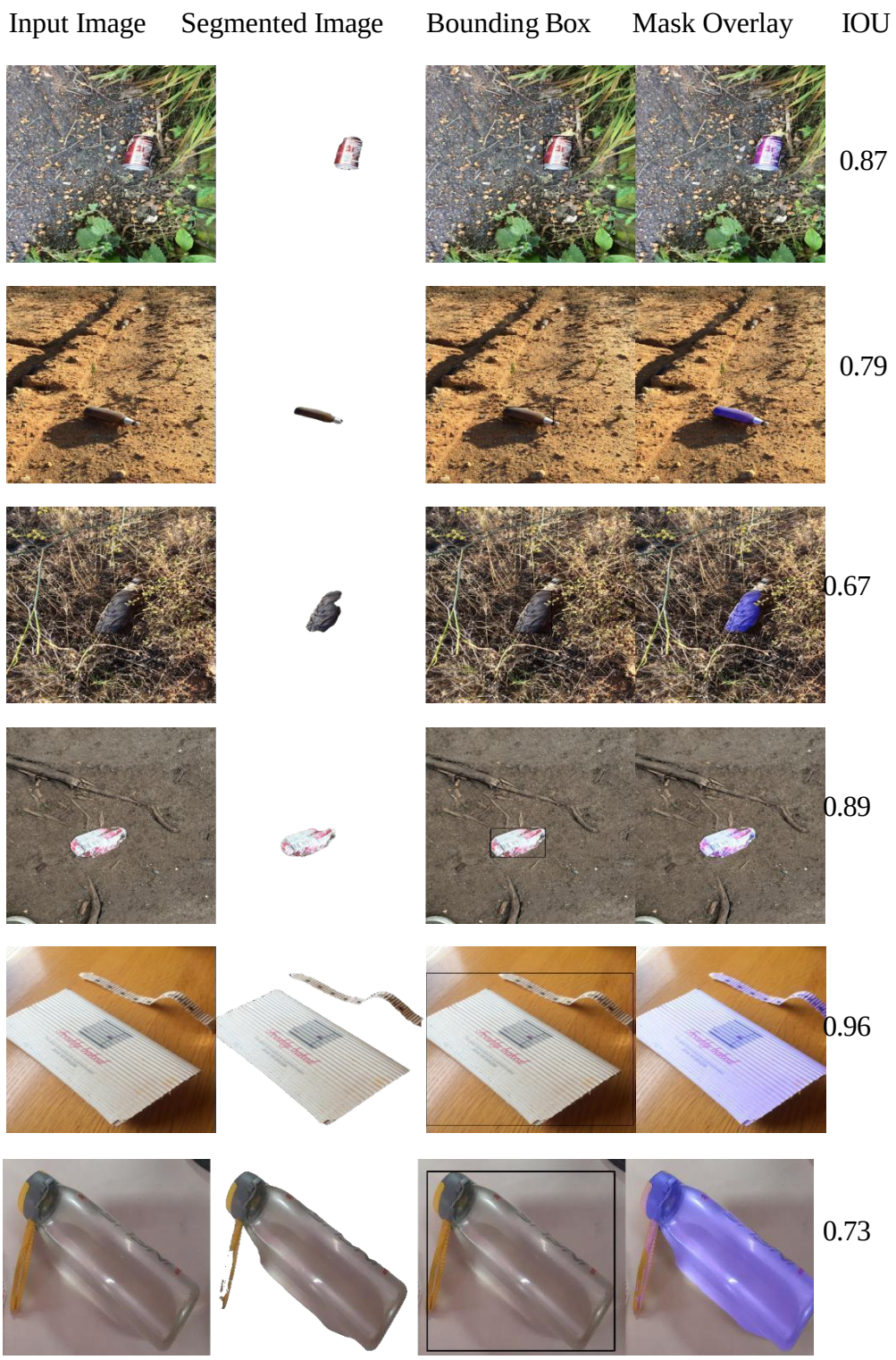

Fig. 5. Trash object Segmentation Process with IoU

DOI: https://doi.org/10.5281/zenodo.15428792

References

[1] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand, "U2-Net: Going deeper with nested U-structure for salient object detection," *Pattern Recognit.*, vol. 106, no. 107404, p. 107404, Oct. 2020.

[2] F. S. Alsubaei, F. N. Al-Wesabi, and A. M. Hilal, "Deep learning-based small object detection and classification model for garbage waste management in smart cities and IoT environment," *Appl. Sci.*, vol. 12, no. 5, p. 2281, Feb. 2022.

[3] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523–3542, Jul. 2022.

[4] S. M. Cheema, A. Hannan, and I. M. Pires, "Smart waste management and classification systems using cutting edge approach," *Sustain. Sci. Pract. Policy*, vol. 14, no. 16, p. 10226, Aug. 2022.

[5] Dina Bashkirova, Mohamed Abdelfattah, Ziliang Zhu, James Akl, Fadi Alladkani, Ping Hu, et al. "ZeroWaste Dataset: Towards Deformable Object Segmentation in Cluttered Scenes," *ArXiv*. 2022.Available: https://doi.org/10.48550/arXiv.2106.02740.

[6] C. Zheng, D. Cao, and C. Hu, "A similarity-guided segmentation model for garbage detection under road scene," *Front Optoelectron*, vol. 15, no. 1, p. 22, May 2022.

[7] Z. Zhou, T. Pan, S. Wu, H. Chang, and O. C. Jenkins, "GlassLoc: Plenoptic grasp pose detection in transparent clutter," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Macau, China, 2019.

[8] L. Donati, T. Fontanini, F. Tagliaferri, and A. Prati, "An energy saving road sweeper using deep vision for garbage detection," *Appl. Sci.*, vol. 10, no. 22, p. 8146, Nov. 2020.

[9] Y. Wu, X. Shen, Q. Liu, F. Xiao, and C. Li, "A garbage detection and classification method based on visual scene understanding in the home environment," *Complexity*, vol. 2021, pp. 1–14, Nov. 2021.

[10] M. Malik, S. Sharma, M. Uddin, C.-L. Chen, C.-M. Wu, P. Soni, and S. Chaudhary, "Waste classification for sustainable development using image recognition with deep learning neural network models," *Sustain. Sci. Pract. Policy*, vol. 14, no. 12, p. 7222, Jun. 2022.

[11] Pedro F Proenca, Pedro Simoes, "TACO: Trash Annotations in Context for Litter Detection," *ArXiv*, 2020

DOI: https://doi.org/10.5281/zenodo.15428792