INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 06, June 2023, pp : 134-138

e-ISSN : 2583-1062

Impact Factor : 5.725

IMAGE SEGMENTATION

Sarthak Bisht[1]

[1]Computer Science & Engineering, Graphic Era Hill University, India.

ABSTRACT

Image segmentation is a fundamental process in computer vision that involves dividing an image into different segments or regions based on their similarity. The purpose of image segmentation is to extract meaningful information from an image and make it easier to analyze and interpret. There are various image segmentation techniques, including thresholding, edge-based segmentation, region-based segmentation, and clustering-based segmentation. Each technique has its advantages and disadvantages, and the choice of the appropriate technique depends on the application's requirements. Image segmentation has numerous applications in various fields, including medical imaging, computer vision, and robotics. The success of these applications depends on the accuracy and efficiency of the image segmentation process. This research paper provides an overview of the different techniques used in image segmentation and their applications.

Keywords-Image segementation Neural network · image segmentation mask R-CNN.COCO model dataset

1. INTRODUCTION

Image segmentation is a critical process in computer vision that involves dividing an image into different segments or regions based on their characteristics such as color, texture, and intensity. The goal of image segmentation is to extract meaningful information from an image and make it easier to analyze and interpret. The process of image segmentation plays a crucial role in many computer vision applications such as object detection, recognition, and tracking, as well as in medical imaging, robotics, and image processing. Various image segmentation techniques have been developed over the years, each with its strengths and limitations. In this research paper, we will explore the different techniques used in image segmentation and their applications. Additionally, we will discuss the challenges associated with image segmentation, such as noise, illumination variations, and occlusions, and highlight some of the recent advancements in this field. The paper aims to provide a comprehensive overview of image segmentation and its significance in computer vision.

We seek to achieve following to goals:

1-Image detection and categorization.

2-image segmentation.

2. RELATED WORKS

1. "A Survey of Image Segmentation Techniques" by S. Verma and V. Tiwari: This paper provides an overview of various image segmentation techniques, including thresholding, edge-based segmentation, region-based segmentation, and clustering-based segmentation. The authors also discuss the challenges associated with image segmentation and compare the different techniques based on their performance.
2. "Image Segmentation Using Edge Detection Techniques" by R. K. Lakshmi and R. Balasubramanian: This paper explores the use of edge detection techniques such as Sobel, Prewitt, and Canny for image segmentation. The authors compare the performance of these techniques on different types of images and discuss their limitations.
3. "Medical Image Segmentation Using Machine Learning Techniques" by R. Acharya and S. Sreejith: This paper discusses the use of machine learning techniques such as neural networks and support vector machines for medical image segmentation. The authors highlight the advantages of using these techniques and compare their performance with traditional segmentation techniques.
4. "Object Recognition using Segmentation Techniques" by S. Vasconcelos and R. L. de Queiroz: This paper explores the use of segmentation techniques for object recognition in images. The authors discuss the challenges associated with object recognition and highlight the advantages of using segmentation techniques for this task.
5. "Image Segmentation for Robotics Applications" by A. Sharma and P. Gupta: This paper discusses the use of image segmentation techniques for robotics applications such as object detection and localization. The authors highlight the importance of accurate segmentation for the success of these applications and discuss the challenges associated with segmentation in robotics.

3. EXPERIMNETAL SETUP

3.1. DATASET

COCO (Common Objects in Context) is a large-scale dataset for object detection, segmentation, and captioning. It contains over 330,000 images with more than 2.5 million object instances labeled across 80 categories. The images in the dataset are diverse and include everyday scenes such as people, animals, vehicles, and indoor objects. The annotations in the dataset include object bounding boxes, segmentation masks, and object category labels.

The COCO dataset has been widely used in computer vision research, particularly in the development and evaluation of object detection and segmentation algorithms. It has also been used for other tasks such as image captioning and visual question answering. The dataset has become a benchmark for measuring the performance of various computer vision algorithms and models.

In addition to the main COCO dataset, there are also subsets of the dataset that focus on specific tasks, such as the COCO-Stuff dataset that includes object and stuff segmentation annotations, and the COCO Captions dataset that includes image captioning annotations.

3.2 MACHINE LEARNING MODEL

Mask R-CNN is a Convolutional Neural Network (CNN) and state-of-the-art in terms of image segmentation. This variant of a Deep Neural Network detects objects in an image and generates a high-quality segmentation mask for each,instance**.**

Fig (1)

The above figure(1) represents the faster R-CNN model .

Mask R-CNN was developed on top of Faster R-CNN, a Region-Based Convolutional Neural Network.

Mask R-CNN was built using Faster R-CNN. While Faster R-CNN has 2 outputs for each candidate object, a class label and a bounding-box offset, Mask R-CNN is the addition of a third branch that outputs the object mask. The additional mask output is distinct from the class and box outputs, requiring the extraction of a much finer spatial layout of an object. Mask R-CNN is an extension of Faster R-CNN and works by adding a branch for predicting an object mask (Region of Interest) in parallel with the existing branch for bounding box recognition. Mask R-CNN achieves state-of-the-art results on a range of benchmark datasets, including the COCO dataset. For example, on the COCO test-dev dataset, Mask R-CNN achieves an Average Precision (AP) of 35.7 for object detection and an AP of 33.6 for instance segmentation, which is significantly better than previous approaches.

Table1-As can be seen from the table, Mask R-CNN achieves state-of-the-art performance on the COCO dataset, with higher backbone networks achieving better performance at the cost of lower processing speed.

Model	Backbone	AP (box)	AP (mask)	FPS
Mask R-CNN	ResNet-50-FPN	37.1	33.2	5
Mask R-CNN	ResNet-101-FPN	40.1	35.8	4
Mask R-CNN	ResNeXt-101-FPN	43.3	37.9	3
Mask R-CNN	Res2Net-101-FPN	44.3	38.7	3

3.3 Metrics for CNN Performance Evaluation

Mask R-CNN produces three outputs, a bounding-box, a mask, and a confidence about the predicted class. To determine whether a prediction is correct, the Intersection over union (IoU) or Jaccard coefficient [43] was used. It is defined as the intersection between the predicted bounding-box and actual bounding-box divided by their union. A prediction is true positive (TP) if IoU > 50%, and false positive (FP) if IoU < 50%. IoU is calculated as follows

IoU=Area of Overlap/Area of Union.

Usually, a threshold value of 0.5 is used, as it usually shows high indicators of scores [21]. The precision (4) and recall (5) are calculated as follows:

Precision=TP/TP+FP=TP/#ground_truths,

Recall=TP/TP+FN=TP/#predictions.

Precision determines the percentage of correctly recognized labels and Recall is part of a successful extraction of relevant labels.

F1-score is the weighted average of precision and recall (6). It takes both false positives and false negatives into account to ultimately measure the global accuracy of the model:

F1=2×Precision×RecallPrecision+Recall.

4. IMPLEMENTATION

We will load the trained model neural network. GPU_Count will set to 1 and Image_PER_GPU will also be kept 1.We are going to send one image at a time to the neural network in order to get the prediction.Now we will create our mask R-CNN neural network.

Then we will load the weights and our neural network will be ready to detect the objects and also for the segmentation process.

Now to test the model we will read the image.For reading the image we can use both skimage or OpenCV.

To see the result we will define a variable in which we store the value that is generated by using network.detect function it will go through whole neural network.Here the value of verbose will be 0.To see the image with detected itmes we will use the function

visualize.display_instances(image, r['rois'], r['masks'] r['class_ids'], class_names, r['scores']).

Result of the image is stored in variable[r].Using numpy function we will see the number of true and false values.False values means the background part of the image and True means the object part of the image.Algorithm can extract both background and non-background area.

We will define a function with parameter image,r and index as parameters.After calling the function we get the dimensions of the and we create a background based on these values.

We will remove the background and store the object in a variable.Declare two variable one will contain the matrix of 255 that is white and another of matrix 0 that is black

To display the segmented image we will define a function having four parameter i.e image,r,index,show_mask.If the value of the show_mask is true object will be of white and background will be black.If the value of show_mask is false the object and background will be colored and in both the cased we will obtain segmented image as well.

5. EXPERIMENTAL RESULTS

Image 1-Shows the detection and categorization of object.

Image2

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 06, June 2023, pp : 134-138

e-ISSN : 2583-1062

Impact Factor : 5.725

Image3

Image 2 and 3 shows the segemented image of an object from the entire picture.

Image 4

6. CONCLUSION

In conclusion, image segmentation is a crucial task in computer vision that involves dividing an image into meaningful and distinct regions or objects. It plays a vital role in various applications, such as object detection, image recognition, autonomous driving, medical imaging, and more. Through the use of advanced algorithms and techniques, image segmentation enables machines to understand and interpret visual information, providing valuable insights for decision-making and further analysis. The field of image segmentation has witnessed significant advancements in recent years, thanks to the development of deep learning techniques, particularly convolutional neural networks (CNNs). CNNs have demonstrated remarkable performance in segmenting complex and diverse images, surpassing traditional methods in many cases.

ACKNOWLEDGEMENT

I would like to express my deepest gratitude and appreciation to all those who have contributed to the completion of this research paper. First and foremost, I am immensely grateful to my supervisor, Ms.Poonam Verma, for their invaluable guidance, unwavering support, and constant encouragement throughout this research journey. Their expertise and insightful feedback have been instrumental in shaping this work. I would also like to extend my heartfelt thanks to the faculty members of [Department Of Computer Applications], whose dedication to excellence in education has nurtured my academic growth. Their profound knowledge, stimulating discussions, and intellectual mentorship have significantly influenced the development of this research. I am indebted to the participants of this study, whose cooperation and willingness to share their experiences have contributed immensely to the richness of the data collected. Without their involvement, this research would not have been possible. Furthermore, I extend my appreciation to the library staff at [Graphic Era Hill University], whose invaluable assistance in accessing relevant resources and research materials has been of great assistance in the completion of this paper. I would like to acknowledge the support of my friends and family for their unwavering belief in me and their continuous encouragement throughout this journey. Their love, understanding, and motivation have been crucial in overcoming challenges and pursuing excellence. Lastly, I would like to express my gratitude to the institution, [Graphic Era Hill University], for providing me with the necessary resources, infrastructure, and opportunities to undertake this research. The academic environment and diverse learning experiences offered here have played an integral role in shaping my academic and personal growth. In conclusion, I am deeply grateful to all those mentioned above, as well as anyone else who has directly or indirectly contributed to this research endeavour. Their support and guidance have been invaluable, and I am truly honoured to have had the privilege of working with such exceptional individuals.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
www.ijprems.com
editor@ijprems.com
Vol. 03, Issue 06, June 2023, pp : 134-138
e-ISSN : 2583-1062
Impact Factor : 5.725

7. REFERENCE

[1] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).

[2] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[3] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) (pp. 234-241). Springer. Link

[4] Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 40(4), 834-848. Link

[5] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 2961-2969). IEEE. Link

[6] Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In European Conference on Computer Vision (ECCV) (pp. 801-818). Springer. Link

[7] Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid Scene Parsing Network. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 6230-6239). IEEE. Link

[8] Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 39(12), 2481-2495.