Semantic Segmentation

Mehran Tamjidi[1], Amir Mousavinia[2]
Faculty of Electrical and Computer Engineering, K. N. Toosi University of Technology, Tehran, Iran

ABSTRACT

Semantic Segmentation is a classification algorithm which partition the scene or image to the limited categories and assign label to each class by deeming all their object or stuff as a single entity.

This paper provides a brief review over semantic segmentation, explaining the concept, applications, approaches and also discusses about novel solutions applying to modify some contemporary algorithms , improve their evaluation metrics and obviate their drawbacks.

It also includes methods for semantic segmentation on 3D images, point clouds and some new concepts of segmentation.

INTRODUCTION

Due to the rapid evolution of autonomous driving systems, indoor robots, UAVs the need for **scene understanding** and **scene parsing** has increased. The automaton needs to percept their environment, detect the objects in it's vicinity and even estimate their dimensions and distance.
segmentation has also found it's way in medical diagonalizing. **Medical scans** are able to detect abnormalities by analyzing MRI or CT scans. pulmonary lobe segmentation is a method to diagnose many pulmonary diseases like COVID-19 [1].
Development of satellite and airborne imagery technologies also provides high resolution aerial images to determine exact boundaries between different regions. Semantic segmentation of **remote sensing** images can analyze the venetian of a territory [2] or even the impact of water stress to an specific crop [3].
Extracting and learning the visual styles by semantic segmentation can obtain **clothing and fashion** industry an opportunity to product clothes with the styles ,that are going to be trend in future [4] and forecast the feature of styles [5].
Semantic segmentation algorithms **Annotating All pixels** with labels determining to witch category they belong ,and partitioning the scene to the limited category. This segmentation are based on **processing uncountable stuff** and does not discriminate the objects within a category.
The semantic segmentation algorithms almost **Analyzing upon the perception** which they obtain about the pixels and it's neighbors, so for a outdoor scene expects road for the nearby pixels of car categories.

Figure 1. input image.

Figure 2. output image by Segnet algorithm [6].

Traditional approaches

Markov Random Field (MRF) is a subset of statistical models, it's edges are undirected and the graph support circular connections. This algorithm assigns random variable to each feature and all of the pixels of the image. The edges determines the dependency of the neighbors. Pixel's variables are the labels specifying to which class they belong.
Conditional Random Field (CRF) is a probabilistic graphical model which widely used for segmentation and image classification, assign arbitrary labels to inputs regardless it's neighbors.

Random walk(RW) is clustering graph based image segmentation algorithm**.**
Support vector machine (SVM) is a kernel based classifier which provide loss function for non linear separatable data also nonlinearly map data to higher dimension.
Random decision forest(RDF) first discussed by [7] sometimes named holistic method uses combination of multiple decision trees are classifiers are trained. This algorithm is faster and have stronger scale of measure of the features.
Traditional Neural network(RDF) state that increasing the hidden layers would model every complex feature model. Efficient classifies and activation functions are needed for segmentation.

Deep learning approaches

Recently deep learning algorithm reach promising results. They are strong in extracting abstract local features. among them convolutional networks, recurrent networks and deep belief networks are common structure. **Transformers** as a new structure of deep neural networks can have better perform and beat previous structures in some cases.

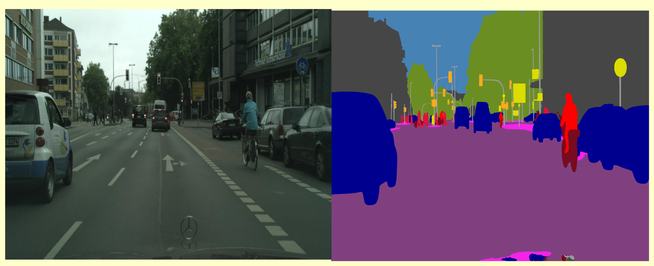

Figure 3. input image.

Figure 4. output image by [8].

Figure 5. Qualitative result of transformation based algorithm (right) with respect to the FCN [9].

Semantic Segmentation On 3D images

The ubiquities inexpensive RGBD cameras and Lidar sensor are wieldy used for scene understanding and 3D modelling .Unlike the regular structure of 3D images ,The 3D point clouds generated by Lidar are scattered in continuous space . Semantic segmentation of point clouds can be used for many applications.

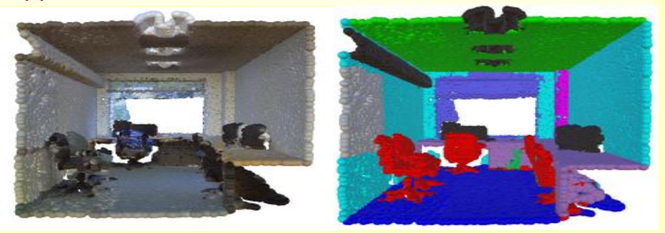

Figure 6. semantic segmentation of 3d point cloud by DGCNN algorithm [10].

Future Trends

Semantic segmentation is not able to distinguish different objects within a class. But the recently developed **panoptic segmentation** carry out semantic and instance segmentation together.

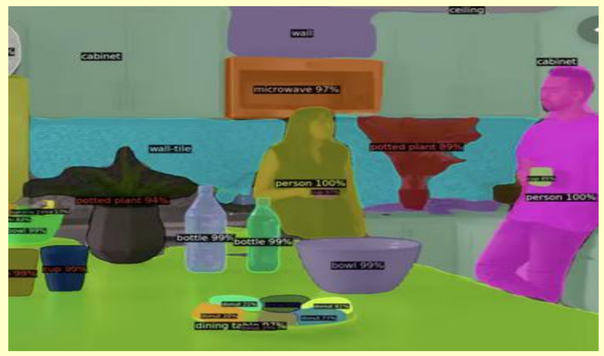

Figure 7. panoptic segmentation

Deep learning algorithms need Lage annotated dataset to perform well. **Few-shot and zero-shot segmentation** can provide solution. zero-shot semantic segmentation, which aims to segment unseen objects, which were not seen during training with only category-level semantic representations provided for unseen categories [11,12].

REFERENCES

1. W. Xie, C. Jacobs, J. -P. Charbonnier and B. van Ginneken, "Relational Modeling for Robust and Efficient Pulmonary Lobe Segmentation in CT Scans," in IEEE Transactions on Medical Imaging, vol. 39, no. 8, pp. 2664-2675, Aug. 2020, doi: 10.1109/TMI.2020.2995108..
2. Ronald Kemker, Carl Salvaggio, Christopher Kanan,Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning,ISPRS Journal of Photogrammetry and Remote Sensing,2018.
3. Devin L. Mangus, Ajay Sharda, Naiqian Zhang,Development and evaluation of thermal infrared imaging system for high spatial and temporal resolution crop water stress monitoring of corn within a greenhouse,Computers and Electronics in Agriculture,Volume 121,2016.
4. Martinsson, John, and Olof Mogren. "Semantic segmentation of fashion images using feature pyramid networks." Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops. 2019.
5. Stiefelhagen, R., & Grauman, K. (2017). Fashion Forward: Forecasting Visual Style in Fashion. ArXiv. https://doi.org/10.48550/arXiv.1705.06394
6. Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "Segnet: A deep convolutional encoder-decoder architecture for image segmentation." IEEE transactions on pattern analysis and machine intelligence 39.12 (2017): 2481-2495.
7. Geurts, Pierre, Damien Ernst, and Louis Wehenkel. "Extremely randomized trees." Machine learning 63 (2006): 3-42.
8. Tao, Andrew, Karan Sapra, and Bryan Catanzaro. "Hierarchical multi-scale attention for semantic segmentation." arXiv preprint arXiv:2005.10821 (2020).
9. Zheng, Sixiao, et al. "Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.
10. Wang, Yue, et al. "Dynamic graph cnn for learning on point clouds." Acm Transactions On Graphics (tog) 38.5 (2019): 1-12.
11. Zhao, Na, Tat-Seng Chua, and Gim Hee Lee. "Few-shot 3d point cloud semantic segmentation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.
12. Cheraghian, Ali, Shafin Rahman, and Lars Petersson. "Zero-shot learning of 3d point cloud objects." 2019 16th International Conference on Machine Vision Applications (MVA). IEEE, 2019.

CONTACT

Mehran Tamjidi
Email: m.tamjidi@email.kntu.ac.ir