International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 12 | Dec 2020 www.irjet.net p-ISSN: 2395-0072

Semantic Segmentation of Satellite Images: A Survey

Tarun Pratap Singh[1], Ravi Ranjan Singh[2], Himanshu[3], Arpit Mishra[4], Nidhi Sharma[5]

[1,2,3,4]Department of Computer Science and Engineering, Delhi Technical Campus, Greater Noida, UP, India
[5]Professor, Department of Computer Science and Engineering, Delhi Technical Campus, Gr. Noida, UP, India

---***---

Abstract - *The incrementing utilization of raw satellite images has many uses for mapping of the earth for coastal and ocean applications. Hazard assessment and natural resource management can also be done via this process. Our paper is a sincere attempt to analyse the different algorithms/models for efficient semantic segmentation of satellite images. The different methods for image segmentation are EM, FCM, SVM, K-Means, Unet, PSPnet. The dataset was obtained by the IEEE-GRSS data fusion contest. The span of satellite images were two American cities namely, Jacksonville, Florida and Omaha, Nebraska, USA. Finally, the evaluation metrics have been provided with the concerned methods.*

***Key Words*: Image Segmentation, Semantic Segmentation, Clustering, EM, FCM, SVM, K-Means, Unet, PSPnet, Segnet.**

1. INTRODUCTION

There are only certain components of the image which are of interest for the research and application of the image. These segments are usually referred to as a foreground (the other part is termed background), which commonly corresponds to the image in distinct and unique features of the area. There is a requirement to derive and divide them to classify and interpret objects. In Image Segmentation, it is necessary that we not only classify the images into different categories as a whole but to segment each pixel in the image accordingly. The process of classifying each segment of an image into different classes is called image segmentation. However, segmentation of the image on a per-pixel level is a very tedious task and traditionally it requires a lot of complex calculations and codes, but with the help of Convolutional Neural Networks (CNN), we not only have been able to achieve exceptionally impressive results but with great efficacy too. A satellite image can be segmented into parts depicting either man-made structures such as buildings and roadways and certain other physical structures such as forests and fields.

For image segmentation there are many models which are used such as EM, FCM, SVM, K-Means, Unet, PSPnet etc. SVM, FCM, K-Means are clustering algorithms that work by grouping similar data points in substantially larger data sets. These algorithms try to fit each pixel into a specific cluster to separate and identify objects. Unet, PSPnet and FCN are convolutional neural networks where FCNs replace the fully-connected layers of standard CNNs which do not contain any Dense layers. The CNN models have provided state of the art results with very impressive scores which are now being used not only in satellite image segmentation but many other different fields as well.

2. LITERATURE SURVEY

Throughout the years multiple contributions have been made in the field of satellite semantic segmentation. Although few of these approaches might be obsolete and are no longer being used but they give us an insight of how the technology has evolved and why it might be better and efficient to use certain algorithms and their benefits. We have reviewed the major technology changes over the years in this paper.

Syed Nazeebur Rehman et al. (2018) has stated an approach to segment the satellite images using Fuzzy C-Means (FCM). FCM allows a piece of data to be divided into two or more clusters, hence helps in pattern recognition very efficiently as shown in figure: 1. The results concluded that the PSNR value of the FCM algorithm was 36.23 [1].

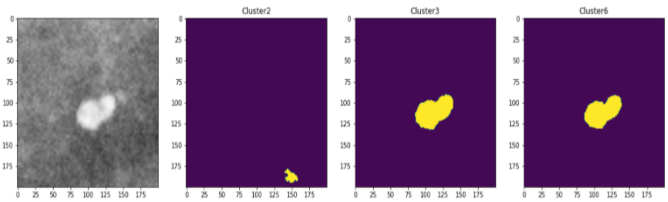
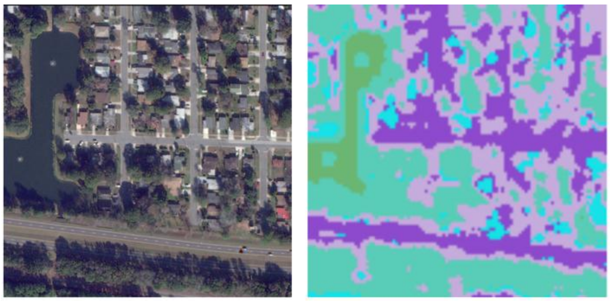


Fig-1: Fuzzy C Means Clustering

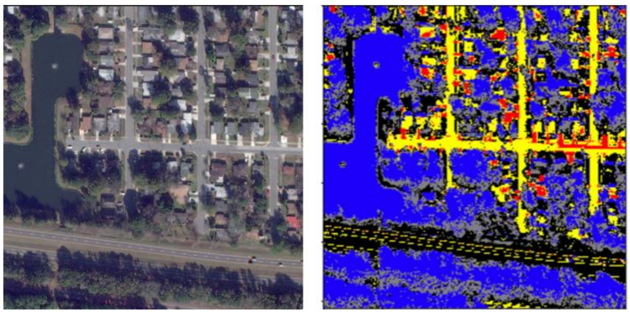
L. Senthin Nathan (2018) has proposed an approach to segment the satellite images using Expectation Maximization (EM) algorithm. He used the median filter to reduce the noise level of the picture before segmenting the images, the EM parameters have been used to determine the latent variable distribution. The results have been shown in figure 2. From the results he concluded that the EM algorithm is an efficient technique and has obtained a PSNR value of 38.65[2].



Original Image Segmented Image

Fig2: Segmentation by EM algorithm

L. Senthin Nathan (2017) has published his work on K-Means Algorithm bases satellite image segmentation. K-means clustering panel n annotations into K-groups where each observation belongs to a neighboring cluster with a bordering mean. The K-Means parameters are used here to find out the latent variable distribution. The results conclude that the K-Means algorithm has a PSNR value of 40.5[3]. The segmented image by the K-Means algorithm is shown in figure 3.



Original Image Segmented Image

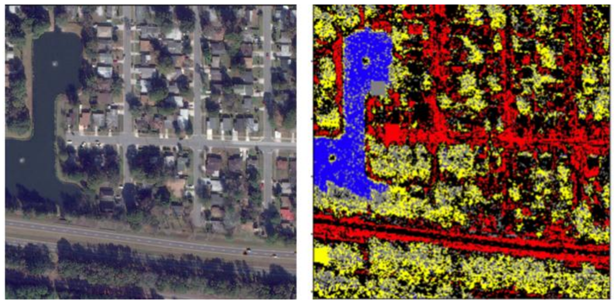
Fig3: Segmentation by K-Means algorithm

The comparison between the PSNR values of FCM, EM and K-Means is given below in figure 4.

Method Used	PSNR Range
EM	38.65
FCM	36.23
K-Means	40.5

Fig4: Comparison of PSNR values

M. Vakalopoulou et al. (2015) has proposed a solution based on Support Vector Machines (SVM). The SVM model has been employed for the classification of the training and the test images. The resulting map calculates a score for each pixel inside the class, the buildings were theme extracted using the MRF-Based Model which improved the final building detection on the map. The overall accuracy from the above method was found to be 74% [4]. In a similar approach using the SVM model O. Benarchid and N. Raissouni [5] has proposed a method to automatically extract buildings using Very High-Resolution Satellite (VHRS) images in which he used an object-based approach of segmentation before classification. The results of the proposed algorithm are shown in figure 5. The accuracy of the stated method was observed to be 83.76%.



Original Image Segmented Image

Fig5: Segmentation by SVM algorithm

Making huge improvements Guillaume Chhor et al. [6] have stated in their paper that they had developed a modified version of the original Unet [7] which was initially used for the biomedical image segmentation and trained the model from scratch as the pre-trained weights were trained on synthetic data which would adversely affect the training. The optimizer was changed to Adam from SGD for faster convergence followed by using the dice coefficient for the loss function instead of cross entropy which was used in the original Unet model. The dropout layer was also removed as no overfitting of the data was observed. The proposed model had achieved a mean_IoU of 0.6 which was less than the state-of-the-art results that had a mean_IoU of 0.7 with Segnet [8] but had performed better than the baseline FCN [9] with mean_IoU of 0.53. The accuracies of the above discussed methods are shown in figure 6.

Models Used	mean IoU	Acc. (Pixel)
Baseline FCN [2]	53.82%	92.79 %
Baseline FCN + MLP[2]	64.67o6	94.42 %
FCN (VGG16 encoder)	66.21%	94.54 %
FCN + MLP (VGG16 encoder)	68.17%	94.95 %
Segnet (VGG16 encoder)	**70.14%**	**95.17** %

Fig6: State of the art results on INRIA dataset

In 2019 Junxing Hu et al. [10] has done a comparative analysis on the different convolutional networks based on segmentation performance, number of parameters, resource consumption and edge prediction. The models that were compared were VGG based models. The U-net and DeepLab had the best performance with the mean_IoU of 0.74 while PSPnet had the second-best mean_IoU of 0.71. Here they concluded that DeepLab is overall best suited for real world situations as the datasets are mostly unbalanced. On the other hand, the Segnet is the smallest model with the least parameters while FCN is the largest one. As per the resource consumption Segnet and U-net both had small FLOP (Floating Point Operations) while the other models have larger FLOP. Unet and DeepLab had excellent edge prediction scores which would later help to restore the shape of the target objects and hence predict a very sharp segmented image. The segmented images along with the original satellite mage has been shown in figure 7, 8, 9, 10, 11.

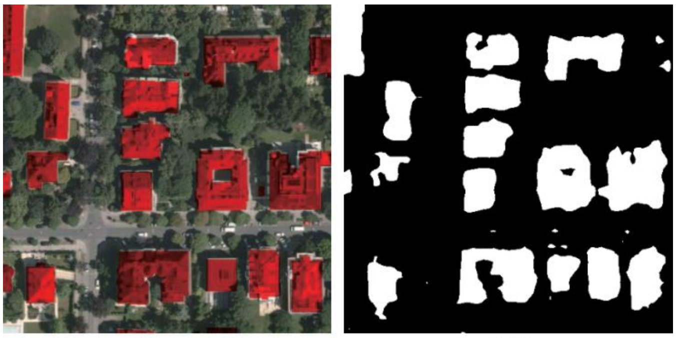


Image and the ground truth FCN

Fig7: Using the FCN model for Segmentation

Image and the ground truth PSPNet

Fig8: Using the PSPnet model for Segmentation

Image and the ground truth SegNet

Fig9: Using the Segnet model for Segmentation

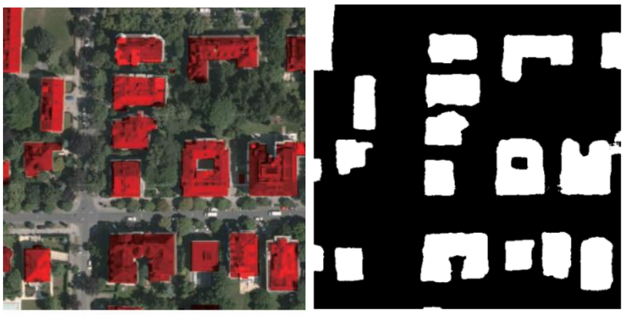


Image and the ground truth U-Net

Fig10: Using the U-net model for Segmentation

Image and the ground truth DeepLab

Fig11: Using the DeepLab model for Segmentation

The mean_IoU and the accuracy of the different models using the images at various locations along with the overall metrics is given below in figure 12.

Method	Austin		Chicago		Kitsap		West Tyrol		Vienna		Overall	
	IoU	Acc	IoU	Acc	IoU	Acc	IoU	Acc	IoU	Acc	IoU	Acc
FCN-8s	50.28	92.30	53.89	87.24	32.09	98.52	56.40	95.84	62.75	88.30	56.19	92.44
U-Net	**78.62**	**96.89**	**70.39**	**92.89**	**66.26**	**99.27**	70.93	97.71	78.28	93.85	74.79	**96.12**
Segnet	70.60	94.74	64.81	89.72	60.55	98.89	71.41	97.28	74.97	91.79	70.10	94.48
DeepLab	76.65	96.54	69.39	92.56	65.78	99.24	**75.01**	**97.98**	**79.24**	**94.06**	**74.86**	96.08
PSPnet	71.69	95.73	66.67	91.62	63.08	99.18	72.07	97.72	76.49	93.12	71.67	95.47

Fig12: Evaluation metrics for the CNN models

3. CONCLUSIONS

In this paper, we discuss and evaluate main image segmentation techniques used for the purpose of satellite image segmentation. Since there are many depending factors for an image like image content, image resolution, dataset, number of channels, etc therefore it is difficult to devise a perfect method for the process of semantic segmentation of satellite images. The traditional image segmentation methods EM, K-means and FCM produced results which were unsatisfactory but when the technology moved to convolutional neural networks (CNN) we observed that the results were very impressive and up to the mark of our expectations. In situations where the datasets were refined and balanced the models like Unet and PSPnet performed exceptionally well although similar results were observed with Segnet which consumed drastically a smaller number of resources than the ones mentioned above. On the other hand, where the datasets are not balanced, Deeplab had a slight edge over the other state of the art models. Hence it was preferred to use Deeplab in real life situations because the datasets are usually unbalanced. It is concluded that the different models or possibly even a hybrid model can be used depending on the situation that arises.

REFERENCES

[1].L. Senthil Nathan, 2018. Expectation Maximization Based Satellite Image Segmentation. Journal of Engineering and Applied Sciences, 13: 9343-9345.

[2].Rehman, Syed & Hussain, Mohameed. (2018). Fuzzy C-Means Algorithm Based Satellite Image Segmentation. Indonesian Journal of Electrical Engineering and ComputerScience.9.332-334. 10.11591/ijeecs.v9.i2.pp332-334.

[3].Senthil Nathan, 2017. K-Means Algorithm based Satellite Image Segmentation. Journal of Engineering and Applied Sciences, 12: 7995-7997.

[4]. Maria Vakalopoulou, Konstantinos Karantzalos, Nikos Komodakis, Nikos Paragios. Building detection in very high resolution multispectral data with deep learning features. Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International, Jul 2015, Milan, Italy. pp.1873-1876, ff10.1109/IGARSS.2015.7326158ff. ffhal-01264084f

[5].O. Benarchid and N. Raissouni, Support Vector Machines for Object Based Building Extraction in Suburban Area using Very High Resolution Satellite Images, a Case Study: Tetuan, Morocco, IAES International Journal of Artificial Intelligence (IJ-AI), Vol. 2, No. 1, March 2013, pp. 43 50.

[6].Chhor, G., & Aramburu, C.B. (2017). Satellite Image Segmentation for Building Detection using U-net [at] stanford.edu.

[7].O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, Computer Science Department and BIOSS Centre for Biological Signalling Studies, University of Freiburg, Germany, 2015.

[8].V.Badrinarayanan, A.Kendall, R.Cipolla, SegNet: ADeepConvolutional Encoder-Decoder Architecture for Image Segmentation

[9].Long, E. Shelhamer and T. Darrell, Fully Convolutional Networks for Semantic Segmentation, University of Berkeley, Proceedings of the IEEE, 2015.

[10].Hu, Junxing & Li, Ling & Lin, Yijun & Wu, Fengge & Zhao, Junsuo. (2020). A Comparison and Strategy of Semantic Segmentation on Remote Sensing Images. 10.1007/978-3-030-32456-8_3.