

# Open Set Semantic Segmentation of Remote Sensing Images

Caio Cesar Viana da Silva\* , Jefersson Alex dos Santos

Department of Computer Science, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

{caiosilva, jefersson}@dcc.ufmg.br

**Abstract**—The development of computational vision approaches that exploit satellite imagery is relatively recent, mainly due to the limited availability of this type of image. In the area of remote sensing, applications that employ computational vision techniques are modeled for classification in closed set scenarios. However, the world is not purely closed set, many scenarios present classes that are not previously known by the algorithm, an open set scenario. Thus, the main objective of this paper is the study and development of semantic segmentation techniques considering the open set scenario applied to remote sensing images. Focusing on this problem, this is the first work to study and develop semantic segmentation techniques for open set scenarios applied to remote sensing images. The main contributions of this paper are: 1) a discussion of related works in open set semantic segmentation, showing evidence that these techniques can be adapted for open set remote sensing tasks; 2) the development and evaluation of four novel approaches for open set semantic segmentation. Our methods yielded competitive results when compared to closed set methods for the same dataset.

**Open Set, Deep Learning, Semantic Segmentation, Remote Sensing**

## I. INTRODUCTION

A source of data exploited by the computer vision domain is satellite or aerial images, which may be comprised of beyond the visible spectrum channels. However, satellite imagery has not been widely used mainly due to its limited availability given that, until recently, they were of exclusive military use.

Within the area of remote sensing, most applications work with closed set scenario-based computer vision techniques, wherein models are conceived to learn and predict the same set of classes, ignoring new unknown labels that may arise. However, the world is not purely closed set, since many scenarios present objects that were not previously known by the model. These other scenarios would be better explored using open set algorithms [1], which are capable of distinguishing new unknown classes from the ones used during training.

Formally, open set classification can be described as a task wherein an image can be labeled as belonging to one of the classes learned by the algorithm or as an unknown class if it belongs to any class not learned.

Even with a variety of possible uses for open set scenario algorithms, this area is not so explored, mainly when compared to the enormous number of closed set methods.

Semantic Segmentation is a task that aims to classify not only the whole image but every pixel in an image accordingly to the classes learned by the algorithm.

Therefore, open set semantic segmentation can be described as a set of techniques that receive an image as input and outputs a prediction for all pixels, either labeling them within a known class or as belonging to an unknown class. In this work, we evolve on the concept of open set semantic segmentation in remote sensing and propose two deep learning methods based on it.

In practice, the main contributions of this work are: (i) first evidences demonstrating the feasibility of combining semantic segmentation and open set concepts; (ii) a new method, called OpenPixel, for open set semantic segmentation based on the Pixelwise network [2]; and (iii) an adaptation of the proposed OpenPixel method, applying a morphological filter; (iv) a new method for open set semantic segmentation based on the Openmax method [3], called OpenFCN and (v) an adaptation of the proposed OpenFCN method, applying a Super Pixel as a preprocessing technique.

Following sections in this paper are organized as follows. Section III explains the methodology adopted in this paper for achieving the objectives previously defined. All experimental configuration and some assumptions needed to reproduce this work are contained in Section IV. Section V reports the results found using the proposed methods and discusses them. Finally, Section VI presents our final remarks and future works.

## II. RELATED WORK

This section presents background knowledge and a literature review on open set and semantic segmentation methods for image classification and semantic segmentation. Most of the open set techniques developed today are adaptations form closed set methods.

### A. Open Set Classification

In [1] the authors describe an open set version of the well-known Support Vector Machine (SVM) algorithm developed to classify scenes. The choice of using SVM was made because it has various alluring characteristics that can help in this scenario: its answers are global and unique; it has a basic geometric understanding, and it does not rely upon the dimensionality of the information space. [4] present and develop a technique for an “open world”, a recognition system that should update new object categories and be robust to these unseen groups, in addition, to have minimum downtime. [4] propose the Nearest Non-Outlier (NNO) algorithm that evolves

\*This work relates to an M.Sc. dissertation

a model efficiently by adding object categories incrementally while detecting outliers and managing open space risk.

Looking for a deep learning solution, [3] show a new model, called OpenMax, that represents an alternative for the SoftMax function as the final layer of the network, which estimates the probability of an input being from an unknown class. Reducing the number of errors made by a deep network when given fooling generated images. Using a shallow approach, [5] propose a method named Open-Set NN (OSNN) and a variation called OSNNcv, both are able to recognize samples from unknown classes during training time and outperform other approaches in the literature.

### B. Semantic Segmentation

Besides its open set characteristics, our method relies heavily on deep convolutional architectures. Convolutional Neural Networks (CNNs) [6]–[10] have been the state-of-the-art method for most Computer Vision classification tasks for the better part of the last decade. The first widely adopted CNN was AlexNet [6], followed by the deeper VGG [7] and GoogLeNet [8] architectures.

Between 2015 and 2017, it was observed that the training of earlier layers was severely hampered in deeper architectures due to the Vanishing Gradient problem. Residual Networks (ResNets) [9] and Densely Connected Convolutional Networks (DenseNets) [10] were designed to deal with this limitation of simply stacking convolutions on top of each other by employing shortcuts for the backward gradient between shallower and deeper layers. CNNs can be used for image segmentation by classifying the central pixel of a region according to its class and iterating this algorithm across all pixel positions, as will be further explained in Section III.

### C. Gaps Explored by the Proposed Methods

The knowledge gap explored by this paper becomes clear when observing the related works presented in this section. None of the existing methods properly perform both open set inference and semantic segmentation. Therefore, as far as the authors are aware, the methods proposed in this work are the first techniques that can semantically segment a remote sensing imagery on an open set scenario.

## III. METHODOLOGY

This section describes in detail the proposed methods, there are: (1) the OpenPixel method presented in Section III-A, (2) Morph-OpenPixel, presented in Section III-B, (3) the method OpenFCN presented in the Section III-C and (4) its variation using Super Pixels, called SLIC-OpenFCN in the Section III-D. All methods are based on existing closed set methods for semantic segmentation, but the OpenPixel technique does a pixel-level classification and applies a threshold to differentiate between known and unknown classes while OpenFCN uses the relation between activation maps of each known class and the FCN network to perform the task.

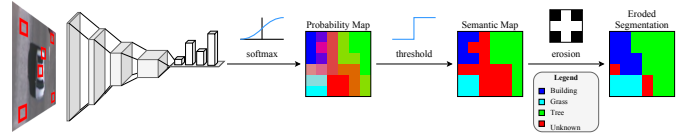


Fig. 1. Simplified OpenPixel and Morph-OpenPixel architecture. The distinction between the networks is that OpenPixel does not contain the last morphological filtering step.

### A. OpenPixel: Pixelwise Open Set Classification

The first method proposed in this paper is an open set adaptation from the closed set Pixelwise algorithm proposed by [2], which had presented good results in remote sensing datasets. This Pixelwise approach consists of the individual treatment of all the pixels present in the images. Since the pixel itself has not enough information to allow its classification, context windows are employed. Precisely, context windows are crops of  $55 \times 55$  pixels with the central one representing the crop class. By iterating over all pixels in an image, it is possible to train a CNN for patch-level classification by using the contextual window.

In order to compose OpenPixel, we added an extra layer at the end of the closed set CNN responsible for thresholding each pixel-wise prediction as well as a layer to filter False Positives, as shown in Figure 1. This network receives an image as input and processes it using three layers, each one composed of convolutional operation, Rectified Linear Unit (ReLU) activation, and max pooling. Then, three Fully Connected (FC) layers further process the activation maps from the convolutional blocks to classify each patch’s central pixel as pertaining to some class.

This architecture is the same as Pixelwise for the closed set scenario, adding a probability threshold after the softmax. To do so, a pixel with a class confidence (given by the softmax) that exceeds a determined threshold is labeled as belonging to that class. However, if the pixelwise probability is inferior to the threshold, the pixel is classified as unknown. As the value of probability given by the softmax varies between 0 and 1, the possible values of threshold also vary between 0 and 1.

### B. Morph-OpenPixel: Morphological Filtering

After the result predicted at the softmax layer and applied the threshold by the network, a post-processing morphological filter is applied at the pixels classified as unknown.

The applied filter can be seen as an erosion done over the unknown pixels, with the only difference being its adaptation for a multi-class context. According to [11], the process of erosion of a set  $A$  by another set  $B$  (both in  $Z^2$ ) is defined as the set of all points in  $z$  such that  $B$ , translated by  $z$ , is contained in  $A$ , as represented on the Equation 1:

$$A \ominus B = \{z | (B)_z \subseteq A\} \quad (1)$$

Since this process of erosion is only applied over the pixels classified as unknown by the network, the technique reduces the amount of false unknown labels created by the uncertainty

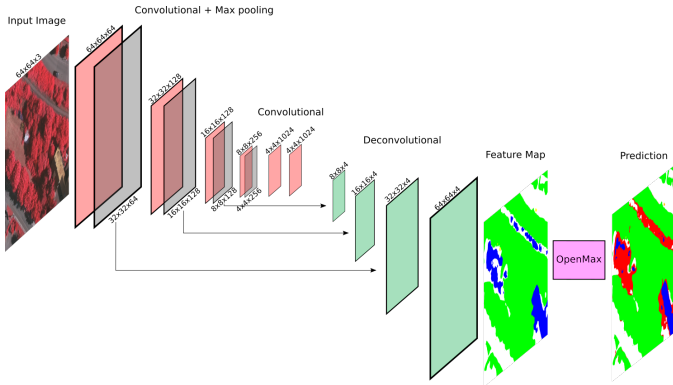


Fig. 2. Simplified architecture view of the OpenFCN method.

of boundary regions. The rightmost modules in Figure 1 shows a didactic example of the application of the morphological filter.

### C. OpenFCN: Open Set Fully Convolutional Networks

The original application of the OpenMax technique was to develop a method that could discern between images generated by computers and real images. The idea described on [4] is that some images can be generated to fool classification networks, simulating features similar to real images. However when comparing the distribution of the classes and their similarities, the OpenMax method could be able to classify an image as a fooling image. In that case, this classification would set the image class as unknown, i.e. not belonging to any known class.

Similarly to most of the other open set techniques, Openmax was based on scene classification, only deciding the general label of the image, not pixel by pixel. So the method OpenFCN uses the same idea of an OpenMax layer, but adapted to an FCN, a pixel-by-pixel approach.

Figure 2 illustrates the architecture of the fully-convolutional neural network used in this work, FCN, adapted for the open set scenario. To achieve the OpenFCN, we substituted the classification layer, softmax, for the OpenMax. This network receives an image as input. This image passes through 6 layers of convolutions and 4 layers of max pooling, to then be deconvoluted 4 times, using information from earlier layers, and then reach the output layer, OpenMax, which is responsible for classifying each pixel and deciding if it belongs to a known or unknown class.

It is valid to observe that even though this work uses an FCN as the network to extract features for each pixel, the architecture of OpenFCN allows the use of any pixel-wise network. The method becomes flexible when observed that the Openmax layer is applied during the testing phase, being isolated from the training.

The classification of the pixels happens in the OpenMax layer. This layer is based on the distance between each evaluated pixel and the distributions of the known classes, being these Weibull distributions. Another way of describing the OpenMax layer is as a classification layer that uses similarity between the activation vector of the pixel, its features, and the

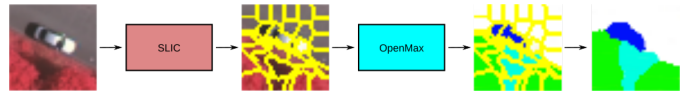


Fig. 3. Example of the process of applying SLIC to the input image, classifying with OpenFCN and using SLIC to return the prediction to the original image. In this example, after the OpenFCN method, white represents the class street, dark blue represents a car, green represents trees and light blue represents grass.

activation vectors of the pixels belonging to known classes to label the pixels.

For the OpenMax layer to correctly classify the pixels, it needs to use the appropriate parameters for the Weibull distribution and calculation of distances. Parameters of Weibull distribution are estimated based on the distance between each correctly classified training example, obtaining a class-specific distance distribution. The exact length of tail for estimating parameters of Weibull distribution is determined during the parameter estimation phase over a small set of the data.

In a different aspect, the alpha rank represents the number of top classes that are considered during the recalibration of the classes. Those top classes are defined as the classes with the biggest variance between the activation maps, in other words, the classes that better represent the diversity of the dataset.

One of the biggest problems that surface when adapting a method based on scene detection to semantic segmentation is the increase in computational time required to analyze each pixel from the image.

This bottleneck is on the stage in which the technique calculates the distances of each pixel to all the class distributions and, after allocating that pixel to a predicted class, redistributes the class distribution, needing to follow a sequential process of finding a pixel class, before evaluating the next one.

### D. SLIC-OpenFCN: OpenFCN Using Super Pixels

The technique utilized in this work, to accelerate the method, was the use of superpixels. To segment, the input image in superpixels was used the Simple Linear Iterative Clustering (SLIC) method [12].

For this segmentation method, different values of superpixels were tested, but the chosen value that increases the speed of the algorithm while maintaining some individual information was a group of 20 pixels for each superpixel. That means that, on average, every 20 pixels from the original image would be represented by a single one in the final image, and this way the method would need to evaluate 20 times fewer pixels for each image. After the prediction by the Openmax layer, the method reverses the superpixel, copying the prediction of the superpixel for each one of the 20 original pixels. This process is represented by one example in Figure 3.

## IV. EXPERIMENTAL SETUP

In this section, we introduce the configuration used during the experiments and needed to guarantee the reproducibility of results. Section IV-A describes the Vaihingen dataset used in our experimental procedure, Section IV-B presents the protocol

for training and testing and, finally, Section IV-C introduces all the metrics used for quantitative evaluation.

### A. Dataset

The Vaihingen dataset<sup>1</sup> contains 33 patches of different sizes, each consisting of a True OrthoPhoto (TOP) extracted from a larger TOP mosaic.

These 33 patches were captured over the city of Vaihingen in Germany by the German Society for Photogrammetry and have a ground sampling distance of 9 cm. The Ground Truth consists of 5 classes: street, building, grass, tree, and car.

The dataset was created to be well controlled and avoid areas without data. To do so, the patches were selected from the central part of the mosaic and not from the boundaries. Even with this approach, some small missing information could occur, to prevent that to happen, interpolation is used to fill all the gaps. The TOP is 8 bit TIFF files with three bands, being three RGB bands, corresponding to the near-infrared, red and green.

### B. Training/Predicting Protocol

For the Vaihingen dataset, we followed the protocol proposed by [13]. Precisely, the dataset is divided into two sets, one for training and one for testing. The testing set consists of images from the patches 11, 15, 28, 30 and 34, whereas the remaining images are used as training.

During the experiments, 4 out of the 5 labels of the dataset are used as known classes (employed to train the model) while the remaining one is exploited as the unknown class. Since this work focus is to demonstrate the concept, it will follow this protocol, but future works can extend it to allow a composition of categories as unknown class.

To better evaluate the proposed method, the protocol also varied the class used as unknown. In each iteration of the experiments, a different class of the dataset was not fed to the algorithm during the training stage, only being seeing during the testing. Figure 4 represents the variation on the unknown class.

The validation set, only used by the OpenPixel method to determine the best threshold values, is composed of a part of images from the training set, excluded from the training.

In the predicting phase, each input image is processed independently by the trained deep model, which outputs a final prediction map in which for each pixel a label indicates whether or not the pixel belongs to a known class, and if it belongs, to which class. This stage uses and presents the one class not seen during learning as the unknown class.

It was analyzed four different contexts with all the networks. The first one was training and testing the networks as the traditional closed set. In this case, the methods are trained and evaluated considering all existing classes of the dataset. This context was only used to show the accuracy of the method without the open set concept.

<sup>1</sup><http://www2.isprs.org/commissions/comm3/wg4/2d-sem-label-vaihingen.html>

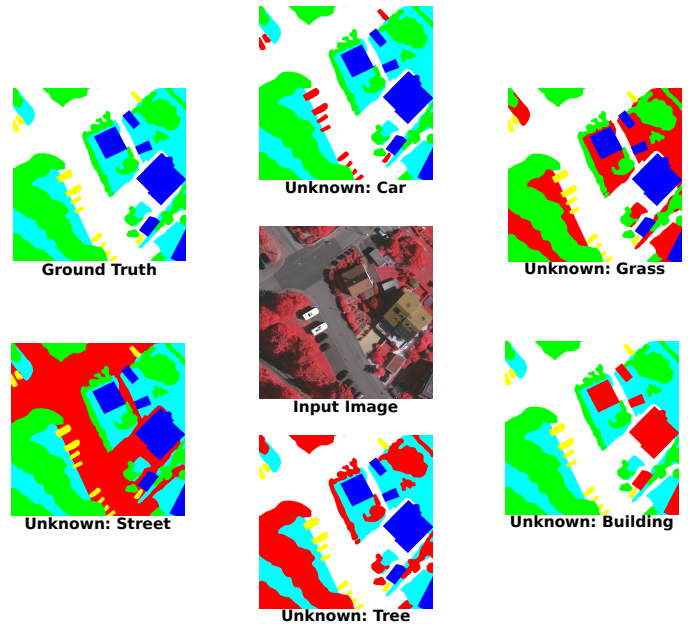


Fig. 4. Representation of the protocol used to vary the unknown class during the experiments.

The second context was training the model as a closed set, but testing it in an open set scenario. In this scenario, the model must classify, during the testing phase, pixels from classes that do not exist training set, resulting in an misclassification of such samples. This scenario, along the next one, shows the relevance of the open set concept for semantic segmentation.

The third one is training and testing the method in open set scenarios. In this case, the network knows it will analyze some pixels from not known classes during the training phase and will be able to classify them as unknown.

The last one is very similar to the third one, the only difference is that it applies the morphological filter to enhance the prediction and mitigate some False Positives on the OpenPixel methodology and uses the SLIC to reduce the time consumption of the OpenFCN method.

### C. Metrics

All results obtained in this work are reported using Cohen's Kappa Index, Overall and Normalized Accuracy scores, given that these metrics take into account the existence of multiple classes and the importance of correct segmenting all of them [14].

## V. RESULTS AND DISCUSSION

The conducted experiments (and results) presented in this section aim to answer the following questions: (1) Is the OpenPixel able to semantic segment a remote sensing image? (2) Is the OpenFCN able to semantic segment a remote sensing image?

### A. OpenPixel Evaluation

Considering the analysis performed on the previous section, we analyze the effectiveness of the proposed method to

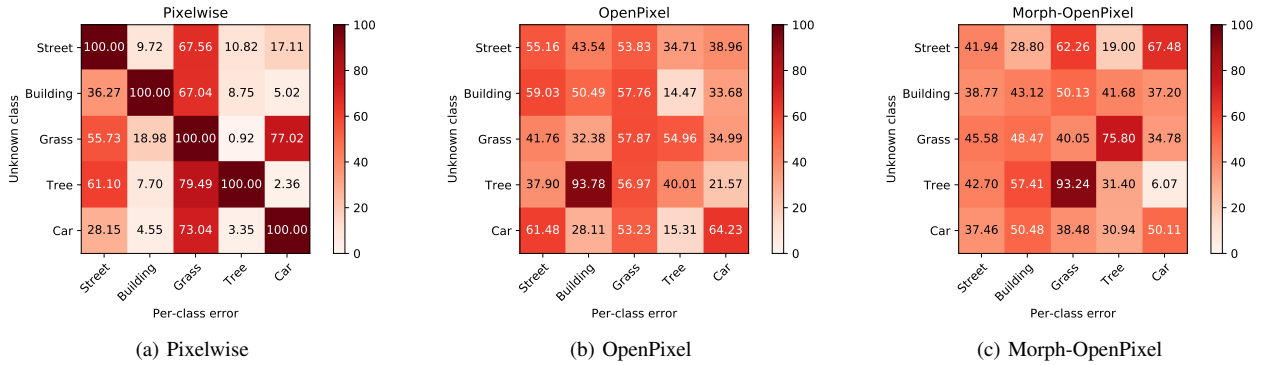


Fig. 5. Per-class error rates for each unknown class according to the closed set baseline (a) and the proposed open set methods (b-c).

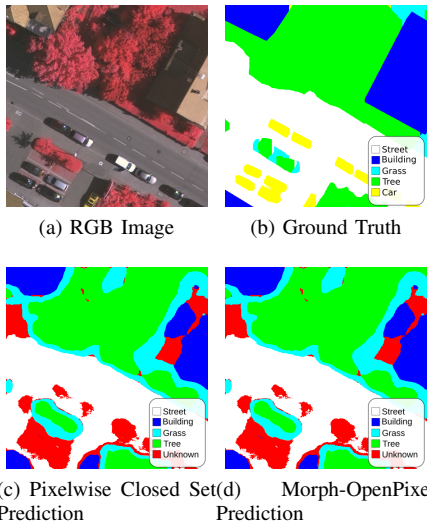


Fig. 6. Example wherein Pixelwise wrongly classifies the unknown class (car) as known, while the Morph-OpenPixel classifies the areas as unknown.

perform open set semantic segmentation. Table I presents the obtained results.

Through the table, it is possible to note that the proposed method, OpenPixel, yielded acceptable results, showing the feasibility of combining semantic segmentation and open set concepts.

TABLE I  
NORMALIZED ACCURACY AND KAPPA INDEX OBTAINED BY THE  
OPENPIXEL METHOD AND BASELINES.

Network	Overall Accuracy	Normalized Accuracy	Kappa
Pixelwise (closed)	55.84%	53.98%	0.5585
OpenPixel	55.78%	53.15%	0.5106
Morph-OpenPixel	<b>57.51%</b>	<b>54.23%</b>	<b>0.5602</b>

Analyzing the results shown in Table I, one can observe that the OpenPixel method trained on an open set configuration achieved similar results to the Pixelwise closed set when tested on an open set scenario.

Even though the open set methods only surpassed the closed set architecture with morphology filtering (Morph-OpenPixel) the simpler version (OpenPixel) also has the benefit of finding unknown classes that are always mislabeled by closed set methods in this scenario.

This advantage is better demonstrated in Figure 5, in which the Error Rate is reported. Observing the error rate of the Pixelwise network, more precisely the main diagonal, one can notice that in all cases the error is 100%, an expected outcome, since the network was not designed to deal with unknown classes, misclassifying all the pixels belonging to this class. When comparing the main diagonal from all three methods, it is possible to understand the advantage of using an open set semantic segmentation technique. Both proposed methods have error rate lower than the Pixelwise network for the unknown classes, while keeping the error rate low on most known classes.

Aside from this, since each class has distinct patterns, it is essential to evaluate the proposed method by varying the unknown class. Results for this experiment are presented in Figure 7. Through the table, it is possible to note that Morph-OpenPixel achieved better metric rates on most of the setups, including the ones in which all the networks had a worse performance.

Visual results are presented in Figure 6, in which the ground truth had known and unknown classes, being the car class (in yellow) the unknown. The Closed Set Pixelwise technique wrongly classified all the pixels belonging to cars, as was expected, since it does not know this class. The prediction resulting from the Morph-OpenPixel, as it can be noticed, has most of the instances of the known classes classified correctly, while still classifying the car pixels as unknown (in red).

### B. OpenFCN Evaluation

Table II that presents the results obtained using the overall and normalized accuracy and kappa index metrics. Analyzing it, it is possible to note that the proposed method, OpenFCN, obtained results that also shows that semantic segmentation can be applied to open set scenarios.

Now, observing the results from OpenFCN at Table II, is possible to analyze that the method outperforms the closed set

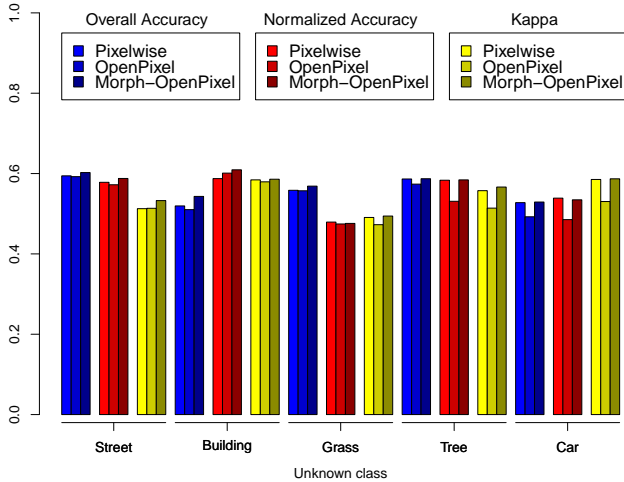


Fig. 7. Overall Accuracy, Normalized Accuracy and Kappa Index obtained by the Pixelwise network and the proposed methods, OpenPixel and Morph-OpenPixel, for each unknown class in the experimental procedure.

TABLE II  
NORMALIZED ACCURACY AND KAPPA INDEX OBTAINED BY THE OPENFCN METHOD AND BASELINES

Network	Weibull Tail Size	Alpha Rank	Overall Accuracy (%)	Normalized Accuracy (%)	Kappa
FCN [15]	-	-	74.76	56.86	0.6648
OpenFCN	125,000	1	47.82	45.91	0.3774
SLIC-OpenFCN	1,000,000	4	80.77	63.92	0.7437
OpenFCN	1,000,000	4	82.27	64.39	0.7630

technique when applied to an open set scenario. Similarly, as the OpenPixel, the method can correctly classify the pixels belonging to known classes, and still classify most of the unknown pixels correctly. Figure 8 shows one example of semantic segmentation done by OpenFCN when compared to the application of the closed set FCN [15] on an open set scenario.

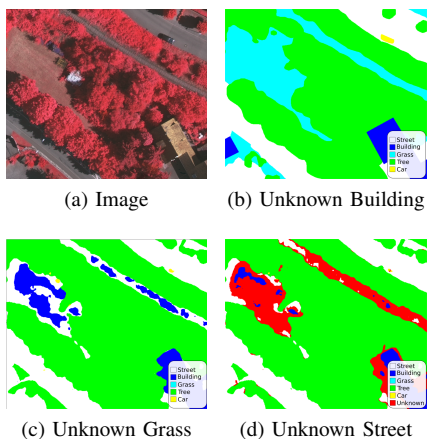


Fig. 8. Example of prediction by the OpenFCN method with the Grass as unknown class. The color white, dark blue, green and yellow represents the class street, building, tree and car, respectively and red represents the predict unknown class, grass, colored as light blue on the ground truth.

The OpenFCN method uses known classes to calculate distances of each pixel being evaluated and uses that to classify them. For this reason, depending on which classes are being used for the algorithm to learn and which classes are set as unknown the results may vary. For this reason, the method needed to be tested against all the cases, and the results for the instance in which each class is set as unknown is presented in Table III.

TABLE III  
OVERALL, NORMALIZED ACCURACY AND KAPPA INDEX OBTAINED BY THE OPENFCN METHOD FOR EACH CLASS AS UNKNOWN

Unknown Class	Overall Accuracy (%)	Normalized Accuracy (%)	Kappa
Street	84.95	68.02	0.7921
Building	83.93	66.08	0.7782
Grass	78.99	59.96	0.7294
Tree	79.86	60.82	0.7343
Car	83.59	67.11	0.7811

The results on Table III shows a lower accuracy and kappa when the classes grass or tree are set as unknown. It is easy to understand the reason when observed the concept of the method. When using grass as an unknown class, the class tree is learned and used to calculate the distances, and since the classes grass and tree have similar features, the method tends to label grass pixels as the class tree, since it learned tree features.

## VI. CONCLUSION AND FUTURE WORKS

The method presented in this paper, dubbed OpenPixel, presented acceptable rates of normalized accuracy when compared to closed set methods on the same dataset. On average, the open set scenario method presented an overall accuracy of 57.51%, a normalized accuracy of 54.23% and a kappa index of 0.4600.

The OpenFCN method presented good results of normalized accuracy when compared to closed set methods on the same dataset and even better when compared to the OpenPixel. On average the open set scenario method presented an overall accuracy of 82.27%, a normalized accuracy of 64.39% and a kappa index of 0.7630.

Observing the experiments and the results presented in this paper, it is possible to affirm that the proposed methods are effective in semantically segmenting pixels belonging to unknown classes, while still correctly classifying pixels from known classes, performing an open set semantic segmentation on remote sensing images.

In conclusion, this paper main contributions are: (1) a discussion of the related works, showing evidence that the semantic segmentation techniques can be applied to open set scenarios; and (2) the development of four methods for open set semantic segmentation. The dissertation also resulted in two publications, one entitled Towards Open-Set Semantic Segmentation of Aerial Images [16], accepted in the 2020 Latin American GRSS ISPRS Remote Sensing Conference, and the second one entitled Fully Convolutional Open Set Segmentation [17] and submitted to Machine Learning Journal.

## REFERENCES

- [1] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T. E. Boult, "Toward open set recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1757–1772, 2013.
- [2] K. Nogueira, M. Dalla Mura, J. Chanussot, W. R. Schwartz, and J. A. dos Santos, "Learning to semantically segment high-resolution remote sensing images," in *2016 23rd International Conference on Pattern Recognition (ICPR)*. IEEE, 2016, pp. 3566–3571.
- [3] A. Bendale and T. E. Boult, "Towards open set deep networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1563–1572.
- [4] A. Bendale and T. Boult, "Towards open world recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015.
- [5] P. R. Mendes, R. M. de Souza, R. d. O. Werneck, B. V. Stein, D. V. Pazinato, W. R. de Almeida, O. A. Penatti, R. d. S. Torres, and A. Rocha, "Nearest neighbors distance ratio open-set classifier," *Machine Learning*, vol. 106, no. 3, pp. 359–386, 2017.
- [6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [7] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [8] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.
- [9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *CVPR*, 2016, pp. 770–778.
- [10] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *CVPR*, 2017, pp. 4700–4708.
- [11] R. C. Gonzalez and P. Wintz, *Digital image processing(Book)*, 1977, no. 13.
- [12] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.
- [13] K. Nogueira, M. Dalla Mura, J. Chanussot, W. R. Schwartz, and J. A. Dos Santos, "Dynamic multicontext segmentation of remote sensing images based on convolutional networks," *TGRS*, vol. 57, no. 10, pp. 7503–7520, 2019.
- [14] R. G. Congalton and K. Green, *Assessing the accuracy of remotely sensed data: principles and practices*. CRC press, 2008.
- [15] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.
- [16] C. C. da Silva, K. Nogueira, H. N. Oliveira, and J. A. d. Santos, "Towards open-set semantic segmentation of aerial images," *arXiv preprint arXiv:2001.10063*, 2020.
- [17] H. Oliveira, C. Silva, G. L. Machado, K. Nogueira, and J. A. d. Santos, "Fully convolutional open set segmentation," *arXiv preprint arXiv:2006.14673*, 2020.