

# Semantic Segmentation Using PSP Network with Attention Mechanism

Devika K. P.<sup>\*1,2</sup>, Reshmi S. Bhooshan<sup>2</sup>

<sup>1</sup>Department of Electronics and Communication, Vidya Academy of Science and Technology, Technical Campus, Kilimanoor, India.

<sup>2</sup>Department of Electronics and Communication, College of Engineering Trivandrum, Trivandrum, India

## ABSTRACT

Semantic segmentation is a challenging problem in computer vision. In recent years, the performance of semantic segmentation has been considerably enhanced by employing cutting edge technique. This paper presents an advanced semantic segmentation methodology that uses the PSPNet (Pyramid Scene Parsing Net-work) architecture augmented with atrous convolution networks and a spatial attention module. The primary objective is to improve segmentation accuracy by integrating spatial attention mechanisms with the PSPNet framework, in association with atrous convolution networks. The spatial attention module selectively highlights pertinent spatial regions within feature maps, enhancing the ability of the model to capture intricate details crucial for precise segmentation. Experimental evaluations are carried out in two datasets: Stanford Background dataset and the Aerial Semantic Segmentation Drone dataset. This improvement underscores the efficacy of integrating spatial attention mechanisms and atrous convolution networks within the PSPNet architecture for semantic segmentation tasks, propelling advancements in the state-of-the-art performance within this domain.

**Keywords:** Semantic segmentation, PSP Network, Spatial attention module

## INTRODUCTION

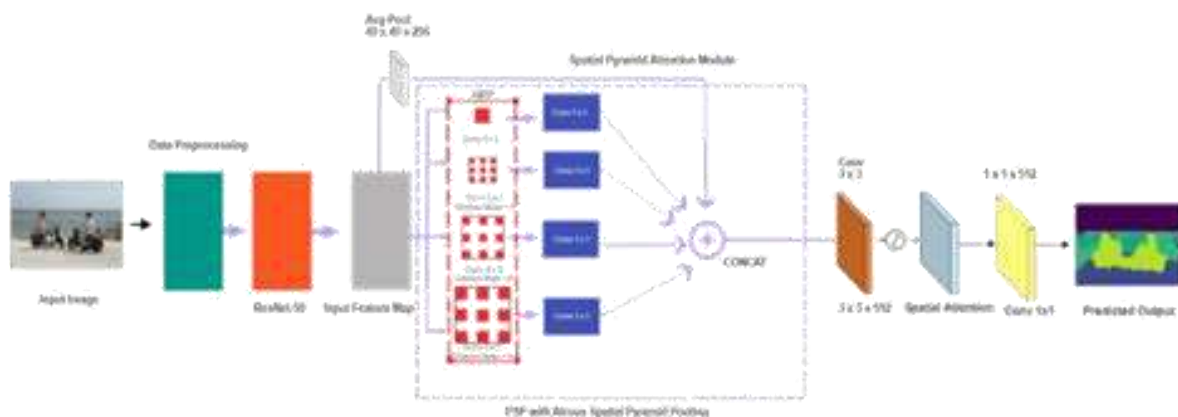
Semantic segmentation is a basic computer vision challenge that involves classifying each pixel in an image into various object such as "person," "car," "building", etc. Semantic segmentation presents a comprehensive knowledge of the scene by dividing it into meaningful parts based on item categories, in contrast to image classification, which gives a single label to a whole image. In computer vision, semantic segmentation [2] plays an important role in different disciplines like autonomous driving and medical imaging. Systems for object detection and recognition are aided by the use of semantic segmentation, and the accuracy of localization and identification is improved by the accurate drawing of object boundaries. It also applies in involving scene understanding, such as autonomous driving, robotics, surveillance systems, and scene parsing, since it helps robots to efficiently understand spatial layouts and semantic contexts. Semantic segmentation enhances diagnosis, treatment planning, and medical research

in medical imaging by analysing MRI, CT images, and histopathology slides. In the field of semantic segmentation, various innovative strategies have evolved to improve accuracy and efficiency. Fully Convolutional Networks (FCNs) [24] replace fully connected layers with convolutional layers to make spatially dense predictions. U-Net, well known for its usage in medical imaging, employs an encoder-decoder design with skip links to preserve localization and contextual information. SegNet has a similar design, but it up samples utilising max-pooling indices from the encoder step. Deep Lab [26] employs atrous convolutions to gather multi-scale contextual information and Atrous Spatial Pyramid Pooling (ASPP) [25] to improve feature extraction. The PSPNet [1] includes a pyramid pooling module for gathering contextual information at multiple sizes, which improves scene interpretation. Semantic segmentation faces several challenges, including handling complex scenes, maintaining fine-grained details, computational efficiency, and dealing with varied object scales. Using PSPNet with atrous convolution and a spatial attention module effectively

**Relevant conflicts of interest/financial disclosures:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

addresses several key challenges in semantic segmentation. Atrous convolution, used in the atrous spatial pyramid pooling (ASPP) module, collects multi-scale information while maintaining resolution, resulting in detailed and high-quality segmentation. The ASPP module is made up of parallel atrous convolutions with varying dilation rates, which assist capture features at different scales and incorporate multi-level contextual information. The paper introduces a spatial attention module to the PSPNet [1] design and Atrous convolution, improving it through the introduction of a unique semantic segmentation technique. The PSPNet, [3] is a well-known architecture designed for semantic

segmentation applications. Its salient feature is its capacity to employ pyramid pooling modules to record contextual information at various sizes, thereby enabling an improved understanding of situations. By including a spatial attention module, this method enhances the capabilities of the PSPNet network and improves its overall performance. The effectiveness of this suggested strategy in improving segmentation accuracy and performance has been extensively tested on the Stanford Background Dataset [6], [11]. The PSPNet obtains context from a variety of receptive fields by dividing the input feature map into sub-regions and carrying out pooling operations with varying kernel sizes.



**Fig. 1 Architecture of Proposed Methodology**

Furthermore, in order to maintain spatial resolution while efficiently broadening the receptive field without adding more parameters, PSPNet uses dilated convolutions [3]. Because it can capture both fine-grained information and broad context feature from images. The experimental results indicate that greater segmentation performance results from combining the PSPNet and atrous network with the attention module based on the acquired values. The remaining of the paper discusses about the following. Section II discuss about the previous related work, Section III gives a brief description about the methodology that is used in this research. Section IV shows the experimental analysis and final results; Section V concludes the work.

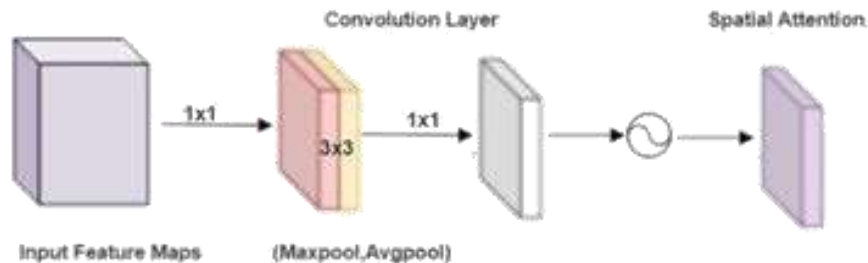
## LITERATURE REVIEW

Semantic segmentation effectively manages contextual information, particularly when attention processes are not optimal. Current study focuses on increasing contextual understanding through attention

mechanisms. To address these issues, new research has proposed novel algorithms that include atrous convolution networks and attention processes into semantic segmentation frameworks. By incorporating attention modules into established frameworks, these solutions can increase contextual comprehension and reactivity to changing lighting circumstances. This section reviews the pivotal contributions and methodologies proposed in recent years, focusing on various network architectures and attention mechanisms that have enhanced the performance and efficiency of semantic segmentation models. In the context of multi-class image segmentation, Guangzhe Zhao et al. [12] introduced an architecture incorporating a bilateral U-Net network model with a spatial attention mechanism. In this model, a lightweight MobileNetV2 as the backbone network for feature hierarchical extraction and proposes an Attentive Pyramid Spatial Attention (APSA) module. However, the model generalization of ASPP (Atrous Spatial Pyramid Pooling) is poor and does not allow

effective segmentation in a dataset with many categories and feature information. To correct detailed information by merging pooling index maps from the encoder with high-level feature maps, Qi Zhao et al. [13] introduced an end-to-end attention-

based semantic segmentation network known as SSAtNet. Here employing and capturing fine-grained details is more complex so, a semantic segmentation [14] algorithm integrating an attention mechanism



**Fig. 2 Spatial Attention Module**

is introduced. The dilated convolution is employed to preserve image resolution and capture detailed information. Recently Spatial Pyramid techniques are used for pixel labeling and which is combined with attention mechanism. Hanchao Li et al. [15] introduced a method, that replace the complicated dilated convolution operations and perform better segmentation. Also a Successive Pooling Attention Module (SPAM), and a Feature Fusion Module (FFM) [16] is used to extract high-level features and low-level features using the initial 11 layers of ResNet50. The importance of semantic segmentation has been highlighted in number of recent works [17], [15]. But does not focus on managing contextual information and segmentation effectiveness for small target object is very poor. For addressing edge splitting and small object disappearance issues in complex scene images, attention modules were chosen to fix. Therefore, they improve the model's ability to focus on crucial information by dynamically weighting different regions of the input dataset. From the previous works, managing contextual information is challenging, especially when attention mechanisms are not operating at their best. And also, Limited adaptability to varied lighting conditions. So Attention modules were chosen to solve the problems because they increase the model's capacity to focus on key information by dynamically weighting various sections of the input data. By doing so, they improve the extraction of critical spatial and contextual information, resulting in more accurate and robust performance.

### 3 Model Architecture

The proposed method improves the PSPNet architecture by adding a spatial attention module to complement the pyramid pooling modules and dilated convolutions. These components make feature extraction easier by gathering contextual data at different scales and extending the receptive area while maintaining spatial accuracy. The spatial attention module improves precise segmentation by emphasising key spatial locations in feature maps. The Fig: 1 illustrates a Spatial Pyramid Attention Module for semantic segmentation, which starts with an input image processed by a feature extraction network, such as a data preprocessing block and a convolutional neural network (CNN) that is ResNet-50. This network collects important elements from the picture, producing an initial feature map that serves as the basis for further actions. This first feature map passes via spatial pyramid pooling (SPP), which is a critical component for collecting context at many sizes. The PSPNet [5] uses this SPP to improve feature representation. In PSPNet, spatial pyramid pooling is achieved via average pooling to acquire global context information by downscaling the feature map and applying atrous convolutions (also known as dilated convolutions) with various dilation rates. On accuracy. Convolutions with rates of 3, 6, and 9 are used to capture characteristics at varied spatial resolutions, emphasising medium-range to wide contextual information. And a 3x3 convolution is used to reduce the dimensionality of the feature map, ensuring both local context capture and

dimensionality reduction. Finally, the up sampled features from all area sizes are concatenated [7] along the channel dimension and fed into a 1x1 convolutional layer [4]. Table 1 shows the detailed architecture of proposed methodology. In this study, a spatial attention module (Fig: 2) [21] is incorporated that enhances the performance of the model for semantic segmentation. By employing this spatial

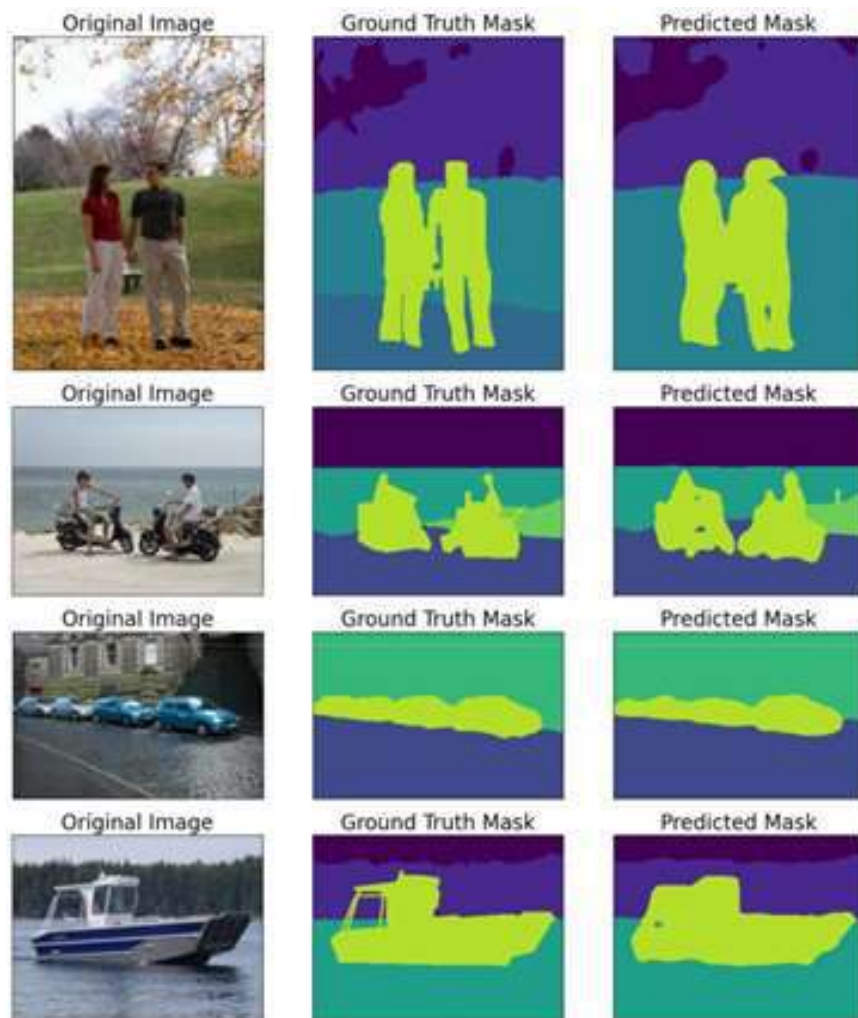
attention module to emphasise relevant spatial regions within the feature maps, the model may prioritise features that are critical for precise segmentation. Overall, the accuracy and performance of the semantic segmentation model are enhanced by the inclusion of the spatial attention module [9], [10] , making it more valuable in real-world applications

**Table 1 Details of Proposed Architecture**

Layer	Input Shape	Output Shape	Filter Size
Input Image	(320,320,3)	(320,320,3)	-
ResNet-50	(320,320,3)	(40,40,2048)	Multiple
Avg Pool	(40,40,2048)	(40,40,256)	-
ASPP Conv 1x1	(40,40,2048)	(40,40,256)	1x1x2048
ASPP Conv 3x3 (r=3)	(40,40,2048)	(40,40,256)	3x3x2048
ASPP Conv 3x3 (r=6)	(40,40,2048)	(40,40,256)	3x3x2048
ASPP Conv 3x3 (r=9)	(40,40,2048)	(40,40,256)	3x3x2048
ASPP Concatenation	(40,40,1024)	(40,40,1024)	-
ASPP Conv 1x1 (red.)	(40,40,1024)	(40,40,512)	1x1x1024
Global Avg Pool	(40,40,512)	(1,1,512)	40x40x512
Conv 3x3	(40,40,512)	(40,40,512)	3x3x512
Upsample to Input	(40,40,512)	(320,320,512)	-
Spatial Attention	(320,320,512)	(320,320,512)	-
Final Conv 1x1	(320,320,512)	(320,320,9)	1x1x512

Using spatial pyramid pooling, the feature maps from various scales are concatenated along the channel dimension. This stage is critical because it merges multi-scale contextual information into a single complete feature map that includes both local and global features at multiple spatial resolutions. The merged feature map is then analysed by a spatial attention mechanism. This technique guarantees that the network focuses on the critical sections of the feature map that are important to the segmentation job, boosting the accuracy and resilience of the segmentation predictions. The incorporation of the PSPNet [19] into this module takes use of its capability in gathering different contextual information, improving the spatial attention mechanism's capacity to refine and emphasise important aspects. The improved feature map generated by the spatial attention mechanism goes

through a final 1x1 convolution step. This layer is critical because it decreases the number of channels in the feature map to match the number of output classes needed for the segmentation process. The 1x1 convolution guarantees that the feature map is properly formatted for producing accurate segmentation predictions. The end result is a predicted segmentation map, in which each pixel in the input picture is assigned a class label matching to the recognised objects or areas. This complete strategy, which combines multi-scale context capture, feature integration, and spatial attention refinement, significantly improves the semantic segmentation process, yielding very accurate and detailed segmentation results. The implication of PSPNet to this architecture dramatically improves the network's capacity to grasp and analyse complicated situations by including a wide variety of contextual information into the segmentation process.



**Fig. 3 Result on Stanford Background Dataset (Original Image, Ground Truth, And Predicted image)**

## 4 Experimental Setup

In this section, the comprehensive evaluation findings of the chosen methodology for evaluating the performance of Stanford Background dataset is shown. Also includes a comparison with different approaches.

### 4.1 Dataset

The performance evaluation of the proposed method is done using two datasets:

#### 4.1.1 Stanford Background Dataset

The performance evaluation of the proposed method was conducted through extensive experiments on the Stanford Background Dataset [11], [20], a well-established bench-mark for semantic segmentation tasks. The collection includes 715 images drawn from four public datasets: LabelMe, MSRC, PASCAL VOC, and Geometric Context. The images were

chosen based on the following criteria: they were of outdoor sceneries, had around 320-by-240 pixels, had at least one foreground item, and had the horizon location inside the image. Amazon Mechanical Turk was used to produce semantic and geometric labels.

#### 4.1.2 Aerial Semantic Segmentation Drone Dataset

The Semantic Drone Dataset [23] focuses on semantic understanding of urban scenes for increasing the safety of autonomous drone flight and landing procedures. The imagery depicts more than 20 houses from nadir (bird's eye) view acquired at an altitude of 5 to 30 meters above ground. A high-resolution camera was used to acquire images at a size of 6000x4000px (24Mpx). The training set contains 400 publicly available images and the test set is made up of 200 private images.

## 4.2 Data Preprocessing

The experimental setup includes thorough data preprocessing, augmentation techniques, and model training strategies to ensure robust and reliable results. Prior to training, the dataset underwent preprocessing by resizing images uniformly, normalizing pixel values, and converting pixel-wise annotations into suitable formats for model training. Augmentation methods such as random flips, rotations and scaling were utilized to diversify the training data, thereby enhancing the model's generalization across various scenarios. In the training phase, best practices such as transfer learning with a pre-trained backbone network, learning rate scheduling, and early stopping are implemented to stabilize training and mitigate overfitting.

### 4.3 Evaluation Metrics

The Performance of the proposed method is evaluated using three metrics- mIoU (Mean Intersection over Union), PA (Pixel Accuracy) and MPA (Mean Pixel Accuracy).

- Mean Intersection over Union (mIoU) : The mIoU metric quantifies the degree of correspondence between the segmentation outcome and the actual value of the original picture.

$$mIoU = \frac{1}{N} \left( \sum_{i=1}^N \frac{X_{ii}}{T_i + \sum_{j=1}^N X_{ji} - X_{ii}} \right)$$

- Pixel Accuracy: The PA is the ratio of the total number of pixels to the number of correctly split

pixels. The following is the precise calculation formula:

$$PA = \left( \frac{\sum_{i=1}^N X_{ii}/T_i}{N} \right).$$

- Mean Pixel Accuracy: The ratio of the sum of correctly classified pixels for each class to the total number of pixels in the image.

$$PA = \left( \frac{\sum_{i=1}^N X_{ii}/T_i}{N} \right).$$

where,

- N is the number of classes,
- $X_{ii}$  is the number of true positive pixels for class i,
- $C_{ii}$  is the number of correctly predicted pixels for class i,
- $T_i$  is the total number of pixels for class i,
- $X_{ji}$  is the number of pixels predicted as class i that are actually class j.

### 4.4 Experimental Results and Analysis

Table 2 presents the results of semantic segmentation [22] on the Stanford Background Dataset, where the mIoU metric is employed to evaluate the accuracy of object shape prediction by the proposed models. Our proposed method is achieving the mIoU of 85.25%. Comparing to previous method in [6], we got the accuracy 7.84% higher than the achieved one.

**Table 2 Performance comparison of**

Method	mIoU (%)
DeepLabV3 [27]	74.33
DeepLabv3+LoAd [28]	75.05
UNET	64.2
Proposed method	85.25

As shown by obtaining a higher mIoU score in Table 3 [22], the proposed method improves previous methods in semantic segmentation tasks. In the

proposed method, we use ResNet-50 [18] as the backbone architecture which captures multi-scale features through its hierarchical structure.

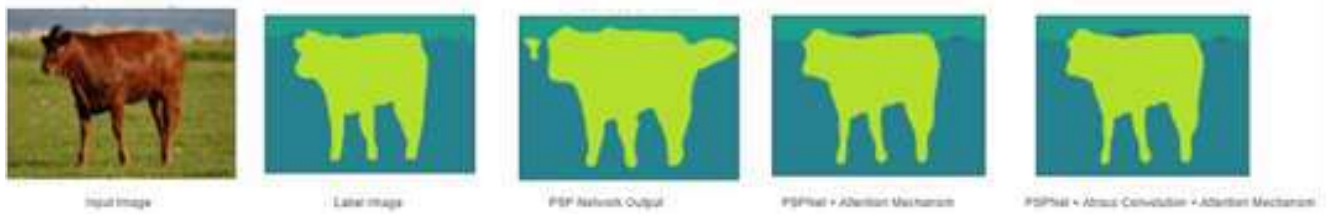


Fig. 4 Results of the three approaches

Table 3 Performance comparison of different methods on semantic segmentation

Method	Backbone	mIoU (%)
FCN-8s [4]	-	65.3
PSPNet [3]	ResNet-101	43.29
PSANet [28]	ResNet-101	43.7
RefineNet [29]	ResNet-101	73.6
DeepLab V2 [30]	ResNet-101	70.4
Proposed method	ResNet-50	85.25

In fig:3 shows the results obtained from the Stanford Background Dataset indicate an improvement in segmentation performance using the proposed methodology and resulting in an accuracy of 83.04%. The Table 4 compares DeepLabv3+, Adversarial, and the proposed PSP approach on the Stanford

Background Dataset, focusing on pixel accuracy across classes and mIoU. The proposed PSP method outperforms the previous methods, with greater accuracy in difficult classes such as Mountain (61.66%) and an overall mIoU of 80.04. This indicates that the PSP technique is more successful for semantic segmentation on this data set.

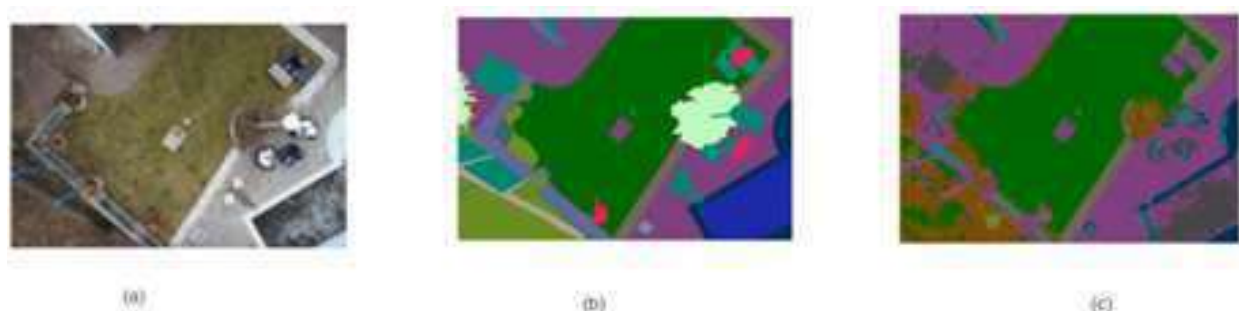


Fig. 5 Result on Aerial Drone Dataset (Original Image, Ground Truth, And Predicted image)

Table 4 Performance Comparison on Stanford Background Dataset

Method	Sky	Tree	Road	Grass	Water	Building	Mountain	Foreground	mIoU%
DeepLabv3+ [27]	89.38	72.21	87.28	77.44	72.70	80.03	48.64	66.92	74.33
Adversarial [6]	89.35	72.54	87.31	77.53	72.78	80.04	49.18	66.69	74.43
Proposed Method	74.04	84.19	92.28	87.72	90.14	78.83	61.66	78.65	80.04

## 5 Ablation Study

In this study on semantic segmentation using the Stanford Background dataset, we performed an ablation study to evaluate the impact of various network enhancements. Initially, using a PSPNet, we achieved a mIoU of 74.3%. To improve accuracy, we

integrated an attention mechanism into the PSPNet, which increased the mIoU to 76%. Further enhancing the network, we combined PSPNet with atrous convolution and an attention mechanism, specifically incorporating a spatial attention module. This approach significantly boosted the mIoU to 85.25%.

**Table 5 Ablation Study Results**

Method	mIoU (%)
PSPNet	74.3
PSPNet + Attention Mechanism	76.0
PSPNet + Atrous Convolution + Attention Mechanism (with Spatial Attention Module)	85.25

Fig: 4 shows the output result for the three approaches. PSP network with attention mechanism and PSP network with atrous convolution and attention module produce remarkable segmentation output. To further analyze the performance of the proposed approach, we compare with another dataset

known Aerial Semantic Segmentation Drone Dataset [23]. Here the experimental analysis done only with PSPNetwork and got the mIoU of 63.0%. And for Stanford Background Dataset on PSP Network achieves the mIoU of 74.3%. Fig: 5 shows the result on Aerial Drone Dataset. Table 6 shows the Comparison of two datasets.

**Table 6 Comparison of Methods on Different Datasets**

Dataset	Method	mIoU (%)
Aerial Drone Dataset	PSPNet	63.0
Stanford Background Dataset	PSPNet	74.3

## CONCLUSION

The findings of this paper underscore the efficacy of leveraging the PSP network in semantic segmentation tasks, augmented with essential mechanisms for feature extraction and context understanding. By integrating a spatial attention module alongside an atrous convolution network, the approach demonstrates substantial improvements in segmentation accuracy and performance. Through experimental analysis, we achieved an impressive segmentation accuracy of 83.04%, showcasing the effectiveness of the incorporated features in capturing contextual information at multiple scales. The utilization of spatial attention mechanisms allows the model to selectively emphasize relevant spatial regions, while atrous convolution facilitates the extraction of contextual features crucial for accurate segmentation. Thereby offering improved accuracy and robustness for various computer vision applications. Future work might include incorporating more advanced attention processes, improving multi-scale context aggregation, and optimising the model for real-time applications.

## REFERENCE

1. Zhao, Qi, et al." Semantic segmentation with attention mechanism for remote sensing images." IEEE Transactions on Geoscience and Remote Sensing 60 (2021): 1-13.
2. T. Pham," Semantic Road Segmentation using Deep Learning," 2020 Applying New Technology in Green Buildings (ATiGB), Da Nang, Vietnam, 2021, pp. 45-48, doi: 10.1109/ATiGB50996.2021.9423307. keywords: Deep learning; Image seg-mentation; Green buildings; Roads; Semantics; Neural networks; Autonomous auto-mobiles; deep learning; semantic segmentation; computer visions; convolutional neu-ral networks.
3. Zhao, Heng Shuang, et al." Pyramid scene parsing network." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
4. Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
5. Shaaban, Aya M., Nancy M. Salem, and Walid I. Al-atabany." A semantic-based scene segmentation using convolutional neural networks." AEU-International Journal of Electronics and Communications 125 (2020): 153364.

6. Luc, Pauline, et al." Semantic segmentation using adversarial networks." arXiv preprint arXiv:1611.08408 (2016).
7. Kim, Dong Seop, et al." ESSN: Enhanced semantic segmentation network by residual concatenation of feature maps." *IEEE Access* 8 (2020): 21363-21379.
8. Niu, Zhaoyang, Guoqiang Zhong, and Hui Yu." A review on the attention mechanism of deep learning." *Neurocomputing* 452 (2021): 48-62.
9. Zhao, Qi, et al." Semantic segmentation with attention mechanism for remote sensing images." *IEEE Transactions on Geoscience and Remote Sensing* 60 (2021): 1-13.
10. Li, Haifeng, et al." SCAttNet: Semantic segmentation network with spatial and channel attention mechanism for high-resolution remote sensing images." *IEEE Geoscience and Remote Sensing Letters* 18.5 (2020): 905-909.
11. Gould, Stephen, Richard Fulton, and Daphne Koller." Decomposing a scene into geometric and semantically consistent regions." 2009 IEEE 12th international conference on computer vision. IEEE, 2009.
12. Zhao, Guangzhe, et al." Bilateral U-Net semantic segmentation with spatial attention mechanism." *CAAI Transactions on Intelligence Technology* 8.2 (2023): 297-307.
13. Zhao, Qi, et al." Semantic segmentation with attention mechanism for remote sensing images." *IEEE Transactions on Geoscience and Remote Sensing* 60 (2021): 1-13.
14. Bai, Wei." An ENet Semantic Segmentation Method Combined with Attention Mechanism." *Computational Intelligence and Neuroscience* 2023 (2023).
15. Li, Hanchao, et al." Pyramid attention network for semantic segmentation." arXiv preprint arXiv:1805.10180 (2018).
16. Sun, Le, et al." SPANet: Successive pooling attention network for semantic segmentation of remote sensing images." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 15 (2022): 4045-4057.
17. Jiao, Yijie, et al." Image semantic segmentation fusion of edge detection and AFF attention mechanism." *Applied Sciences* 12.21 (2022): 11248.
18. Yu, Yang, et al." Multi-scale spatial pyramid attention mechanism for image recognition: An effective approach." *Engineering Applications of Artificial Intelligence* 133 (2024): 108261.
19. Sang, Haiwei, Qiuha Zhou, and Yong Zhao." PCANet: Pyramid convolutional attention network for semantic segmentation." *Image and Vision Computing* 103 (2020): 103997.
20. Kim, Dong Seop, Yu Hwan Kim, and Kang Ryoung Park." Semantic segmentation by multi-scale feature extraction based on grouped dilated convolution module." *Mathematics* 9.9 (2021): 947.
21. Chen, Boyu, et al." Spatiotemporal convolutional neural network with convolutional block attention module for micro-expression recognition." *Information* 11.8 (2020): 380.
22. Minaee, Shervin, et al." Image segmentation using deep learning: A survey." *IEEE transactions on pattern analysis and machine intelligence* 44.7 (2021): 3523-3542.
23. Nedeveschi, Sergiu." A Critical Evaluation of Aerial Datasets for Semantic Segmentation." 2020 IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP). IEEE, 2020.
24. Ozturk, Ozan, Batuhan Saritürk, and Dursun Zafer Seker." Comparison of fully convolutional networks (FCN) and U-Net for road segmentation from high resolution imageries." *International journal of environment and geoinformatics* 7.3 (2020): 272-279.
25. Chen, Jin, Chuanya Wang, and Ying Tong." AtICNet: semantic segmentation with atrous spatial pyramid pooling in image cascade network." *EURASIP Journal on Wireless Communications and Networking* 2019 (2019): 1-7.
26. Wang, Zuoshuai, et al." Multi-scale dense and attention mechanism for image semantic segmentation based on improved DeepLabv3+." *Journal of Electronic Imaging* 31.5 (2022): 053006-053006.
27. Jamali-Rad, H. and Szabo, A., 2021. Lookahead adversarial learning for near real-time semantic segmentation. *Computer Vision and Image Understanding*, 212, p.103271.
28. Zhao, Hengshuang, et al." Psanet: Point-wise spatial attention network for scene parsing."

Proceedings of the European conference on computer vision (ECCV). 2018.

29. Lin, Guosheng, et al.” Refinenet: Multi-path refinement networks for high-resolution semantic segmentation.” Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
30. Chen, Liang-Chieh, et al.” Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs.” IEEE transactions on pattern analysis and machine intelligence 40.4 (2017): 834-848.

**HOW TO CITE:** Nuradeen Abdullahi Yusuf, Danlami Dauda\*, Saudat Bello Adamu, Effect of Entrepreneurship Education on Entrepreneurial Motivation Among Students of Federal Polytechnics in North-West Nigeria, *Int. J. Sci. R. Tech.*, 2025, 2 (11), 647-656. <https://doi.org/10.5281/zenodo.17682633>