

## INFLUENCE OF ADDITIONAL SPECTRAL BANDS FOR EPIPHYTE SEGMENTATION ON DRU-NET

Anjana K Menon<sup>a</sup>, Sajithvariya V V<sup>a,\*</sup>, Ramesh Sivanpillai<sup>b</sup>, Sowmya V<sup>a</sup>, Gregory K Brown<sup>c</sup>,  
K P Soman<sup>a</sup>

<sup>a</sup> Center for Computational Engineering and Networking, Amrita School of Engineering, Coimbatore,  
Amrita Vishwa Vidyapeetham, India - (ORCID: 0000-0002-3756-6724) - anjanakmenon1305@gmail.com, (ORCID:  
0000-0003-3944-8155) - vv\_sajithvariya@cb.amrita.edu,  
(ORCID: 0000-0003-3745-6944) - v\_sowmya@cb.amrita.edu, kp\_soman@cb.amrita.edu

<sup>b</sup> Wyoming GIS Center, University of Wyoming, Laramie, WY 82072, USA - (ORCID: 0000-0003-3547-9464) - sivan@uwyo.edu

<sup>c</sup> Department of Botany, University of Wyoming, Laramie, WY 82072, USA - gkbrown@uwyo.edu

**KEY WORDS:** DRU-Net, epiphyte, segmentation, RGB, grayscale, spectral bands, illumination

### ABSTRACT:

Dense Residual U-Net (DRU-Net) is a neural network used for image segmentation. It is based on the U-Net architecture and is a combination of modified ResNet as the encoder and modified DenseNet as the decoder blocks. DRU-Net captures both the local and contextual information. Previous studies on DRU-Net have not tested the influence of the spectral resolution of the images. In an earlier study, the DRU-Net was trained with grayscale images for epiphyte segmentation. The network trained and tested with grayscale images underperformed while varying the illumination and occupancy of the target in the frame. In this study, the same network was trained and tested with RGB images for assessing the increase in overall learning. The performance of the network in segmenting epiphytes under conditions such as good/poor illumination and high/low target occupancy was analyzed. Dice and Jaccard scores were used as evaluation metrics. The DRU-Net model trained with RGB images had an improvement of 20% over the grayscale model in both average Dice and average Jaccard scores of the target class. Based on the higher Dice and Jaccard scores, adding additional spectral information improves DRU-Net learning. The increased computation time required for training DRU-Net with RGB images will result in better output. This model could be further used for identifying multiple epiphytes in images with poor illumination and different occupancy conditions.

### 1. INTRODUCTION

In image processing, the image is divided into distinct parts with comparable features. As a result, image segmentation may entail grouping pixel areas based on form or color similarity, or separating the target from the background (Padma et al., 2022). Color is vital for the capacity to distinguish artificial and natural objects, and its significance in visual object identification models should be considered (Bramão et al., 2011).

For image segmentation, Ronneberger et al. (2015) proposed transfer-ring feature maps from the encoder to the decoder section using skip connections created with U-Net (Ronneberger et al., 2015). This study demonstrated good performance with a small number of images (Albishri et al., 2022). U-Net integrates low-level detail and high-level semantic information, resulting in promising image segmentation results (Zhang et al., 2018). A study used U-Net architecture for segmentation of lungs in chest x-ray (CXR) images to identify different pulmonary diseases such as pneumoconiosis, covid-19, and tuberculosis. It uses CXR images with low contrast and opacification. In terms of the Dice Coefficient and Jaccard Index, the U-Net performed well on lung segmentation for images with low contrast and opacification (Alam et al., 2021). Several image processing techniques discussed in Panicker et al. (2020) help in visualizing the features to detect the target animal in an unevenly illuminated and occluded image, improving the proposed system's recognition power (Panicker et al., 2020).

Dense Residual U-Net (DRU-Net) was introduced in Jafari et al.'s (2020) work in which it was used for medical image segmentation to segment skin lesions and brain MRI. The Dense Residual U-Net is derived from a basic U-Net architecture. Hence, the<sup>1</sup>

DRU-Net is assumed to be capable of producing good results with fewer image data. It is a basic yet effective network block that outperforms ResNet, DenseNet, and Attention network-based techniques in image segmentation. Compared to ResNet, DenseNet, and Attention network approaches, DRU-Net requires fewer model parameters (Jafari et al., 2020).

Sajjadian et al. (2022) investigated how image processing may be used to identify building components to verify compliance with building standards and detect anomalies. For this, DRU-net was used to segment various elements such as bricks, concrete, and insulation from architectural images. The performance of DRU-Net outperformed U-Net and RU-Net in terms of the Dice score, recall, and precision measures. But the architecture failed when the number of pixels of a certain class was lower than other classes (Sajjadian et al., 2022).

Menon et al. (in press) demonstrated the use of DRU-Net to identify epiphytes in grayscale images. The purpose of this study was to reduce the data volume and segment the epiphytes with very minimal images for training. The performance of the trained model was evaluated with 7 test images and the average target class Jaccard score was 0.594 and Dice score was 0.726. Though these scores are high, they are not sufficient. The model did not perform well when the target epiphyte was in different illumination and different occupancy in the images. To improve the identification and performance metrics, images with more spectral bands i.e., RGB images were used for training the model.

The objective of the study is to evaluate the influence of RGB features in epiphyte segmentation.

\* Corresponding author

## 2. MATERIALS AND METHODOLOGY

In this study, the target epiphyte is *Werauhia Kupperiana*, which belongs to the Bromeliad family. The epiphytes were captured at the Braulio Carrillo National Park in Costa Rica (V V et al., 2019; Aswin et al., 2021). The RGB images in this dataset exhibit target epiphytes at various distances and the images were taken at different times of the day. Different illumination and partially obstructed target epiphytes were present in the images. In the dataset, there were 132 images featuring the target epiphyte. The images were divided into training, validation, and testing sets. The images used for training, validation, and testing were 107, 5, and 20 respectively. Images of epiphytes from the dataset are shown in Figure 1 and corresponding masks are shown in Figure 2.



Figure 1: Example images from the dataset in which *Werauhia Kupperiana* is present with a) poor illumination b) good illumination and occupancy c) low occupancy



Figure 2: Binary annotated ground-truth corresponding to the images in above figure. The target epiphyte is in white and the background is in black

### 2.1 Pre-Processing

All input images were scaled to  $512 \times 512$  pixels to save processing time and ensure uniformity in terms of dimension. The images had three spectral bands (red, blue, and green). For a supervised algorithm, the images had to be masked and annotated. A total of 119 images featuring the target plant were labeled using the LabelMetool as part of previous research (Anivilla et al., 2020). The test set is handpicked such that it contains samples with poor/good illumination and high/low occupancy. These images were annotated using LabelMe in this work. The labels were binarily annotated in white and black, where the background is in black and the target epiphyte is in white. For the dataset to be symmetric and the features to have a similar range for the gradients, the input images were normalized using zero mean.

### 2.2 DRU-Net for epiphyte segmentation

The U-Net is the base architecture for the Dense Residual U-Net (DRU-Net). In addition to the original U-net architecture, each layer has a batch normalization (BN) operation, a well-known approach for achieving quicker convergence and enabling stable



Figure 3: The encoder block of DRU-Net (Jafari et al., 2020)

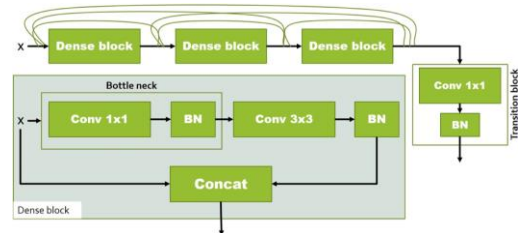


Figure 4: The decoder block of DRU-Net (Jafari et al., 2020)

network training. The encoder (Figure 3) and decoder (Figure 4) blocks of the U-Net are derived from ResNet and DenseNet respectively. ResNet adds a shortcut link from the input to the output of a Conv-ReLU, which differs from standard Conv-ReLU operation. An extra connection is proposed between the output of the first Conv-BN operations and the final Conv-BN output with a summing operation for feature map aggregation, taking into account the benefits and shortcomings of ResNet and DenseNet networks. So, when the gradients in the second Conv approach zeros, this extra shortcut link allows the parameters in the first conv layer to be updated which helps in the back propagation of the gradients. Concatenation is used to combine the input and output for each layer in our method's encoder path. Following that, the merged feature map is sent into the next layer. This concatenation is used not just to include more information from each layer's input, but also to make feature map dimensions directly compatible with feature maps in the following layer without the need for additional parameters. Instead of cropping the input channels to make the dimension consistent with the layer output for a summing operation, a conv 1x1 is employed in the decoder block.

The feature learning process is made more efficient by modifications such as aggregation of feature maps using summation and concatenation of the input to the output at each layer. Since the dataset has images of varying quality and target occupancy, the features learned by DRU-Net will help in obtaining robust features. The output feature map from the network is capable to look for both global and local features. The edges of the target epiphyte and its textural features could be better learned using the feature map aggregation at the output of the decoder block and from the residual blocks of the network. These advantages motivated us to use this network for epiphyte segmentation and also apart from this we study the influence of spectral bands features in epiphyte segmentation.

### 2.1 Evaluation Metrics

The evaluation metrics used in this work are the dice and the Jaccard score which are commonly used for image segmentation tasks. The Jaccard score calculates the average of the intersection over the union of the labeled segments for each class. As a result, the Jaccard score considers both false alarms and missing values for every class. The dice score is a metric that counts how many positives are found, and also penalizes the false positives (Csurka and Larlus, 2013).

The input to the DRU-Net is RGB images from the dataset. The objective of this work is to understand the capability of the network to segment targets under different conditions when trained with images with spectral information and grayscale images.

The jaccard/dice scores for the grayscale and RGB images (n = 20) were compared with a paired t-test at the alpha level of 0.05.

### 3. RESULTS

The parameter settings used for training the DRU-Net are shown in Table 1.

Parameter	Value
Batch size	2
Loss function	Cross Entropy
Optimizer	Adam
Initial learning rate	0.001

Table 1: Parameter settings for training DRU-Net (Jafari et al., 2020)

The network was trained with 100, 200, and 500 epochs. The loss saturated and was not decreasing further after 500 epochs. The models were saved at each epoch to understand the variations in the learning of the network.

The previous work (Menon et al., in press) shows the potential of DRU-Net in segmenting epiphytes from grayscale images and the segmentation score Jaccard and dice for that model are given in Table 2. The performance of DRU-Net in segmenting the target epiphyte in RGB images was compared with that of grayscale images using the dice and the Jaccard score as shown in Table 2.

Epochs	Grayscale		RGB	
	Dice (Class 1)	Jaccard (Class 1)	Dice (Class 1)	Jaccard (Class 1)
100	0.3941	0.2815	0.6576	0.5412
200	0.4671	0.3478	0.6647	0.5467
500	0.4982	0.3720	0.6649	0.5475

Table 2: Average Dice and Jaccard score for class 1 of the gray and RGB models for 100, 200, and 500 epoch model

Table 2 shows the average dice and Jaccard scores of class 1 (target class) for 20 test images. The scores show that the model was able to segment the target epiphyte in RGB images better than in grayscale images. The intermediate models are saved here to capture the minor changes in model performance. The model was settled at 500 epochs.

To understand DRU-Net’s performance and influence of spectral bands in images with the target under different conditions of illumination and occupancy, the test set was categorized into good illumination and high occupancy poor illumination and high occupancy good illumination and low occupancy poor illumination and low occupancy. The below images show the performance of the model with samples of images from each category.

	INPUT IMAGE	GROUND TRUTH	PREDICTED IMAGE (Gray scale)	PREDICTED IMAGE (RGB)
a				
	<b>Dice</b>		0.951	0.956
	<b>Jaccard</b>		0.906	0.922
b				
	<b>Dice</b>		0.707	0.930
	<b>Jaccard</b>		0.547	0.868

Figure 5: Images that have the target epiphyte under good illumination and high occupancy

Figure 5 has the target with high occupancy and is accurately segmented. The tree’s bark was segmented greatly by the grayscale model whereas the bark was less segmented in the RGB model.

	INPUT IMAGE	GROUND TRUTH	PREDICTED IMAGE (Gray scale)	PREDICTED IMAGE (RGB)
a				
	<b>Dice</b>		0.619	0.801
	<b>Jaccard</b>		0.448	0.669
b				
	<b>Dice</b>		0.628	0.861
	<b>Jaccard</b>		0.458	0.756

Figure 6: Images that have the target epiphyte under good illumination and high occupancy

Figure 6 shows the images that have the target epiphyte under good illumination and low occupancy. The dried leaves in the epiphyte Figure 6.b) were segmented in the grayscale model due to a lack of color information and it was overcome by the RGB model. The dice and Jaccard scores for these images were not so high because of the less occupancy of the target epiphyte in these images.

Figure 7 shows images that have target epiphyte under poor illumination and high occupancy. Even though the target had high occupancy Figure 7.b), both RGB and grayscale models could not segment the target in the first image where there were multiple epiphytes at different distances from the camera. And also, the poor illumination affected the segmentation of both models but the RGB model had better scores when compared to the grayscale model. In Figure 8, the targets with low occupancy were not fully segmented by both models. According to the dice and Jaccard scores, the poor illumination in Figure 8.b) degraded both RGB and grayscale models’ ability to segment the target epiphytes accurately.









	INPUT IMAGE	GROUND TRUTH	PREDICTED IMAGE (Gray scale)	PREDICTED IMAGE (RGB)
a				
	Dice		0.709	0.883
	Jaccard		0.549	0.791
b				
	Dice		0.300	0.280
	Jaccard		0.177	0.162

Figure 7: Images that have the target epiphyte under good illumination and high occupancy









	INPUT IMAGE	GROUND TRUTH	PREDICTED IMAGE (Gray scale)	PREDICTED IMAGE (RGB)
a				
	Dice		0.712	0.795
	Jaccard		0.553	0.660
b				
	Dice		0.393	0.468
	Jaccard		0.245	0.305

Figure 8: Images that have the target epiphyte under good illumination and high occupancy

The RGB model could segment the epiphytes better than the grayscale model.

The findings shown in Table 3 is obtained by doing a paired t-test comparing the gray scale and RGB model dice and jaccard scores.

	Dice	Jaccard
p-value	0.00004	0.00002

Table 3: Paired t-test results for final model (500 epochs)

The Jaccard scores obtained for the RGB images were significantly higher ( $p < 0.001$ ) than the scores obtained for the grayscale images. Similar results were obtained for the dice scores.

This shows that addition of spectral bands in images is beneficial and significant in segmentation of epiphytes under different illumination and occupancy conditions.

#### 4. CONCLUSION

The DRU-Net was able to learn local and contextual information when additional spectral bands were present in the images. The epiphytes were better segmented when the model was trained and tested with RGB images when compared to the grayscale images trained model. The RGB model had an almost 20% increase in the average dice and Jaccard scores than the grayscale model for 500 epochs. DRU-net could accurately segment epiphytes under different illumination and occupancy conditions with fewer images in the dataset. The additional computation time for RGB images is worth using for epiphyte segmentation using DRU-Net.

The learning of textural features of the target epiphyte by the network can be further studied. The segmentation of the target epiphyte under poor illumination and low occupancy accurately can be considered for future work.

#### REFERENCES

- Alam, M. S., Wang, D. and Sowmya, A., 2021. Image data augmentation for improving performance of deep learning-based model in pathological lung segmentation. In: 2021 Digital Image Computing: Techniques and Applications (DICTA), pp. 1–5.
- Albshiri, A. A., Shah, S. J. H., Kang, S. S. and Lee, Y., 2022. Amunet: automated mini 3d end-to-end u-net based network for brain claustrum segmentation. *Multimedia Tools and Applications*.
- Anivilla, S., Sajithvariya, V., Sowmya, V., Soman, K. P., Sivanpillai, R. and Brown, G., 2020. Identifying epiphytes in drone photos with a conditional generative adversarial network (c-gan), ASPRS Annual Conference, Vol. XLIV-M-2-2020, pp. 99–104.
- Aswin, S., Sajithvariya, V., Sivanpillai, R., Sowmya, V., Brown, G. K., Shashank, A. and Soman, K. P., 2021. Effect of annotation and loss function on epiphyte identification using conditional generative adversarial network. In: 2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), pp. 1–6.
- Bramão, I., Reis, A., Petersson, K. M. and Faísca, L., 2011. The role of color information on object recognition: A review and meta-analysis. *Acta Psychologica* 138(1), pp. 244–253.
- Csurka, G. and Larlus, D., 2013. What is a good evaluation measure for semantic segmentation?, *BMVC*, Vol. 26.
- Jafari, M., Auer, D., Francis, S., Garibaldi, J. and Chen, X., 2020. Dru-net: An efficient deep convolutional neural network for medical image segmentation, *International Symposium on Biomedical Imaging*, pp. 1144–1148.
- Menon, A., Sajithvariya, V., Sivanpillai, R., Sowmya, V., Brown, G. K. and Soman, K. P., In press, Epiphyte segmentation using drunet, *Internal Conference on Intelligent Computing and Communication (ICICC)*.
- Padma, T., Kumari, C. U., Yamini, D., Pravallika, K., Bhargavi, K. and Nithya, M., 2022. Image segmentation using mask r-cnn for tumor detection from medical images. In: 2022 International Conference on Electronics and Renewable Systems (ICEARS), pp. 1015–1021.
- Panicker, S. A., Kumar, R., Ramachandran, A. and Padmavathi, S., 2020. Analysis of image processing techniques to segment the target animal in non-uniformly illuminated and occluded images, *Inventive Communication and Computational Technologies, Lecture Notes in networks and Systems*, pp. 15–26.

Ronneberger, O., Fischer, P. and Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, International Conference on Medical Image Computing and Computer-Assisted Intervention, pp 234–241.

N. Navab, J. Hornegger, W. M. Wells and A. F. Frangi (eds), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Springer International Publishing, Cham, pp. 234–241.

Sajjadian, S. M., Jafari, M. and Chen, X., 2022. Automatic architectural drawing labelling using deep convolutional neural network. In: J. R. Littlewood, R. J. Howlett and L. C. Jain (eds), Sustainability in Energy and Buildings 2021, Springer Singapore, Singapore, pp. 69–78.

Sajithvariyar, V V, S., Gopalakrishnan, E., Sowmya V., Soman, K P., Sivanpillai, R. and Brown, G., 2019. Opportunities and challenges of launching uavs within wooded areas, ASPRS Annual Conference and International Lidar Mapping Forum.

Zhang, Z., Liu, Q. and Wang, Y., 2018. Road extraction by deep residual u-net. IEEE Geoscience and Remote Sensing Letters 15(5), pp. 749–753.