

LAND USE LAND COVER MAPPING USING UAS IMAGERY: SCENE CLASSIFICATION AND SEMANTIC SEGMENTATION

Monica Rajkumar¹, Sudhagar Nagarajan², Peter DeWitt³

¹mrajkumar2020@fau.edu

²snagarajan@fau.edu

^{1,2}Department of Civil, Environmental and Geomatics Engineering, Florida Atlantic University, 777 Glades Road, FL 33431.

³Program Manager of Jupiter Inlet Lighthouse ONA at Bureau of Land Management, pdewitt@blm.gov

KEY WORDS: Machine Learning, Deep Learning, CNN, UAS, LULC, Augmentation, Scene Classification, Segmentation.

ABSTRACT:

Land use and Land cover classification plays a vital role in understanding the changes happening on the surface of the earth. Vegetation classification can be performed by incorporating various deep learning models using Convolution Neural Network approach. The primary purpose of this research is to test the performances of pre-existing Deep Convolution Neural Network (CNN) models for vegetation classification in the Jupiter Inlet Lighthouse Outstanding Natural Area (JILONA). Specifically, this study focuses on examining of the capacity of Scene and Pixel classification technique using certain 3 band and 5 band combinations. Eight well known scene-based deep convnets, namely VGG19, ResNet152V2, InceptionV3, EfficientNetB5, Xception, InceptionResNetV2, MobileNetV2, DenseNet201, and one important pixel classification model, UNet have been used for vegetation mapping in the North and South part of the JILONA. Among all the scene classification models, EfficientNetB5 and DenseNet201 outperformed all other models with an accuracy of over 97%. Xception and ResNet152V2 model exhibited 95% accuracy, whereas the accuracy of remaining models ranged from 85% to 92%. The classification map achieved through UNet pixel-based method, resulted with an accuracy of about 87% and an even better 91% when all 5 bands (Blue, Green, Red, RedEdge, Near-infrared) were used for training.

1. INTRODUCTION

Jupiter Inlet Lighthouse Outstanding Natural Area (JILONA) is a region covered by various kinds of vegetation, wildlife and other cultural components that rely on adequate land cover. Vegetation acts as a purifier and combats climatic change, hence it plays a critical role in the biosphere (Biello, 2008). With the growing importance and significance of Land Use Land Cover (LULC) classification, there are many effective methods to carry out the vegetation mapping, namely satellite-based, aircraft-based, terrestrial-based, etc. Recently, deep learning, an asset of machine learning, is being considered as the spotlight of remote sensing field. This is because CNN had made a series of breakthroughs in various remote sensing applications such as classification, object detection and segmentation with its superior performance.

The objective of this study is to demonstrate the capability of various deep learning models to perform LULC classification using UAV Imagery and hence determine the best CNN model to perform such vegetation classification. This is important when a particular site needs to be examined for improvements and preservation. Therefore, the primary aim of this research is to examine the power of deep CNN for LULC classification based on UAV Imagery, to investigate the generalization capacity of the existing CNN models, compare the efficiency of certain well known deep CNN scene-based models, which includes, VGG19, ResNet152V2, InceptionV3, InceptionResNetV2, MobileNetV2, DenseNet201, Xception, EfficientNetB5 and perform a study on UNet, which is known to be an efficacious pixel-based classification model. Thus, this study results in mapping the vegetation cover by a comprehensive and elaborative analysis using Unmanned Aerial remote sensing data acquired during 2018 and 2020.

2. MATERIALS AND METHODS

Though various techniques exist for vegetation mapping using different resolutions of satellite images (Xie et al., 2008), the traditional classification approach is not sufficient for community level mapping. Using high-resolution multispectral sensors can improve the image classification process (Cingolani et al., 2004). This research is to use the UAS Imagery data collected from Micasense RedEdge camera for classifying the land cover classes. The overall workflow for this research is shown in Figure 1. Being a subset of machine learning, Deep learning deals with various algorithms that includes neural networks, which was inspired by the structure and function of a human brain. In all the data science techniques including Deep learning algorithms, performance will be based on the amount of data. In other words, accuracy improves when the quantity of training data is higher. Deep learning is often referred to as “Feature learning”, which is the ability of extracting features from raw data automatically. Deep learning exhibits maximum precision whenever the inputs and outputs are analog. The analog data includes imagery with pixels, text data, audio files etc., This is because, the model gains knowledge from the input data through a general-purpose learning procedure, without the intervention of human engineers. One of the deep learning packages and a high-level API for TensorFlow, ‘Keras applications’, consists of various models that have pre-trained weights which can be further used for different processes including feature extraction and prediction. Certain other python modules such as pandas, scikit-learn, matplotlib, seaborn, pickle etc., were used for this study. The efficiency of some of the Keras deep learning models in vegetation classification along the JILONA region is studied using a scene-based classification approach. Apart from this, a pixel-based classification approach is also performed for this study to come up with a classification map. This approach showcases and compares different

combination of bands available from the Micasense RedEdge camera. From these set of band information, three major combinations (1) Red, Green, Blue; (2) Green, Red, Near Infrared; (3) Blue, Green, Red, RedEdge, Near Infrared were identified, and the accuracy of the trained models was investigated.

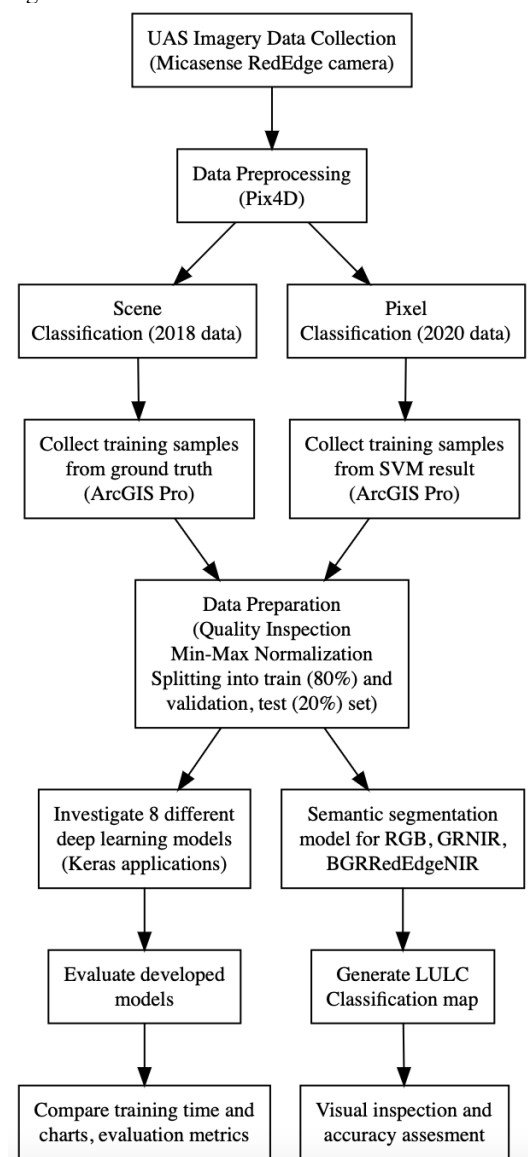


Figure 1: Research Methodology / Workflow

The raw data from the Micasense RedEdge camera was processed using Pix4D Mapper, in order to generate an orthomosaic imagery. This was followed by the process of collecting training samples for both scene-based and pixel-based classification methods. Training samples were collected separately based on the requirement of the model. Following this, the collected samples were utilized to train the deep learning models, so that the accuracy of the models could be compared. While training, the process of augmentation was carried out in order to increase the number of training samples. Adam optimizer, SoftMax activation function, categorical cross entropy loss functions were all examined during the model training procedures. In addition to this, for both scene-based and pixel-based classification methods, 80% of the total dataset was considered to be the training set whereas the remaining 20% was allocated to validation and testing sets equally. Moreover,

the acquired ground truth data was also considered to test the accuracy of the developed pixel-based classification model. The evaluation metrics such as accuracy, intersection over union (iou) and dice co-efficient were analysed while testing the model. The output from the scene-based classification model would be a name of a class that the input imagery belongs to, based on the probability, whereas for pixel-based classification model, the result would be the class code for each pixel. Hence, using pixel-based method, the classification map of whole region can be determined easily and accurately.

2.1 Scene Based Classification

Scene classifier is a deep learning technique in which scenes from the collected imagery data are classified categorically. When Object based method classifies prominent objects in foreground, scene classification utilizes the layout of available objects within the given image along with the ambient context. It is often referred to as scene analysis or scene recognition.

2.1.1 VGG19

VGG19 is one of the variants of Visual Geometry Group (VGG) model. VGG19 model comprises of 19 layers, which includes 16 convolution layers, 3 fully connected layers, 5 MaxPool layers and 1 SoftMax layer. Similar kinds of variants include VGG11, VGG16, etc., It shows more accurate results when deep convolution neural layers are used. Spatial padding, Max pooling and convolving were all used in training this model, which was followed by Rectified linear unit (ReLu) activation function in order to make the model classify effectively by introducing non-linearity. The fully connected layer was fed into SoftMax function, which is the final layer in this model.

2.1.2 ResNet152V2

ResNet is referred to as Residual Neural Network, which is useful for image classification tasks. It supports various architectural configurations that helps in attaining satisfactory balance between work speed and quality. It introduces a structure called residual learning unit, that alleviates the degradation of deep neural networks. The structure of the unit looks like feedforward network, producing the best classification accuracy without increasing the model's complexity. Among all the variants of ResNet family, ResNet152V2 is said to have the best accuracy.

2.1.3 InceptionV3

One of the popular networks called GoogLeNet was extended further and that gave rise to InceptionV3. By making use of transfer learning, it achieved greater performance in several applications. InceptionV3 followed GoogLeNet and thereby proposed a new inception model by concatenating multiple different sized convolution filters, forming a new filter. This approach of creating a new filter reduces the count of parameters involved that needs to be trained and thereby decreases the computational complexity.

2.1.4 InceptionResNetV2

The combination of Inception structure and Residual connection resulted in the model called InceptionResNetV2. In this model, various multiple sized convolution filters are fused together through residual connections to form an Inception-ResNet block. By using these types of residual connections, degradation problems occurring due to deep structures can be avoided, thereby reducing the training time.

2.1.5 DenseNet201

In Dense Convolutional Network (DenseNet), each layer obtains necessary information from all the preceding layers. This information gets passed on to its own feature-maps to all the subsequent layers. Through the process of concatenation, each layer in this model receives “collective knowledge”. Hence, it possesses strong gradient flow, parameter and computational efficiency, diverse features and thereby maintains low complexity features. By utilizing all complexity features, it tends to result in smoother decision boundaries.

2.1.6 Xception

Xception, which is an extension of Inception, replaces the standard inception modules with deepwise separable convolutions. It is one of the convolutional neural networks that contains 71 deep layers.

2.1.7 MobileNetV2

Since MobileNetV2 uses inverted residual blocks with bottlenecking features, it still differs from MobileNet with respect to the number of parameters. Having lower parameters count, MobileNetV2 supports any input size which is greater than 32×32 . It provides much better accuracy based on the size of the input image. Larger the size of the image, greater the performance. Two types of blocks exist in MobileNetV2. The major difference between these two blocks is the value of stride. For each block, there are three layers namely convolution layer with ReLU, depth wise convolution layer and convolution layer without non-linearity. Hence, this model is more effective in extracting features on object detection and during the process of semantic segmentation. It is a type of convolutional neural network that outperforms on mobile devices.

2.1.8 EfficientNetB5

Efficient Net is said to be the next state-of-the-art network, which has a smaller number of parameters. It has variants ranging from B0 to B7. The resolution of EfficientNetB5 is 456 and training the model is relatively faster. All the existing variants of EfficientNet models are being scaled from Efficient Net-B0 but uses different compound co-efficient. This consistently reduces the number of parameters and FLOPS through an order of magnitude.

2.2 Pixel Based Classification

In this method, each individual pixel is utilized in the classification task in order to define land cover class, as there exists a difference among each pixel of the imagery. Floristic plant communities in tropical savanna were classified based on the range of large spatial scale by the integration of SPOT 5 and Landsat 5 TM images with auxiliary data (Donna et al., 2012). Each individual pixel, or group of pixels, will have different brightness values, texture etc., in different bands, which is the major feature that allows the model to predict the land cover class it originally belongs to. These features and their related statistics derived from each individual pixel allows the trained model to differentiate them.

2.2.1 UNet

UNet, which is one of the popular fully convolutional networks (FCN), comprises of contraction and expansion paths. Contraction path helps in extracting more advanced features and reduces the size of feature maps, whereas expansion path covers

the size of segmentation map. UNet is basically a semantic segmentation approach, which is the process of associating each and every pixel of an image with a corresponding class label.

3. STUDY AREA AND DATASET

The geographic region of interest for this study is chosen to be Jupiter Inlet Lighthouse Outstanding Natural Area (JILONA) as shown in Figure 2. It is located in the northern Palm Beach County along the Atlantic coast of South Florida. It is one of the parts of Bureau of Land Management's 27-million-acre National Conservation Lands and it is the only complete unit in the east of the Mississippi River. It is adjoined by the Loxahatchee River and the Indian River Lagoon and is only half mile away from the Atlantic Ocean. It is situated 14 miles north of West Palm Beach and it is approximately a 2-hour drive from the north of Miami via Interstate 95. The total land cover area is about 120 acres of open space.

The data for this research work has been collected from very high-resolution drone imagery. The Unmanned Aircraft System (UAS) data used in this study was collected using drones with Micasense Multispectral Sensor Package in two different years. For Scene classification approach, the three band (RGB) data collected during 2018 was used, whereas for Pixel classification, five bands (Blue, Green, Red, Red edge, Near-infrared) data collected during September 2020 was utilized.



Figure 2: Study area - JILONA

UAS has several advantages such as low device cost, ultra-high spatial resolution, enhanced temporal resolution, lesser survey time and ability to study inaccessible areas. The Micasense RedEdge sensor is commonly preferred for its calibrated narrow bands, durability, ruggedness, and flexibility to integrate with several types of drones. At about 230g in weight, the Micasense RedEdge consists of DLS 2 light sensors and cables. Accurate

and high-resolution measurements of various spectral bands such as blue, green and RedEdge and near-infrared is facilitated by the calibration and global shutter imagers of the sensor. The sensor captures the blue spectral band at 475 nm center wavelength and 32 nm bandwidth, green band at 560 nm center wavelength and 27 nm bandwidth, red band at 668 nm wavelength and 14 nm bandwidth, red edge at 717 nm wavelength, and 12 nm bandwidth, and the near-IR band at 842 nm center wavelength and 57 nm bandwidth. The Micasense RedEdge is easily one of the best and effective sensors available today because of its optimized GSD (Resolution) of 8cm per pixel (per band) at 120m (-400ft) AGL, DLS 2 light sensors, presence of global shutter which can help avoid image distortion even at a capturing rate of 1 per second (all bands), and the ability to generate plant health indexes and RGB color images that are aligned with all bands. Apart from being compact and lightweight, the sensor's wide 47.2 degrees HFOV and various interfaces like Serial, 10/100/1000 ethernet, removable Wi-Fi, external trigger, GDS and SDHC, easily adds on to the fact that Micasense RedEdge is indeed one of the best sensors in recent times. It is also very convenient to use with a variety of UAS due to the availability of various triggering options such as Timer mode, overlap mode, external trigger mode, and manual capture mode (Micasense RedEdge, 2022).

4. SCENE CLASSIFICATION

4.1 Training

The drone data collected in the year 2018 was used for scene classification method, which consists of only three band information namely Red, Green, and Blue. Both the North and South JILONA imagery was merged in order to perform scene-based classification. Training samples were collected from the mosaicked imagery of North and South region of JILONA. Cabbage palm, Grassland, Ground, Mangrove, Parking Lot, Road, Sand pine, Scrub oak and Sea grape were the nine classes that were chosen. All the collected training samples were then subjected to certain augmentation techniques such as rescaling, horizontal flipping, vertical flipping, rotation range, brightness range, shear range, and zoom range. Image augmentation is a technique in which existing data gets altered, so that more amount of data samples gets generated for model training process. Some of the results of augmentation process are shown in Figure 3. Finally, the whole dataset was split accordingly, so that 80% was allocated to training, 10% for validation and the remaining 10% for testing purposes. Each model was trained for about 50 epochs.

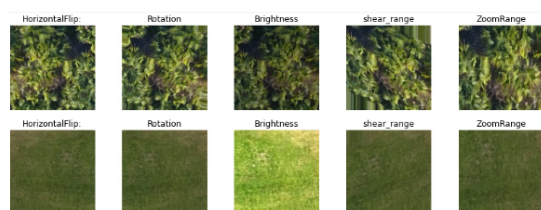


Figure 3: Augmented scenes

4.2 Validation and Testing

Once the model is trained using the extracted dataset, it needs to be validated and tested. For both training and testing purposes, 256*256 image patch size was used. Figure 4 shows the time taken for training each scene-classification model in the GPU utilizing Tesla T4 with 170 GB RAM in 64-bit Windows OS. The average accuracy of all the models was nearly 93%.

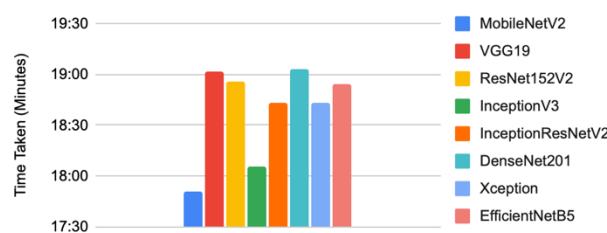


Figure 4: Training time for scene classification models

4.3 Evaluation Metrics

In order to evaluate the developed scene classification models, certain evaluation metrics have been used, which are discussed in detail in this section. This is done to measure the quality of any deep learning model or statistical analysis. The types of evaluation metrics include classification accuracy, confusion matrix, logarithmic loss, Cohen's kappa co-efficient, etc. The ratio of number of correct predictions to the total number of input data samples is called classification accuracy, whereas the false classifications are often referred to as log loss or logarithmic loss. Confusion matrix describes the complete performance of the machine learning model. Evaluation metrics helps in determining whether the model is operating correctly and optimally.

4.3.1 Accuracy and Loss Charts

The accuracy, validation accuracy, loss, and validation loss for all the explored models can be visualized clearly in Figure 5. Through these charts, one can determine whether the model is learning properly, or it has any issues such as overfitting or diverse probability values, underfitting, cramming values etc., so that model can be insisted to minimize its loss and thereby maximizing the accuracy.

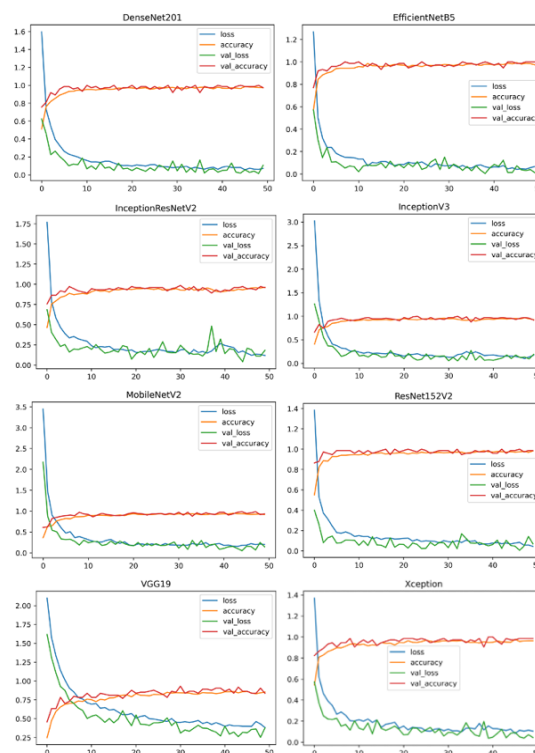


Figure 5: Accuracy, Loss for scene models

4.3.2 Confusion Matrix

Figures 6 and 7 shows the confusion matrix and accuracy and the cohen's kappa of all the investigated models respectively. EfficientNetB5 and DensNet201 are proven to be the best scene-classification models for JILONA region, with an accuracy of nearly 98%. In contrast, VGG19 scored the least accuracy percentage of 85.5%. Likewise, MobileNetV2, ResNet152V2, InceptionV3, InceptionResNetV2, Xception exhibited an accuracy of about 90.6%, 95.1%, 91%, 91.8%, 95.4% respectively.

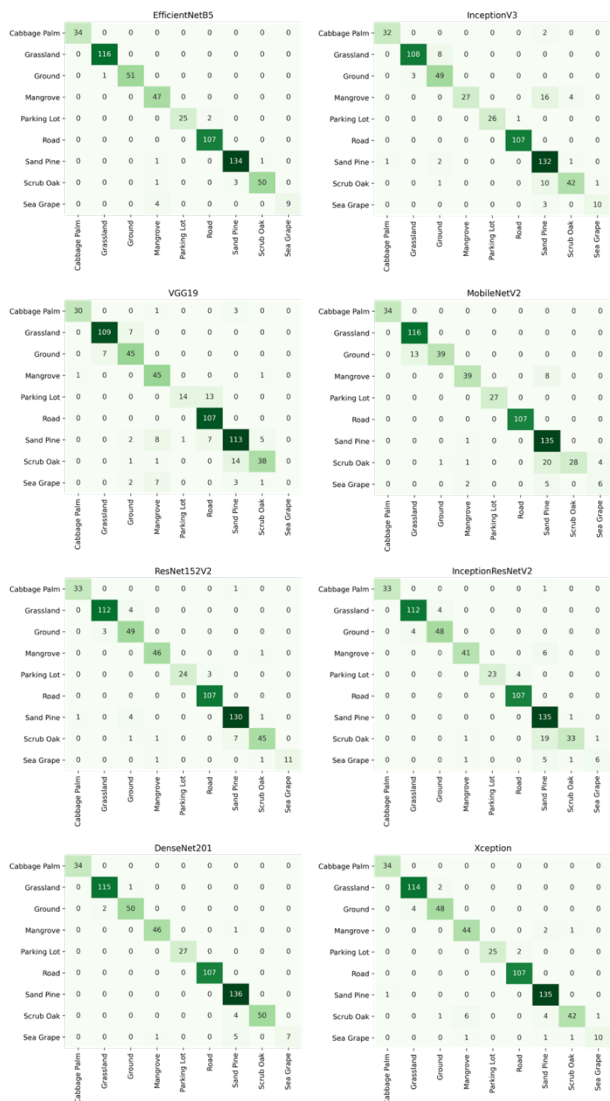


Figure 6: Confusion matrices for scene models

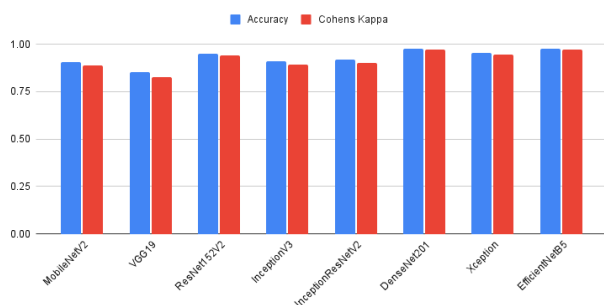


Figure 7: Model accuracy and kappa co-efficient

5. PIXEL CLASSIFICATION

5.1 Training

Figures 8 and 9 depicts the training data for pixel classification approach. Nearly 30% of the training data has been collected separately for North and South region of JILONA through ArcGIS Pro from the existing classification output from Support Vector Machine (SVM). This is done since a classification map is needed as an input for the UNet methodology. SVM classifier is a supervised learning method, which is found to be more memory efficient and effective in high dimensional spaces (Kesavan, 2021). The existing datasets are then artificially expanded through a process called "Image augmentation", to train the deep learning model. Seven classes namely, Oakscrub, Sandpine, Grass, Ground, Palm, Shadow, Water were considered for South region of JILONA, whereas eight classes were considered for North, which includes Mangrove trees in addition to the seven classes for south. With the help of the existing SVM results, the UNet model has been developed and trained for various epochs using different approaches of band combinations.

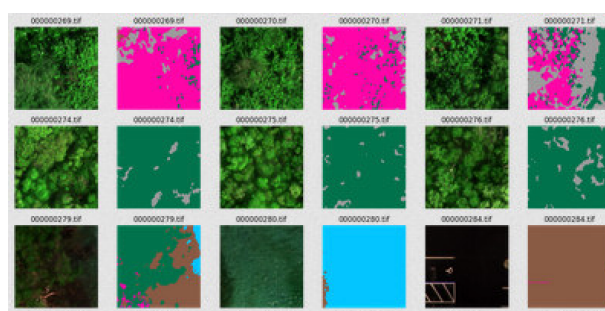


Figure 8: JILONA North training data

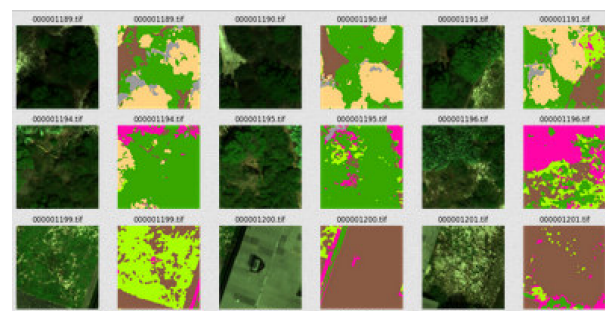


Figure 9: JILONA South training data

The three chosen combination of band are (1) Red, Green and Blue; (2) Green, Red, Near infrared; (3) Blue, Green, Red, Red edge, Near infrared. Several iterations of the model were trained in order to determine the best model for semantic segmentation in the JILONA region. Different sets of epochs and hyperparameter tunings were investigated during this study. Meanwhile, Quality Assurance Quality Check (QA-QC) was performed manually for both the North and South training dataset, before using it as input for the model. QA-QC helped in removing a lot of misclassified images, which improved the prediction accuracy.

5.2 Experiments

Two types of UNet models namely, CustomUNet and SatelliteUNet were examined during this study with three

different band combinations. Certain other hyper parameters such as number of layers, learning rate, batch size, optimizer were also investigated in detail in this research work. In addition to that, it is also proven that the models outclass upon using augmented images for training North and South region of JILONA separately. Several epochs of training were carried out to determine the improvement of model. All the assessments that were made during this study and the best performing model, out of all demonstrations is shown in Table 1.

Parameters	Investigated values	Best model
Models	CustomUNet, SatelliteUNet	CustomUNet
Band Combinations	RGB, GRNIR, BGRRedEdgeNIR	BGRRedEdgeNIR
No. of Layers	4, 5, 6	5
Learning rate	0.01, 0.001, 0.0001	0.01
Batch size	16, 32, 64	64
Optimizer	Adam, SGD	Adam
Epochs	100, 500, 1000, 1500, 2000, 2500	2500
Augmentation	Yes, No	Yes
Image	Mosaicked, North-South separately	North-South separately

Table 1: Investigated Hyperparameters

5.3 Validation and Testing

Once the training is completed for all the developed models, testing was performed for randomly chosen points, which has ground verified class. Individual accuracy of each class for the best model is shown in Figure 10.

	Total	Pass	Pass%
Ground	224	203	90.62
SandPine	197	180	91.37
OakScrub	236	220	93.22
Grass	218	210	96.33
Palm	266	243	91.35
Shadow	69	44	63.77
Water	15	15	100.00

	Total	Pass	Pass%
SandPine	279	272	97.49
OakScrub	288	201	69.79
Palm	251	225	89.64
Ground	316	291	92.09
Water	31	28	90.32
Grass	82	29	35.37
Mangrove	398	378	94.97
Shadow	134	124	92.54

Figure 10: North (Right) & South (Left) Test accuracy

5.4 Evaluation Metrics

Overall accuracy for three major results of this study is being discussed in Table 2. It is evident that, the Custom UNet model outperforms while using all the five-band information with 87.02% of accuracy for North imagery and 91% of accuracy for South imagery of JILONA region. The overall accuracy was calculated based on the individual accuracies explained in section 5.3. Overall Accuracy is the ratio of number of correctly classified pixels to the total number of pixels.

Methodology	North	South
Custom UNet - RGB	86.96%	88%
Custom UNet - GRNIR	85.89%	88%
Custom UNet - BGRRedEdgeNIR	87.02%	91%

Table 2: Accuracy of proposed model

5.5 Classification map

The model that was most accurate was found to be the one that uses all the five bands for classification. The prediction results of both the north and south region are shown in the Figures 11 and 12.



Figure 11: JILONA south classification using all 5 bands



Figure 12: JILONA north classification using all 5 bands

6. CONCLUSIONS

In this study, Scene and Pixel based Classifications were performed for JILONA region. The performance of eight well known scene-based models namely VGG19, ResNet152V2, InceptionV3, EfficientNetB5, Xception, InceptionResNetV2, MobileNetV2, DenseNet201 were compared. In addition, various combinations of hyperparameters were tuned in order to capture the finest working model for pixel-based classification. Finally, LULC map was generated from the determined best UNet segmentation model. This model was developed using the UAS imagery, which is cost effective as it utilizes a low-cost sensor. The findings of this research are that among all the evaluated models, EfficientNetB5 outperformed in scene classification with nearly 98% of prediction accuracy. Similarly, CustomUNet with all 5-band combination trained using Adam optimizer, 64 as batch size, 5 layers and 0.01 learning rate, outclassed all other investigated models for pixel-based classification or semantic segmentation. It is also notable that accuracy percentage of UNet model, got elevated after performing the Quality Inspection for training images.

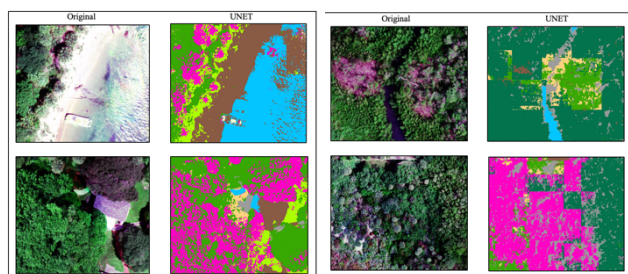


Figure 13: Misclassified areas

The current study is restricted to a specific study area, which can be extended in the future. Also, the comparison of the classification methods can be carried out for various other band combinations of the same (or) different sensors. Further, an in-depth analysis can be performed to improve the prediction accuracy of the UNet model. Similarly, hyperparameters can be tuned in such a way, so that both the scene and pixel classification model improve. Moreover, there are some patch issues in the classification map, which can be seen in Figure 13, that needs to be resolved in future.

7. ACKNOWLEDGEMENTS

The author would like to thank the Bureau of Land Management for providing all the necessary support. A special thanks to the Department of Civil, Environmental and Geomatics Engineering at Florida Atlantic University for permitting the author to carry out this research. In addition, the author wishes to express sincere gratitude to all colleagues for the required academic assistance.

8. AUTHOR CONTRIBUTIONS

Conceptualization, MR and Dr. SN.; Data Collection, PD and Dr. SN; Analysis and interpretation of result, MR; Investigation, Dr. SN; Methodology, MR; Supervision, Dr. SN; Validation, MR, Dr. SN and PD; Report writing, MR.

9. FUNDING

This research work received an external funding. It was supported by the Grant# L19AC00180 from the Bureau of Land Management.

10. CONFLICTS OF INTEREST

The authors declare no conflict of interest with any other researchers or entities.

REFERENCES

- Adão, Telmo, et al. "Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry." *Remote Sensing*, vol. 9, no. 11, 2017, p. 1110., doi:10.3390/rs9111110.
- Abdishaqur: Land use/Land cover classification with Deep Learning. <https://towardsdatascience.com/land-use-land-cover-classification-with-deep-learning-9a5041095ddb> 2018. (Accessed 24 July 2021).
- Bannari, A., et al. "A Review of Vegetation Indices." *Remote Sensing Reviews*, vol. 13, no. 1-2, 1995, pp. 95–120., doi:10.1080/02757259509532298.
- Biello David. "Old-Growth Forests Help Combat Climate Change", 2018, <https://www.scientificamerican.com/article/old-growth-forests-help-combat-climate-change/>
- Bradski, G.: The OpenCV Library. In: Dr. Dobb's Journal of Software Tools (2000).
- Chollet, Francois, Keras Applications. Available online: <https://keras.io/api/applications/> (Accessed on 23 June 2021).
- Christopher Benjamin Anderson. The CCB-ID approach to tree species mapping with airborne imaging spectroscopy. <http://dx.doi.org/10.7717/peerj.5666> (Accessed on 23 July 2021).
- Cingolani AM, Renison D, Zak MR, et al. Mapping vegetation in heterogeneous mountain rangeland using Landsat data: an alternative method to define and classify land-cover units, 2004. *Remote Sens Environ* 92:84–97
- Cohen's kappa co-efficient. Available online: <http://1boringgoldman.com/index.php/2012/09/24/take-notice/> (Accessed on 22 July 2021).
- Data Science - <https://towardsdatascience.com/review-densenet-image-classification-b6631a8ef803> (Accessed on 16 July 2021).
- Deep Learning. <https://www.deeplearningbook.org/> (Accessed on 5 June 2021).
- Deep Learning with Python by François Chollet (Accessed on 21 June 2021).
- DenseNet Model. <https://towardsdatascience.com/paper-review-densenet-densely-connected-convolutional-networks-acf9065dfebf> (Accessed on 23 July 2021).
- DenseNet-201 - Available online: <https://www.mathworks.com/help/deeplearning/ref/densenet201.html> (Accessed on 20 July 2021).
- Donna Lewis, Stuart Phinn & Kirrilly Pfitzner. Pixel-based image classification to map vegetation communities using SPOT5 and Landsat5 Thematic Mapper data in a tropical savanna, northern Australia, *Canadian Journal of Remote Sensing*, 38:5, 570-585, 2012, DOI:10.5589/m12-047

- Evaluation Metrics. Available online: <https://deeptai.org/machine-learning-glossary-and-terms/evaluation-metrics> (Accessed on 21 July 2021).
- Foody, Giles M. "Sample Size Determination for Image Classification Accuracy Assessment and Comparison." *International Journal of Remote Sensing*, vol. 30, no. 20, 2009, pp. 5273–5291., doi:10.1080/01431160903130937.
- Gillies, Sean u.a.: Rasterio: geospatial raster I/O for Python programmers. 2013. <https://github.com/mapbox/rasterio>.
- He, Kaiming ; Zhang, Xiangyu ; Ren, Shaoqing ; Sun, Jian: Deep Residual Learning for Image Recognition. In: arXiv preprint arXiv:1512.03385 (2015). – URL <https://arxiv.org/abs/1512.03385>.
- Hunter, J. D.: Matplotlib: A 2D graphics environment. In: *Computing in Science & Engineering* 9 (2007), Nr. 3, S. 90–95. Available Online: <https://matplotlib.org>.
- Image Augmentation. Available online: <https://www.analyticsvidhya.com/blog/2021/03/image-augmentation-techniques-for-training-deep-learning-models/> (Accessed on 23 July 2021).
- Josh King, Vayu Kishore, Filippo Ranalli. Scene Classification with Convolutional Neural Networks (CNN). Available online: <http://cs231n.stanford.edu/reports/2017/pdfs/102.pdf> (Accessed on 12 July 2021).
- Kaiming He, Xiangyu Zhang, et al. Deep Residual Learning for Image Recognition (2015): arXiv:1512.03385v1 (Accessed on 22 July 2021).
- Long Nguyen, Dongyun Lin, et al. Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation (Accessed on 21 July 2021).
- Lv, Qi, et al. "Urban Land Use and Land Cover Classification Using Remotely Sensed SAR Data through Deep Belief Networks." *Journal of Sensors*, vol. 2015, 2015, pp. 1–10., doi:10.1155/2015/538063.
- Masoud Mahdianpari, Yun Zhang, et al. Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery: <http://dx.doi.org/10.3390/rs10071119> (Accessed on 22 July 2021).
- Masoud Mahdianpari, Bahram Saelhi, Mohammad Rezae, Fariba Mohammadimanesh and Yun Zhang. Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery, 2018 MDPI, Basel, Switzerland.
- Micasense Incorporated. 2022. Micasense Red Edge-MTM Multispectral Camera – Integration Guide, <https://support.micasense.com/hc/en-us/articles/360011389334-RedEdge-MX-Integration-Guide>
- Milas, Anita Simic, et al. "Different Colours of Shadows: Classification of UAV Images." *International Journal of Remote Sensing*, vol. 38, no. 8-10, 2017, pp. 3084–3100., doi:10.1080/01431161.2016.1274449.
- MobileNetV2 model from TensorFlow API of Keras application https://www.tensorflow.org/api_docs/python/tf/keras/applications/mobilenet_v2 (Accessed on 11 July 2021).
- Nagarajan, Sudhagar, et al. "UAS Based 3D Shoreline Change Detection of Jupiter Inlet Lighthouse ONA after Hurricane Irma." *International Journal of Remote Sensing*, vol. 40, no. 24, 2019, pp. 9140–9158., doi:10.1080/01431161.2019.1569792.
- Neubert, M. & Meinel, G. "Evaluation of segmentation programs for high resolution remote sensing applications." *Proceedings of the Joint ISPRS/EARSel Workshop*, 2003.
- Pandian Kesavan, 2021. Machine Learning Approach for Vegetation Classification Using UAS Multispectral Imagery: Florida Atlantic University (Accessed on 23 July 2021).
- Reis, Mariane, et al. "Towards a Reproducible LULC Hierarchical Class Legend for Use in the Southwest of Pará State, Brazil: A Comparison with Remote Sensing Data-Driven Hierarchies." *Land*, vol. 7, no. 2, 2018, p. 65., doi:10.3390/land7020065.
- ResNet152: <https://supervise.ly/explore/models/res-net-152-image-net-4228/overview> (Accessed on 22 July 2021).
- Semantic Segmentation. Available online: <https://paperswithcode.com/task/semantic-segmentation> (Accessed on 25 July 2021).
- SVM Classifier. Available online: <https://scikit-learn.org/stable/modules/svm.html> (Accessed on 23 July 2021)
- UNet Architecture. Available online: <https://towardsdatascience.com/review-u-net-biomedical-image-segmentation-d02bf06ca760> (Accessed on 20 July 2021).
- UNet Segmentation advantages. Available online: <https://jean-charles-nigretto.medium.com/advantages-of-u-net-for-image-segmentation-8ce869d28b4d> (Accessed on 24 July 2021).
- VGG model, Available online: <https://iq.opengenus.org/vgg19-architecture/> (Accessed on 21 June 2021).
- Wan, Liang, et al. "Combining UAV-Based Vegetation Indices and Image Classification to Estimate Flower Number in Oilseed Rape." *Remote Sensing*, vol. 10, no. 9, 2018, p. 1484., doi:10.3390/rs10091484.
- Watts, Adam C., et al. "Unmanned Aircraft Systems in Remote Sensing and Scientific Research: Classification and Considerations of Use." *Remote Sensing*, vol. 4, no. 6, 2012, pp. 1671–1692., doi:10.3390/rs4061671.
- Xception Model. Available online: <https://www.mathworks.com/help/deeplearning/ref/xception.html> (Accessed on 24 July 2021)
- Xie, Zhuli, et al. "Classification of Land Cover, Forest, and Tree Species Classes with ZiYuan-3 Multispectral and Stereo Data." *Remote Sensing*, vol. 11, no. 2, 2019, p. 164., doi:10.3390/rs11020164.
- Yichun Xie et al. "Remote Sensing Imagery in Vegetation Mapping: a review.", 2008, *Journal of Plant Ecology*, volume 1, number 1, pages 9–23., March 2008 DOI: 10.1093/jpe/rtm005.