Detect Deforestation using Satellite Images In Sri Lanka (Galenbindunuwewa Region)

A Data Science Project Proposal presented by

HERATH, H.M.G.U.K. (S/17/374)
HISHAM, M.H.M. (S/17/379)
PUBUDIKA, L.W.P. (S/17/452)
THAMISHKA, W.P.K. (S/17/490)

to the Department of Statistics and Computer Science
of the

FACULTY OF SCIENCE

for the course of Machine Learning
(DSC4173)
in partial fulfillment of the requirement for the award
of the degree of

B.Sc. (Hons) Degree in Data Science

Table of Contents

1. Abstract

The degradation of Sri Lanka’s Mountain forests due to deforestation is endangering critical ecosystem services. This study leverages advanced techniques in remote sensing and deep learning to effectively classify areas impacted by deforestation in the Galenbindunuwewa region. Employing state-of-the-art models such as FCN-based U-Net, spatial information is retained during image analysis. Through intensive model training using satellite data from recent deforestation events and meticulous hyperparameter optimization, the accuracy of the algorithms is comprehensively evaluated. This study classified landscape affected by human-induced deforestation efficiently using high-resolution remote sensing and deep-learning. Land and forest cover maps was used as based data to construct accurate deep-learning dataset. Sites were classified into forest and non-forest areas. Accuracy of the model is 95.39%. Thus, deep-learning modeling has great potential for estimating human-induced deforestation. The finding of this study will contribute to more efficient monitoring of damaged forests and the determination of policy priorities for forest area restoration.

2. Introduction

Forests play an indispensable role in providing essential ecosystem services that include the sequestration of carbon, the production of life-sustaining oxygen, and the safeguarding of landscapes against erosive forces. Yet, the escalating issue of deforestation poses a significant menace to these crucial ecological functions. In Sri Lanka, the encroachments of illegal farming, rampant cattle grazing, and unchecked deforestation have inflicted profound harm upon the nation's mountainous forests, triggering a cascade of ecological imbalances and biodiversity loss. These challenges are further exacerbated by the inefficiency of traditional methods employed to assess the extent of deforested areas, methods that are both resource-intensive and time-consuming.

However, The evolution of remote sensing technology has ushered in a transformative era for landscape analysis. It empowers us with the capability to efficiently and cost-effectively monitor the alarming progression of deforestation. Nonetheless, even as technology enables us to better observe our changing world, the fine task of forest classification remains intricate, often necessitating both expert insight and meticulously curated training data.
Despite the vast public benefits provided by forests, these ecosystems are frequently damaged by human-induced large-scale deforestation, with consequences to their individual and industrial values. Remote sensing-based forest studies conducted in the early 2000s involved mainly the analysis of forest vitality using the normalized different vegetation index or object-based tree classification. With the rapid development of technologies such as data mining and machine learning since 2010, recent studies have integrated forest classification monitoring and detection using remote sensing and machine learning, a branch of artificial intelligence (AI). ML methods used in this effort include the classification and regression trees, decision tree, support vector machine, artificial neural network, and random forest algorithms. Disadvantages of these methods include the need to create accurate learning data based on expertise in tasks such as tree characteristic extraction. Deep-learning emerged as a solution to this problem. Most deep-learning studies of deforestation have evolved the analysis of high-resolution satellite images and models based on convolutional neural networks (CNNs).

In a decisive move towards addressing these intertwined challenges, This study utilizes the powerful combination of state-of-the-art deep learning algorithms and high-resolution satellite imagery, dependently working together. The primary focus lies in classifying the regions of the Galenbindunuwewa mountainous area in Sri

Lanka that have borne the brunt of deforestation's impact. Employing state-of-the-art models like the FCN-based U-Net and also used custom CNN method to study endeavors to safeguard the integrity of spatial information during image processing.

This effort includes carefully choosing the best deep learning methods, creating datasets from detailed remote sensing information, and adding precise location details to improve how we divide different areas. We then teach the selected models using recent data from deforested regions captured by satellites and carefully adjusting the settings. This allows us to thoroughly examine how well the algorithms work through rigorous testing.

With the main goal of understanding the challenges of deforestation-affected landscapes, this study focuses on the Galenbindunuwewa area in Sri Lanka. The findings are expected to help make informed decisions, support sustainable land management, and play a key role in restoring and safeguarding forest ecosystems. This study, the FCN-based U-Net deep-learning algorithm were used, based on literature review. Next, study areas were selected using google earth pro and mask images were constructed using OpenCV python. Third, the dataset was divided into training and testing sets, and the training data were applied to the deep-learning algorithm. Hyperparameters were tuned in the algorithm learning process, and the accuracy of the optimum test was evaluated. Finally, the results were analyzed to determine the applicability of the algorithm.

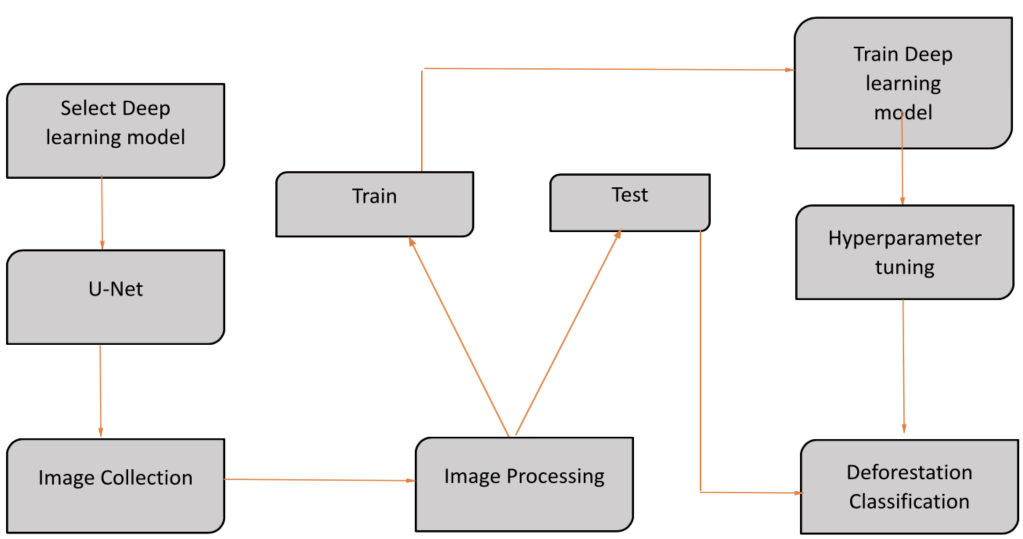


Figure 1. Flow chart in this study

3. Literature review

3.1 In the study titled "Classification of Landscape Affected by Deforestation Using High-Resolution Remote Sensing Data and Deep-Learning Techniques" by Seong-Hyeok Lee et al. (2020), the authors explore the potential and challenges posed by the availability of multisource remote-sensing data from optical-based satellite sensors for monitoring deforestation in the Amazon Biome. This research focuses on leveraging deep-learning techniques and semantic segmentation to address the intricate task of change detection using remote sensing imagery.

The Amazon forest, a vital and environmentally sensitive region, serves as the backdrop for investigating the effectiveness of various convolutional neural network (CNN) architectures in identifying land cover changes caused by deforestation. This literature review delves into these architectures and their roles in this context.

SegNet emerges as a prominent CNN architecture, addressing the intricate balance between spatial information preservation and computational efficiency. Through its ingenious utilization of max-pooling indices during encoding and decoding stages, SegNet expedites training and yields superior results, particularly along object boundaries.

U-Net, a widely embraced architecture, showcases its prowess through skip connections, harmonizing features from different stages of resolution. By preserving minute details otherwise lost in pooling operations, U-Net expedites model convergence, empowering accurate change detection amidst Amazon's diverse landscapes.

In the Amazon forest change detection context, these CNN architectures have proven their mettle in deforestation identification. Leveraging the prowess of Landsat-8 and Sentinel-2 imagery, experts have employed photo-interpretation techniques for model training and evaluation. The complex Amazonian landscape, characterized by multifaceted deforestation dynamics, offers an ideal platform for assessing these architectures' performance and suitability.

3.2 The study "Satellite Imagery for Deforestation Prediction using Deep Learning" addresses the urgent need for effective deforestation detection. It proposes using advanced machine learning, specifically deep convolutional neural networks (CNNs), to predict and identify instances of deforestation through satellite imagery analysis. This approach offers a promising solution to the pressing environmental issue of deforestation.

The authors emphasize deforestation's impact on climate change and ecosystems. Their research focuses on a robust machine learning model to identify deforested areas from satellite images, aiding prevention efforts.

The study outlines a standard deep learning approach, including data collection, preprocessing, training, testing, and deployment. The importance of preprocessing, involving data resizing and normalization, is highlighted for improved training optimization.

Training involves choosing between training from scratch or using pre-trained models like MobileNet, DenseNet, and ResNet. Pre-trained models significantly reduce training time and leverage features from large datasets. Fine-tuning adapts the model for specific tasks.
Testing and evaluation assess model performance using metrics like accuracy, precision, and ROC curves. Graphical representations compare original and ResNet pre-trained models.

Future work includes exploring models with less data, memory optimization, data augmentation, and extending capabilities to forest fire detection and instance segmentation. The study underscores the role of deep learning in deforestation prediction, contributing to environmental preservation and climate change mitigation.

4. Methodology

Method 1:

1. Data Preprocessing

Data Collection and Loading:
The initial step involves collecting a diverse dataset of satellite images portraying both forested and deforested regions. OpenCV is employed to read the images, taking into account the specified `image_extension` and `image_type`. A loop iterates through the dataset, loading each image and ensuring data integrity.

Image Cropping and Resizing:
To establish consistency in dimensions, the `image_patch_size` parameter is defined. Satellite images are cropped to dimensions that are multiples of `image_patch_size` to ensure uniformity. The Python Imaging Library (PIL) is used for cropping and resizing images, while preserving their aspect ratios.

Patch Extraction and Normalization:
The cropped images are converted into NumPy arrays for further processing. Using the `patchify` function, each image is divided into smaller patches to facilitate analysis. Subsequently, min-max scaling is applied to normalize pixel values of individual patches. These normalized image patches are then stored in the `image_dataset` list

Mask Loading and RGB Conversion:
Corresponding mask images indicating deforested regions are loaded using the specified `image_extension` and `image_type`. The mask images are converted to the RGB color space using OpenCV to maintain consistency in color representation.

Mask Cropping and Patching:
Similar to satellite images, mask images are cropped and resized to align with dimensions of corresponding image patches. By utilizing the `patchify` function, individual mask patches are generated and aligned with image patches. These mask patches are stored in the `mask_dataset` list.

2. Dataset Creation and Preprocessing

Pairing Image and Mask Patches:
The creation of a compatible dataset for training involves associating each normalized image patch with its corresponding mask patch. This pairing ensures that image and mask patches are synchronized for model training.

Label Encoding and Categorization:
To enable multi-class segmentation, mask patches are converted into categorical labels. Through label encoding, pixel values are transformed into categorical values representing classes—deforested and non-deforested.

3. Deep Learning Model Architecture

Deep-learning algorithms are currently being used in various remote sensing and spatial information studies. The CNN is the most widely used computer vision-related deep-learning algorithm. CNNs reinforce the pixel characteristics of input images through a convolutional process and perform and iterative process of condensing and pooling reinforced characteristics in a feature map, ultimately producing fully connected layers that may be applied to a neural network. The spatial characteristics of input mages and locational characteristics of objects are not involved in this process.

FCN-based algorithms have been used recently to solve this problem of deep-leaning algorithms based on computer vision. In FCNs, fully connected layers are replaced with convolutional layers to overcome the limitations of the CNN model, such as the loss of image location information and fixation of input images. Due to these characteristics, FCNs can not only classify objects, but also semantically divide them.

Various semantic segmentation models have emerged recently to supplement FCN methods. Among them, U-Net is effective in terms of learning speed and accuracy. The architecture of CNN consists of an encoder and decoder process. The encoder process consists of image compression and feature extraction using the rectified linear unit (ReLU) during activation. Upon its completion, the decoder process restores the image. Image spatial information is maintained during the decoder process because image restoration is performed using the same pooling layer as in the encoder process. When image reconstruction is complete, the image is classified using the softmax function.

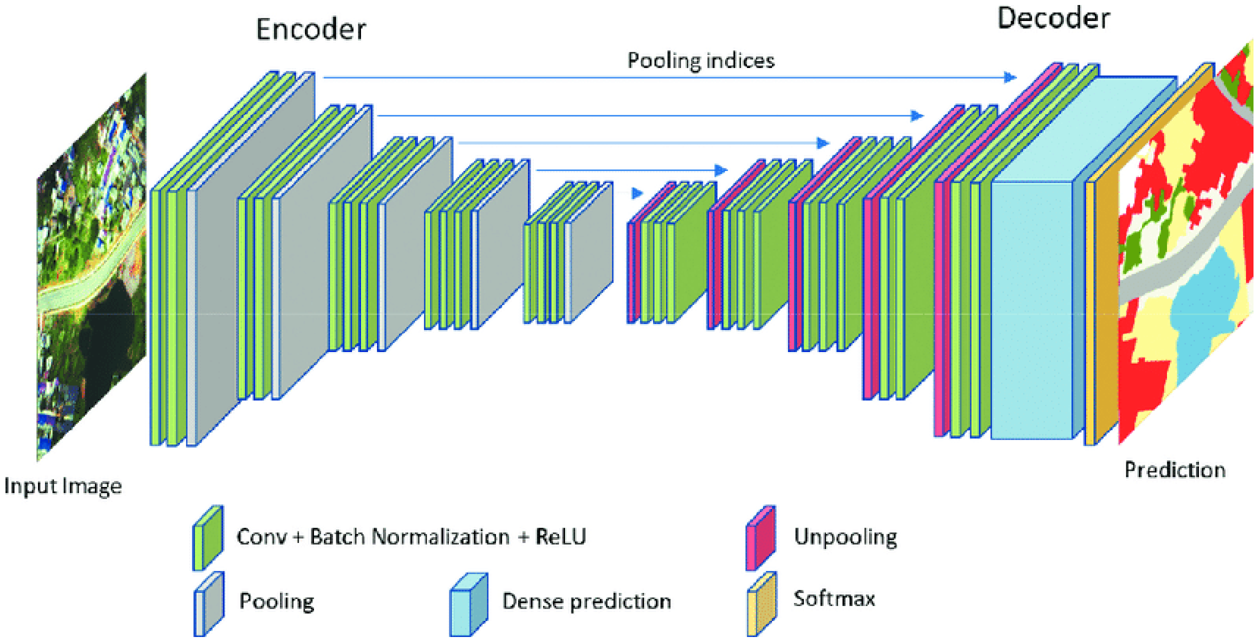


Figure 2. The Architecture of a CNN model (Source: ResearchGate)

U-Net was developed based on the FCN and is applied mainly for the segmentation of a small number of medical images. The U-Net model architecture resembles the letter U, with a contracting path on the left and expensive path on the right. The contracting path uses an image patch, with the N*N*C, C channel as input layers. In each path, sub-sampling is performed using convolutional layers, ReLU activation functions, and max pooling. In the expensive path, U-Net has two definitive characteristics: the copy-and-crop step, which brings source information to the contracting path using a skip connection, and convolution layer without fully connected layers in the image restoration stage. In the network, input images are mirrored to predict the boundary value of the patch. U-Net uses input data in patch units instead of a sliding window, thereby improving its speed. This algorithm accurately captures the context of the image through concatenation using the copy-and-crop function while solving the FCN issue of localization.

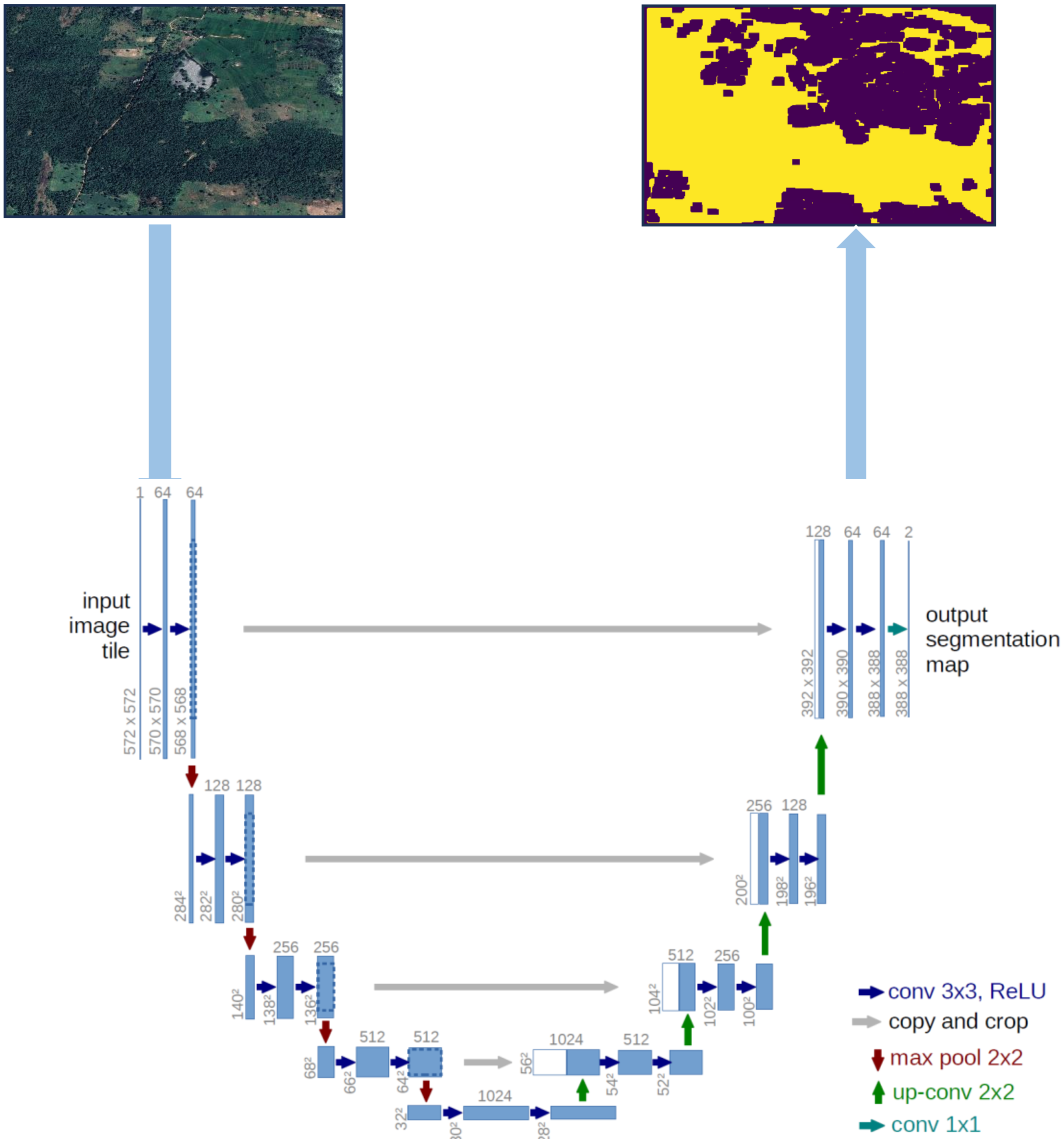


Figure 3. U-net architecture

The following figure indicates the basic input and the output of the U-Net model.

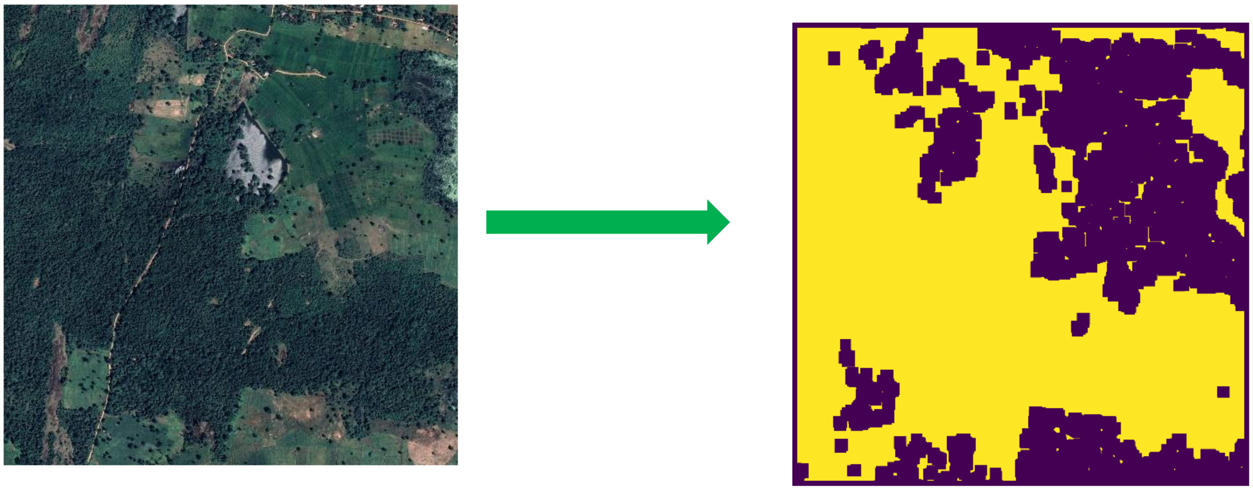


Figure 4. input and output

The following table shows the color separation of the forest areas and non-forest areas of the U-Net model.

Classification	Color	(RGB) - (HEX)
Forest		(253,231,36) - (#fde724)
Non-forest		(68,1,84) – (#440154)

Table 1. color separation for the two classes for U-Net model

4. Model Training and Evaluation

Data set splitting:
The paired data set is divided into two training and testing subsets to facilitate model training and evaluation.

Loss Function and Optimization:
We define two loss functions, Dice Loss and Categorical Focal Loss, crucial for training the model. These functions guide the optimization process, ensuring the model learns to accurately segment deforested areas. The total loss is a combination of both, striking a balance between precision and robustness.

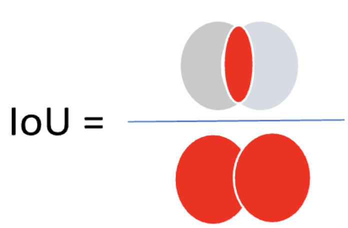
Dice loss – especially used when the data is imbalanced between foreground and background pixels. Dice coefficient, which is a measure of similarity between two sets of binary data.

Dice coefficient = D (A, B) = $\frac{2|A\ intersection\ B|}{|A|+|B|}$; A and B are two sets of binary data, such as ground truth mask and the predicted mask.
Dice loss = L (A, B) =1 – D (A, B). The dice loss ranges from 0 to 1, where 0 means no error and 1 means maximum error.
Focal loss – design to address the problem of class imbalance in classification. Focal loss = -α * (1 - p) ^γ * log(p); Where; α: A balancing factor that assigns different weights to positive and negative examples. It can be used to adjust the influence of the minority class, p: The predicted probability of the true class, γ: A focusing parameter that adjusts the rate at which the loss for well-classified examples is down-weighted. A higher value of γ increases the focus on hard examples.

Total loss – same contribution for both losses. **Total loss = dice loss + (weight*focal loss).** The choice of the weight for each loss function depends on the specific problem and dataset.

Model Training:
With the model architecture defined and loss functions established, we compile the model and initiate the training process. Using the segmented image data from the training set (X_train, y_train), we train the model over multiple epochs. During each epoch, the model refines its internal parameters to minimize the loss functions, enhancing its ability to identify deforested regions.

Performance Evaluation:
Throughout the training process, we monitor the model's performance using metrics such as accuracy and the Jaccard coefficient (IoU). These metrics provide valuable insights into how well the model is learning and generalizing to new data. The validation set (X_test, y_test) allows us to assess the model's accuracy on unseen data, revealing its true potential for deforestation detection.



5.Analysis and Visualization:

Loss Trends:
Visual representations of training and validation loss trends over epochs are plotted to assess model convergence and identify potential overfitting.

IoU Trends:
Graphical representation of training and validation IoU scores are provided to display the model's capacity to accurately detect deforested regions.

Method 2:

1.Image Preprocessing:

Load and visualize the image:
Load the input image using OpenCV and display its height and width to ensure data integrity.

Reshape the images:
Resize the image to a standard patch size (256x256) to ensure consistent dimensions for further processing.

Convert color format:
Convert the image from BGR to RGB format to prepare it for analysis.

Mask creation:
Create a mask for the image by specifying lower and upper color thresholds to segment forested and deforested areas.

2.Image segmentation using U-Net

U-Net model Architecture:
Define the architecture of the U-Net model using encoder and decoder blocks. Each block consists of convolutional layers, max-pooling, and dropout to extract features and reduce overfitting.

Data preparation:
Load and preprocess a dataset containing forest and deforestation images and their corresponding masks. Normalize the pixel values of both images and masks in the range [0, 1].

Model Compilation:
Compile the U-Net model using the Adam optimizer and binary cross-entropy loss function, suitable for segmentation tasks.

Model Training:
Train U-Net model using the preprocessed dataset. Monitor the training process by evaluating the performance on validation data.

Model Evaluation:
Evaluate the trained U-Net model's accuracy, loss, and IoU score (Intersection over Union) on a separate testing dataset.

3.Image Classification using CNN

In this phase, we switch to image classification using a custom Convolutional Neural Network (CNN).

CNN Model Architecture:
We design a CNN model specifically customized for binary image classification. Construct convolutional layers followed by dense layers with dropout to avoid overfitting.

Data Preparation:
Prepare a dataset for binary classification containing masked images labeled as forest or deforestation. Normalize pixel values of images in the range [0, 1].

Model Compilation:
Compile the CNN model using the Adam optimizer and binary cross-entropy loss function, appropriate for binary classification.

Model Training:
Train the CNN model on the prepared dataset, monitoring performance on validation data. Record accuracy and loss metrics over epochs.

Model Evaluation:
Evaluate the trained CNN model's accuracy and loss on an independent testing dataset.

4.Model Visualization

U-Net Performance:
We visualize the U-Net model's performance by generating sample predictions. Original images, true masks, and predicted masks are displayed, allowing us to assess the segmentation accuracy visually.

CNN Confusion Matrix:
To evaluate the CNN model's classification performance, we generate and display a confusion matrix. This matrix provides insights into the model's ability to correctly classify forest and deforestation images.

5. Result and Discussion

Method 1

Model learning was performed before optimum hyperparameter estimation. The hyperparameters were estimated with consideration of the number of iterations, batch size, and patch size. We used Adam optimizer algorithm; an extension to stochastic gradient descent that has recently seen broader adoption for deep learning applications, updating network weights iteratively based on training data. All hyperparameters consider the change in ‘training loss value’ and test accuracy to avoid overfitting. Iteration refers to learning trials; we began with 5 iterations and then increased this in steps of 100 considering the training loss value and test accuracy. A patch size of 256*256 was used. Initial weights were set to 0.5 for both classes. After constructed dataset learning and hyperparameter tuning, the final classification of the deforestation area was performed.

	Epochs = 50	Epochs = 100	Epochs = 200
Loss	0.5625	0.5476	0.5575
Accuracy	0.9379	0.9539	0.9444
Jaccard_coef	0.8612	0.8942	0.8744
Val_loss	0.6090	0.6124	0.6151
Val_accuracy	08960	0.8996	0.8901
Val_jaccard_coef	0.7968	0.8064	0.7753

Table 2. accuracy and loss of the U-Net model

Intersection over union (IoU) values were derived from the calculated hyperparameters. Obtained the Jaccard coefficient for the both training and validation set. The jaccard coefficient is a similarity matrix commonly used in image segmentation tasks to evaluate the performance of the model. The jaccard coefficient is a value between 0 and 1, where 1 indicates a perfect match between the true and predicted target labels, and 0 indicates no overlap between the 2 labels.

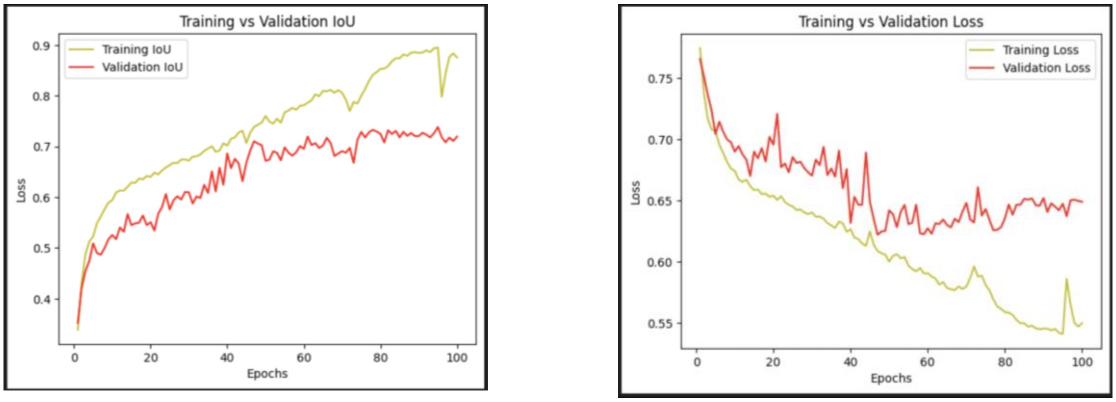


Figure 5. optimum loss and IoU of U-Net model

We selected the optimum number of iterations as 100. The overall accuracy of the U-Net model is 95.39%, and the jaccard coefficient is 0.8942. Also, we obtained 89.96% accuracy for the validation set. Validation Jaccard coefficient is 80.64. These jaccard values are pretty good and hence we can conclude that the model performs well on test data. The figure 5 indicates the loss and IoU values for both training and validation sets separately. The figure 6 shows the performance of the model clearly.

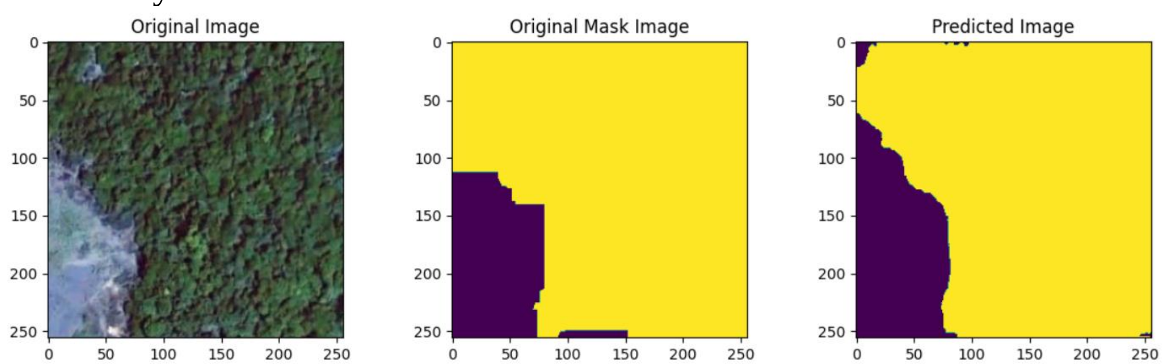


Figure 6. final output of the U-Net model

Method 2

Testing the U-Net model to the images segmentation:

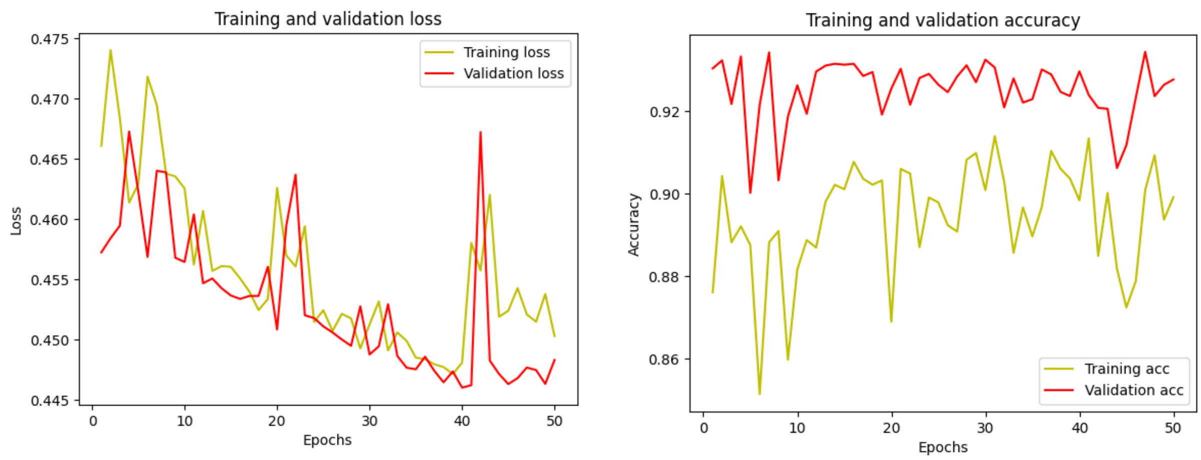


Figure 7.optimum loss and IoU(Method2)

We selected the optimum number of iterations as 100. The overall accuracy of the U-Net model is 92.76%, and loss is 0.4483 . The figure 5 indicates the loss and IoU values for both training and validation sets separately. The figure 6 shows the performance of the model clearly.

U-Net Performance:

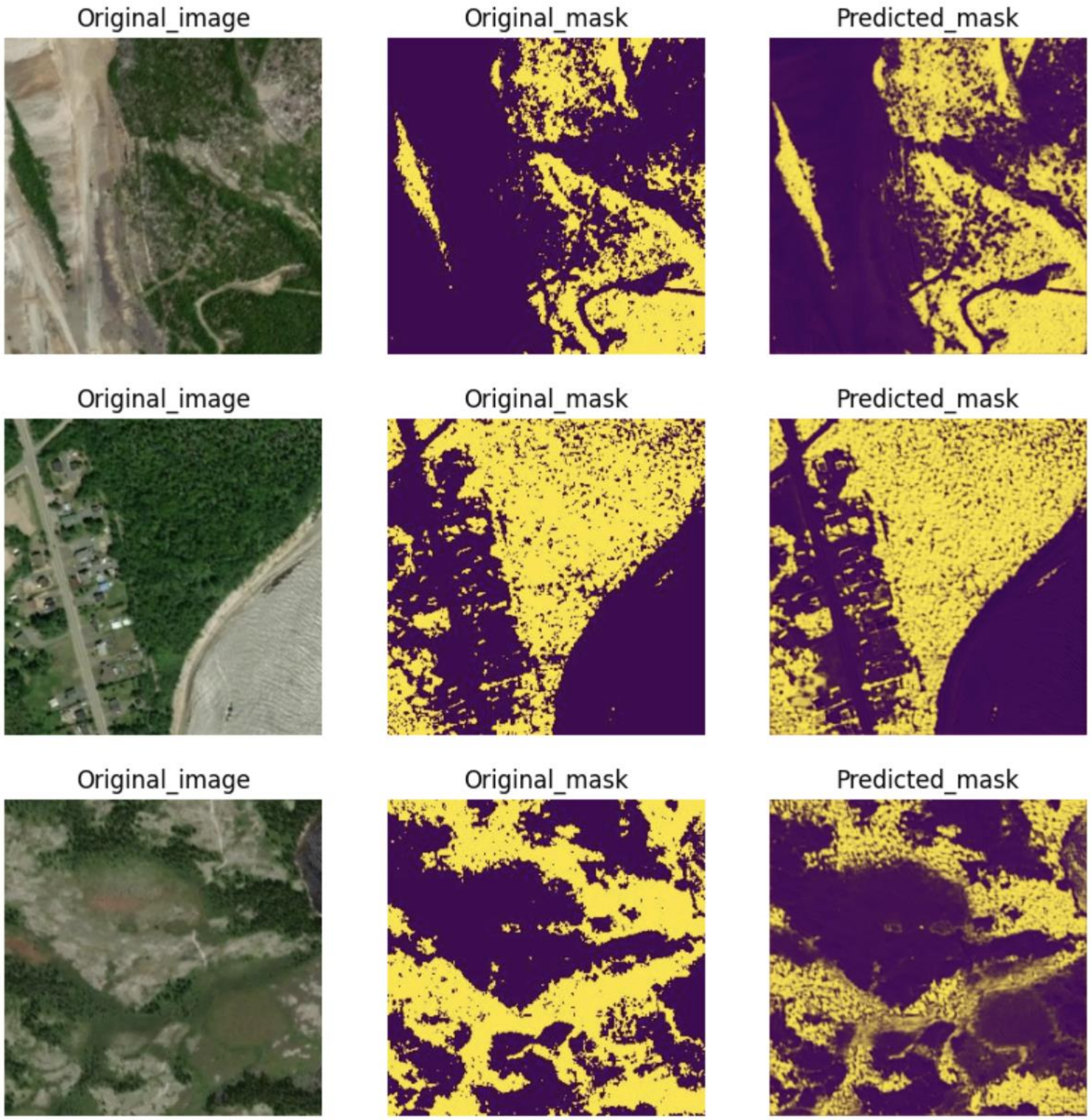


Figure 8. mask images and predicted images given satellite image using U-net

Here we visualize the U-Net model's performance by generating sample predictions. Original images, true masks, and predicted masks are displayed, allowing us to assess the segmentation accuracy visually.

Confusion matrix to evaluate the CNN model's classification performance:

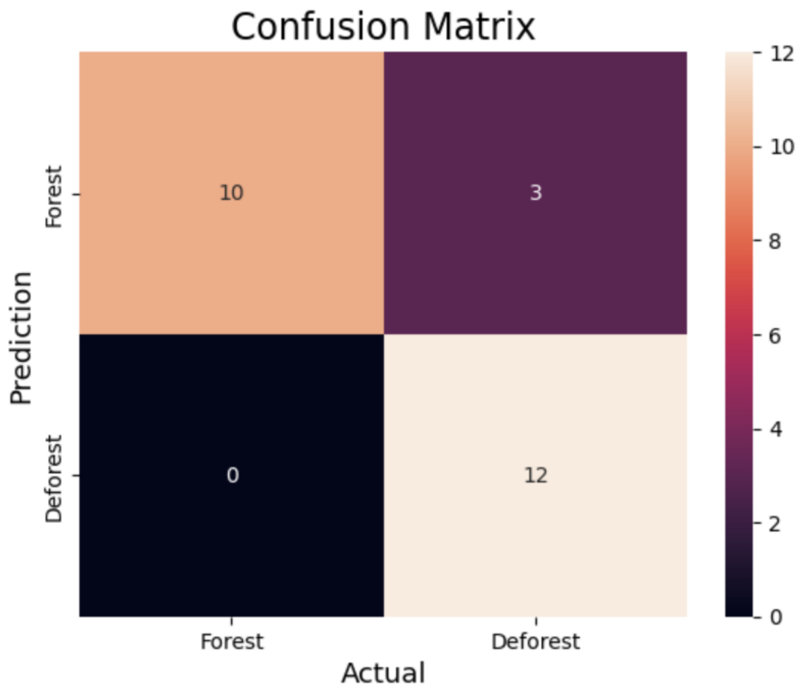


Figure 9. Confusion matrix

True Positives (TP): There were 10 instances of forested areas correctly classified as "forest."
False Negatives (FN): There were 3 instances of deforested areas that were mistakenly classified as "forest."
True Negatives (TN) :There were 12 instances of deforested areas correctly classified as "deforest."
False Positives (FP) : There were no instances of forested areas being classified as "deforest."

Accuracy = (TP + TN) / (TP + TN + FP + FN) = (10 + 12) / (10 + 12 + 0 + 3 + 0) = 22 / 25 = 88%.
This reflects the overall correctness of our model's predictions.
Precision = TP / (TP + FP) = 10 / (10 + 0) = 100%.
This reflects the model that predicts "forest," it is still highly likely to be correct.
Recall = TP / (TP + FN) = 10 / (10 + 3) = 77%.
Our model identified 77% of the actual forested areas.

6. Conclusion

Method 1:

In this study, landscapes affected by human-induced deforestation from high-resolution satellite images were classified using the FCN-based U-Net deep-learning algorithm. To ensure efficiency, precise training dataset were collected such as forest and non-forest areas. To this end, satellite images were preprocessed, and labeling data were created. Training and validation datasets were applied to the U-Net model to estimate hyperparameters, considering batch size, patch size, and number of iterations. For U-Net, the hyperparameter estimation was optimal at a batch size 16 and patch size 256*256 with 100 iterations.

The overall accuracy of the model was 95.39%, however, the testing set showed low accuracy because of the limited resources to train the model. Training dataset can be constructed, and deep-learning algorithms like U-Net applied, for the interpretation of high-resolution satellite images in various ways. Moreover, since the limited resources, hyperparameter tuning was difficult in this study.

This study has several limitations. First, more advanced deep-learning algorithms have been developed since this study was conducted and which may be improved upon the accuracy of our classification results. Second, larger datasets are needed to accurately train the AI algorithms and more clearly distinguish the various features. We faced lots of difficulties when we made the mask image for each and every ground truth image. Obviously, it is another separate task, therefore, we used a simple method to get the mask images using the OpenCV python module. It affects the accuracy of the model. Therefore, we anticipate that further research using improved algorithms and larger datasets will result in better estimation of the causes of deforested areas.

Method 2:
In conclusion, this project endeavor has led us through a comprehensive journey in harnessing cutting-edge deep-learning techniques for the critical tasks of image segmentation and classification in the context of deforestation detection. Through the meticulous implementation of the FCN-based U-Net and a custom CNN model.

Our methodology, meticulously crafted and guided by a systematic approach, encompassed various stages. The preprocessing of satellite images, creation of corresponding masks, and the strategic division of datasets for training and validation formed the foundation of our investigation. Leveraging state-of-the-art frameworks, such as TensorFlow and OpenCV, we harnessed the potential of deep-learning networks to not only segment images with an impressive accuracy of 92.76%, but also to accurately classify forested and deforested areas, achieving an overall accuracy of 88%.

The outcomes derived from our models, as clarified by the earlier discussed confusion matrix, offer valuable insights into how well they performed. We effectively recognized instances where our models made correct predictions (true positives and true negatives) and instances where they made mistakes (false positives and false negatives), allowing us to understand their strengths and areas for enhancement. The high precision score of 100% indicates the models' ability to accurately predict forested regions, and the recall score of 77% demonstrates their capability to identify a considerable portion of real forested areas. This showcases their effectiveness in action.

7. Acknowledgement

We would like to express our gratitude to everyone who contributed to the completion of this group project. Thank you to our project supervisor for their guidance, our group members for their teamwork, and all those who supported us with their insights and assistance.

8. References

1. *Seong-Hyeok Lee 1, Kuk-Jin Han 1, Kwon Lee 2, Kwang-Jae Lee 3, Kwan-Young Oh 3 and Moung-Jin Lee 1* (2020). Classification of Landscape Affected by Deforestation Using High-Resolution Remote Sensing Data and Deep-Learning Techniques.

2. *Chitra, N. T., Anusha, R., Kumar, S. H., Chandana, D. S., Harika, C., & Kumar, V. U. (2021).* Satellite Imagery for Deforestation Prediction using Deep Learning. In 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 522-525). Madurai, India. doi: 10.1109/ICICCS51141.2021.94320.

9. Appendices

- Method 1: https://github.com/udayaKherath/Deeplearning/tree/main/method1
- Method 2: https://huggingface.co/spaces/HishuJanu/DetectDeforestation/tree/main
- Streamlit Output for Method2: https://huggingface.co/spaces/HishuJanu/Detect-Deforestation