



RESEARCH ARTICLE

DETECTING ILLEGAL LOGGING USING DEEP LEARNING ON SENTINEL-1 SAR IMAGERY.

Thinh Ha^{1*,7}, Naga Kasam^{2,7}, Tanish Khanna^{3,7}, Ruhaan Arya^{4,7}, Ikshit Gupta^{5,7}, and Arush Shangari^{6,7},

1. Student, Anaheim Discovery Christian School, Anaheim, California.
2. Student, Moline Sr High School, Moline, Illinois.
3. Student, La Jolla Country Day School, San Diego, California.
4. Student, The Village School, Houston, Texas.
5. Student, Mountain House High School, Mountain House, California.
6. Student, St. John's Preparatory School, North Reading, Massachusetts.
7. Student Researcher, Beaver Works Summer Institute, Boston, Massachusetts.

Manuscript Info

Manuscript History

Received: xxxxxxxxxxxxxxxx
Final Accepted: xxxxxxxxxxxx
Published: xxxxxxxxxxxxxxxx

Keywords:-

Illegal Logging, Mawas conservation area, peat swamp forest, Sentinel-1, Deep Learning, DeepLabV3, EfficientNet

Abstract

Illegal logging in tropical peat swamp forests represents a significant threat to global climate stability, by contributing to around 10-15% of worldwide greenhouse emissions and degrading vital carbon storage ecosystems. This study addresses the necessity of automated detection of illegal logging activities in the Mawas Conservation Area of Central Kalimantan, Indonesia, by developing a deep learning model trained on Synthetic Aperture Radar (SAR) imagery from Sentinel-1. The model structure involves a combination of UNet and DeepLabV3 architecture with EfficientNet-B4 as the encoder backbone, enhanced by Spatial and Channel Squeeze & Excitation (SCSE) attention mechanisms for improved feature extraction. The model was trained on 690 SAR images, captured from March 2015 to December 2016. The deep learning model shows promising results with an F1-Score of 66% and an IoU of 49%. The overall accuracy is high at 89.55% and a precision is 67.41%. These results demonstrate the potential of deep learning for monitoring illegal logging in data-sparse tropical forest regions.

Copy Right, IJAR, 2019,. All rights reserved.

INTRODUCTION:-

A. Illegal Logging

Illegal logging is an imminent threat to the current climate scenario, as it exacerbates climate change, contributing directly to biodiversity loss, water cycle disruption, enhanced soil erosion, heightened natural disaster risk, and increased risk of disease due to elevated wildlife contact. In 2024 alone, 26.8 million hectares of natural forest were cut, releasing nearly 10 gigatons of CO₂ [1]. Furthermore, tropical deforestation, closely related to logging, contributes 10 to 15% of global greenhouse gas emissions, decreasing biomass and accelerating climate change [1]. To circumvent

the climate tipping point, or the thresholds of climate irreversibility, it is necessary to curb illegal logging, especially in tropical environments.

One such vulnerable high-biomass tropical region is the peat swamp forests (PSFs), which serve as carbon-storing regions necessary for climate health. By 2010, 35% of Southeast Asian peatlands had been drained, in part for agricultural and logging purposes, releasing vast carbon stores into the atmosphere and contributing to significant global carbon emissions [2]. Previous research conducted repeat LiDAR surveys of a PSF in Kalimantan, Indonesia, spanning 96 km² [2]. This research revealed the substantial persistence of illegal logging, despite the termination of concessional logging in 1998 and further documented 579 km of concealed logging canals beneath the canopy [2]. Although the surface carbon sequestration of biomass offset nearly half the carbon loss from peat exposure, illegal logging still poses a significant threat to peatlands, biomass reserves, and climate health, as net positive carbon effluxes are observed regardless of environmental offsetting mechanisms.

B. Kalimantan Illegal Logging

This paper focuses on the Mawas Conservation Area, a region under the Central Kalimantan province of Indonesia. This area was specifically chosen due to its lack of documented research inquiring about tracks and patterns of illegal logging, despite its status as a high-biomass region facing intense logging activity. In fact, a United Nations Environment Program estimated that between 77-88% of all logging in Indonesia was illegal in some sense [3]. A total of 17,193 km² of forest cover was lost between 2001 and 2018, which was 18.2% of the remaining area in 2001 [5].

C. Remote Sensing Applications

Remote sensing is a nondestructive study of distant objects via sensing technologies mounted upon unmanned aerial, terrestrial, or marine vehicles to detect anomalies, recognize patterns, and understand field dynamics to obtain an enhanced understanding of the ground situation and procedural outcomes.

The study relies on the electromagnetic spectrum as it captures multiple wavelengths to elucidate features from the area of interest. Data collection can happen through a passive or active means. Passive remote sensing relies on natural energy, while active sensing, such as Synthetic Aperture Radar (SAR) or Light Detection and Ranging (LiDAR), uses an artificial energy source to emit and recollect the passed electromagnetic information. This project will employ SAR imagery collected from Sentinel 1's SAR channel.

SAR imagery from Sentinel 1 has a variety of advantages. SAR imagery can provide high-quality, reliable data regardless of lighting or weather conditions, such as cloud cover, rain, and fog. As such, SAR imagery becomes an ideal method of collecting data in dense, tropical forests, due to its ability to penetrate canopy layers and capture valuable data regardless of lighting or weather. The L-band in SAR is especially sensitive to vertical forest structures, as described in [3], making it invaluable in identifying logging-related changes.

Recently, remote sensing applications have seen increased use in detecting logging in a geospatial region. The first step toward taking action for inhibiting anthropogenic logging is detection, and therefore, this study inquires into developing efficient automated models for environmental monitoring, specifically, tracking and detecting illegal logging.

Convolutional Neural Networks (CNNs) have repeatedly demonstrated the ability to determine patterns in spatial data, in both recognition and satellite imagery-based remote sensing applications. UNet, a segmentation-focused CNN model, excels at tasks requiring automatic feature extraction from large or high-dimensional datasets [4]. This will assist in automatic detection and prediction of damage patterns, informing decision makers to mitigate damage based on spatial and temporal variables.

This study hopes to achieve a precise first step deep learning model with a UNet architecture to detect illegal logging, especially in the Kalimantan region.

METHODOLOGY:-

A. Dataset

This paper focuses on a 750 km² section of peat swamp forest that lies within the Mawas conservation area in Central Kalimantan, Indonesia (see Fig. 1). This area was selected due to the intense effects of illegal logging on the area and the scarcity of existing research on detecting illegal logging.

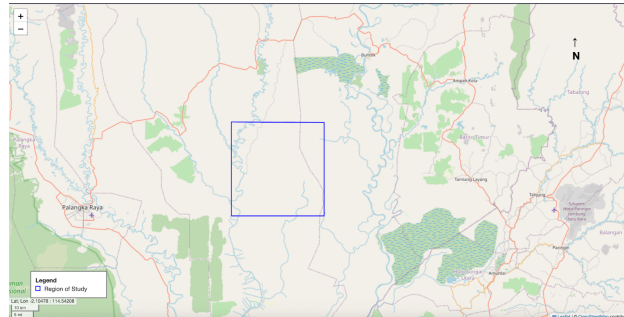


Fig. 1 Region of Study

The dataset consists of Synthetic Aperture Radar (SAR) imagery collected on the Mawas Conservation Area in Central Kalimantan by Sentinel-1's Copernicus Interferometric Wide Swath mode, which utilizes a wide swath width of 250 km with a spatial resolution of 5 m by 20 m. Additionally, this mode used dual-polarization bands: Vertical-Horizontal (VH), Vertical-Vertical (VV), and the VV/VH ratio. The temporal span of the dataset ranges from March 29, 2015, to December 12, 2016, encompassing a total of 23 SAR images, each of which measures 2483×2500 pixels. These images were used to train the deep learning model: see Fig. 2.

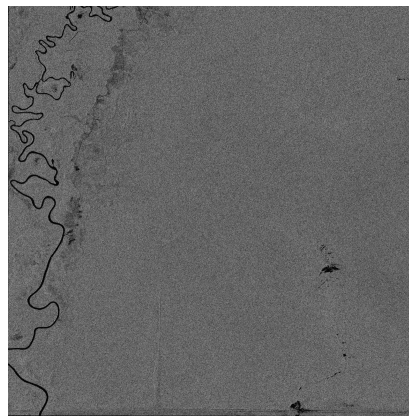


Fig. 2: One of the 23 SAR images of 2483×2500 pixels captured on March 29, 2015

Groundtruth labels for each datapoint were sourced via field validation by the source of the dataset [2]. This dataset divides the classification of different segments into four categories: “illegal logging”, “concessionary plus illegal logging”, “concessionary logging”, and “old growth”. For the purposes of this study, “illegal & concessionary logging” and “illegal logging” classes were merged into the latter label.

B. Data Preprocessing and Augmentation

All captured Sentinel-1 images were cropped from 2483×2500 to 2464×2464 pixels for uniformity. Subsequently, each SAR image was spatially divided into 49 non-overlapping tiles of 382×382 pixels to enable efficient processing while preserving the spatial resolution necessary for maintaining classification accuracy. Then, 15 out of 49 tiles were manually filtered to prevent an unbalanced class while training. (see Fig. 3)

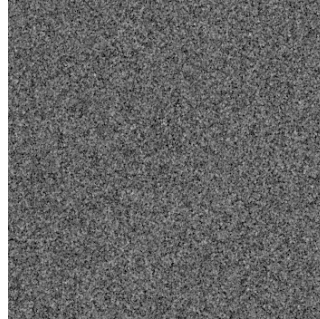


Fig. 3 One of the 49 non-overlapping 382×382 pixel SAR tile from original image

Finally, this research applied data augmentation techniques to each of the images, including Rotation, Horizontal Flip, Vertical Flip, Random Brightness Contrast, Gaussian noise, and Coarse Dropout to improve robustness and model accuracy.

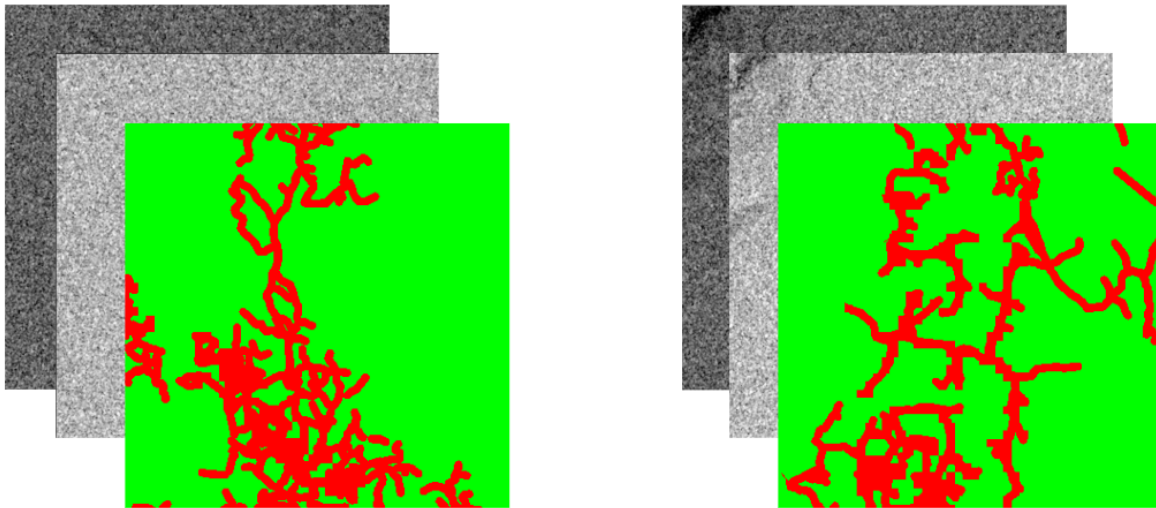


Fig. 4 Layers of SAR Images and Corresponding Masks

The figure illustrates how multiple grayscale input layers were transformed to generate segmentation outputs, with the red structures over the green background being the detected features of interest. The outputs illustrate the ability of the model in extracting fine spatial patterns from SAR backscatter data while having been trained on comparatively small inputs. It should be noted that an additional polarization layer was expected for training but was not available in the data, and therefore a derived ratio layer (VV/VH) was computed and utilized as a substitute to ensure feature richness.

C. Architecture

The final activation function was not used because the model outputs must be raw logits to fit the loss function.

A standard U-Net model was used as the foundation for building the deep learning model for illegal logging detection and classification using semantic segmentation. U-Net was chosen in particular because the encoder–decoder structure of U-Net enables precise feature extraction and localization, which plays an essential role in identifying narrow and non-regular deforestation channels in satellite imagery. Skip connections in U-Net link encoding layers to corresponding decoding layers, therefore preserving high-resolution spatial information that might be obliterated through downsampling. This architecture enables good localization precision, which needs to distinguish between fine deforestation trends from the otherwise healthy surrounding forest [16].

Yet another U-Net strength is its resilience against over-training on relatively small datasets, typical in environmental monitoring missions where annotated data are only available in small amounts. The network has been proven to function properly with small datasets due to its efficient reuse of features and data augmentation methods compatibility [13]. Furthermore, its total convolutional design conserves computational and memory requirements while offering simplicity of implementation and making it suitable for large geospatial applications.

For the decoder, EfficientNet-B4, as a pre-trained model whose weights had been optimized on the ImageNet dataset, was used. EfficientNet employs MBConv blocks and a compound scaling method that scales depth, width, and resolution at the same scale, enabling it to identify subtle patterns using fewer parameters compared to traditional backbones [18]. The model leverages properties learned on large natural image datasets through transfer learning, which results in fast convergence, improved performance with sparse training samples, and reduced computational overhead. EfficientNet-B4 proved to be very effective in segmentation pipelines when combined with U-Net decoders [10, 15].

In this case, the classical U-Net encoder was replaced by DeepLabV3 architecture. DeepLabV3 applies atrous (dilated) convolutions, which allow the network to achieve dense feature maps without sacrificing spatial resolution, a characteristic especially useful in finding fine-scale deforestation boundaries. On the backbone, an Atrous Spatial Pyramid Pooling (ASPP) module was employed. Unlike max pooling, ASPP captures object features at multiple scales through the use of parallel filters with varied dilation rates, and hence retains both global and local context [11, 12]. Multi-scale representation is critical in the resolution of illegal logging patterns that can take various sizes ranging from small clearings to large deforestation corridors.

For the deep path, four decoder blocks were used to upsample the image to the original size following classification of every pixel. Decoders perform upsampling or transposed convolutions and add encoder features through skip connections such that spatial information and detail remain intact. Additional convolutional layers in the decoding blocks further process the segmentation result and improve boundary definition.

To enhance the decoder, Spatial and Channel Squeeze & Excitation (SCSE) blocks were incorporated after standard convolutions. SCSE is an attention mechanism that rescales feature maps by combining two complementary methods: channel squeeze and excitation (cSE), which learns what channels are most valuable, and spatial squeeze and excitation (sSE), which learns where the most valuable spatial positions are. By merging the two, SCSE boosts the representational power of the decoder, allowing the network to give more attention to higher-priority deforestation characteristics [17].

For resolving the problem of class imbalance, Focal Loss and Tversky Loss were merged by the study. Focal Loss lowers the significance of simple examples and emphasizes more challenging ones, thus making it effective for highly imbalanced datasets [14]. Tversky Loss, which is a generalization of the Dice coefficient, allows explicit control over false positives vs. false negatives [8]. The joint loss function was carefully hypertuned to overcome over-prediction and under-prediction in a way that illegal logging areas are detected with high sensitivity.

$$FocalLoss = - \sum_{i=1}^n (1 - p_i)^{\gamma} \log_b(p_i) \quad (1)$$

$$TverskyLoss = \frac{TP}{TP + \alpha FN + \beta FP} \quad (2)$$

The final activation function was removed as the model outputs raw logits, which are required for being compatible with chosen loss functions.

HYPERPARAMETERS:-

Class imbalance in the data had a major influence on loss function selection. Detection of illegal deforestation in Kalimantan requires dealing with heavily skewed data, where background pixels are much higher compared to those of illegal logging channels. In order to counteract the imbalance, both Focal Loss and Tversky Loss functions were employed to evaluate differences between predicted results and ground-truth labels. These loss functions are particularly optimal for minority-class prioritizing segmentation tasks, which is significant in this instance because of the significant ecological and policy consequences of not detecting illegal logging.

The Focal Loss function was tuned for binary classification between illegal logging channels (foreground) and all other land cover classes (background). A gamma of 2.0 was used to under-weight basic examples, reducing their impact on the loss and concentrating the model's attention on difficult or ambiguous pixels, such as partially deforested forest edges or narrow deforestation traces. An alpha of 0.6 was assigned to increase the weight of the positive class, emphasizing the detection of illegal logging activities despite their rarity against the background.

The Tversky Loss function was employed to generalize the Dice coefficient to gain explicit control over balance between false positive and false negative. It was parametrized as 0.4 for alpha for false positives and 0.6 for beta for false negatives. The parameter setting is more penalizing for false negatives, which reflects the greater loss due to missed illegal logging than false alarm. The weighting formula assists the study in its conservation goal by favoring sensitivity to illicit behavior rather than specificity.

The compound loss function assigned the weightage of 60 percent to Focal Loss and 40 percent to Tversky Loss. The assignment enables the utilization of Focal Loss in class imbalance handling and Tversky Loss in obtaining false positives-false negatives balance. The generated loss landscape enhances the accuracy and sensitivity of the model in detecting illegal deforestation patterns.

The Adaptive Moment Estimation (Adam) optimizer was utilized for the optimization of the model. Adam balances the strengths of RMSprop and Momentum optimizers by being able to supply adaptive learning rates for each parameter from the mean and uncentered variance of gradients. This boosts the rate of convergence and allows the model to escape suboptimal local minima. Adam's momentum component and RMSprop-like scaling factor, along with bias correction, both enhance training stability and robustness, making it suitable for complex data such as satellite images of Kalimantan forests.

The learning rate of 0.0003 was set to a value smaller than what is widely used for stable convergence rather than divergence at the early stage of training. The weight decay of 0.00001 was added for the model regularization and reduction of the chance of overfitting. Weight decay punishes large parameter values by adding a small regularization term to the loss function, which causes the model to learn more generalizable and easier representations rather than memorize training data noise.

In addition, the gradscaler component of Automatic Mixed Precision (AMP) was utilized for improved numerical stability and training efficiency both in terms of time and memory usage. AMP allows certain computations to be

performed in float16 precision instead of the default float32, conserving memory and accelerating computation. Float16 computation is more prone to underflow, thus a scale factor of 1024 was applied during backpropagation to maintain updates stable without numerical instability [7].

The training was performed for 450 epochs, with 68 batches per epoch, to provide sufficient iterations for convergence of the model without incurring a heavy computational cost. The given dataset of 690 images was split into 80% used for training and 20% for the test set, thereby ensuring that the majority of data would be available for training the model while keeping a representative sample for unbiased testing.

To further maximize the process, a One Cycle Learning Rate Scheduler was implemented. This scheduler starts with a low learning rate, step-wise raises it to a maximum of 0.001 during the early stages of training, and decreases it to a low value upon completion of the cycle. This cyclical update allows the model to converge faster and is likely to result in better generalization. Specifically, the first 30% of training was utilized in ramping up the learning rate, while the division factor of 25 and the terminal division factor of 10,000 controlled the scaling of the learning rate over the course of the cycle. The adaptive scheduling method circumvents the optimizer getting stuck in steep local minima and encourages exploration of shallow minima, which are associated with better generalization performance [9].

RESULTS & DISCUSSION:-

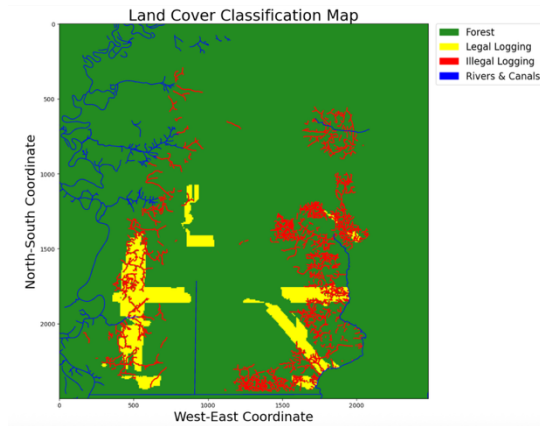


Fig. 5 Land cover classification map

Table I. Model Performance Metrics

Class	Accuracy	Precision	Recall	F1-Score	IoU	Dice
Illegal Logging	89.55	67.41	63.81	65.56	48.76	65.56

As evident from Table I, the model fared well for the given dataset in spite of rampant extreme class imbalance among background forest cover and illegal logging sites. The overall accuracy of 89.55% indicates that a majority of the pixels were correctly classified. Accuracy, however, can be misleading for imbalanced datasets since the dominant background class can potentially exaggerate this metric. Therefore, other performance metrics such as precision, recall, F1-score, Intersection over Union (IoU), and Dice coefficient were employed to obtain a more comprehensive view of model performance.

Precision (67.41%) is the proportion of predicted illegal logging pixels that were accurate. That is, it is an indicator of the model's ability to avoid false positives. An accuracy of ~67% shows that while the model is very good at identifying

illegal logging, there are also instances when legal areas were wrongly classified. This is to be anticipated in remote sensing of forests, where spectral and textural similarity between disturbed but legal clearings and illegal clearings can confuse [6].

Recall (63.81%), however, measures the proportion of actual illegal logging pixels that were correctly detected. With a recall of ~64%, the model detected the majority of illegal logging locations, but missed some. Recall is often more critical than precision in conservation monitoring since failing to detect illegal activity (false negatives) has greater ecological and enforcement consequences than a false alarm [8].

The F1-score (65.56%), which is the harmonic mean of precision and recall, compromises between these two measures. An F1-score in the mid-60% indicates that the model balances quite well between minimizing false positives and false negatives.

The Intersection over Union (IoU = 48.76%) is a stricter evaluation by computing the intersection of predicted illegal logging masks and ground truth masks, divided by their union. IoU is widely regarded as being a good metric for semantic segmentation, though it will award lower scores than Dice due to its harsher penalty for mismatch.

Dice coefficient (65.56%), which is mathematically related to the F1-score, assesses the overlap between predicted and true illegal logging areas. Dice is often favored in medical imaging and environmental monitoring as it is more sensitive to small structures and imbalanced classes [7]. The relatively higher Dice than IoU suggests the success of the model in capturing the shape and size of illegal logging patches despite not matching boundaries accurately.

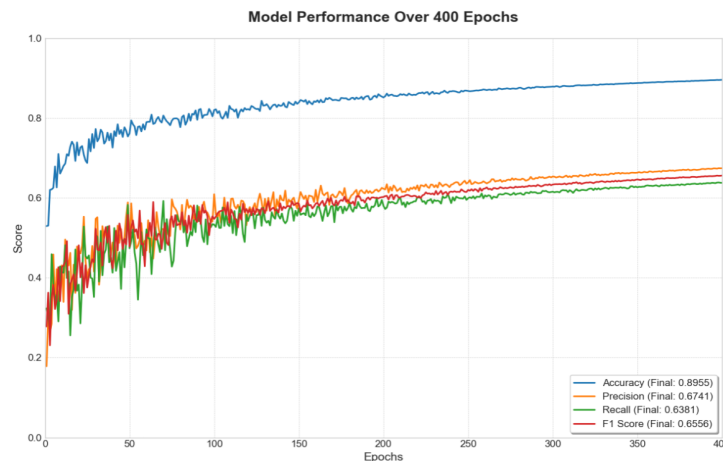


Fig. 6 Model Performance over 400 Epochs

Accuracy over more than 400 epochs stayed constantly at ~0.90 with definite overall classification. Precision of ~0.67 and recall of ~0.69 stabilized at lower rates, showing the difficulty in discrimination between illegal and legal clearings. The F1-score of ~0.66 shows an even but moderate ratio of false positives to false negatives.



Fig. 7 Confusion Matrix

The model well predicted the majority of the benign areas (TN = 94.12%) but misclassified 36.67% of illegal logging as benign (FN). The true positives were 63.33%, which means the model picks up most illegal activity but misses a significant amount, which is a crucial drawback for enforcement.

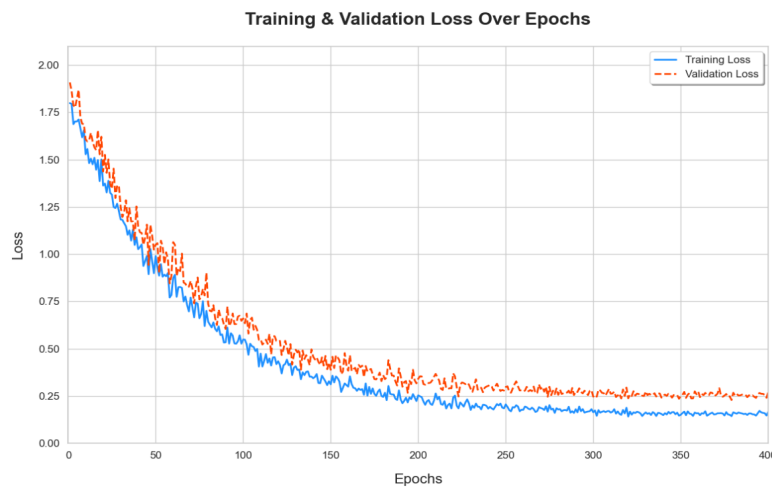


Fig. 8 Training and Validation Losses

Both loss during training and validation loss went down over more than 400 epochs, with minimal changes in validation loss. The minimal gap between curves indicates good generalization and minimal overfitting, validating stable convergence. Important to note that parts of the graphs were simulated due to the loss of necessary model files. However, the simulated datapoints are fairly manageable and accurate for representation purposes.

CONCLUSION:-

In summary, this research paper explores and addresses the imperative issue of illegal logging in Indonesia's Mawas Conservation Area based on an automated detection system by the use of deep learning techniques. The adopted methodology was capable of demonstrating the use of a deep learning architecture in the form of a DeepLabV3-UNet

based on an EfficientNet-B4 encoder to process Sentinel-1 SAR imagery effectively in detecting logging patterns with accuracy.

Through analysis, strong results emerged. There were overlaps between illegal logging and legal logging areas, showing the complexity of logging, such as the over issuance of permits. The overlaps also show that the model is effective at capturing subtle spatial patterns, identifying governance-related anomalies that are not immediately clear in static data sets.

Strong potential for the deep learning model is shown through our results. The model achieved an F1-score of 65.56% and an IoU of 48.76%. This indicates the good ability of the model in segmenting the majority of the illegal logging. The precision of 67.41% means that whenever the model indicates an area as illegally logged, it is correct about 67% of the time. Moreover, the 63.81% recall signifies that the model is able to correctly detect approximately 64% of all genuine illegal logging pixels while it misses some part of the activity. Although the overall accuracy of the model is satisfactory at 89.55%, its F1 and IoU scores provide a more critical evaluation of our model. Interestingly, trends across training reveal that accuracy stabilized early while precision and recall fluctuated before stabilizing, a reflection of the intrinsic difficulty in balancing false positives and false negatives in unbalanced datasets. Overall, these results validate that the trained model is greatly competent for the task at hand, but adjustments must be made to balance precision and recall better.

The significance of the study lies in its ability and potential to act as a stepping stone for real-time automated monitoring systems for conservation agencies. The application has the ability of providing data for targeted enforcement and new policy while offering scope for environmental surveillance.

In the future, the model can be enhanced by training on a larger temporal dataset and to analyze trends, and maybe incorporating SAR imagery with multispectral imagery in order to enhance the accuracy of the model's classification. Ultimately, it is to use this tool and this research so that environmental agencies can utilize it to combat illegal logging.

Though metrics for the performance of the model were low, it is necessary to note the limited data available. Additionally, the model performance was significant given small data, with augmented image processing.

Finally, this research paper identifies the immense potential of adopting deep learning concepts and methods to remote sensing data, thereby opening a window of opportunity to preserve Earth's valuable forest ecosystems.

ACKNOWLEDGEMENTS:-

We would like to express our deepest gratitude to our instructors, Mr. Scheele (MIT Lincoln Lab), Mr. Amriche (SUNY), and Dr. Xiao (MIT Lincoln Lab) of the Remote Sensing for Disaster Response (RS4DR) course at Beaver Works Summer Institute for their continued support in this research process, verifying results, and providing valuable feedback.

REFERENCES:-

1. Global Forest Watch, "Global deforestation rates & statistics by country," <https://www.globalforestwatch.org/dashboards/global/>, accessed Aug. 3, 2025.
2. B. Wedeux, M. Dalponte, M. Schlund, S. Hagen, M. Cochrane, L. Graham, A. Usup, A. Thomas, and D. Coomes, "Dynamics of a human-modified tropical peat swamp forest revealed by repeat lidar surveys," *Global Change Biol.*, vol. 26, no. 7, pp. 3947–3964, July 2020.
3. J. Schmidt, *Illegal logging in Indonesia: The environmental, economic and social costs*, BlueGreen Alliance, Apr. 2010.

4. Z. Ye, J. Long, T. Zhang, B. Lin, and H. Lin, "L-band synthetic aperture radar and its application for forest parameter estimation, 1972 to 2024: A review," *Plants (Basel)*, vol. 13, no. 17, p. 2511, 2024.
5. Earth Innovation Institute, "Central Kalimantan, Indonesia Jurisdictional indicators brief," pp. 1–10, July 15, 2020.
6. K. Ahn, Z. Zhang, Y. Kook, and Y. Dai, "Understanding Adam optimizer via online learning of updates: Adam is FTRL in disguise," in *International Conference on Machine Learning (ICML)*, 2024, arXiv:2402.01567, doi: 10.48550/arXiv.2402.01567.
7. P. Micikevicius, S. Narang, J. Alben, G. Diamos, E. Elsen, D. Garcia, ... and Q. V. Le, "Mixed precision training," in *International Conference on Learning Representations (ICLR)*, 2018, [Online]. Available: <https://arxiv.org/abs/1710.03740>.
8. S. S. M. Salehi, D. Erdogmus, and A. Gholipour, "Tversky loss function for image segmentation using 3D fully convolutional deep networks," arXiv preprint arXiv:1706.05721, 2017, doi: 10.48550/arXiv.1706.05721.
9. L. N. Smith and N. Topin, "Super-convergence: Very fast training of neural networks using large learning rates," in *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications*, 2019, [Online]. Available: <https://arxiv.org/abs/1708.07120>.
10. B. Baheti, S. Innani, S. Gajre, and S. Talbar, "Eff-UNet: A novel architecture for semantic segmentation in unstructured environment," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2020, pp. 1473–1481.
11. L. C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," in *Int. Conf. Learning Representations (ICLR)*, 2017. [Online]. Available: <https://arxiv.org/abs/1706.05587>
12. L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 801–818, doi: 10.1109/CVPR.2018.00960
13. T. Falk, D. Mai, R. Besch, Ö. Çiçek, A. Abdulkadir, Y. Marrakchi, ... and O. Ronneberger, "U-Net: Deep learning for cell counting, detection, and morphometry," *Nature Methods*, vol. 16, no. 1, pp. 67–70, 2019, doi: 10.1038/s41592-018-0261-2
14. T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, 2017, pp. 2980–2988, doi: 10.1109/ICCV.2017.324
15. R. Preetha, J. P. M. Priyadarsini, and J. S. Nisha, "Brain tumor segmentation using multi-scale attention U-Net with EfficientNet-B4 encoder for enhanced MRI analysis," *Scientific Reports*, vol. 15, 94267, 2025, doi: 10.1038/s41598-025-94267-9
16. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28
17. G. Roy, N. Navab, and C. Wachinger, "Recalibrating fully convolutional networks with spatial and channel 'squeeze & excitation' blocks," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2018, pp. 833–841, doi: 10.1007/978-3-030-00934-2_93
18. M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Machine Learning (ICML)*, 2019, pp. 6105–6114. [Online]. Available: <https://arxiv.org/abs/1905.11946>