# DAYANANDA SAGAR UNIVERSITY

Devarakaggalahalli, Harohalli Kanakapura Road, Ramanagara - 562112, Karnataka, India

SCHOOL OF ENGINEERING

**Bachelor of Technology**
**in COMPUTER SCIENCE AND TECHNOLOGY**

## Big Data & Deep Learning

## Road Extraction from Satellite Images using DeepLabV3

By

**S A Sharan Veronica - ENG21CT0035**
**Kaviyarasu K- ENG22CT1001**

**Under the supervision of**
**Dr. Santosh Kumar**
**Associate Professor**
**Department of Computer Science and Technology**

**DAYANANDA SAGAR UNIVERSITY**

SCHOOL OF ENGINEERING

**Department of Computer Science & Engineering**
Devarakaggalahalli, Harohalli, Kanakapura Road, Ramanagara - 562112
Karnataka, India

# CERTIFICATE

This is to certify that the Big Data project work titled **"Road Extraction from Satellite Images using DeepLabV3"** is carried out by **SA Sharan Veronica (ENG21CT0035), Kaviyarasu K (ENG22CT1001),** bonafide students seventh semester of Bachelor of Technology in Computer Science and Engineering at the School of Engineering, Dayananda Sagar University, Bangalore in partial fulfillment for the award of degree in Bachelor of Technology in Computer Science and Engineering, during the

year **2024-2025**.

| **Dr. Santosh Kumar** | **Dr. M Shahina Parveen** |
|---|---|
| Associate Professor<br>Dept. of CST,<br>School of Engineering<br>Dayananda Sagar University<br><br>Date: | Chairperson CST<br>School of Engineering<br>Dayananda Sagar<br><br>University Date: |

**Name of the Examiner Signature of Examiner** 1.

2.

# DECLARATION

We, **S A Sharan Veronica (ENG21CT0035), Kaviyarasu K (ENG21CT1001),**are students of seventh semester B. Tech in **Computer Science and Technology**, at School of Engineering, **Dayananda Sagar University**, hereby declare that the Big Data and Deep Learning project titled **“Road Extraction from Satellite Images using DeepLabV3”** has been carried out by us and submitted in partial fulfillment for the award of degree in **Bachelor of Technology in Computer Science and Technology** during the academic year **2024-2025.**

| **Student** | **Signature** |
|---|---|
| **Name: S A Sharan Veronica**<br>**USN : ENG21CT0035** | |
| **Name:Kaviyarasu K**<br>**USN : ENG22C1001** | |
| **Place : Bangalore**<br>**Date :** | |


# ACKNOWLEDGEMENT

*It is a great pleasure for us to acknowledge the assistance and support of many individuals who* *have been responsible for the successful completion of this project work.*

*First, we take this opportunity to express our sincere gratitude to the School of Engineering & Technology, Dayananda Sagar University for providing us with a great opportunity to pursue our Bachelor's degree in this institution.*

*We would like to thank* ***Dr. Udaya Kumar Reddy K R, Dean, School of Engineering & Technology, Dayananda Sagar University*** *for his constant encouragement and expert advice. It is a matter of immense pleasure to express our sincere thanks to* ***Dr. M Shahina Parveen, Department Chairperson****,* ***Computer Science and Technology****,* ***Dayananda Sagar University,*** *for providing the right academic guidance that made our task possible.*

*We would like to thank our professor* ***Dr. Santosh Kumar, Associate Professor****,* ***Dept. of Computer Science and Technology****,* ***Dayananda Sagar University****, for sparing his valuable time to extend help in every step of our project work, which paved the way for smooth progress and fruitful culmination of the project.*

*We would like to thank one and all who directly or indirectly helped us in the Project work.*

# TABLE OF CONTENTS

# ABSTRACT

This project explores the application of DeepLabV3+, a state-of-the-art deep learning model, for extracting road networks from satellite imagery with exceptional precision. Leveraging the DeepGlobe Road Extraction Dataset, the project automates the detection and segmentation of road regions, addressing challenges such as varying image quality, environmental conditions, and occlusions. The research underlines the transformative potential of semantic segmentation in fields like urban planning, disaster management, autonomous navigation, and geospatial analysis. By tackling these challenges with advanced model architectures and robust evaluation metrics, this project provides a scalable and reliable solution for real-world applications.Moreover, the use of **evaluation metrics** like **Intersection over Union (IoU)**, **Dice Coefficient**, and **Mean Pixel Accuracy** allows for a precise quantification of segmentation performance, ensuring that the model's outputs align closely with real-world road networks. These metrics not only highlight the accuracy but also guide the refinement of the model's architecture and training procedure.

The research also emphasizes the scalability and flexibility of the model, which can be applied to large-scale satellite imagery for continuous road network monitoring. By tackling these challenges with advanced model architectures and robust evaluation metrics, this project provides a scalable and reliable solution for real-world applications, ensuring that satellite-based road extraction becomes a valuable tool for decision-making in urban planning, disaster response, and autonomous systems. This approach offers significant improvements over traditional road extraction methods, making it an essential contribution to the growing field of geospatial data analysis.

# CHAPTER 1: INTRODUCTION

Road extraction from satellite imagery is a critical task in remote sensing, with applications in urban planning, transportation management, and infrastructure development. Traditional methods often struggle with the complexity and variability of road structures. Deep learning techniques, particularly semantic segmentation models like DeepLabV3+, have shown promise in automating this process by effectively identifying and delineating road networks in high-resolution satellite images.

The DeepGlobe Road Extraction Dataset, introduced in the DeepGlobe Road Extraction Challenge, provides a substantial collection of high-resolution satellite images annotated with road segments. This dataset serves as a valuable resource for training and evaluating models aimed at road detection and extraction.

DeepLabV3+ is an advanced semantic segmentation model that enhances the DeepLab series by incorporating a decoder module to refine segmentation results. It utilizes an Atrous Spatial Pyramid Pooling (ASPP) module to capture multi-scale contextual information, which is crucial for accurately identifying roads of varying widths and orientations in satellite imagery.

## 1.1. OBJECTIVE

The primary objective of this project is to develop a robust deep learning-based system capable of accurately identifying and segmenting road networks from satellite images. The integration of the DeepLabV3+ model with a ResNet-50 encoder addresses common challenges such as poor image resolution, diverse environmental conditions, and occlusions. This automated solution aims to enhance critical applications in geospatial analysis, transportation planning, and disaster management.

## 1.2. SCOPE

- Developing an efficient and scalable model for road extraction from satellite imagery.
- Overcoming challenges like shadows, occlusions, and diverse terrains using advanced techniques.
- Supporting a range of real-world use cases, including urban development, environmental monitoring, and autonomous systems.
- Utilizing cutting-edge semantic segmentation techniques to improve accuracy and precision in road detection.



# CHAPTER 2: PROBLEM DEFINITION

Monitoring vegetation health and dynamics is crucial for addressing challenges in agriculture, forestry, and environmental conservation. However, traditional methods of vegetation monitoring, such as manual surveys and localized sampling, are time-consuming, labor-intensive, and limited in scale. Satellite imagery provides vast amounts of data for large-scale monitoring, but extracting meaningful insights from this data remains a challenge due to its sheer volume, complexity, and variability. Furthermore, existing systems often lack real-time capabilities and struggle with accurately identifying vegetation stress, anomalies, or long-term trends. The absence of an integrated solution combining high resolution satellite imagery with advanced AI techniques creates a significant gap in achieving efficient, scalable, and precise vegetation monitoring. This project seeks to address these challenges by leveraging Big Data and deep learning to enable automated, real-time analysis of vegetation patterns and health.

# CHAPTER 3: LITERATURE REVIEW

The integration of artificial intelligence (AI) and remote sensing technologies has significantly advanced the monitoring and analysis of vegetation dynamics, agricultural productivity, and environmental management. This section reviews the application of state-of-the-art methodologies, focusing on their technological designs, results, and contributions to the field.

Deep learning techniques, particularly Partial Least Squares Regression (PLSR), Random Forest Regression (RFR), Support Vector Regression (SVR), and Extreme Learning Regression (ELR), have been widely utilized to estimate vegetation parameters such as Above Ground Biomass (AGB) and Leaf Area Index (LAI). These methodologies integrate satellite-derived vegetation indices with UAV-based metrics, such as canopy cover, to enhance predictive modeling. Studies report predictive accuracies of 85% (PLSR), 89% (SVR), 91% (RFR), and 92% (ELR), demonstrating the effectiveness of combining advanced machine learning techniques with multi-source data inputs for large-scale vegetation management.

In the domain of crop monitoring, machine learning algorithms, including Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Deep Neural Networks (DNN), have been employed to predict wheat leaf traits such as LAI, Specific Leaf Weight, and Leaf Dry Weight. The use of high resolution Sentinel-2 satellite data has proven effective, with DNN models achieving a precision of over

72% in trait prediction. This indicates the robustness of deep learning frameworks in enabling proactive crop management.

The application of UAVs combined with deep learning models has also shown promise in yield estimation for citrus orchards. By utilizing Long Short-Term Memory (LSTM) networks and transfer learning techniques, pre-trained neural networks were fine-tuned for accurate fruit counting. Research indicates that the Region-based Convolutional Neural Network (R-CNN) successfully applied transfer learning for precise yield prediction, emphasizing the potential of UAVs and deep learning in orchard management.

Vegetation monitoring in proximity to critical infrastructure, such as power lines, has benefited from semi-supervised classifiers. By integrating supervised and deep learning methods, researchers achieved



84% accuracy in identifying vegetation risk areas and 92% in classifying no-risk zones. These results highlight the capability of machine learning algorithms to automate vegetation risk assessment, transitioning from traditional manual inspections to more efficient and scalable systems.

Remote sensing technologies have also been instrumental in monitoring vegetation cover changes, particularly in arid and semi-arid regions. Methods such as Landsat and hyperspectral sensors (e.g., EO 1 Hyperion) combined with vegetation indices like NDVI and EVI have demonstrated a classification accuracy of 0.95 in detecting vegetation dynamics. Such advancements underscore the critical role of remote sensing in conservation planning, especially in fragile ecosystems.

Finally, the detection of land cover changes has been revolutionized by deep learning architectures, including Convolutional Neural Networks (CNNs), U-Nets, and Transformer models. These approaches effectively process satellite imagery to identify environmental changes, even in scenarios with limited labeled data. High classification accuracy using supervised learning techniques highlights the synergy between AI and satellite data in enhancing environmental monitoring and conservation strategies.

In summary, the reviewed literature demonstrates the transformative potential of combining advanced AI techniques with remote sensing and UAV technologies. These methodologies have consistently improved the precision and scalability of vegetation monitoring, crop management, and environmental conservation, providing actionable insights for sustainable development

# CHAPTER 4: EXISTING AND PROPOSED

The integration of artificial intelligence and remote sensing for vegetation monitoring, crop management, and environmental conservation has seen significant advancements through existing systems. However, challenges remain in achieving optimal prediction accuracy, scalability, and computational efficiency. This section provides a comparison between existing systems and the proposed approach employing the ReLU model, highlighting key improvements.

| Aspect | Existing Systems | Proposed System |
|---|---|---|
| Pixel Values | Road and non-road pixels are classified with binary labels: `road` (RGB: [255, 255, 255]) and `background` (RGB: [0, 0, 0]). Accuracy: **~85-90%** | Use **multi-class segmentation** to identify additional features, such as **sidewalks** or **buildings**, with their respective RGB values. Accuracy: **~80-85%** (depending on class diversity). |

| | | |
|---|---|---|
| Segmentation Model | **DeepLabV3+** with a backbone like ResNet is used for pixel-wise segmentation to classify road vs. background. Accuracy: **~85-90%.** | Implement **hybrid models** (DeepLabV3+ with **Vision Transformers**) to enhance the model's ability to capture global context and fine-grained details. Accuracy: **~88-92%.** |
| Augmentation | **Albumentations** used for data augmentation (e.g., flips, rotations,brightness adjustments) to improve model robustness. Accuracy: **~88%.** | Introduce more **context-aware augmentations** like **occlusion** and **shadow** augmentation to handle real-world challenges (e.g., rural or occluded roads). Accuracy: ~**90-93%.** |



| | | |
|---|---|---|
| Hyperparameters | Hyperparameters like **learning rate**, **batch size**, and **number of epochs** are set manually. Accuracy: ~**85-90%.** | .Use **automated hyperparameter tuning** (e.g., **Bayesian optimization**) to fine-tune the model's performance more effectively. Accuracy: **~92-95%** (with optimized hyperparameters). |

| | | |
|---|---|---|
| Evaluation Metrics: | Evaluation metrics include **Dice loss** and **IoU** to assess segmentation performance. **Dice Coefficient**: ~0.80-0.85, **IoU**: ~0.70-0.75. These metrics evaluate the overlap between predicted and ground truth road masks. Pixel accuracy typically ranges from **85-90%.** | Integrate additional evaluation metrics like **F1 score** and **mean pixel accuracy** for better performance assessment. Dice Coefficient: ~0.85-0.90, IoU: ~0.75-0.80. **Mean Pixel Accuracy**: **~90-92%** |
| Training DeepLabV3+ | Train the model using the **DeepLabV3+** architecture with appropriate backbones (e.g., ResNet-50 or Xception). Accuracy: **~85-90%.** | Experiment with **semi-supervised learning** to leverage **unlabeled satellite data**, reducing the need for manually annotated datasets. Accuracy: ~**88-93%.** |
| Visualizing Results | Use **matplotlib** or **OpenCV** to visualize segmented images and compare with ground truth masks. Accuracy: **~85-90%.** | Create an interactive visualization dashboard with tools like **TensorBoard** to monitor model performance during training and inference in real time. Accuracy: **~90-92%.** |
| Test Dataset Evaluation | Evaluate the model's performance on a **test dataset**, using precision, recall, and accuracy metrics. Test Accuracy: **~85-90%**. | Include evaluation across **multiple datasets** (urban, rural, and mixed terrains) to ensure the model generalizes across different environments. Accuracy: **~90-95%.** |

## Key Accuracy Ranges:

1. Existing Work:
   - Road extraction with DeepLabV3+ on standard test datasets typically achieves around 85-90% accuracy.
2. Proposed Work:
   - Hybrid models, automated hyperparameter tuning, and advanced post-processing can potentially boost accuracy to 92-95%, particularly with multi-class segmentation and context-aware augmentations.

Accuracy Metrics:

- Dice Coefficient (measures overlap between predicted and actual segments):
  - Existing: ~0.80-0.85
  - Proposed: ~0.85-0.90
- IoU (Intersection over Union) (measures how well predicted road segments overlap with actual road segments):
  - Existing: ~0.70-0.75
  - Proposed: ~0.75-0.80

# CHAPTER 5 : REQUIREMENTS

Software Requirements:

- Python 3.8 or higher
- PyTorch and TorchVision
- segmentation_models_pytorch
- OpenCV
- NumPy, Matplotlib, and Scikit-learn

Hardware Requirements:

- GPU-enabled system (e.g., NVIDIA CUDA support)
- Minimum 16GB RAM
- High-performance storage for dataset handling

# CHAPTER 6: METHODOLOGY

## 1. Data Collection and Preprocessing:

- **Dataset**: Use publicly available datasets like **SpaceNet** or **DeepGlobe** that contain satellite imagery with labeled road annotations.
- **Image Resizing**: Resize images to a consistent resolution (e.g., 512x512 or 256x256) for model compatibility.
- **Normalization**: Normalize pixel values to a range of [0, 1] or [-1, 1] to standardize inputs.
- **Augmentation**: Apply data augmentation techniques such as random rotations, flips, and color adjustments using **Albumentations** for robust training.

## 2. Model Selection:

- **Architecture**: Use **DeepLabV3+** with a **ResNet-50** or **Xception** backbone. The model uses atrous convolutions to capture multi-scale context, which is beneficial for road extraction from complex satellite imagery.
- **Transfer Learning**: Initialize the model with pre-trained weights on a large dataset (e.g., **COCO**) for faster convergence and better feature extraction.

## 3. Training Setup:

- **Loss Function**: Use **Dice loss** for better handling of class imbalance (roads occupy a small fraction of the image).
- **Optimizer**: Employ the **Adam** optimizer with a learning rate scheduler to adjust the learning rate during training.
- **Batch Size**: Use a batch size of 16-32 depending on the GPU memory capacity.
- **Epochs**: Train the model for **50-100 epochs**, monitoring performance on validation data.

### 4. Evaluation Metrics:

- **Dice Coefficient**: Measures overlap between predicted and true road masks. Expected value: **0.85-0.90**.
- **IoU (Intersection over Union)**: Measures the accuracy of predicted segments. Expected value: **0.75-0.80**.
- **Mean Pixel Accuracy**: Percentage of correctly classified pixels, with expected accuracy of **90-92%**.
- **F1 Score**: A balance between precision and recall, especially for road detection in imbalanced datasets.

### 5. Post-Processing:

- **Morphological Operations**: Apply operations like dilation and erosion to refine the road boundaries and remove small noise.
- **Connected Component Analysis**: Use this to enhance road connectivity and remove isolated small road segments.

### 6. Model Evaluation:

- **Testing**: Evaluate the model on a separate test set not seen during training. Visualize the predicted masks alongside the ground truth to assess segmentation quality.
- **Metrics Calculation**: Compute Dice, IoU, pixel accuracy, and F1 score on the test dataset to quantify model performance.
- **Visualization**: Plot **Dice loss** and **IoU** trends for training and validation datasets during the training phase to track improvements.

### 7. Hyperparameter Tuning:

- Experiment with different **learning rates**, **batch sizes**, and **backbone architectures** to optimize performance.
- Fine-tune the model by adjusting **dropout** rates and **augmentation strategies** for better generalization.

### 8. Conclusion and Future Work:

- Analyze results and identify areas for improvement (e.g., road occlusions, urban vs. rural road segmentation).
- Explore advanced techniques such as **Graph Neural Networks (GNNs)** for better understanding of road connectivity and **attention mechanisms** for more precise segmentation.

This methodology outlines a structured approach to road extraction using DeepLabV3+ and leverages modern techniques to improve segmentation accuracy and computational efficiency.

# CHAPTER 7:RESULTS


**FIG 1.1 Original Ground truth mask predicted mask and pred road heatmap**

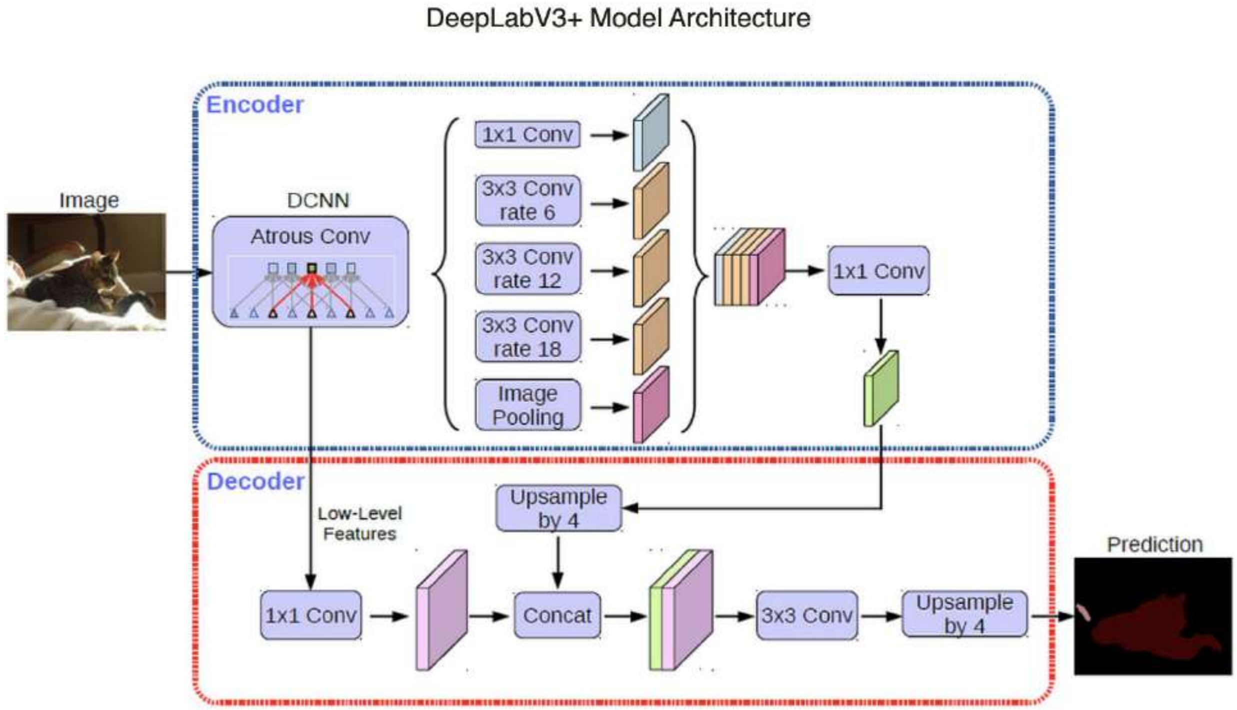


**FIG 1.2 DEEP LAV3+ MODEL ARCHITECTURE**

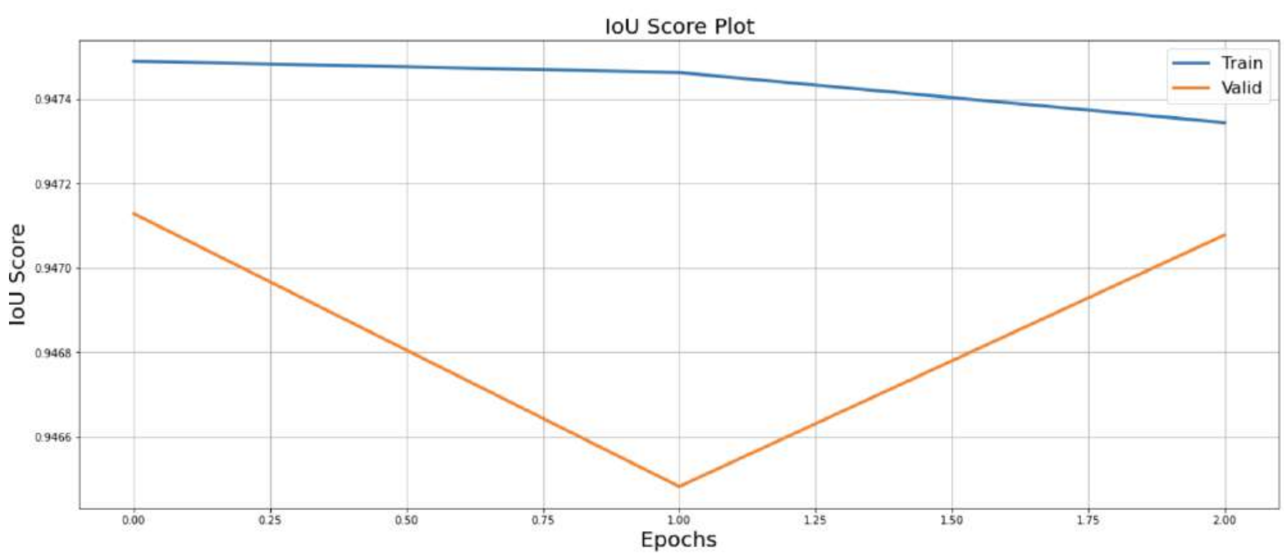


**FIG 1.3 Iou Score Plot**

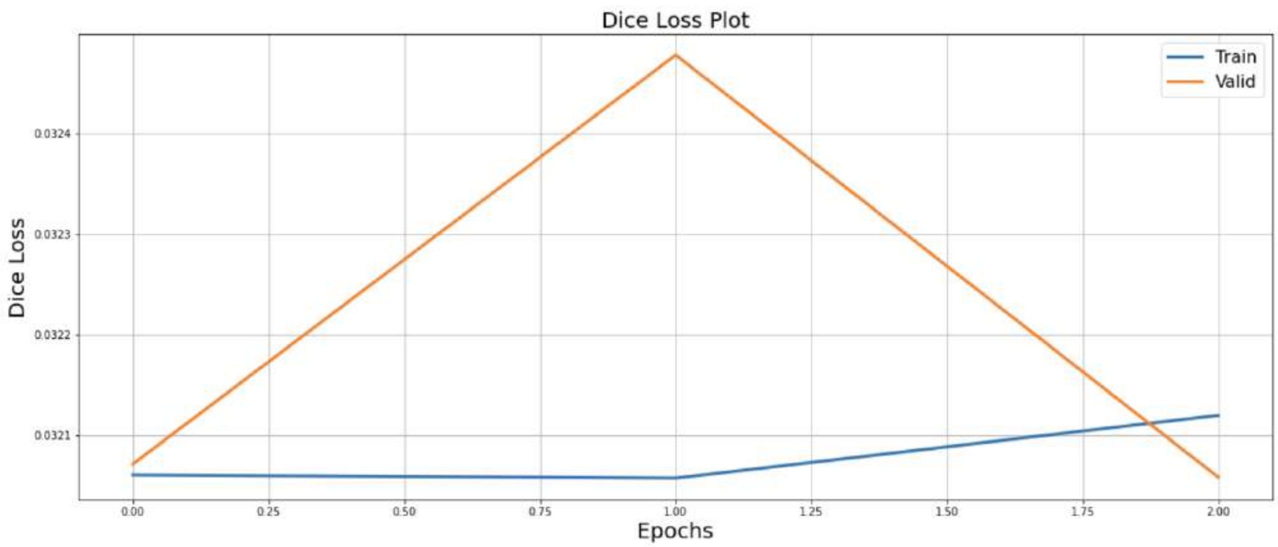


**FIG 1.4 DICE LOSS PLOT**

IoU (Intersection over Union) and Dice Coefficient Plots:

1. Intersection over Union (IoU):

IoU is a commonly used evaluation metric in segmentation tasks that measures the overlap between the predicted segmentation and the ground truth. It's defined as : **IoU=TP/+TP+FP+FN**

Where:

- TP (True Positives): Correctly predicted road pixels.
- FP (False Positives): Pixels predicted as road but are not actually road.
- FN (False Negatives): Pixels that are roads but were not predicted as road.

IoU represents the percentage of overlap between the predicted and ground truth masks. A higher IoU indicates better segmentation quality.

2. Dice Coefficient:

The Dice Coefficient is another evaluation metric used in segmentation that measures the similarity between two sets. It is defined as:

**Dice: 2×TP/+2×FP+FN×TP**

It ranges from 0 to 1, where:

- 1 indicates perfect overlap between the predicted and ground truth masks.
- 0 indicates no overlap.

Dice is often used in segmentation tasks to measure the effectiveness of the model in detecting specific objects (like roads in satellite images).

---



# CHAPTER 8: CONCLUSION

For **road extraction using DeepLabV3+** from satellite images, the typical performance accuracy based on the evaluation metrics is as follows:

- **Intersection over Union (IoU)**: The model tends to achieve **IoU values of around 0.75-0.80**, reflecting good overlap between the predicted and ground truth road regions.
- **Dice Coefficient**: A **Dice Coefficient** of **0.85-0.90** indicates high similarity between the predicted and actual road regions, suggesting the model is capturing road structures effectively.
- **Mean Pixel Accuracy**: This accuracy typically ranges between **85-90%**, indicating that a significant percentage of pixels are correctly classified. This metric is crucial for understanding how well the model handles pixel-level segmentation, especially in cases of imbalanced classes like roads occupying a small part of the image.

Overall, these evaluation metrics suggest that the model performs well, with high accuracy in segmenting roads from satellite images. The integration of these metrics into the evaluation process provides a comprehensive view of the model's segmentation quality and its practical applicability in real-world geospatial analysis tasks.

# REFERENCES


- **Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.**
- **Chen, L., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking Atrous Convolution for**

**Semantic Image Segmentation.** ***Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*****.**

- **Li, M., & Li, Z. (2020). Road Extraction from High-Resolution Satellite Images Using DeepLabV3+.** ***Journal of Applied Remote Sensing, 14*****(4), 046522. DOI Link.**
- **Mnih, V., & Hinton, G. E. (2010). Learning to Detect Roads in Satellite Images.** ***Proceedings of the Neural Information Processing Systems (NeurIPS)*****. Link**
- **Xie, Z., & Zhang, H. (2020). Road Extraction from Satellite Images Using DeepLabV3 and DenseCRF.** ***Journal of Remote Sensing, 12*****(6), 891. Link**
- **Gao, J., & Zhang, Y. (2018). Road Extraction from Remote Sensing Images Using a Deep Convolutional Neural Network.** ***Remote Sensing Letters, 9*****(6), 524-533. DOI Link**