इसरो isro

BHARATIYA ANTARIKSH HACKATHON 2025

Powered by H2S

Team Name: Optimizers

Team Leader Name: Prakhar Asthana

Problem Statement : Novel Approaches for Optimizing Deep Learning in Earth Observations with Imbalanced Data

Team Members

Team Leader:

Name: Prakhar Asthana
College: IIT Guwahati

Team Member-1:

Name: Manthan Bagade
College: IIT Guwahati

Team Member-2:

Name: Pushpendra Singh
College: IIT Guwahati

Team Member-3:

Name: Krishang Vyas
College: IIT Guwahati

Brief about the Idea:

We present a novel deep learning pipeline specifically designed to enhance segmentation of rare classes such as wetlands and water bodies in high-resolution LISS-IV satellite imagery.

We approach integrates five key innovations:

- **Spatial-Aware SegFormer:** Adds an auxiliary saliency head to guide the network's focus toward rare-class regions and inform class-aware sampling.
- **Adaptive Rare-Class Balanced Tversky Loss:** A custom loss that blends per-class Tversky weighting with focal modulation, tuned to class frequency and hard pixels.
- **Class-Adaptive Sharpness Minimization:** A novel optimizer that adapts sharpness-aware training to emphasize rare-class signal within each batch.
- **Rare-Class Sampling:** During training, 50% of crops will be drawn from high-saliency regions predicted by the auxiliary head, ensuring rare-class visibility.
- **Model Babysitting:** Continuous metrics monitoring and hyperparameter tuning to maximize rare-class recall without the loss of generalization.

Existing Approaches:

- Rely on cross-entropy or standard Dice loss, which fail on rare classes.
- Ignore class imbalance at the optimizer and sampling level.
- Use backbones that don't focus attention on minority regions.

Our Approach:

- Uses a spatial-aware transformer with an auxiliary saliency head that drives attention on rare classes and helps in rare-class sampling.
- Uses a custom loss, optimizer and sampling technique that is desgined specifically for rare-class recall.

How are we solving the Core Problem:

- Our loss function ensures that rare classes give steeper gradients.
- Our Optimizer ensures that the gradients converge to a flatter minima for better generalization across classes.
- Our Sampling Technique using the saliency head ensures that rare classes are seen frequently during training.

Our Changes in the Original Segformer Architecture

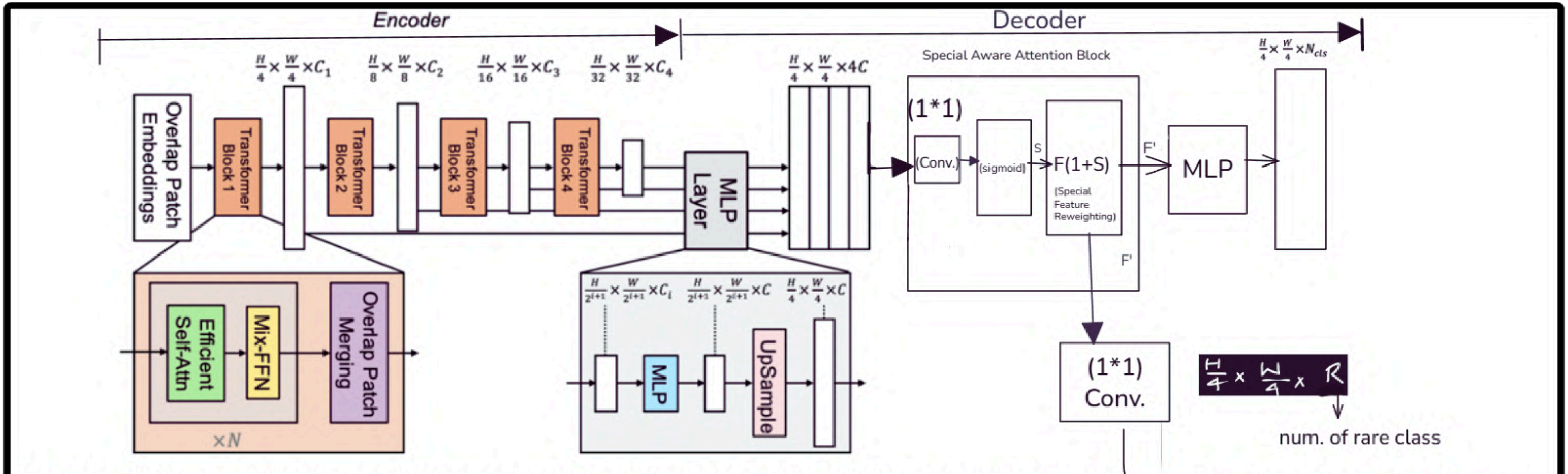


Figure 2: **The proposed SegFormer framework** consists of two main modules: A hierarchical Transformer encoder to extract coarse and fine features; and a lightweight All-MLP decoder to directly fuse these multi-level features and predict the semantic segmentation mask. "FFN" indicates feed-forward network.

Proposed Architecture

What This Architecture Enables:

- Direct focus on rare regions via learned saliency.
- Data efficient learning through saliency-guided sampling.
- Robust convergence with rare aware optimization.
- Each component helps solve class imbalance without the loss of precision of the majority classes.

Loss Function

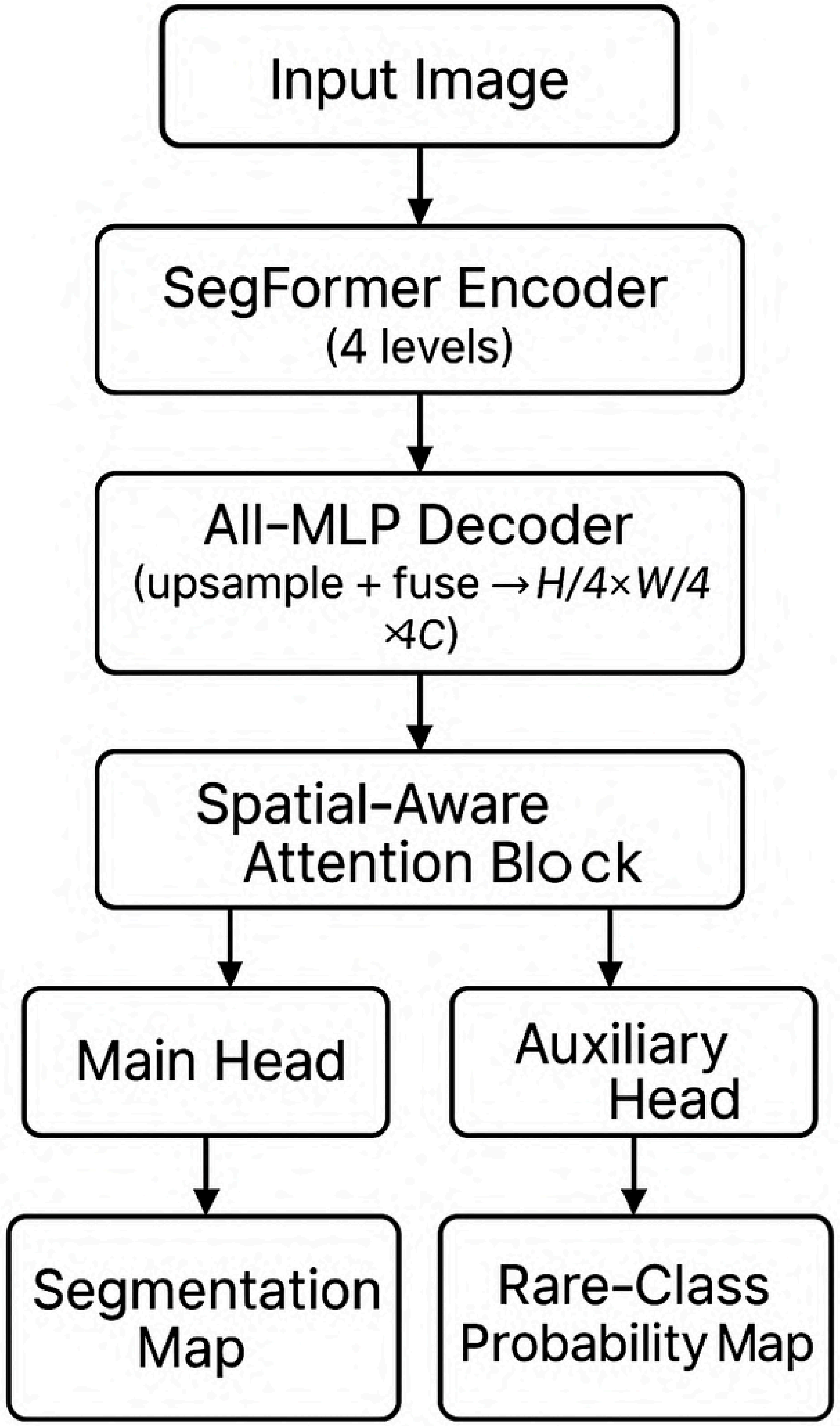
$$L_{ARBT} = \sum_{c=1}^{C} w_c \cdot \left(1 - \frac{\sum_i P_i^c G_i^c}{\sum_i P_i^c G_i^c + \alpha_c \sum_i P_i^c (1-G_i^c) + (1-\alpha_c) \sum_i (1-P_i^c) G_i^c + \epsilon}\right) + \lambda \sum_{i,c} |P_i^c - G_i^c|^\gamma$$

Optimizer

$$\epsilon = \epsilon_0 \cdot (1 + \kappa \cdot (1 - p_r))$$

$$\text{Ascent step:} \quad \theta^+ = \theta + \epsilon \cdot \frac{\nabla_\theta \mathcal{L}}{\|\nabla_\theta \mathcal{L}\| + \epsilon'}$$

$$\text{Descent step:} \quad \theta \leftarrow \theta - \eta \cdot \nabla_\theta \mathcal{L}(\theta^+)$$



The Detailed Explanations of the Loss Function and Optimizer are mentioned in later slides

Powered by

Wireframe

Technologies to be used in the solution:

In this project, many techniques and frameworks are employed, including Python, PyTorch, SegFormer transformers, and CNN-based attention mechanisms for high-resolution semantic segmentation. Geospatial tools such as GDAL and OpenCV are used for processing LISS-IV satellite imagery and generating precise land cover masks. To enhance rare-class detection, the pipeline integrates a spatial-aware attention block, dual-head architecture, and custom loss functions like ARBT, along with a Class-Adaptive Sharpness Minimization (CASM) optimizer.

Model training and evaluation can be conducted on Kaggle or Google Colab using GPU acceleration, mixed precision training, and saliency-guided sampling for data efficiency. Visualization and analysis are supported by TensorBoard, Matplotlib, and QGIS for both quantitative metrics and qualitative inspection. This integrated use of deep learning, geospatial analysis, and training optimization ensures a robust, scalable, and rare-class sensitive segmentation system.

We will build a simple web interface using Flask for the backend and HTML, CSS, and JavaScript for the frontend. The interface will allow users to upload satellite images, view segmentation results, and download predicted masks. The model will be integrated into the Flask server, making it easy to run inference and display results dynamically. This setup ensures easy deployment and accessibility through any web browser.

Mathematical model of Solution

Loss Function: Adaptive Rare-Class Balanced Tversky Loss (ARBT)

Problem: Standard losses (e.g., Cross Entropy, Dice) bias heavily toward majority classes, under-penalizing rare-class errors.

We define our total loss as a combination of class-weighted Tversky loss and a focal regularization term:

$$\mathcal{L}_{\text{ARBT}} = \sum_{c=1}^{C} w_c \cdot \left(1 - \frac{\sum_i P_i^c G_i^c}{\sum_i P_i^c G_i^c + \alpha_c \sum_i P_i^c (1 - G_i^c) + (1 - \alpha_c) \sum_i (1 - P_i^c) G_i^c + \epsilon}\right) + \lambda \sum_{i,c} |P_i^c - G_i^c|^\gamma$$

- P_i^c Predicted probability at pixel i for class c
- G_i^c Ground truth one-hot encoding at pixel i for class c
- f_c Frequency of class c in the training dataset
- $\alpha_c = 1 - \left(\frac{f_c}{\max_k f_k}\right)^\eta$ (class-adaptive balance)
- $w_c = \frac{1}{(f_c + \epsilon)^\delta}$ (class-frequency-based weight)
- λ Scaling factor for focal penalty
- γ Exponent to emphasize hard pixels
- ϵ Small constant to avoid division by zero

Mathematical model of Solution

Optimizer: Class-Adaptive Sharpness Minimization (CASM)

Problem: Standard optimizers do not adapt based on rare-class presence in a batch. SAM improves generalization, but treats all pixels equally, making it ineffective in rare-class learning.

We extend SAM by adapting the perturbation scale based on the rare-class density in the batch:

$$\epsilon = \epsilon_0 \cdot (1 + \kappa \cdot (1 - p_r))$$

The optimizer proceeds with a two-step update:

$$\text{Ascent step:} \quad \theta^+ = \theta + \epsilon \cdot \frac{\nabla_\theta \mathcal{L}}{\|\nabla_\theta \mathcal{L}\| + \epsilon'}$$

$$\text{Descent step:} \quad \theta \leftarrow \theta - \eta \cdot \nabla_\theta \mathcal{L}(\theta^+)$$

θ	Model parameters	κ	Scaling factor controlling sensitivity to p_r
$\mathcal{L}$	Training loss (ARBT + auxiliary loss)	ϵ'	Small constant for denominator stability
p_r	Fraction of rare-class pixels in the batch	η	Learning rate
ϵ_0	Base perturbation scale	θ^+	Temporarily perturbed parameters (SAM ascent)

Powered by H2S

Mathematical model of Solution

Component: Spatial-Aware Attention Block + Auxiliary Head + Rare-Class Guided Sampling

Problem: Rare-class regions are spatially small and sparsely distributed. Standard decoders and sampling strategies often miss them entirely during training.

We introduce a saliency-driven spatial attention block to emphasize rare-class regions in the feature map:

$$S = \sigma\left(\mathrm{Conv}_{1\times 1}(F)\right), \quad F' = F \cdot (1 + S)$$

The auxiliary head is supervised to predict rare-class pixel locations using a binary loss:

$$\mathcal{L}_{\mathrm{aux}} = \frac{1}{B}\sum_{b=1}^{B}\sum_{r\in\mathcal{R}} \mathrm{BCE}\left(S_b^{(r)}, ⊮(Y_b = r)\right)$$

During training, we use the output saliency map to guide sampling:

$$\text{Sample 50\% crops from top-}k\text{\% regions in } S$$

F	Decoder feature map $(B, 4C, H/4, W/4)$	Y_b	Ground truth label map for batch b
S	Saliency map from 1×1 conv + sigmoid	$\mathcal{R}$	Set of rare class indices
F'	Reweighted feature map used by heads	$S^{(r)}$	Predicted rare-class probability map
σ	Sigmoid activation function	$⊮(Y_b = r)$	Binary mask of class r

Powered by H2S

BHARATIYA ANTARIKSH HACKATHON
2025

THANK YOU