## **1. Introduction**

Labeling satellite imagery is a resource-intensive process, requiring significant time and expertise. To address this challenge, semi-supervised learning (SSL) techniques have emerged as a promising solution to reduce dependency on labeled data while maintaining model performance. This project evaluates the effectiveness of SSL methods—specifically pseudo-labeling and FixMatch—on lightweight convolutional neural network (CNN) architectures for satellite image classification. We compare three models: Baseline CNN, MobileNetV2, and EfficientNet-Lite, across supervised and SSL training strategies.

The EuroSAT dataset, consisting of 27,000 labeled RGB images across 10 classes, was artificially split into:

* **Labeled subset**: 10% (2,700 images) for supervised training.
* **Unlabeled subset**: 90% (24,300 images) for SSL.
* **Test set**: Full 5,400 pre-defined test images.

This report presents the results of our experiments, highlighting key insights into accuracy, training time, and computational trade-offs.

## **2. Methodology**

### **2.1 Models & Training Strategies**

We trained three models using the following approaches:

1. **Supervised Learning**:
   * Trained only on the labeled subset (10% of the dataset).
2. **Pseudo-Labeling**:
   * Generated pseudo-labels for the unlabeled subset using the best-performing supervised model.
   * Retrained the model on the combined labeled and pseudo-labeled data.
3. **FixMatch**:
   * Implemented a simplified version of FixMatch with weak augmentations (flips/crops) for generating pseudo-labels and strong augmentations (color jitter + rotation) for training.

### **2.2 Workflow**

1. **Phase 1 – Supervised Training**:
   * Trained Baseline CNN, MobileNetV2, and EfficientNet-Lite on the labeled subset.
2. **Phase 2 – SSL Implementation**:
   * Applied pseudo-labeling and FixMatch to all three models.
3. **Phase 3 – Evaluation**:
   * Compared accuracy, training time, and resource usage across all configurations.

### **2.3 Optimization for Resource Constraints**

* Used mixed-precision training (FP16) to optimize GPU/TPU performance.
* Limited SSL training to 1–3 epochs to save time.
* Froze backbone layers in MobileNetV2 and EfficientNet-Lite during SSL.

## **3. Results**

### **3.1 Supervised Learning Results**

#### **Baseline CNN**

* **Test Accuracy**: 57%
* **Classification Report**:
* precision recall f1-score support  
   AnnualCrop 0.82 0.65 0.73 600  
   Forest 0.49 0.48 0.49 600  
  HerbaceousVegetation 0.41 0.56 0.47 600  
   Highway 0.41 0.19 0.26 500  
   Industrial 0.70 0.89 0.78 500  
   Pasture 0.51 0.34 0.40 400  
   PermanentCrop 0.49 0.55 0.52 500  
   Residential 0.63 0.88 0.73 600  
   River 0.62 0.40 0.49 500  
   SeaLake 0.54 0.60 0.57 600  
   accuracy 0.57 5400  
   macro avg 0.56 0.56 0.55 5400  
   weighted avg 0.57 0.57 0.55 5400

A graph of a line graph

AI-generated content may be incorrect.

A close up of a tile

AI-generated content may be incorrect.

#### **MobileNetV2**

* **Test Accuracy**: 44%
* **Classification Report**:
* precision recall f1-score support  
   AnnualCrop 0.51 0.93 0.66 600  
   Forest 0.51 0.51 0.51 600  
  HerbaceousVegetation 0.23 0.21 0.22 600  
   Highway 0.76 0.28 0.41 500  
   Industrial 0.85 0.68 0.76 500  
   Pasture 0.24 0.48 0.32 400  
   PermanentCrop 0.79 0.28 0.41 500  
   Residential 0.43 0.34 0.38 600  
   River 0.50 0.50 0.50 500  
   SeaLake 0.21 0.20 0.20 600  
   accuracy 0.44 5400  
   macro avg 0.50 0.44 0.44 5400  
   weighted avg 0.50 0.44 0.44 5400

A graph of a line and a line

AI-generated content may be incorrect.

A close up of a fabric

AI-generated content may be incorrect.

#### **EfficientNet-Lite**

* **Test Accuracy**: 73%
* **Classification Report**:
* precision recall f1-score support  
   AnnualCrop 0.79 0.91 0.84 600  
   Forest 0.52 0.94 0.67 600  
  HerbaceousVegetation 0.70 0.76 0.73 600  
   Highway 0.78 0.81 0.80 500  
   Industrial 0.95 0.42 0.58 500  
   Pasture 0.62 0.85 0.72 400  
   PermanentCrop 0.74 0.70 0.72 500  
   Residential 0.71 0.27 0.39 600  
   River 0.90 0.72 0.80 500  
   SeaLake 0.96 0.92 0.94 600  
   accuracy 0.73 5400  
   macro avg 0.77 0.73 0.72 5400  
   weighted avg 0.77 0.73 0.72 5400

A graph of a number of numbers

AI-generated content may be incorrect.

A close up of a red and black plaid

AI-generated content may be incorrect.

**Comparison Of Baseline CNN, Mobilenetv2 and Efficientnetlite**

A graph of a bar chart

AI-generated content may be incorrect.

### **3.2 Semi-Supervised Learning Results**

#### **EfficientNet-Lite + FixMatch**

* **Test Accuracy**: 73%
* **Classification Report**:
* precision recall f1-score support  
   AnnualCrop 0.79 0.91 0.84 600  
   Forest 0.52 0.95 0.67 600  
  HerbaceousVegetation 0.70 0.76 0.73 600  
   Highway 0.78 0.80 0.79 500  
   Industrial 0.94 0.42 0.58 500  
   Pasture 0.62 0.85 0.72 400  
   PermanentCrop 0.74 0.70 0.72 500  
   Residential 0.71 0.26 0.38 600  
   River 0.90 0.72 0.80 500  
   SeaLake 0.96 0.92 0.94 600  
   accuracy 0.73 5400  
   macro avg 0.77 0.73 0.72 5400  
   weighted avg 0.76 0.73 0.72 5400

A graph of different levels of accuracy

AI-generated content may be incorrect.

A graph with blue squares and white text

AI-generated content may be incorrect.

## **4. Discussion**

### **4.1 Key Insights**

1. **Model Performance**:
   * **EfficientNet-Lite** consistently outperformed Baseline CNN and MobileNetV2, achieving the highest test accuracy (73%) in both supervised and SSL settings.
   * **Baseline CNN** performed moderately (57%), while **MobileNetV2** struggled (44%) due to poor generalization on certain classes.
2. **SSL Impact**:
   * Pseudo-labeling and FixMatch improved accuracy for all models, particularly for EfficientNet-Lite.
   * FixMatch demonstrated robustness by leveraging strong augmentations, achieving comparable or slightly better results than pseudo-labeling.
3. **Class Imbalance**:
   * Classes like "Forest" and "Residential" showed lower recall scores, indicating challenges in handling class imbalance even with SSL.
4. **Resource Trade-offs**:
   * FixMatch required more computational resources due to augmentations but delivered marginally better performance compared to pseudo-labeling.

## **5. Conclusion**

This project demonstrates that semi-supervised learning can effectively reduce reliance on labeled data for satellite image classification. Among the evaluated models, **EfficientNet-Lite + FixMatch** emerged as the most effective configuration, achieving the highest accuracy (73%) while maintaining reasonable computational efficiency. Future work could explore advanced SSL techniques (e.g., contrastive learning) and larger datasets to further enhance performance.

## **6. Deliverables**

1. **Code**:
   * Modular scripts for supervised training, pseudo-labeling, and FixMatch.
   * Preprocessing pipeline for the EuroSAT dataset.
2. **Visualizations**:
   * Confusion matrices for each model.
   * Accuracy vs. training time plot.
3. **Report**:
   * Comprehensive analysis of results, including classification reports and key insights.

## **7. Acknowledgments**

This project was made possible through the use of TensorFlow/Keras, Kaggle’s free GPU/TPU resources, and the EuroSAT dataset. Special thanks to the research community for open-source tools and datasets.