## Land Type Classification Using Satellite Images

---

### Project Overview

This project (Project 7: Land Type Classification using Sentinel-2 Satellite Images) focuses on developing an automated Deep Neural Network (DNN) model for Land-Use and Land-Cover (LULC) Classification. Accurate classification of major land types is essential for environmental monitoring, urban planning, and resource management.

### Key Project Components

- **Dataset:** The EuroSAT RGB dataset containing 27,000 satellite images sized 64 × 64 × 3 across 10 LULC classes.
- **Methodology:** Transfer Learning using a fine-tuned Convolutional Neural Network (CNN).
- **Milestone Goal:** Deliver a trained, optimized, and deployed real-time classification model.

---

### Milestone 1 — Data Preparation & Exploration (EDA)

The first milestone focused on collecting, exploring, and preprocessing satellite images.

### Data Exploration

- EDA confirmed that the EuroSAT dataset is well-balanced across all 10 classes, reducing bias during model training.

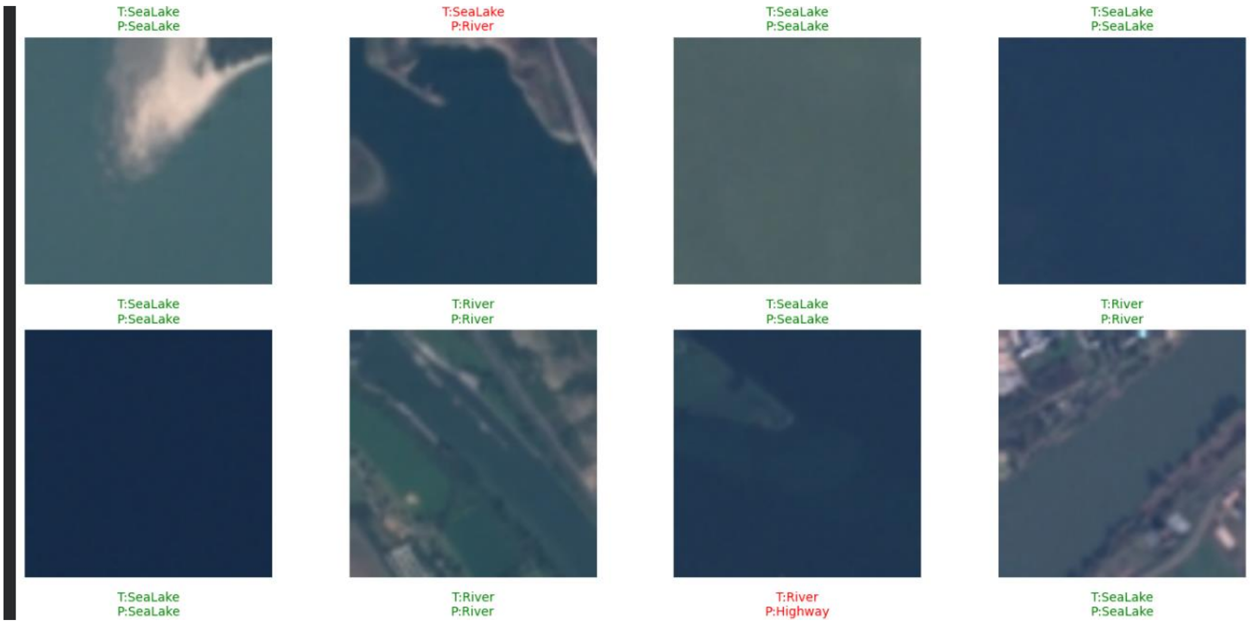
### Spectral Proxy Analysis

- **Vegetation Proxy Score (G / (R + B))**
  Highest values observed for:
  • Pasture (0.580)
  • Forest (0.561)
- **Blue Proxy Score (B / (R + G))**
  Correctly identified water bodies due to high blue reflectance:
  • SeaLake (0.823)
  • River (0.669)

### Preprocessing Steps

- Images normalized and resized

- Split into training, validation, and test sets
- **Data Augmentation:** Random rotation, zoom, and flipping to enhance robustness




---

### Milestones 2 & 3 — Model Development, Training, and Optimization

These milestones involved building, training, and fine-tuning the classification model.

#### Model Architecture

- Transfer Learning using the **EfficientNetB0** architecture for computational efficiency and high performance.

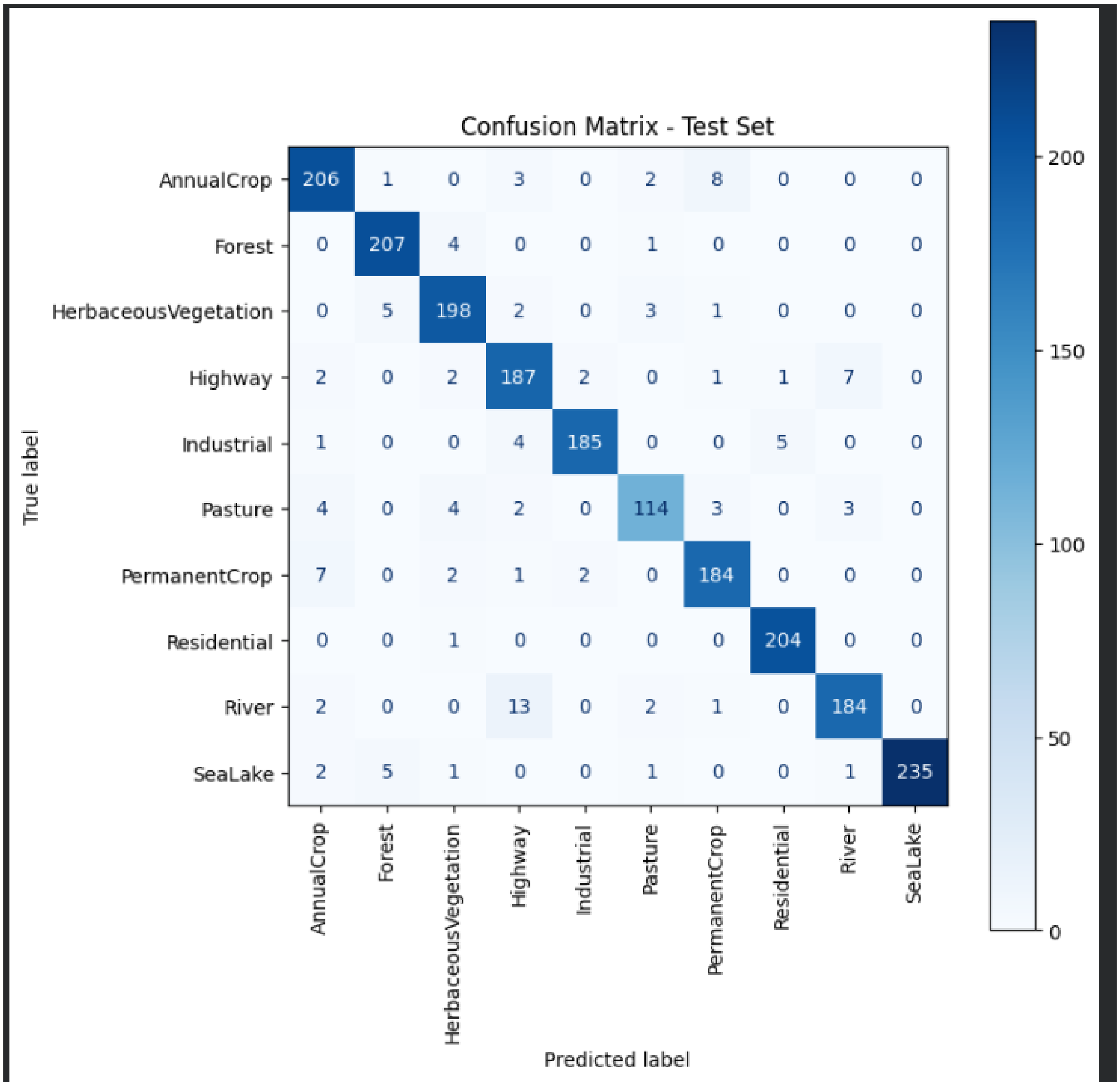
#### Two-Phase Training Strategy

- **Phase 1: Feature Extraction**
  EfficientNetB0 base frozen; only the classification head trained.
- **Phase 2: Fine-Tuning/Optimization**
  Layers after index 220 unfrozen for deeper tuning.

#### Hyperparameter Tuning

- Applied a very low learning rate (**$1 \times 10^{-5}$**) for fine adjustments during tuning

- Early Stopping and Batch Normalization used to prevent overfitting



```python
print("Classification report (per class):")
print(classification_report(y_true, y_pred, target_names=class_names))

per_class_acc = cm.diagonal() / cm.sum(axis=1)
for cls_name, acc in zip(class_names, per_class_acc):
    print(f"Accuracy for {cls_name}: {acc:.4f}")
```

```
Classification report (per class):
                      precision    recall  f1-score   support

          AnnualCrop       0.92      0.94      0.93       220
              Forest       0.95      0.98      0.96       212
HerbaceousVegetation       0.93      0.95      0.94       209
             Highway       0.88      0.93      0.90       202
          Industrial       0.98      0.95      0.96       195
             Pasture       0.93      0.88      0.90       130
       PermanentCrop       0.93      0.94      0.93       196
         Residential       0.97      1.00      0.98       205
               River       0.94      0.91      0.93       202
             SeaLake       1.00      0.96      0.98       245

            accuracy                           0.94      2016
           macro avg       0.94      0.94      0.94      2016
        weighted avg       0.95      0.94      0.94      2016

Accuracy for AnnualCrop: 0.9364
Accuracy for Forest: 0.9764
Accuracy for HerbaceousVegetation: 0.9474
Accuracy for Highway: 0.9257
Accuracy for Industrial: 0.9487
Accuracy for Pasture: 0.8769
Accuracy for PermanentCrop: 0.9388
Accuracy for Residential: 0.9951
Accuracy for River: 0.9109
Accuracy for SeaLake: 0.9592
```

### Model Evaluation and Critical Analysis

The final model was evaluated on unseen test data, producing strong metrics.

### Overall Performance

- **Test Accuracy:** 94.44%
- **Macro F1-score:** 0.94
- **High-Class Performance:** Residential class achieved 99.51% accuracy.

### Critical Technical Insight: River vs. Highway Confusion

A key source of misclassification occurred between the River and Highway classes:

- **River Accuracy:** 91.09%
- **Highway Accuracy:** 92.57%

**Explanation:**
Both rivers (especially narrow ones) and highways appear as thin, winding linear features in 64 × 64 RGB images. This visual similarity reduces distinguishability for the CNN, with limited texture and spectral cues.

---

### Milestones 4 & 5 — Deployment, Demonstration, and Conclusion

#### Deployment

- The model was deployed using an interactive **Streamlit** web application.
- The app runs locally for demonstration and is maintained on GitHub.

#### Future Deployment

- Planned cloud deployment using **Microsoft Azure** for scalable, production-ready access.

#### Live Demonstration

The deployed Streamlit app will be demonstrated to showcase real-time land-type classification.

#### Conclusion

The project successfully met all milestones and resulted in a highly accurate and deployable LULC (Land-Use and Land-Cover) classification system.

#### Future Work

- Integrate Near-Infrared (NIR) bands from Sentinel-2 to improve separation between visually similar classes (e.g., River vs. Highway).

## Authors

This project was completed as part of the **Land Type Classification Using Satellite Images** research initiative. The following contributors participated in the development, analysis, implementation, and documentation of the system:

**Project Authors:**

• Alaa Mamdouh
• Mohamed Abdelmoneim
• Iyad Mahdy

• Mariam El-Khameesy
• Ahmed alaa eldin

**Project Supervisor:**

• *ENG/ Ahmed Moustafa*

**Institution:**

DEPI (Microsoft Machine Learning Track)

Round 3