Applied Artificial Intelligence
COMP 6721

CLASSIFICATION OF SATELLITE IMAGES AND AERIAL DATA USING CNN MODELS

BEAVAN JOE MATHIAS - 40274832

HARSHIL PRAJAPATI - 40266103

SHAMITA DATTA - 40276530

SLADE JUSTIN FERRAO - 40275410

Introduction

Satellite image processing involves analyzing satellite images to classify land cover and detect patterns useful for urban planning, disaster management, and environmental monitoring. It enhances decision-making through automated land-cover classification by leveraging advancements in machine learning and deep learning techniques.

Through this project, we aim to classify satellite images into predefined categories using Convolutional Neural Networks (CNNs) and transfer learning. The models utilized include AlexNet, ReXNet, and MobileNet, which are well-suited for extracting spatial and semantic features from high-resolution imagery.

CNN Models

Convolutional Neural Networks (CNNs) are deep learning models specifically designed for analyzing visual data. They excel at extracting spatial features and patterns from images, making them ideal for tasks like image classification, object detection, and segmentation.

AlexNet: Foundational CNN architecture with 5 convolutional layers; suited for smaller datasets.

ResNet18: Employs residual learning to tackle vanishing gradient problems, enabling deeper architectures.

MobileNetV2: Optimized for efficiency with depth-wise separable convolutions; ideal for resource-constrained environments.

Transfer Learning

Transfer learning is a deep learning technique where a pre-trained model, developed for a specific task, is reused as the starting point for a new, related task. It allows models to leverage knowledge gained from large datasets and apply it to smaller, domain-specific datasets.

Why use transfer learning

- Instead of training from scratch, transfer learning uses pre-trained models, significantly speeding up the process.
- Requires fewer computational resources compared to training large models from scratch.
- Improves accuracy by enhancing performance, as the model already understands basic visual features.
- Enhances generalization with limited labeled data

Transfer Learning Models

DenseNet121

- Dense connectivity allows efficient feature reuse, achieving 99.48% accuracy on RSI-CB.
- Particularly effective for high-level feature extraction.

EfficientNetB0

- Balances depth, width, and resolution using compound scaling.
- Achieved 92.95% accuracy on RESISC45 with lower computational costs.

Datasets Used

EuroSAT:

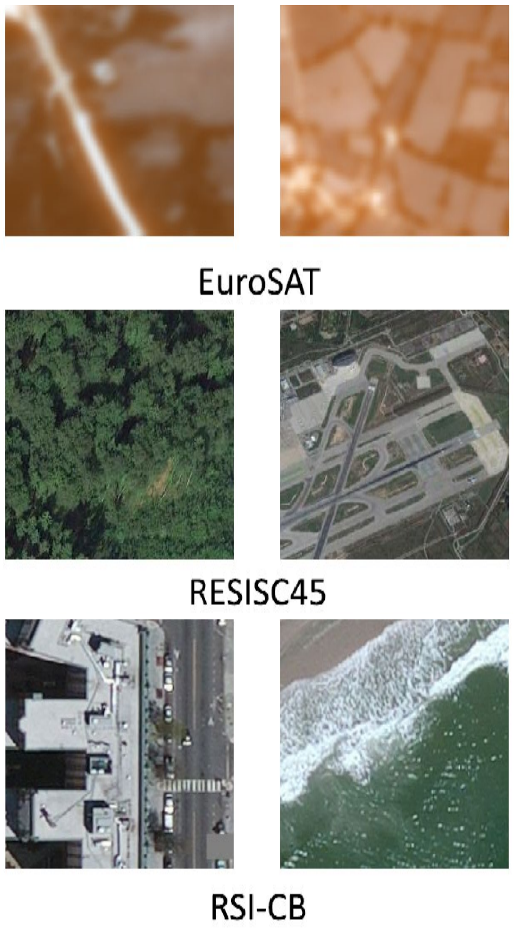
- 27,000 images across 10 land-cover classes (e.g., residential, forest).
- Derived from Sentinel-2 satellite imagery with preprocessing steps like resizing to 224×224 and normalization.

RESISC45:

- 31,500 images spanning 45 classes (e.g., windmills, ports).
- Challenges include fine-grained distinctions in visually similar scenes.

RSI-CB:

- 22,000 images grouped into 5 major land-cover classes.
- Focuses on hierarchical scene classification.



Hyperparameters and Optimization

Hyperparameters:

- Learning Rate: 0.001 (adjusted during experimentation).
- Batch Size: 64 (optimized for validation performance).
- Epochs: Determined dynamically using early stopping.

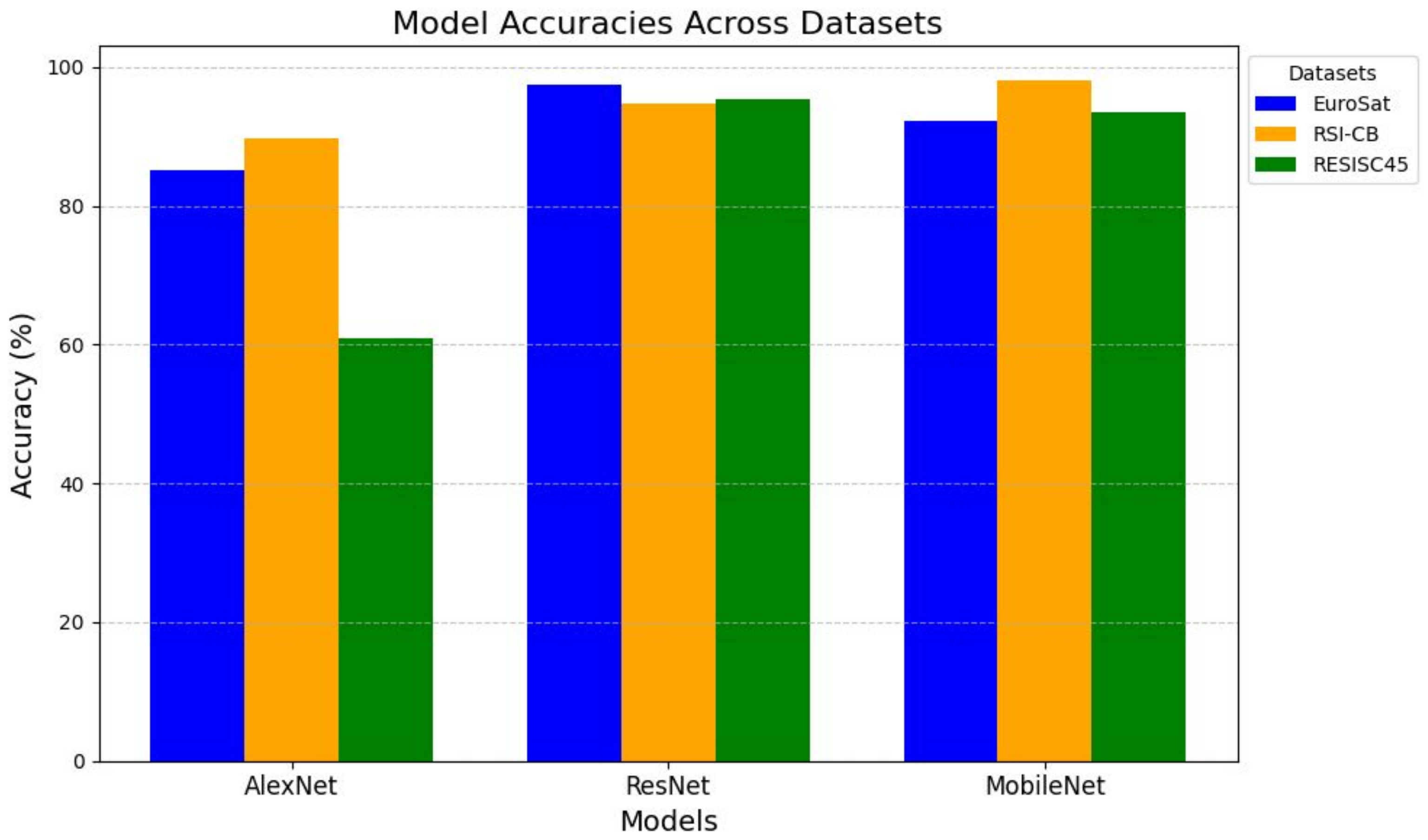
Optimization Techniques:

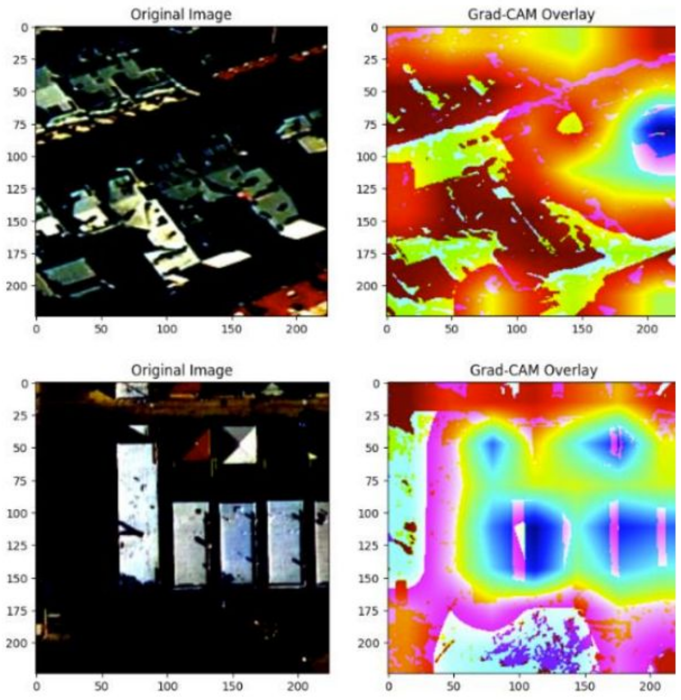
- Data Augmentation (e.g., random rotations, flips).
- Early Stopping: Halted training when validation loss plateaued.
- Learning Rate Scheduler: Adapted learning rate to enhance convergence.

Training the Models

- Images are resized to **224x224 pixels** for compatibility with pre-trained CNN models and **normalized** to ensure consistent input.
- Data augmentation techniques like **random rotations** and **flips** were applied to improve generalization.
- Models trained for multiple epochs using the **Adam optimizer** with a learning rate of **0.001**.
- Metrics tracked: **Training Loss**, **Validation Loss**, and **Accuracy** after each epoch.
- **Validation Phase**: The model switches to evaluation mode to disable dropout and batch normalization updates.
- **Grad-CAM Visualization**: Generated for interpretability, highlighting image regions that contributed most to predictions.
- After training, models were tested on the **unseen test dataset** to assess generalization.
- Metrics used: **Accuracy**, **Precision**, **Recall**, **F1-Score**, **Confusion Matrix**

Training Accuracies & GradCam





Results

Model/Dataset	*EuroSat*		*RESISC45*		*RSI-CB*	
	Acc	*F1*	*Acc*	*F1*	*Acc*	*F1*
AlexNet	0.804	0.801	0.453	0.449	0.834	0.831
MobileNet	0.824	0.814	0.767	0.767	0.967	0.967
ResNet	0.873	0.876	0.800	0.795	0.934	0.924

Table 1. Results for Main Models.

Model/Dataset	*RSI-CB*		*Model/Dataset*	*RESISC4*	
	Acc	*F1*		*Acc*	*F1*
DenseNet121	0.994	0.994	EfficientNetB0	0.929	0.928

Table 2. Results for Transfer Learnings.

References

- X. Zhu, L. Xu, and J. Wang, "A deep learning approach for land use and land cover classification using remote sensing and GIS," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 167, pp. 92 -104, 2020.
- C. Zhang, L. Xie, and J. Liu, "Land -Cover Classification Using Sentinel -2 and Deep Learning," Remote Sensing, vol. 13, no. 8, pp. 5108, 2021.
- PyTorch. “ResNet18.” [Online]. Available: https://pytorch.org/vision/main/models/generated/torchvision.models.resnet18.html. [Accessed: Sept. 27, 2024].
- S. Bhavsar, “AlexNet Architecture Explained,” Medium, Jun. 17, 2021. [Online]. Available: https://medium.com/@siddheshb008/alexnet -architecture - explained -b6240c528bd5. [Accessed: Sept. 27,2024].
- “MobileNetV2,” Papers With Code. [Online]. Available: https://paperswithcode.com/method/mobilenetv2. [Accessed: Sept. 27, 2024].
- “EuroSAT Dataset,” Kaggle. [Online]. Available: https://www.kaggle.com/datasets/apollo2506/eurosat -dataset. [Accessed: Sept. 27, 2024].
- H. Lei, “RSI -CB,” GitHub. [Online]. Available: https://github.com/lehaifeng/RSI -CB. [Accessed: Sept. 27, 2024].
- “RESISC45,” TensorFlow. [Online]. Available: https://www.tensorflow.org/datasets/catalog/resisc45. [Accessed: Sept. 27, 2024].

THANK YOU!!!