**Project Proposal**

**Land Type Classification Using Sentinel-2 Satellite Images and Deep Neural Networks**

**Abstract**

Accurate land-use and land-cover (LULC) classification is a critical component of modern environmental monitoring and spatial planning. With the growing availability of high-resolution Earth observation data, advanced machine learning techniques such as Deep Neural Networks (DNNs) provide new opportunities to automate and enhance the precision of land-type classification. This project focuses on developing a DNN model that classifies various land types—including agriculture, water, urban areas, deserts, roads, and vegetation—using multispectral imagery from the European Space Agency’s Sentinel-2 mission. The report outlines the technical approach for dataset preparation, model design, training, and evaluation, as well as the expected outcomes for supporting sustainable resource management, urban planning, and climate research.

**1. Introduction**

Land-use classification from satellite imagery has long been an important task in remote sensing, supporting a wide range of applications from agriculture and forestry to environmental policy. Traditional classification techniques such as Maximum Likelihood or Support Vector Machines rely heavily on handcrafted features and often underperform in complex landscapes. Deep learning approaches, in contrast, automatically extract hierarchical spatial and spectral features from images, allowing them to achieve state-of-the-art performance in computer vision and remote sensing tasks.

The Sentinel-2 mission, launched by the European Space Agency, provides free multispectral imagery with 13 spectral bands and spatial resolutions ranging from 10 m to 60 m. These images are particularly well suited for large-scale land-cover mapping. The integration of Sentinel-2 data with DNN architectures such as Convolutional Neural Networks (CNNs) or Vision Transformers (ViTs) offers the potential to classify land types more efficiently and accurately than traditional methods.

**2. Problem Definition**

Accurate land-type classification remains challenging due to factors such as heterogeneous land cover, atmospheric effects, seasonal variations, and limitations in annotated datasets. In many regions, especially developing countries, up-to-date land-cover maps are unavailable or inconsistent. Manual interpretation of satellite images is time-consuming and prone to human error. Therefore, an automated system that leverages deep learning to classify land types from Sentinel-2 imagery can significantly improve accuracy and scalability, providing critical data for decision-makers and researchers.

**3. Objectives**

The primary objective of this project is to develop a deep neural network capable of classifying land types from Sentinel-2 images. Specific goals include:

* Preprocessing and organizing Sentinel-2 imagery for use in a deep learning pipeline.
* Training a DNN to classify six major land-type categories: agriculture, water, urban areas, desert, roads, and trees.
* Evaluating the model using appropriate metrics such as accuracy, precision, recall, and F1-score.
* Producing land-cover maps and validating the results against reference datasets or field data.
* Demonstrating the applicability of the trained model for practical uses in environmental monitoring and urban planning.

**4. Methodology**

**4.1 Data Collection and Preprocessing**

A rady made data set will be used from kaggle called “DeepGlobe Land Cover Classification Dataset.”. This dataset was obtained from Land Cover Classification Track in DeepGlobe Challenge .

### Data

* The training data for Land Cover Challenge contains 803 satellite imagery in RGB, size 2448x2448.
* The imagery has 50cm pixel resolution, collected by DigitalGlobe's satellite.
* The dataset contains 171 validation and 172 test images (but no masks).

### Label

* Each satellite image is paired with a mask image for land cover annotation. The mask is a RGB image with 7 classes of labels, using color-coding (R, G, B) as follows.
  + Urban land: 0,255,255 - Man-made, built up areas with human artifacts (can ignore roads for now which is hard to label)
  + Agriculture land: 255,255,0 - Farms, any planned (i.e. regular) plantation, cropland, orchards, vineyards, nurseries, and ornamental horticultural areas; confined feeding operations.
  + Rangeland: 255,0,255 - Any non-forest, non-farm, green land, grass
  + Forest land: 0,255,0 - Any land with x% tree crown density plus clearcuts.
  + Water: 0,0,255 - Rivers, oceans, lakes, wetland, ponds.
  + Barren land: 255,255,255 - Mountain, land, rock, dessert, beach, no vegetation
  + Unknown: 0,0,0 - Clouds and others

**4.2 Model Design**

A Convolutional Neural Network (CNN) architecture will be implemented to exploit both spectral and spatial information. The model may be based on existing architectures such as ResNet, EfficientNet, or U-Net, depending on the desired output resolution. Transfer learning will be applied using pre-trained weights from large-scale image datasets to accelerate convergence and improve generalization.

**4.3 Training and Validation**

The dataset will be divided into training, validation, and testing subsets (typically 70/15/15). Data augmentation techniques such as rotation, flipping, and brightness adjustments will be employed to enhance robustness. The model will be trained using cross-entropy loss and optimized with the Adam or SGD optimizer. Validation performance will be monitored to prevent overfitting, and hyperparameters such as learning rate and batch size will be tuned using grid search.

**4.4 Evaluation Metrics**

Model performance will be assessed through accuracy, precision, recall, F1-score, and confusion matrices. Class-wise performance will be analyzed to identify categories that require additional training samples or feature adjustments.

**5. Results and Evaluation Plan**

The expected outcome of the training phase is a deep neural network capable of distinguishing major land-use types with high accuracy (> 90% on test data). Preliminary evaluation will be performed on existing benchmark datasets such as EuroSAT to establish baseline performance. The trained model will then be tested on new Sentinel-2 scenes from different geographic regions to verify generalization. Results will be visualized through classified land-cover maps and compared with existing LULC maps for validation.

**6. Applications and Expected Outcomes**

The developed model will provide a scalable, automated solution for land-type classification that can be integrated into GIS platforms or deployed as a web-based mapping service. Key applications include:

* **Urban planning:** Monitoring expansion of built-up areas and infrastructure development.
* **Environmental management:** Tracking deforestation, water resource depletion, and desertification.
* **Agricultural analysis:** Estimating crop coverage and detecting land-use change.
* **Disaster assessment:** Evaluating the impact of floods or fires on land cover.

Beyond these direct uses, the methodology can be adapted for time-series analysis and change detection, enabling long-term environmental monitoring.

**7. Conclusion**

This project proposes a comprehensive workflow for automated land-type classification using Sentinel-2 satellite imagery and deep neural networks. By combining multispectral data with advanced deep learning architectures, the approach promises significant improvements in accuracy, efficiency, and scalability over traditional classification techniques. The resulting model will support a wide range of applications in environmental monitoring, urban development, and sustainable resource management. Future work may extend the framework to multi-temporal analysis, 3D land modeling, and integration with other satellite data sources such as Landsat or PlanetScope for enhanced precision.