*Vishal Bansal, Aditya Wadhwa, Krish Manwani, Aadi Chopra, Pranjal Seluriyal  
University of Petroleum and Energy Studies (UPES), Dehradun, India  
Under the Guidance of: Amit Sir*

**Title: Land Cover Classification Using K-Means Clustering on Google Maps Images**

**Abstract**

In this research, we propose a simple and beginner-friendly model to identify different types of land on Earth's surface, such as forests, water, urban areas, etc., using color images from Google Maps. We use an unsupervised machine learning algorithm called **K-means clustering** to group similar-looking pixels together based on their colors. This helps us estimate how much area is covered by each land type without needing labeled data. The model is built using Python and Google Maps API to fetch satellite images, and then we analyze those images using basic image processing techniques.

**1. Introduction**

Understanding land use and land cover is important for planning and environmental monitoring. Normally, such analysis is done using satellite images and complex tools, which can be difficult for beginners.

To make it easier, we use Google Maps images because they are easily available and already pre-processed. We apply an unsupervised machine learning method—**K-means clustering**—to group pixels of similar colors, helping us estimate the types of land in a region.

**2. Objective**

* To estimate different land cover types using satellite images.
* To apply **unsupervised learning** (no labeled data needed).
* To make this process simple using tools like Python and Google Maps API.

**3. Technologies Used**

* **Google Static Maps API**: To download satellite images of a selected location.
* **Python**: Programming language for model implementation.
* **Libraries**:
  + NumPy: For matrix operations.
  + Matplotlib: For visualizing data.
  + OpenCV: For image processing.
  + PyQt: To create a basic user interface (optional).

**4. Methodology**

**Step 1: Image Collection**

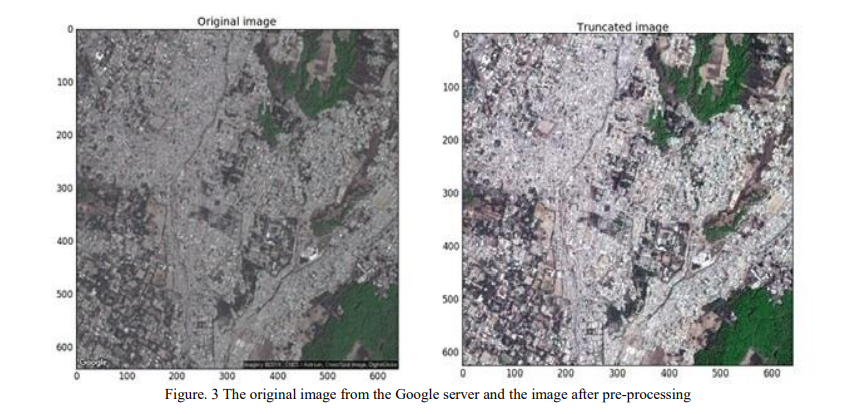
We use **Google Static Maps API** to download satellite images of a selected location using latitude, longitude, zoom level, and image size.

**Step 2: Image Pre-processing**

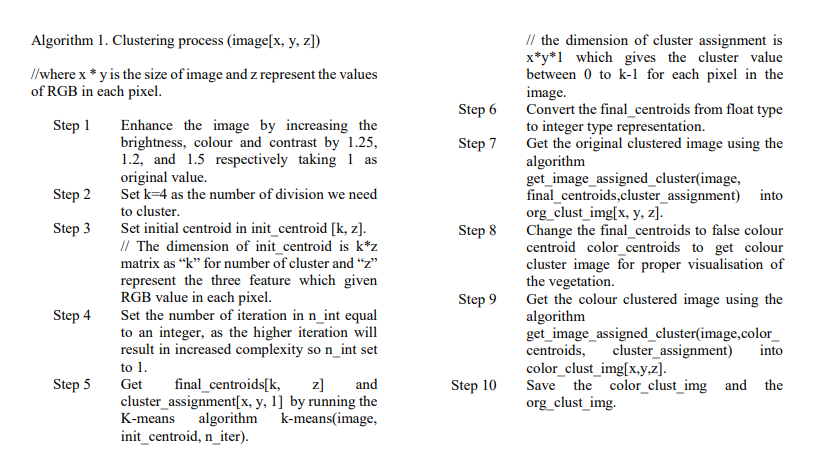
We enhance the image quality by adjusting:

* Brightness
* Contrast
* Color

This makes the image clearer and helps the machine learning algorithm work better.

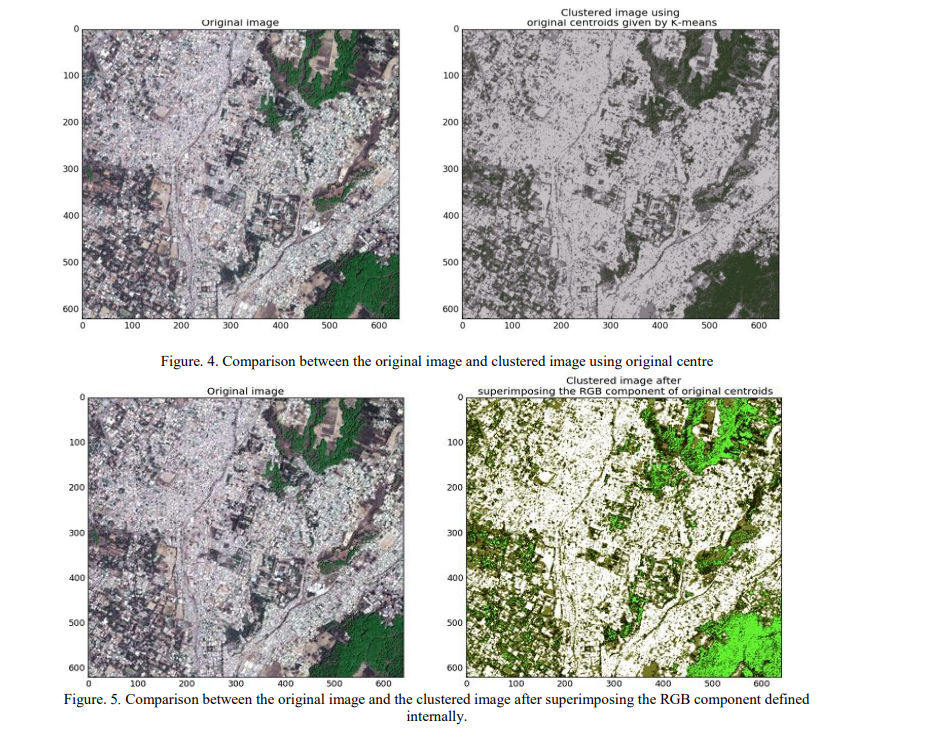


**Step 3: Applying K-Means Clustering**

* Each pixel in the image has three color values: **Red**, **Green**, and **Blue (RGB)**.
* We apply **K-means**, which groups the pixels into **K clusters** (in our case, we chose K = 4).
* These clusters represent different types of land like:
  + Vegetation
  + Urban areas
  + Water bodies
  + Barren land
  + 

**Step 4: Visualizing Results**

We replace each pixel with the color of its cluster, producing a new image where different land types are clearly visible in different colors.



**Step 5: Area Estimation**

Based on the number of pixels in each cluster, and the real-world area covered by the map, we estimate how much area (in sq. km) each land type covers.

**5. Why K-Means?**

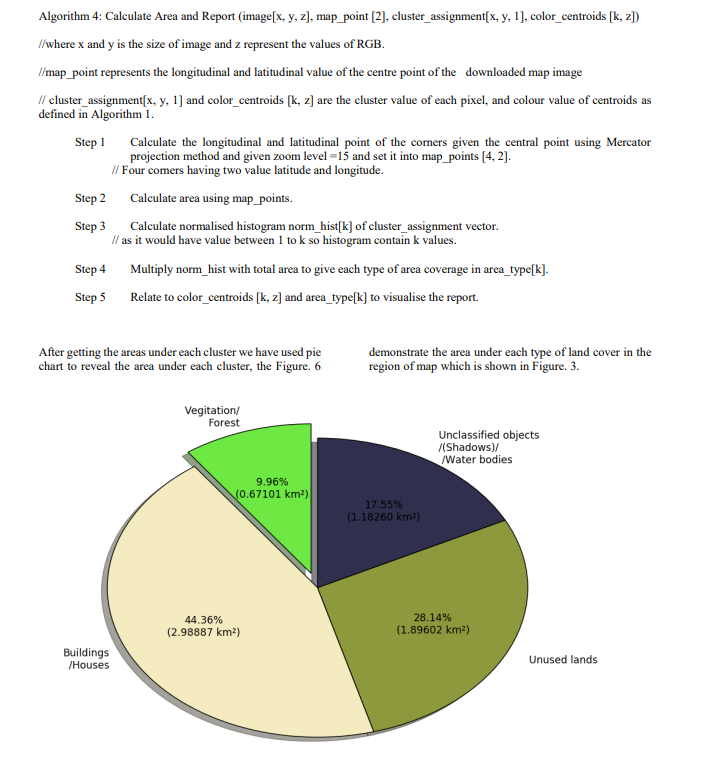
* It's **easy to understand** and implement.
* It doesn't need labeled data (unsupervised).
* It’s fast and works well on RGB image data.

**6. Results**

The model successfully identified different land types in a selected region. It also calculated the area under each category using pixel counts and map scaling.

**7. Conclusion**

Our project shows that it's possible to use basic machine learning and freely available tools to estimate land cover types. This model is a good starting point for beginners in machine learning and geospatial analysis.



**8. Future Scope**

* Use better clustering methods like **Gaussian Mixture Models** for more accuracy.
* Work with images that contain more data (like infrared or multi-band images).
* Increase the number of land classes for detailed analysis.

**9. Use of Machine Learning in Google Maps**

Machine Learning (ML) plays a major role in enhancing Google Maps. It helps improve accuracy, automate mapping processes, and provide real-time smart features. Below are key use cases where ML is applied — especially those closely related to our project.

**1. Land Cover Classification**

Google uses satellite and aerial imagery, combined with ML techniques such as **unsupervised clustering (like K-means)** and **deep learning (like CNNs - Convolutional Neural Networks)** to:

* **Automatically detect** land types such as forests, water bodies, urban areas, etc.
* **Update map data** by analyzing changes in imagery over time.
* **Support urban planning** and environmental studies by mapping large areas efficiently.

🧠 **Algorithm used**:

* **K-Means Clustering**: For basic land segmentation using pixel color groups (like in our project).
* **CNNs**: For more advanced, high-accuracy object detection and scene segmentation in large-scale mapping.

Our project is a simplified version of this, using K-means to classify land from Google Maps images into 4 basic types (vegetation, buildings, unused land, water/shadow).

**2. Traffic Prediction**

Google Maps predicts traffic conditions using real-time and historical data, powered by ML.

🧠 **Algorithms used**:

* **Recurrent Neural Networks (RNNs)** and **Long Short-Term Memory (LSTM)** networks for time-series traffic data.
* These models learn from past trends and current conditions to predict:
  + Delays
  + Congestion spots
  + Estimated arrival times (ETAs)

**3. Route Optimization**

Google Maps suggests the **fastest, safest, and most efficient routes** based on live conditions.

🧠 **Algorithms used**:

* **Graph-based algorithms** like **Dijkstra’s Algorithm** for pathfinding.
* Enhanced with **Reinforcement Learning** and **A/B testing** to personalize route suggestions based on user behavior and past outcomes.

**4. Place Recognition and Business Listings**

ML is used to detect and classify new or updated places (like restaurants, ATMs, and landmarks).

🧠 **Algorithms used**:

* **Image classification models** (like CNNs) to recognize place types from satellite or Street View imagery.
* **Natural Language Processing (NLP)** models to extract information from user reviews and business descriptions.

**5. Street View Image Analysis**

Google uses **Computer Vision**, a subfield of ML, to analyze Street View images.

🧠 **Algorithms used**:

* **Object Detection models (e.g., YOLO, SSD)** to:
  + Read road signs and extract text.
  + Detect traffic lights, crossings, and buildings.
* **Privacy models** to automatically **blur faces and license plates**.

**🔗 Summary**

| **Feature** | **ML Technique Used** | **Purpose** |
| --- | --- | --- |
| Land Cover Classification | K-Means, CNNs | Classify terrain from satellite images |
| Traffic Prediction | RNN, LSTM | Predict congestion and travel time |
| Route Optimization | Dijkstra’s Algorithm, Reinforcement Learning | Find fastest and safest routes |
| Place Recognition | CNN, NLP | Identify and update place information |
| Street View Analysis | Object Detection (YOLO, SSD) | Detect and protect visual data |