AURORA 2.0

BEYOND THE HORIZON

Adaptive Mining Activity Monitoring using Sentinel-2 Time Series

Team Name:
Marauders

Team Members:
Vasudev Sajeev, Disha Hebbar, V Kavyanjali, Sahana Bhagwat

Contents

1 Abstract

This report addresses the challenge of regulating mining activity, which is complicated by variations in mine types and weather conditions. The proposed solution consists of a pipeline that learns the typical spectral signatures of each mine from its historical data. An unsupervised K-Means clustering algorithm is employed to cluster pixels, subsequently categorizing these clusters into "excavation-like" and "non-excavation-like" groups.

Spatiotemporal features are integrated into the clustering algorithm to mitigate the effects of seasonal variations on the analysis. The system monitors mining sites and generates binary masks highlighting excavated pixels for each image. Through post-processing of these binary masks, excavated pixels are converted into vector polygons representing excavated regions. These polygons are utilized to achieve the project objectives, including calculation of total excavated area and assessment of intersections with designated no-go zones.

2 Problem Understanding

The problem addressed in this work is the regulation and monitoring of mining agencies to ensure that mining operations remain within legal boundaries and avoid no-go zones. The primary challenge lies in accounting for the diverse types of mines (open-cast coal, limestone, bauxite, sand, etc.), each possessing distinct mine-specific spectral signatures. This diversity renders conventional rule-based monitoring methods ineffective.

Therefore, it is necessary to develop a model that learns the unique excavation spectral signatures of each mine without employing hard-coded thresholds, and is capable of identifying newly excavated pixels with a quantifiable level of confidence.

Additionally, the system must track and monitor excavation activities over time, generating time-series plots of excavation patterns. A normalized metric is required to enable comparison of excavation intensities across different mining sites.

The system must ensure that no-go zones are not breached during excavation operations, incorporating early-warning alerts when mining activities approach these zones, and detecting new or expanding excavation within no-go zones.

3 Methodology

3.1 Inputs

- Sentinel-2 Level-2A Multispectral Time Series Data obtained from Google Earth Engine via Python API to streamline the workflow.
- Geometries of Legal Mine Boundaries for all mines under consideration.
- No-Go Zone Polygons synthetically generated to demonstrate model effectiveness.

3.2 Proposed Solution Overview

The proposed solution consists of a pipeline through which each mine is processed. A model is developed that learns from each mine's historical data to adapt to the mine's excavation signals and other spectral signatures. Consequently, each mine operates a distinct instance of the same model.

This approach addresses the challenge of adaptive excavation signature learning by directly learning from each mine's unique spectral characteristics.

The model employs an Unsupervised Learning framework, as labeled data is unavailable. The task can be conceptualized as clustering pixels within the mine into those corresponding to excavated areas and those representing other land cover types.

4 Training Pipeline

4.1 Phase 1: Data Preparation

4.1.1 Step 1: Mine Selection and Data Filtering

- Mine selection is performed by filtering Sentinel-2 data using spatial bounds corresponding to the mine's geometry.
- Cloud masking is implemented using the Scene Classification Layer (SCL) to ensure data quality.
- A representative training time period is defined for training the clustering algorithm.

4.1.2 Step 2: Feature Engineering

This component constitutes the most critical aspect of the project, as the model's clustering effectiveness depends entirely on the quality of spatiotemporal features selected to accurately differentiate between excavated and non-excavated pixels.

Spectral Indices:

- **NDVI** – Normalized Difference Vegetation Index (effective for detecting vegetation removal associated with mining)
- **SWIR** – Shortwave Infrared (capable of detecting exposed rocks and soil resulting from mining operations)
- **NBR** – Normalized Burn Ratio (changes in NBR can indicate surface disturbances)
- **BSI** – Bare Soil Index (mining activities expose bare soil, which this index can highlight)

For temporal features, a window length comprising a specified number of images is selected, and the following statistics are calculated:

- Rolling medians
- Rolling means
- Rolling variance

These temporal features facilitate differentiation between seasonal changes and excavation activity. The optimal window length is determined through empirical testing.

4.2 Phase 2: K-Means Clustering

4.2.1 Step 3: Initial Clustering

The objective is to apply K-Means Clustering to the training data, which comprises the selected spatiotemporal features of each pixel within the mine at each time t within the training period bounds, resulting in K clusters.

Determining K:

- If K is insufficient: excavated pixels merge with non-excavated pixels, rendering the model non-functional.
- If K is excessive: the data becomes noisy with multiple "excavation" clusters, which is undesirable.
- Optimal solution: selection of the minimum value of K that yields distinct excavation clusters for each mine.
- Practical approach: implementation of a fixed value of K (determined through testing) that is sufficiently large to avoid aforementioned issues.

4.2.2 Step 4: Cluster Classification

Following clustering into K clusters, it is necessary to identify which clusters represent "excavated pixels" and which represent "non-excavated pixels".

Process:

1. The training data is clustered into K clusters based on spatiotemporal features, yielding K centroids.
2. Given adequate spatiotemporal features, a significant separation exists between excavated and non-excavated pixels, such that "excavated pixel cluster" centroids are spatially distant from "non-excavated pixel cluster" centroids.
3. The K centroids are re-clustered into 2 meta-clusters ($K = 2$) using an additional K-Means Clustering operation.
4. The meta-cluster representing "excavated pixels" is identified by examining mean NDVI values, with the lower-NDVI cluster corresponding to excavation.

Benefits:

- This methodology partitions all K clusters into "excavated pixels" clusters and "non-excavated pixel" clusters.
- If multiple mining techniques are employed within a site, even if separated into distinct clusters, they exhibit sufficient similarity to other "excavated" clusters to be detected by the model.

4.3 Training Pipeline Flow Chart

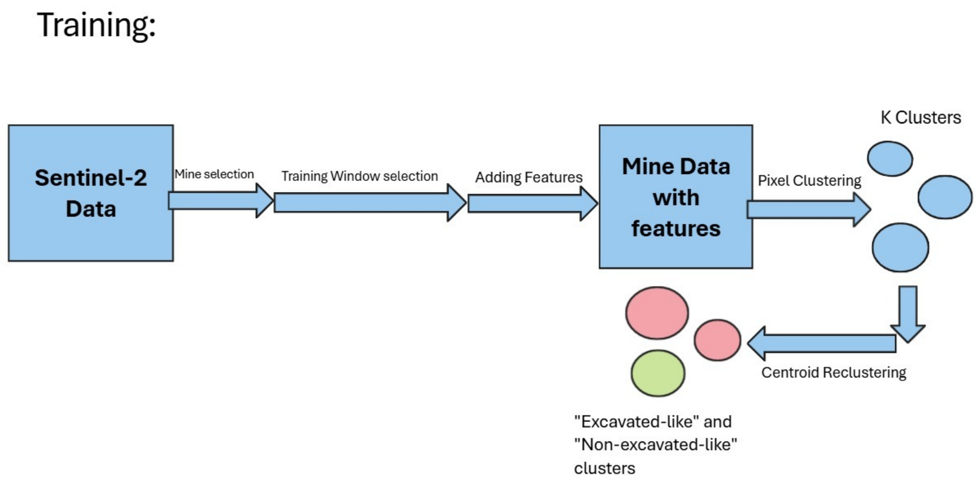

Figure 1: Training Pipeline Flow Chart

5 Monitoring Pipeline

5.1 Phase 3: Real-time Monitoring and Detection

5.1.1 Step 5: Cluster Assignment for New Data

For each new time t:

1. The cluster membership of each pixel at that time is predicted.
2. Each pixel is classified as "excavated" or "non-excavated" based on its cluster assignment.

5.1.2 Step 6: Candidate Excavation Detection

Pixels are categorized into distinct states:

- **Candidate Excavated Pixels:** Pixels transitioning from non-excavated to excavated clusters.
- **Returned to Normal:** Pixels transitioning from excavated to non-excavated clusters are reclassified as normal non-excavated pixels.

This mechanism enables detection of new excavation behavior.

Handling Seasonal Changes: Seasonal changes may cause pixels to temporarily appear excavated. All seasonal changes are transient, whereas excavation represents a permanent alteration (assumed to persist for a substantial time period following initial excavation).

Confidence Mechanism:

- For each timeframe during which a pixel continues to appear "excavated" (remaining within an "excavated" cluster), a confidence counter is incremented.
- Upon exceeding a specified time threshold while maintaining an excavated-like state (sufficient duration to exclude seasonal changes), the pixel is classified as a "confirmed excavated pixel".

- Within this model, confidence represents the persistence of a change, ensuring it reflects a permanent alteration such as excavation rather than a temporary variation.

Confidence Update Rules:

- If a "candidate excavated pixel" returns to a "non-excavation cluster" before confidence reaches the threshold, the change is attributed to seasonal variation, and the confidence value is reset to zero.
- If a pixel is designated as a "confirmed excavated pixel," it retains this classification even if post-excavation vegetative growth causes the pixel to transition to a "non-excavated pixels" cluster, as the excavation itself remains permanent.

5.2 Phase 4: Post-Processing

5.2.1 Step 7: Binary Mask Cleaning

Following clustering during the monitoring phase, a binary mask of "excavated pixels" is produced. This binary mask requires refinement and conversion into polygons to ensure data utility and noise reduction.

Morphological Operations: Utilizing pre-existing skimage functions:

- **Opening:** Elimination of small false positives that may arise from K-Means Clustering.
- **Closing:** Filling of holes that may exist within excavated regions due to false negatives.
- These operations ensure that the binary mask coherently forms polygons.

5.2.2 Step 8: Polygon Extraction

1. Boundary tracing of formed polygons is performed.
2. Raster data is converted into vector polygons.
3. Polygons below a minimum size threshold are removed to ensure accuracy.
4. Smoothing operations are applied.

Result: Clean polygons representing excavated regions suitable for subsequent analysis.

5.3 Phase 5: Analysis and Reporting

5.3.1 Step 9: Quantification and Violation Detection

- Total area of excavated pixels is calculated from polygon geometries.
- Intersection analysis with no-go zones is conducted.
- Alerts are generated when candidate excavated pixels are detected within no-go zones.
- Alert logs are maintained, recording the initial date of excavation detection within no-go zones.

5.3.2 Step 10: Visualization and Output Generation

For each time t, the following outputs are generated:

- Binary mask of "excavated pixels" across the mine site.
- Spatial maps incorporating:
  - Detected excavated region polygons overlaid on satellite imagery.
  - Legal boundaries displayed.
  - No-go zones highlighted.

Time-Series Outputs:

- Time-series plot depicting the area of "candidate excavated pixels" over time.
- Time-series plot depicting the area of "confirmed excavated pixels" over time.
- Time-series plot of excavated area within legal boundaries.
- Time-series plot of excavated area within no-go zones.

These multiple visualizations provide transparency and clarity regarding the model's performance.

5.4 Monitoring Pipeline Flow Chart

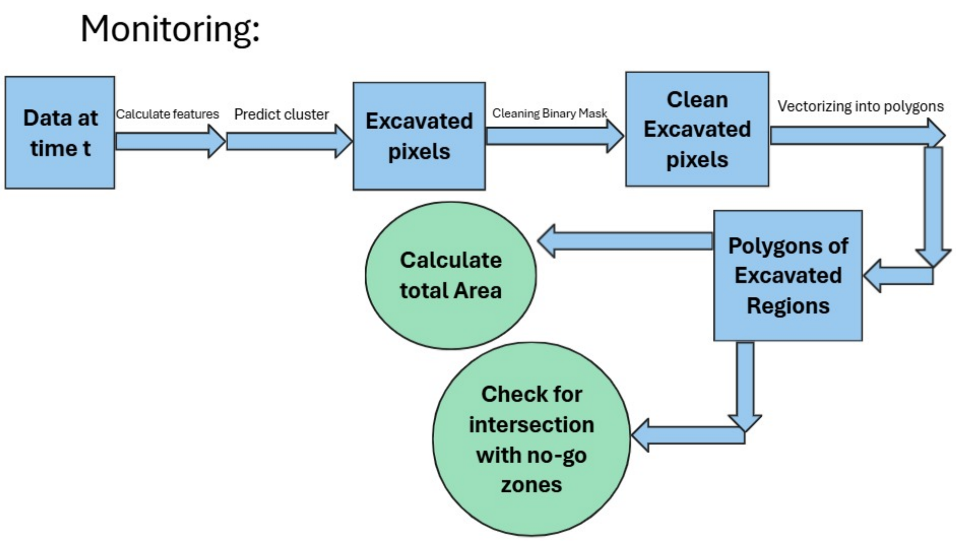

Figure 2: Monitoring Pipeline Flow Chart

6 User Interface Outputs

A user interface is provided that enables mine selection and provides the following functionalities:

- Visualization of time series plots depicting excavated area within the legal boundaries of mines and within no-go zones.

- Generation of spatial maps for specified timestamps where detected excavated regions are highlighted with overlays of legal boundaries and no-go zones.
- Access to alert logs documenting no-go zone violations over time.

7 Conclusion

The proposed clustering model demonstrates generalizability and adaptability across diverse mine types. The system exhibits resistance to seasonal and other transient changes. It successfully detects and monitors mining activity, quantifying excavation through time-series profiles of total excavated area.

The system is capable of identifying and quantifying excavation activity within no-go zones, providing violation alerts. It is readily scalable to new mining sites without reconfiguration, as it learns from the spectral history of each individual mine.