ARTISANAL AND SMALL-SCALE GOLD MINING DETECTION IN THE AMAZON FOREST USING CONTEXTUAL DATA

Selma Dissing, Jan-Christoph Kalo

University of Amsterdam
selma.dissing@student.uva.nl, j.c.kalo@uva.nl

ABSTRACT

Artisanal and small-scale gold mining (ASGM) is a major driver of land cover change in the Amazon, often challenging to detect due to its spectral similarity with other surface features. This study investigates whether incorporating contextual geospatial data from OpenStreetMap (OSM) alongside Sentinel-2 imagery can improve ASGM detection in Venezuela's Bolívar state. Rasterized OSM-derived semantic masks are appended as additional input channels to the satellite imagery and processed through a CNN. This setup enables joint learning of spectral and contextual features, allowing for a more accurate and reliable distinction between ASGM sites and spectrally similar land uses.

1. INTRODUCTION

Artisanal and small-scale gold mining (ASGM) is a largely informal mining practice where individuals or small groups extract gold using rudimentary techniques. It is the primary driver of illegal mining in the Amazon rainforest [1]. Remote sensing has emerged as a key tool for monitoring ASGM [2]. However, current mining detection models rely solely on spectral information, which presents challenges in distinguishing ASGM sites from other land-use changes, such as dried riverbeds or bare soil patches that exhibit similar spectral characteristics [3]. This limitation highlights the importance of integrating complementary data sources that can provide additional contextual information.

One promising way to address these limitations is the integration of contextual geospatial data. A prominent example is OpenStreetMap (OSM) [4], which contains detailed, freely accessible information on infrastructure, transportation networks, and settlements. This study explores whether incorporating such contextual data alongside satellite imagery can improve the detection of ASGM activity by providing additional cues to distinguish it from spectrally similar land-use changes. To explore this issue in detail, this study focuses on the southeastern Venezuelan state of Bolívar, one of the most ecologically significant and mining-affected regions within the Amazon basin. We investigate whether integrating contextual geospatial data from OSM improves the performance of ASGM detection using satellite imagery. Sentinel-2 provides spectral signals related to land disturbance. At the same time, OSM contributes spatial context, such as the presence of rivers, roads, or buildings, which can help distinguish ASGM sites from visually similar land uses. To test this, rasterized OSM-derived semantic masks are appended as additional input channels to the Sentinel-2 imagery and processed together through a convolutional neural network (CNN). This setup enables the model to jointly learn spectral and contextual patterns, allowing for an evaluation of the added value of OSM data for ASGM classification.

2. RELATED WORK

Satellite Imagery for Remote Sensing. Researchers have utilized various geospatial data sources, notably optical satellite imagery, to detect land disturbances associated with ASGM [5]. These approaches build on established remote sensing techniques for LULC change detection, widely applied to monitor deforestation, urban growth, agriculture, and natural disasters [6–9]. Sentinel-2 imagery is particularly effective for identifying ASGM impacts such as vegetation loss, exposed soil, and sediment-laden water, observable via changes in surface reflectance and water clarity [10–12]. However, optical imagery alone has limitations: ASGM sites often resemble agricultural or natural disturbances spectrally, leading to frequent misclassifications [3].

Image Classification with Contextual Data. Incorporating contextual information into image classification improves model performance by enabling reasoning. Applications in product recognition and medical imaging benefit from spatial and semantic context, which enhances the detection of small objects and anomalies [13, 14]. Fusing remote sensing with spatial features such as land types and surroundings provides semantic cues that enhance accuracy [12, 15]. For instance, Gomez et al. [16] used proximity to transport and water networks to detect ASGM activity using SAR imagery and manually derived features. In contrast, our study employs a framework that integrates multispectral Sentinel-2 imagery with geographic OSM data. OSM, a crowd-sourced dataset, offers layers such as roads, rivers, land-use, and buildings. For example, [17] used OSM building footprints to classify

informal settlements via spatial clustering and machine learning. Similarly, [18] aligned OSM features with remote sensing data for LULC classification and infrastructure detection. Despite its utility, OSM poses challenges due to variable quality and coverage, particularly in remote regions such as the Amazon Basin.

3. METHODOLOGY

This study builds on an open-source pipeline for ASGM detection using Sentinel-2 satellite imagery provided by Earth Genome [19]. While the baseline system includes labeled sampling points, a patch extraction process using Google Earth Engine (GEE), and a CNN for binary classification, this work extends it by incorporating contextual geospatial information from OSM, particularly features such as highways, waterways, buildings, land-use, and aeroways. The proposed method processes Sentinel-2 patches with added channels of rasterized OSM-derived semantic masks in the CNN architecture. This setup enables the model to learn spectral and spatial-contextual patterns, with the aim of improving performance.

Sampling Points and Patch Extraction. The study focuses on the Bolívar region of southeastern Venezuela. This region was selected due to its high density of mining operations and the availability of labeled ASGM data.

The baseline process handles Sentinel-2 patch extraction [19], which converts each sampling point into a standardized satellite image patch. Using consistent projection and spatial resolution, a square tile is generated and aligned with Sentinel-2 imagery for each sampling point. Sentinel-2 image data is retrieved from GEE for 2021, corresponding to the period when the sampling points were published on GitHub. The extracted bands include B1–B4, B5–B7, B8, B8A, B9, B11, and B12, covering the visible, red-edge, near-infrared, and shortwave-infrared regions. Band 10 is excluded as it is primarily used for cloud detection and is not relevant for land cover analysis. To reduce the impact of clouds, cloud shadows, and other atmospheric effects, a median composite is generated by averaging all observations across the year. This approach enhances image quality and consistency, particularly in cloud-prone regions such as the Amazon Basin.

The resulting image patches are uniformly sized with the 12 spectral channels. Each patch covers an area, centered around a sampling point labeled as either mining or non-mining. This patch captures the mining activity and the surrounding area, which can provide additional context through OSM data. The dataset includes 156 mining patches and 323 non-mining patches.

OSM Data. Vector-based geospatial features were extracted from OSM. These features were selected based on their relevance to ASGM operations. They comprise: (1) **highway**, including roads, tracks, and footpaths indicating access to mining areas; (2) **waterway**, such as rivers and streams, often used in alluvial mining; (3) **building**, representing nearby informal or residential structures; (4) **landuse**, comprising tagged areas potentially overlapping with ASGM zones; and (5) **aeroway**, which may denote airstrips supporting remote mining operations. See Figure 1 for the spatial distribution of the OSM features in the Bolívar region.

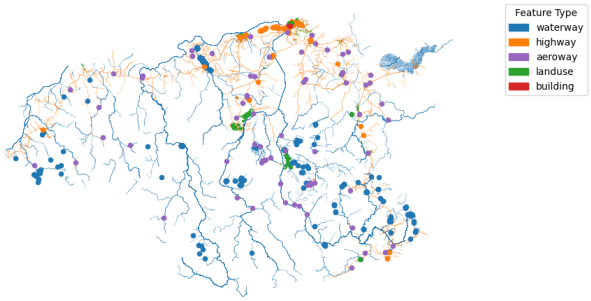

Fig. 1. Spatial Distribution of OSM Features in Bolívar.

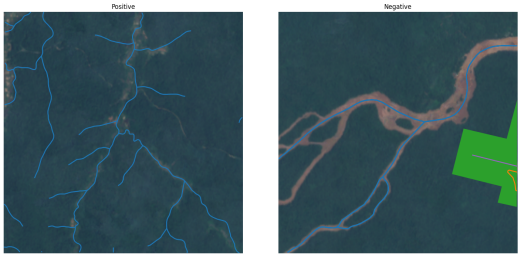

Fig. 2. Visualizations of random patches with OSM overlays.

Rasterization of OSM data. For each image patch, a corresponding subset of OSM features is extracted by clipping the vector data to the same extent. Each feature has its own channel, where a pixel with values for that feature receives a value of 1, and a pixel without that feature receives a value of 0. This ensures that both inputs have identical footprints, allowing the OSM data to be rasterized and stacked with the Sentinel-2 bands before model training. Of the 156 mining patches, 95 contain one OSM feature, representing 60.9% coverage. Among the 323 non-mining patches, 245 include OSM annotations, resulting in 75.85% coverage. This variation reflects the inconsistent availability of OSM data in remote regions, such as Bolívar, where mapping is often sparse.

Model. The baseline model is adapted from Earth Genome's open-source ASGM detection pipeline [19], which employs a lightweight CNN. The architecture consists of nine convolutional layers, each with 32 filters and 3×3 kernels, using 'same' padding and ReLU activation functions. These layers are grouped into three convolutional blocks, each followed by max-pooling layers (2×2 or 3×3) to progressively reduce the spatial resolution. The convolutional feature extraction stage is followed by three dense layers with 64, 64, and 32 neurons. Dropout layers with a rate of 0.3 are included

after each dense layer to prevent overfitting. The final output layer is a sigmoid-activated neuron for binary classification.

4. EVALUATION

4.1. Experimental Setup.

To prevent spatial leakage, the dataset is split using DBSCAN clustering on patch center coordinates, with clusters randomly assigned to train (70%), validation (15%), and test (15%) sets. Patches with excessive masking are filtered out, and Sentinel-2 reflectance values are normalized to the range of [0, 1]. All models use the same CNN architecture trained for up to 160 epochs using the Adam optimizer (3×10^{-4}), binary cross-entropy loss, and a batch size of 16. Data augmentation (random rotations, shearing, zooming, flips) is applied via Keras' `ImageDataGenerator`. Experiments vary by patch size (48px vs. 256px), class imbalance weighting, negative sampling strategy (156 points curated from Earth Genome vs. 780 points randomly sampled), and inclusion of OSM features.

4.2. Results and Discussion.

Ten ASGM detection experiments were conducted under varying patch sizes, negative sampling strategies, class imbalance handling, and the inclusion of OSM features. Performance was evaluated using F1 (with bootstrap confidence intervals), precision, recall, and PR-AUC to capture threshold-dependent and threshold-independent model performance. Table 1 summarizes performance across these settings.

Patch Size. In curated settings, both 256px (Exp 1, F1 = 0.867) and 48px (Exp 3, F1 = 0.711) models achieved strong performance without OSM. Adding OSM improved performance at 256px (Exp 2, F1 = 0.923; –2 FN, –5 FP), but only slightly at 48px (Exp 4, F1 = 0.788; –2 FP). This indicates that OSM features are more effective with larger patches that capture broader spatial context.

Class Weighting. With curated negatives, class weighting (Exp 5) increased FP (+7) and lowered precision (0.594), dropping F1 to 0.728. OSM (Exp 6) further raised FP (+10) and reduced F1 to 0.611, well below the unweighted OSM model (Exp 2). In random-negative settings, class weighting destabilized training: Exp 9 performed poorly (F1 = 0.129; FP = 62), while Exp 10 achieved F1 = 0.898 with perfect recall but lower precision.

Negative Sampling. With random negatives and no class weighting, performance remained strong: Exp 7 (F1 = 0.948) and Exp 8 (F1 = 0.792) both achieved high recall with few FP. Class weighting caused instability, leading one model to collapse (Exp 9) and another to trade precision for recall (Exp 10). Overall, curated negatives yielded more stable results, while random negatives better reflected real-world variability but increased uncertainty.

Confidence intervals highlight variability across experimental setups. Models trained with curated negatives and larger patches showed narrower confidence intervals, while random-negative sampling produced wider intervals, reflecting greater variability and task difficulty.

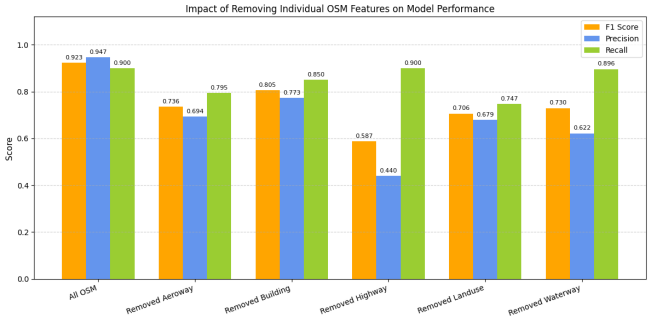

Fig. 3. Impact of removing OSM feature layers on model performance, reported as mean F1, precision, and recall. Results are shown relative to the full OSM model (Exp2).

Ablation Study. Figure 3 shows how the exclusion of different OSM feature layers influences performance. While most features contribute positively, certain layers have a more pronounced effect on the F1 score and precision than others.

Removing highway features resulted in the largest performance decline (F1 = 0.587), underscoring their importance in distinguishing ASGM sites. Land-use removal also reduced performance, while aeroway, waterway, and building layers had moderate effects. Overall, the study shows that multiple OSM layers contribute to detection, with highways providing the most critical contextual cues.

Limitations. This study focused exclusively on the Bolívar region in Venezuela, which limits the generalizability of the findings to other Amazonian settings. Model performance is strongly dependent on OSM coverage, which is uneven in remote areas and may bias results toward data-rich regions. A temporal mismatch between Sentinel-2 imagery from 2021 and OSM features from 2025 could also distort observed associations. In addition, the use of all 12 Sentinel-2 bands, including less informative ones such as B1 and B9, indicates the need for a more critical assessment of feature selection and modeling choices.

5. CONCLUSION

This study demonstrates that integrating contextual geospatial data from OSM with Sentinel-2 imagery can improve ASGM detection, particularly when larger patches and curated negatives are used. The most consistent benefit of OSM was a reduction in false positives, which helped distinguish mining from spectrally similar areas, although this sometimes came at the cost of recall. Notably, in the most challenging setting with random negatives and class weighting, OSM features enabled strong performance, whereas the spectral-only model failed to generalize. Overall, these results show that con-

Exp	OSM	Patch	Neg Type	Class Weight	F1 [95% CI]	PR-AUC	FP	FN	Precision	Recall
1	No	256	Curated	No	0.867 [0.739, 0.960]	0.865	6	0	0.769	1.000
2	Yes	256	Curated	No	0.923 [0.827, 1.000]	0.972	1	2	0.947	0.900
3	No	48	Curated	No	0.711 [0.529, 0.851]	0.850	7	5	0.682	0.750
4	Yes	48	Curated	No	0.788 [0.625, 0.913]	0.782	3	5	0.833	0.750
5	No	256	Curated	Yes	0.728 [0.565, 0.849]	0.866	13	1	0.594	0.950
6	Yes	256	Curated	Yes	0.611 [0.444, 0.744]	0.677	23	1	0.452	0.950
7	No	256	Random	No	0.948 [0.800, 1.000]	0.900	1	0	0.900	1.000
8	Yes	256	Random	No	0.792 [0.533, 0.957]	0.909	3	1	0.727	0.889
9	No	256	Random	Yes	0.129 [0.028, 0.237]	0.069	62	4	0.075	0.556
10	Yes	256	Random	Yes	0.898 [0.727, 1.000]	0.928	2	0	0.818	1.000

Table 1. Performance metrics across ASGM detection experiments with varying use of OSM data, patch sizes, negative sample types, and class weighting settings. Bootstrap confidence intervals are shown for F1-scores.

textual information can substantially enhance detection accuracy, but its effectiveness depends on data quality, temporal alignment, and the training setup. Future work should extend evaluation across regions, develop a tailored model, incorporate historical OSM data to address temporal mismatches, and explore advanced fusion strategies that more effectively align spectral and contextual features.

Acknowledgements

This research is supported by the European Union's Horizon Europe research and innovation programme within the ENEXA project (grant Agreement no. 101070305).

References

[1] S. Wang *et al.*, "Evaluating the Feasibility of Illegal Open-Pit Mining Identification Using Insar Coherence," *Remote Sensing*, vol. 12, no. 3, 2020.

[2] M. A. Alessi, P. G. Chirico, and M. Millones, "Artisanal Mining River Dredge Detection Using SAR: A Method Comparison," *Remote Sensing*, vol. 15, no. 24, p. 5701, 2023.

[3] W. Han *et al.*, "A survey of machine learning and deep learning in remote sensing of geological environment: Challenges, advances, and opportunities," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 202, pp. 87–113, 2023.

[4] OpenStreetMap contributors, *OpenStreetMap: The Free Wiki World Map*, 2025.

[5] P. Kozińska and J. Górniak-Zimroz, "A review of methods in the field of detecting illegal open-pit mining activities," *IOP Conference Series: Earth and Environmental Science*, vol. 942, no. 1, p. 012 027, 2021.

[6] J. Wang *et al.*, "Machine learning in modelling land-use and land cover-change (LULCC): Current status, challenges and prospects," *Science of The Total Environment*, vol. 822, p. 153 559, 2022.

[7] S. Chen *et al.*, "Review of drivers of forest degradation and deforestation in Southeast Asia," *Remote Sensing Applications: Society and Environment*, vol. 33, p. 101 129, 2024.

[8] Z. Shao *et al.*, "Urban sprawl and its impact on sustainable urban development: A combination of remote sensing and social media data," *Geo-spatial Information Science*, vol. 24, no. 2, pp. 241–255, 2021.

[9] M. Kucharczyk and C. H. Hugenholtz, "Remote sensing of natural hazard-related disasters with small drones: Global trends, biases, and research opportunities," *Remote Sensing of Environment*, vol. 264, p. 112 577, 2021.

[10] S. Camalan *et al.*, "Change Detection of Amazonian Alluvial Gold Mining Using Deep Learning and Sentinel-2 Imagery," *Remote Sensing*, vol. 14, no. 7, p. 1746, 2022.

[11] S. A. Mehta *et al.*, "A Characterization of Land-use Changes in the Proximity of Mining Sites in India," *ACM J. Comput. Sustain. Soc.*, vol. 2, no. 1, 10:1–10:23, 2024.

[12] R. Balaniuk, O. Isupova, and S. Reece, "Mining and Tailings Dam Detection in Satellite Imagery Using Deep Learning," *Sensors*, vol. 20, no. 23, p. 6936, 2020.

[13] J. Leng *et al.*, "Realize your surroundings: Exploiting context information for small object detection," *Neurocomputing*, vol. 433, pp. 287–299, 2021.

[14] M. E. Tschuchnig and M. Gadermayr, "Anomaly Detection in Medical Imaging - A Mini Review," in *Data Science – Analytics and Applications*, 2022, pp. 33–38.

[15] A.-W. Moomen *et al.*, "Assessing the strategic applications of remote sensing for addressing illicit artisanal and small-scale gold mining activities," *GeoJournal*, vol. 89, no. 3, p. 92, 2024.

[16] A. Fonseca Gomez, "Detecting Artisanal Small-Scale Gold mines with LandTrendr multispectral and textural features at the Tapajós river basin, Brazil.," M.S. thesis, University of Twente, 2021.

[17] B. Ayo, "Integrating OpenStreetMap data and Sentinel-2 imagery for classifying and monitoring informal settlements," M.S. thesis, Universidade Nova de Lisboa, 2020.

[18] H. Li *et al.*, "Leveraging OpenStreetMap and Multimodal Remote Sensing Data with Joint Deep Learning for Wastewater Treatment Plants Detection," *Int. Journal of Applied Earth Observation and Geoinformation*, vol. 110, p. 102 804, 2022.

[19] Earth Genome, *Gold Mine Detector*, 2025.