Data Engineering and Applications

Research Paper

Multi-Source Data Fusion and AutoGluon-Tabular Learning for Accurate Open-Pit Mine Boundary Delineation

Kinga Pajor[1, *] and Bronisław Herdzik[1]

[1] Faculty of Information Technology, Warsaw University of Technology, Warsaw, 00-661, Poland
*Corresponding author: kinga.pa@pw.edu.pl

Abstract. This study proposes an AutoGluon-Tabular workflow for open-pit mine boundary delineation from multi-source remote-sensing mosaics. The method converts optical, SAR, terrain, texture and contextual mosaic descriptors into object-level records, uses automated stacked tabular learning to estimate boundary probability, and reconstructs vector boundaries through calibrated probability gradients. The workflow is designed for mining scenes where active pits, waste dumps, haul roads, bare soil and shadowed highwalls have overlapping spectral responses, making boundary extraction different from ordinary mine-area classification. By combining object-level predictors with uncertainty-aware vector cleaning, the approach aims to reduce fragmented boundaries and focus manual review on ambiguous segments. Experiments on 18 mining districts and 42,600 labelled objects show an F1-score of 0.924, an IoU of 0.873, a boundary F1-score of 0.846 within a 20 m tolerance, and a mean boundary offset of 7.9 m. Compared with random forest, XGBoost, LightGBM-only and compact U-Net baselines, the proposed workflow improves geometric boundary quality while retaining interpretable feature groups and manageable computational cost.

Keywords: *Open-Pit Mine Boundary, AutoGluon-Tabular, Multi-Source Remote Sensing, Mosaic Feature Fusion*

Received on 09 March 2023, Accepted on 27 June 2023, Published on 05 July 2023

Introduction

Open-pit mining has expanded in many mineral-producing areas, creating rapidly changing excavation faces, waste dumps, haul-road networks and exposed benches that require regular spatial monitoring. Wang et al. [1] used optical remote-sensing observations and transfer learning to support automatic identification and dynamic monitoring of open-pit mines. However, the boundary of an open-pit mine is not merely a land-cover class; it is a geometric transition among active extraction, auxiliary mining land, natural bare ground, reclaimed surfaces and transportation facilities. Optical indices help distinguish vegetation loss and exposed rock [2], radar backscatter remains useful under cloud or illumination constraints [3], and DEM derivatives can reveal slope breaks and terrace-like structures close to pit walls [4]. Multi-source mosaics therefore offer a richer observation base than single-scene imagery, but they also introduce seasonality, sensor-specific noise and mixed spatial support [5].

Existing mine mapping studies have generally used open-pit areas as the pixel-wise classification target, object-based image analysis target, or semantic segmentation mask. Yu et al. [6] showed that open-pit mine land-use classification can combine segmentation-derived units and random forest learning when labelled samples are limited. Object-based methods reduce salt-and-pepper noise by using segmented image regions as the basic mapping units [7]. Du et al. [8] developed an open-pit mine extraction model based on very high-resolution remote sensing images, indicating that deep convolutional models can strengthen spatial representation when dense masks are available; however, such models may still be sensitive to variations in mine lithology, bench geometry, atmospheric conditions and surrounding land use. Engineers in their work often need a method that can handle diverse predictors, tolerate missing features, provide calibrated probability scores, and be retrained after new mosaic products are released [9]. Automated machine learning for tabular data is suitable in this context because it searches among various model families, ensemble complementary learners, etc., reduces manual tuning, and maintains interpretability of feature spaces [10].

The main problem that this paper addresses is how to convert remote-sensing mosaics into a stable table for boundary demarcation. A pit boundary is represented by candidate image objects located near the transition between the mined and non-mined areas here. All objects have attributes such as spectral, radar, terrain, texture and context contrast, and the learning objective is to determine whether the object is in the mine-boundary belt. This formulation is not the same as the general mine-area classification; it also needs to consider edge fidelity, uncertainty calibration and vector reconstruction. It can also use the following models for supervised learning: gradient boosting, tree ensembles, neural tabular models and stacked generalization. Determine if such automated tabular learning can compete with more specialized remote-sensing models and maintain engineering transparency and computational economy.

The remainder of this paper is organized as follows. Section 2 reviews mine boundary mapping, multi-source remote-sensing fusion and automated tabular learning for geospatial applications. Section 3 describes the construction of object-level predictors from optical, SAR, terrain and contextual mosaic layers. Section 4 presents the uncertainty-guided AutoGluon-Tabular modeling strategy and the conversion from probability estimates to vector boundaries. Section 5 reports the study data, comparative experiments, ablation analysis, uncertainty diagnostics and engineering implications. Section 6 concludes the study and outlines future research directions.

Related Work

Open-Pit Mine Mapping from Remote Sensing

Remote-sensing studies of open-pit mines have progressed from manual interpretation to semi-automatic and automatic extraction of mining disturbance. Early operational workflows often used spectral thresholds and visual correction, but later methods incorporated vegetation indices, bare-soil indicators and shortwave-infrared responses to separate exposed minerals from surrounding land covers [11]. Object-based image analysis has good trade-off performance of pixel-level detail and polygon-level manageability; that is, the segmented region can store shape, compactness, texture and local context [12]. This object perspective is very suitable for mining landscapes, as the repeated spatial units of haul roads, benches, spoil heaps and pit floors are larger than single pixels but smaller than administrative mining blocks.

Many of the workflow's principal deficiencies are related to the separability of spectra. Bright quarry floors, arid bare land, urban construction sites and river sediment bars can all have similar optical signatures with open-pit surfaces [13]. SAR data have more roughness and moisture information, but due to radar geometry and speckle, the response at steep pit walls is unclear. DEM-derived slope, relief and curvature features can be used to describe the excavated shape of open-pit mines, especially near the boundary of a steep slope [14]. However, these features are rarely sufficient on their own because the mine boundary often has a gradual transition, reclaimed slopes and service areas that are not always abrupt. Therefore, a multi-source strategy is needed, but fusion should be based on boundary behavior rather than general land-cover classification.

Automated Tabular Learning for Geospatial Classification

Tabular machine learning is still widely used in geospatial classification because remote sensing objects can be expressed as heterogeneous predictors with different scales, physical meanings and missing-data patterns. Random Forests and Gradient Boosting Machines are therefore popular choices that can achieve good accuracy and have relatively simple interpretations in land-cover mapping [15]. Automated machine learning extends the above idea by automatically selecting model types, tuning hyperparameters and combining models through ensembling without manual configuration [16]. For engineering applications, this will reduce the development time and maintain the familiar tabular structure for feature construction and error inspection.

AutoGluon-Tabular is suitable for remote sensing mosaics as it supports stacked ensembles, bagging, model calibration and heterogeneous learners under a reproducible training pipeline [17]. These attributes are necessary when mining samples are imbalanced and boundary objects account for a small proportion of the scene. However, automated tabular learning has been applied to open-pit boundary delineation very rarely. Many studies have reported the classification accuracy of mine or non-mine areas, but fewer have connected probability calibration with vector boundary quality and fragmentation reduction [18]. Therefore, in this paper, the boundary candidates are defined as object-level records, and the probability-to-vector stage is included in the overall engineering workflow.

Multi-Source Mosaic Representation

Mosaic Sources and Object Candidates

The input data are cloud-screened optical mosaics, SAR backscatter mosaics and terrain derivatives registered to a common grid. Optical mosaics provide blue, green, red, near-infrared and shortwave-infrared reflectance, and SAR mosaics provide VV, VH and ratio layers. A DEM stack includes elevation, slope, local relief, plan curvature and profile curvature. Candidate objects are generated by multiresolution segmentation of a compact stack containing shortwave-infrared reflectance, normalized bare-ground response, SAR ratio and slope. The division scale is set to maintain the bench edge and road-width transition, and it is not too small. Figure 1 shows the full data-to-object construction path once in this article.

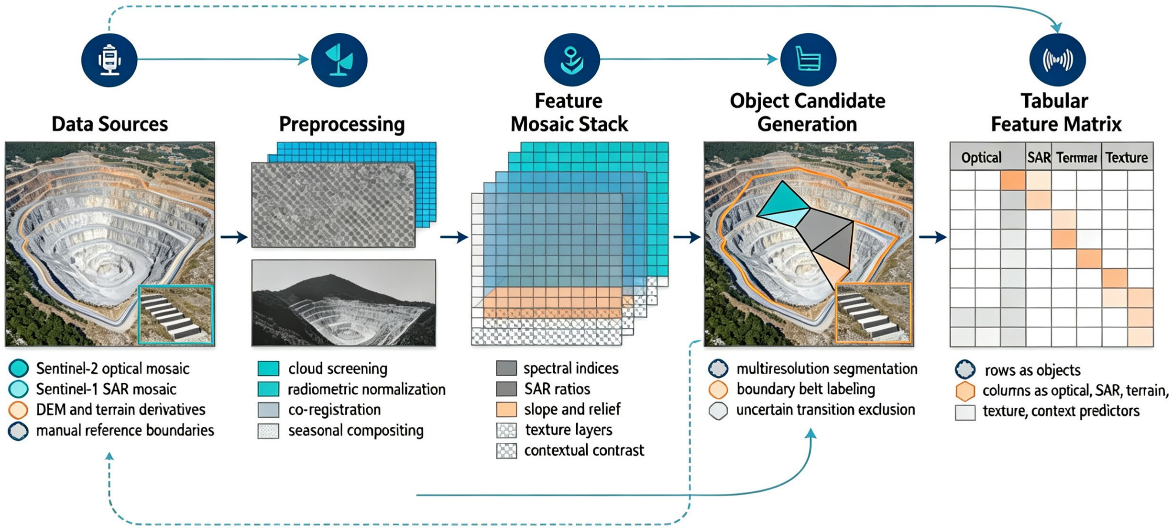

Figure 1. Multi-source remote sensing mosaic preparation and object-candidate construction workflow.

$$x_i = [s_i, r_i, t_i, g_i, c_i] \quad \text{Eq.(1)}$$

In this expression, x_i denotes the feature vector of object i, s_i contains spectral statistics, r_i contains SAR descriptors, t_i contains terrain variables, g_i contains texture descriptors, and c_i contains contextual contrast variables. Mean, standard deviation, interquartile range and edge contrast are computed for each physically meaningful layer. Context variables are measured by comparing each object with a 60 m annular neighborhood, which helps detect abrupt transitions between excavated land and the surrounding matrix.

$$NDMI = \frac{NIR - SWIR1}{NIR + SWIR1} \quad \text{Eq.(2)}$$

Moisture-sensitive and bare-ground-sensitive indices are included because active pits often expose dry mineral surfaces, and reclaimed or inactive sections may have regrown vegetation and moisture. The set of features does not include normalized indices; it also includes raw bands and cross-source interactions, and automated tabular learners can use non-linear thresholds and feature combinations.

Feature Normalization and Label Geometry

Labels are generated from manually checked mine-boundary polygons. To train a boundary model rather than a general mine-mask model, polygons are converted into a belt representation around the reference boundary. Objects whose centroids fall within the belt are positive boundary samples, while objects outside the belt and away from uncertain transition zones are negative samples. This design avoids penalizing objects that are inside active mine interiors but do not express boundary characteristics. Numeric features are winsorized at the 1st and 99th percentiles within each mosaic year, missing SAR observations are encoded with availability indicators, and all continuous predictors are standardized inside the training fold.

$$y_i = \mathbf{1}\{d(o_i, B) \leq \tau\} \qquad \text{Eq.(3)}$$

Here, y_i is the boundary label, $d(o_i, B)$ is the shortest distance between object centroid o_i and the reference boundary B, and τ is the boundary-belt tolerance. A value of 20 m is used in the main experiment because it is equal to two 10 m pixels and meets the positional uncertainty expected in mixed optical-SAR mosaics.

Boundary-Oriented Predictors

In total, the 86 variables in the final predictor table are divided into 28 optical statistics, 12 SAR statistics, 13 terrain descriptors, 21 texture variables and 12 contextual contrast features. Texture variables are calculated from the grey-level co-occurrence matrix of shortwave-infrared and SAR-ratio layers, including contrast, entropy, homogeneity and angular second moment. Boundary-oriented contrast is calculated for slope, shortwave-infrared reflectance, NDMI and VV/VH ratio because these attributes often change at the boundaries of pits, waste dumps and surrounding ground. Feature correlation is examined after object generation, but correlated predictors are not aggressively eliminated because stacked ensembles can distribute the weight of related variables and maintain the redundancy necessary for seasonal noise.

Uncertainty-Guided AutoGluon Boundary Delineation

AutoGluon-Tabular Ensemble Configuration

The modeling stage uses AutoGluon-Tabular to train a binary classifier on object-level records. The candidate learner pool includes LightGBM, CatBoost, XGBoost, random forest, extremely randomized trees, k-nearest neighbors and neural tabular networks. Bagging is enabled to stabilize estimates under mine-wise heterogeneity, and stack levels are limited to avoid excessive training time. Class weights are applied because positive boundary objects form a narrow belt and are less frequent than background objects. The weight ratio is selected from validation behavior rather than fixed only by class frequency, because excessive weighting can improve recall while reducing boundary compactness.

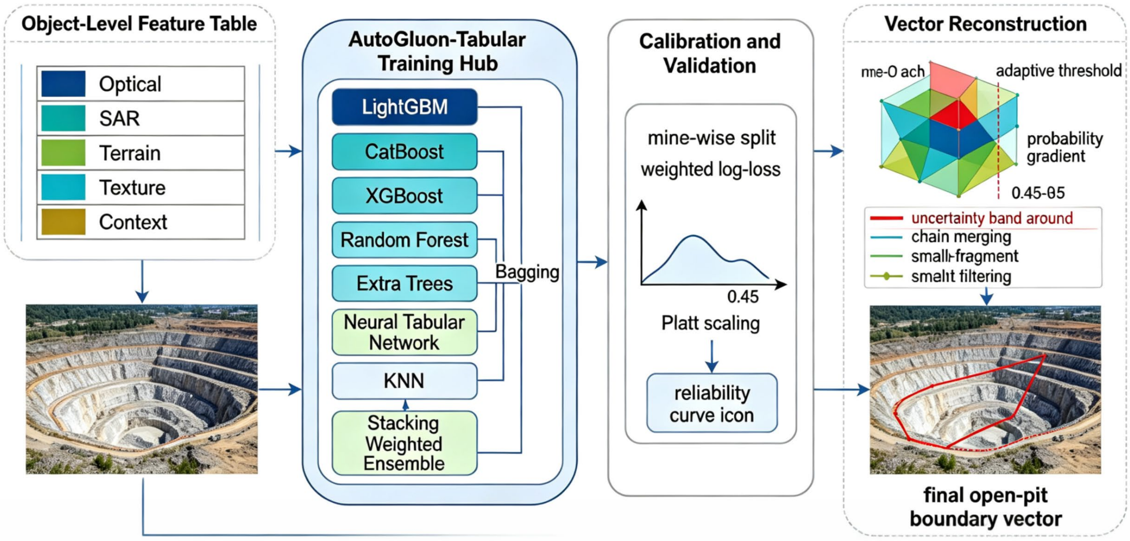

Figure 2. Uncertainty-guided AutoGluon-Tabular modeling and probability-to-vector boundary reconstruction framework.

$$p_i = P(y_i = 1 \mid x_i; M) \qquad \text{Eq.(4)}$$

The output p_i is the estimated probability that object i belongs to the mine-boundary belt under the ensemble model M. A calibrated probability is required for the following vector reconstruction to use probability gradients instead of hard class labels. A validation set is divided by mine district to reduce the impact of overconfident predictions in unseen lithological and geomorphic environments through calibration.

$$L = -\sum_{i} w_{y_i}[y_i log(p_i) + (1 - y_i) log(1 - p_i)] \quad \text{Eq.(5)}$$

Weighted log-loss gives more weight to the positive boundary objects and still penalizes false expansions in non-mining areas. Based on the validation results, instead of setting the weight ratio according to the class frequency, one is chosen; otherwise, a large weight would increase recall but reduce the geometric compactness of the reconstructed boundary.

Probability Calibration and Vector Reconstruction

Predict the object probability and then map it to the mosaic geometry. Candidate boundary chains are generated by selecting connected object edges with a probability exceeding an adaptive threshold and a local probability gradient that indicates a transition from mine interior to exterior. Remove small isolated chains, and if the end points are close and the direction is the same as in the surrounding terrain break, merge adjacent chains.

$$q_i = \frac{1}{1 + exp\big(-(az_i + b)\big)} \quad \text{Eq.(6)}$$

The calibrated score q_i is obtained from the raw ensemble logit z_i using validation-fitted parameters a and b. A simple calibration form is used to maintain the rank order after correcting for systematic overconfidence. Vector reconstruction then uses q_i and the neighborhood difference of adjacent objects.

$$G_{ij} = \frac{|q_i - q_j|}{dist(o_i, o_j) + \epsilon} \quad \text{Eq.(7)}$$

A large value of G_{ij} indicates a sharp probability transition between neighboring objects. In regions with relatively high accuracy probabilities and strong local gradients, boundary chains are used to reduce false boundaries in homogeneous ore layers.

Evaluation Metrics

The evaluation uses both object-level and boundary-level metrics. Object-level precision, recall, F1-score and area-weighted IoU describe classification quality. Boundary quality is measured by tolerance-based boundary F1-score, mean boundary offset and fragment count. These metrics are necessary because a model can obtain high object accuracy while still producing fragmented or spatially shifted boundary vectors.

$$F1 = \frac{2PR}{P + R} \quad \text{Eq.(8)}$$

Precision P measures the share of predicted boundary objects that match the reference belt, while recall R measures the share of reference boundary objects recovered by the model.

$$IoU = \frac{A_{pred} \cap A_{ref}}{A_{pred} \cup A_{ref}} \quad \text{Eq.(9)}$$

IoU is reported for reconstructed boundary-support areas after rasterizing the vector output at 10 m resolution. Boundary offset is computed as the mean nearest-neighbor distance between predicted and reference boundary vertices after resampling both polylines to a 5 m interval.

Experimental Data and Analysis

Study Data and Experimental Protocol

The 18 open-pit mining areas in the experiment cover coal, iron, copper, limestone and rare-metal production areas. Mosaics include dry-season and wet-season optical composites, ascending and descending SAR composites, and a 30-m DEM resampled to the analysis grid. Seasonal optical compositing is sensitive to changes in clouds, shadows and vegetation over mining areas [19]. Manually set boundary references and cross-reference with high-resolution interpretation layers and historical mining permits. The final object table had 42,600 labelled objects, consisting of 8,940 positive boundary objects and 33,660 background objects. A mine-wise split was employed, with 11 districts for training, 3 for validation and 4 for testing, and adjacent objects belonging to the same mine were not present in different partitions.

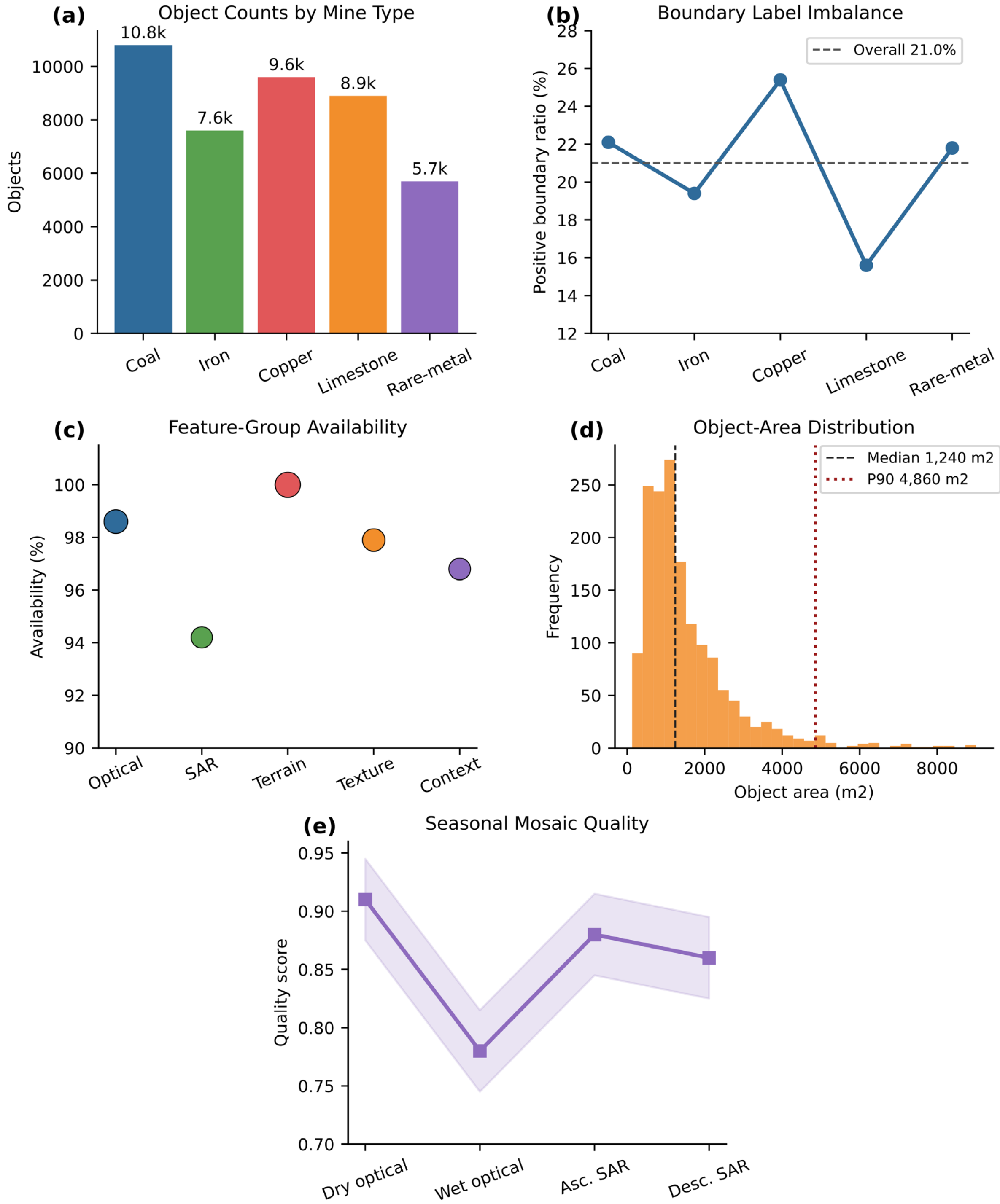

Figure 3. Study data composition and mosaic quality diagnostics: (a) The number of objects counted by mineral type; (b) Positive boundary ratio classified by mineral type; (c) feature-group availability; (d) object-area distribution; (e) seasonal mosaic quality scores.

Figure 3(a) reports the sample composition by mine type, where coal and limestone sites contribute 10,800 and 8,900 objects respectively, while rare-metal sites contribute 5,700 objects. Figure 3(b) shows that positive boundary ratios range from 15.6% in limestone mines to 25.4% in copper mines, indicating a moderate imbalance rather than an extreme anomaly-detection setting. Figure 3(c) maps feature-group availability, with optical predictors complete for 98.6% of objects, SAR predictors complete for 94.2%, and terrain predictors complete for all objects. Figure 3(d) is the distribution of object areas; the median object area is 1,240 m² and the 90th percentile is 4,860 m². Figure 3(e) shows the seasonal mosaic quality scores, and the mean usable-pixel

https://doi.org/10.64972/dea.2023.v2i3.3671d:1-13

score of the dry-season optical composite is 0.91; that of the wet-season composite is 0.78. Together, these diagnostics show how the dataset constrains model behavior before classification: the smaller rare-metal subset increases the risk of under-representation for narrow benches; the limestone class encourages background-dominant learning if class weights are not employed; incomplete SAR availability justifies the missingness indicator; and the long-tailed object-size distribution warns that large homogeneous pit-floor segments can inflate area-based scores while degrading boundary geometry.

Feature distributions confirm that no single source can define the boundary reliably. Shortwave-infrared reflectance separates fresh excavation from vegetated surroundings in many coal and limestone mines, but it overlaps with dry riverbeds and construction land. SAR ratio is more stable across seasons, yet high backscatter occurs on rough natural slopes as well as highwalls [20]. Terrain variables increase the detection of bench edges, but waste dumps and natural scarps can also show similar slope responses. The above interaction is shown in Figure 4(a); the median standardized slope contrast for true boundary objects is 1.38, but it is 0.64 for confusing non-boundary objects near roads. Figure 4(b) shows that NDMI contrast is relatively large for coal mines, and the interquartile range is 0.22-0.47. Figure 4(c) is the only heat-map diagnosis in the paper, and it shows that SWIR2 texture, VV/VH ratio and local relief form a correlated block associated with rough exposed surfaces.

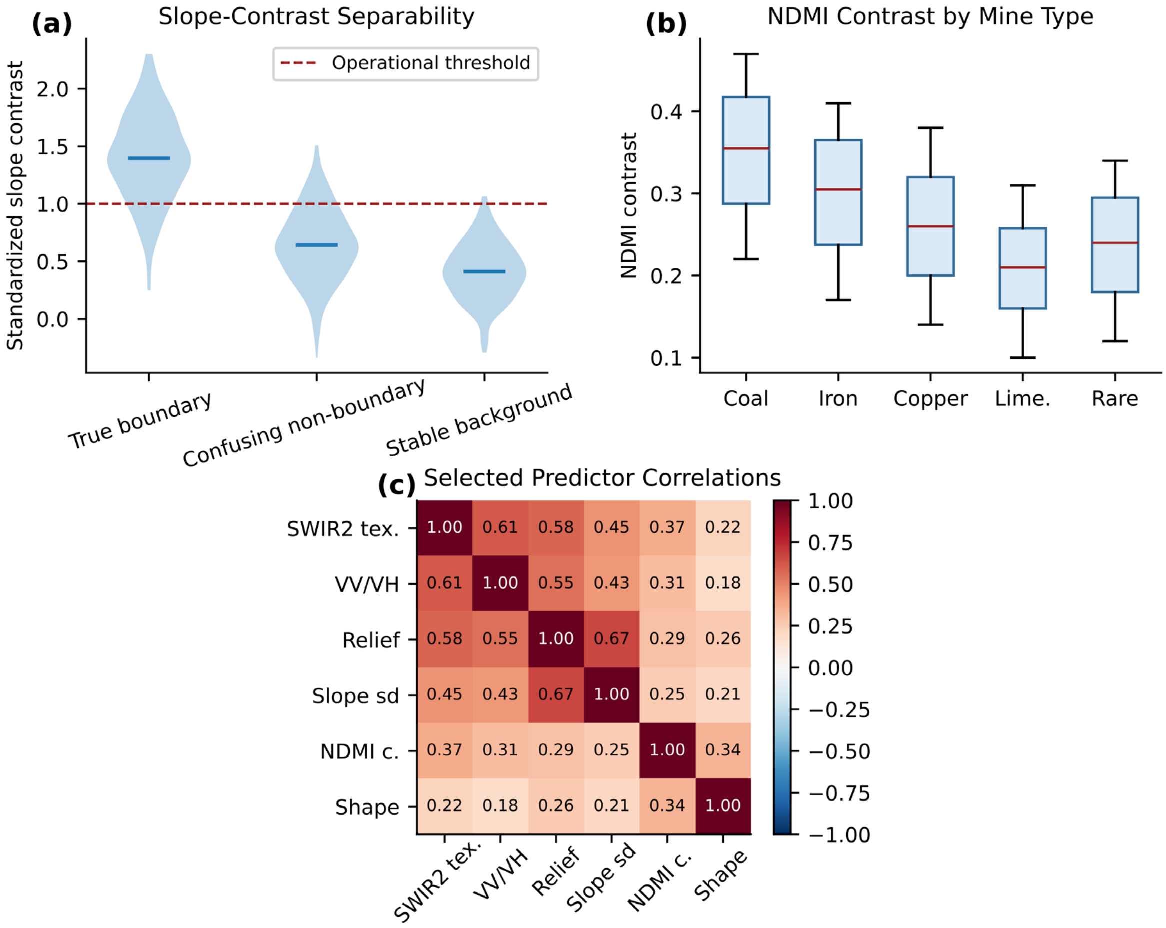

Figure 4. Feature separability and cross-source relationships: (a) slope-contrast distributions; (b) Comparison of NDMI by mineral type; (c) correlation heat map for selected predictors.

A Sampling Protocol was designed to evaluate the Generalization of boundaries rather than memorizing local textures. Each mine district was regarded as a spatial group in the partitioning, and only the validation set was used for threshold selection, calibration and ensemble configuration. The two large coal mines, the limestone quarry and the copper mine with a steep terraced highwall were in the test set. The three realistic stress conditions in this composition are: repeated road-like linear objects around haul corridors, high-reflectance bare rock outside the permitted pit, and shadowed slopes with incomplete optical contrast. During the testing process, the width of the positive band did not increase, so even if the prediction visually followed the overall boundary of the mining area, it would be considered an error as long as the deviation exceeded the tolerance range. This strict rule explains why the boundary F1-score is lower than the object-level F1-score for all methods and makes the reported geometric metrics more suitable for cartographic updates.

Comparative Model Performance

The proposed AutoGluon-Tabular ensemble was compared with random forest, XGBoost, LightGBM-only, support vector machine, and a compact U-Net trained on rasterized feature stacks. Dense segmentation models can be effective for remote sensing boundaries but usually require spatially exhaustive masks and careful transfer control [21]. Figure 5(a) shows that AutoGluon achieved the highest object-level F1-score at 0.924, compared with 0.866 for random forest, 0.891 for XGBoost, 0.902 for LightGBM-only, 0.831 for support vector machine, and 0.878 for U-Net. It is also a relatively small classification gain. Figure 5(b) shows the boundary F1-score in a 20m tolerance, and AutoGluon achieved 0.846; it was also 5.8 percentage points higher than the best baseline. As shown in Figure 5(c), the mean boundary offset was 7.9 m, and random forest and U-Net produced offsets of 12.8 m and 11.6 m, respectively. Figure 5(d) is the processing time per 100km^2; AutoGluon took 18.4 minutes after feature extraction, and U-Net training and inference on the same hardware took 43.7 minutes. Based on the above results, the stacked tabular ensemble improves the quality of geometric boundaries without the need for dense segmentation annotations and computations.

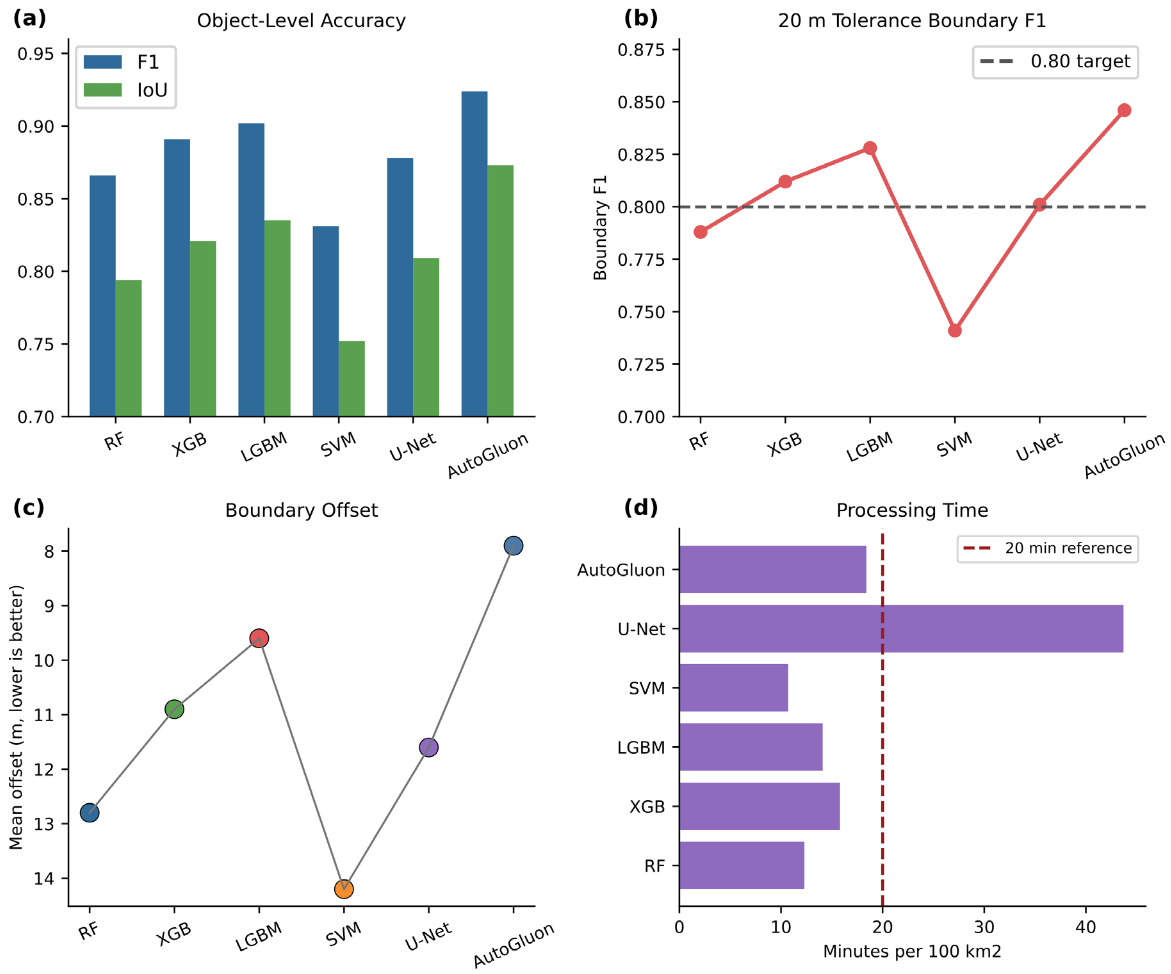

Figure 5. Comparative performance across baseline models: (a) object-level F1-score and IoU; (b) tolerance-based boundary F1-score; (c) mean boundary offset; (d) processing time per 100 km^2.

The confusion structure is spatially meaningful. Most false positives occur at waste-dump toes, quarry access roads and bright construction sites near the mining area. AutoGluon reduces these cases by using contextual contrast and terrain variables together, rather than treating bare surfaces as sufficient evidence. Calibration diagnostics are usually used to determine whether probabilistic classifiers can support downstream thresholding and human inspection [22]. In the test districts, commission error declined from 13.9% with LightGBM-only to 9.8% with the stacked ensemble, while omission error remained at 8.1%.

Figure 6(a) shows precision-recall curves for five mine types, with copper and coal mines reaching average precision values of 0.953 and 0.941. Limestone sites are harder, with average precision of 0.902, because quarry benches and surrounding bare limestone share similar reflectance. Figure 6(b) reports calibration curves; the expected calibration error decreases from 0.071 before calibration to 0.028. Figure 6(c) plots the threshold-response surface, where the best boundary F1-score appears between probability thresholds of 0.54 and 0.62. Figure 6(d) shows fragment counts after applying different minimum-chain filters, and the selected 35m filter reduces fragments from 412 to 287 while retaining 96.4% of the reference boundary length. The separated precision-recall curves show that a single global threshold cannot be taken as the equivalent evidence in all environments. The calibration panel shows that the raw ensemble scores are overconfident in the middle probability range, and the threshold panel indicates why lower thresholds accept weak road-side candidates and higher thresholds reject discontinuous but valid pit-wall segments. The fragment curve shows the vector improvement, and most of the reference boundary length has not been lost.

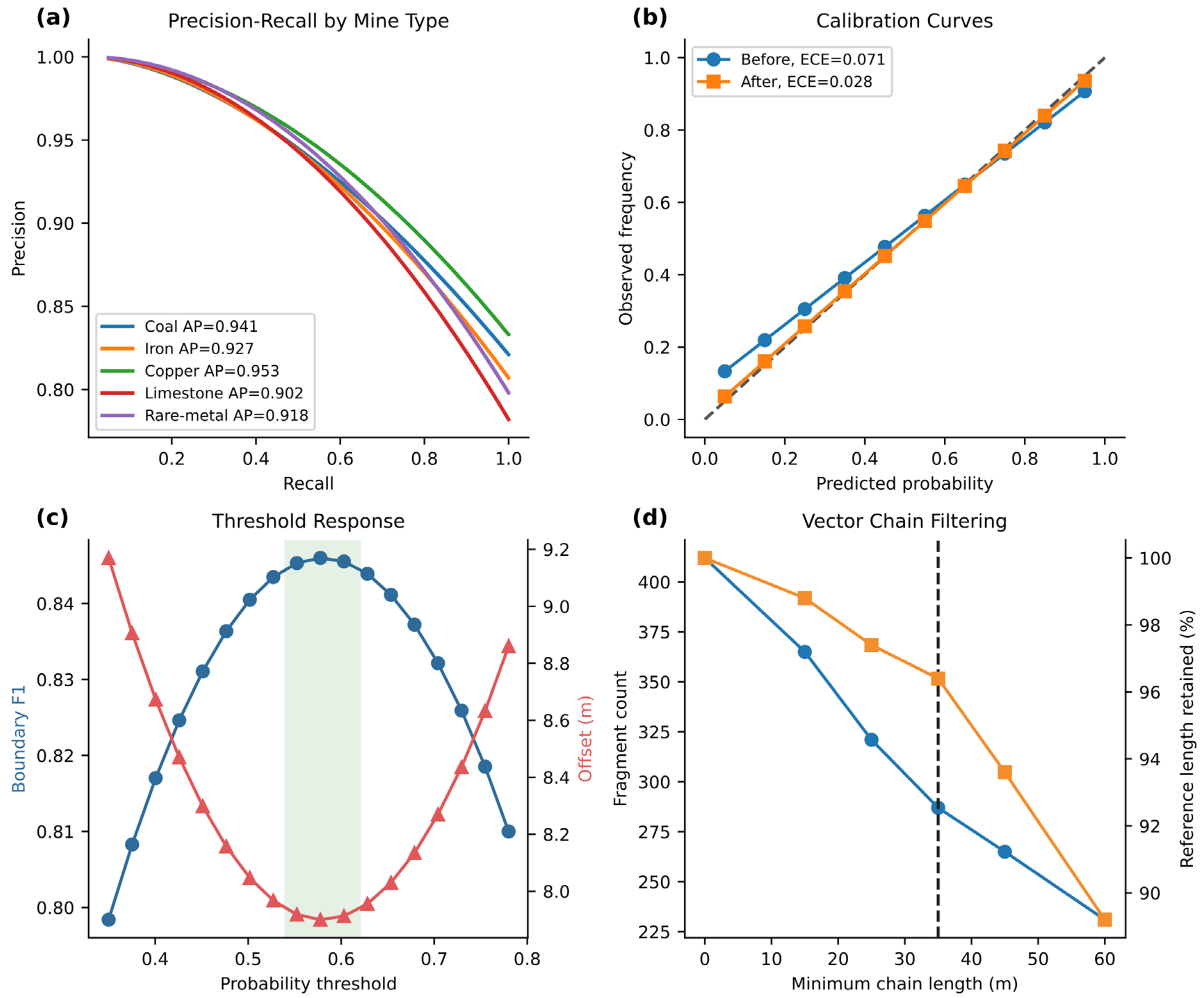

Figure 6. Probability behavior and boundary reconstruction diagnostics: (a) Precision-Recall curve categorized by mineral type; (b) probability calibration before and after scaling; (c) threshold response for boundary F1-score and offset; (d) fragment count under minimum-chain filtering.

The model rank did not change when the tolerance distance was set to 10m or 30m. At the strict 10m tolerance, AutoGluon achieved a boundary F1-score of 0.731; LightGBM-only and XGBoost were at 0.684 and 0.671, respectively. At 30m, the corresponding scores were 0.903, 0.861 and 0.849, and thus it can be seen that the proposed method improves both near-exact alignment and broader perimeter recovery. The greatest improvement was observed in mines with alternating benches and service roads, and single learners showed an over-fitting of one visual cue. The stacked ensemble assigned high probability only when several cues agreed, for example high local relief, strong shortwave-infrared texture, and a contrast between object and annular neighborhood. This behavior reduced the boundary extension into road surfaces and had a spectral response similar to active pit floors, but with a weaker terrain gradient and neighborhood contrast.

Ablation, Error Sources, and Engineering Implications

Ablation experiments quantify the contribution of each feature group. Terrain discontinuities are frequently used to describe excavated landforms and slope breaks in open-pit mines [23]. After excluding SAR variables, the boundary F1-score was 0.812 and the reduction was most significant in shadowed highwall areas. Excluding terrain variables, it was reduced to 0.804; the mean boundary offset increased from 7.9m to 10.7m, and therefore, slopes and relief are required for geometric alignment. Removing the texture variables reduced this amount; the boundary F1-score dropped to 0.827 and the number of road-related false positives increased by 3.4 percentage points. Contextual contrast variables reduced the commission error most significantly near spoil piles; without them, the commission error increased from 9.8% to 14.6%. Permutation-based feature importance is often used to examine the impact of variables in tree-dominated remote-sensing classifiers [24]. Figure 7(a) presents this ablation pattern across feature groups, and Figure 7(b) ranks permutation importance, where local relief contrast, SWIR2 entropy, VV/VH ratio, NDMI contrast, and slope standard deviation are the five most influential predictors. These two panels indicate that terrain primarily controls geometric placement, while radar and shortwave-infrared texture describe exposed roughness that separates pit walls from smoother bare ground.

Error inspection shows the location of the workflow failure. Reclaimed Pits with some vegetation have a low boundary probability due to a gradual spectral transition between the mine and its surroundings. Water-filled pit lakes have high optical contrast but weak terrain texture in the area near the shore; therefore, the vector boundary is shifted inwards. SAR layover produces a strong backscatter outside the true pit wall in steep mountainous mines, especially for slopes facing the sensor [25]. Uncertainty-guided quality control has been used to focus manual inspection on samples where the model's confidence is low [26]. Figure 7(c) gives boundary-offset distributions for four difficult landscape settings: reclaimed pits have a median offset of 9.4 m, mountainous mines 10.2 m, arid quarries 7.5 m, and coal benches 6.8 m. Figure 7(d) links uncertainty to manual correction effort; objects with calibrated probability between 0.45 and 0.65 represent 18.7% of the test set but contain 53.2% of all errors. The offset distributions show that reclaimed pits and mountainous mines are not merely lower-accuracy cases; their errors are more dispersed, so a few severely shifted segments can dominate manual correction effort. The uncertainty-workload relationship shows why ambiguous objects deserve priority review, while high-confidence segments still require occasional inspection in complex waste-dump margins.

Training the final ensemble took 31.6 minutes on a 16-core workstation, and inference over 100 km^2 required 18.4 minutes after the object features were computed. Vector cleaning was faster than 4 minutes per district. Human-in-the-loop mapping workflows often perform better when automated models provide interpretable evidence instead of just final labels [27]. When the boundary probabilities contained uncertainty flags, the manual correction time for a simulated quality-control exercise in that district was reduced from 42.5 minutes to 29.1 minutes because analysts could concentrate on the uncertain areas rather than scanning the entire perimeter. The model is also sufficiently interpretable for operational acceptance, and feature importance, calibration plots and error maps can be added to the production report. Mosaic standardization for sensors and years is also not implemented. If a new sensor exhibits different radiometric characteristics, the tabular feature distribution should be re-evaluated before using the trained model in a new area [28].

A quality-control experiment has been conducted to determine whether calibrated probabilities are superior to binary masks. The three priority classes of the predicted vectors for inspection by analysts were high-confidence accepted segments, ambiguous segments, and low-confidence rejected segments near the candidate boundary. Ambiguous parts accounted for less than a fifth of the object table but accounted for over half of the observed errors. Therefore, the relatively small number of reviews can be used to focus on areas requiring human judgment. Uncertainty flags can also distinguish between the two kinds of edits. Low-probability gaps in the otherwise stable pit wall generally required line completion, and high-probability clusters outside the pit needed to be deleted or locally tightened. The above working divisions cannot be obtained from the general hard classification map.

https://doi.org/10.64972/dea.2023.v2i3.3671d:1-13

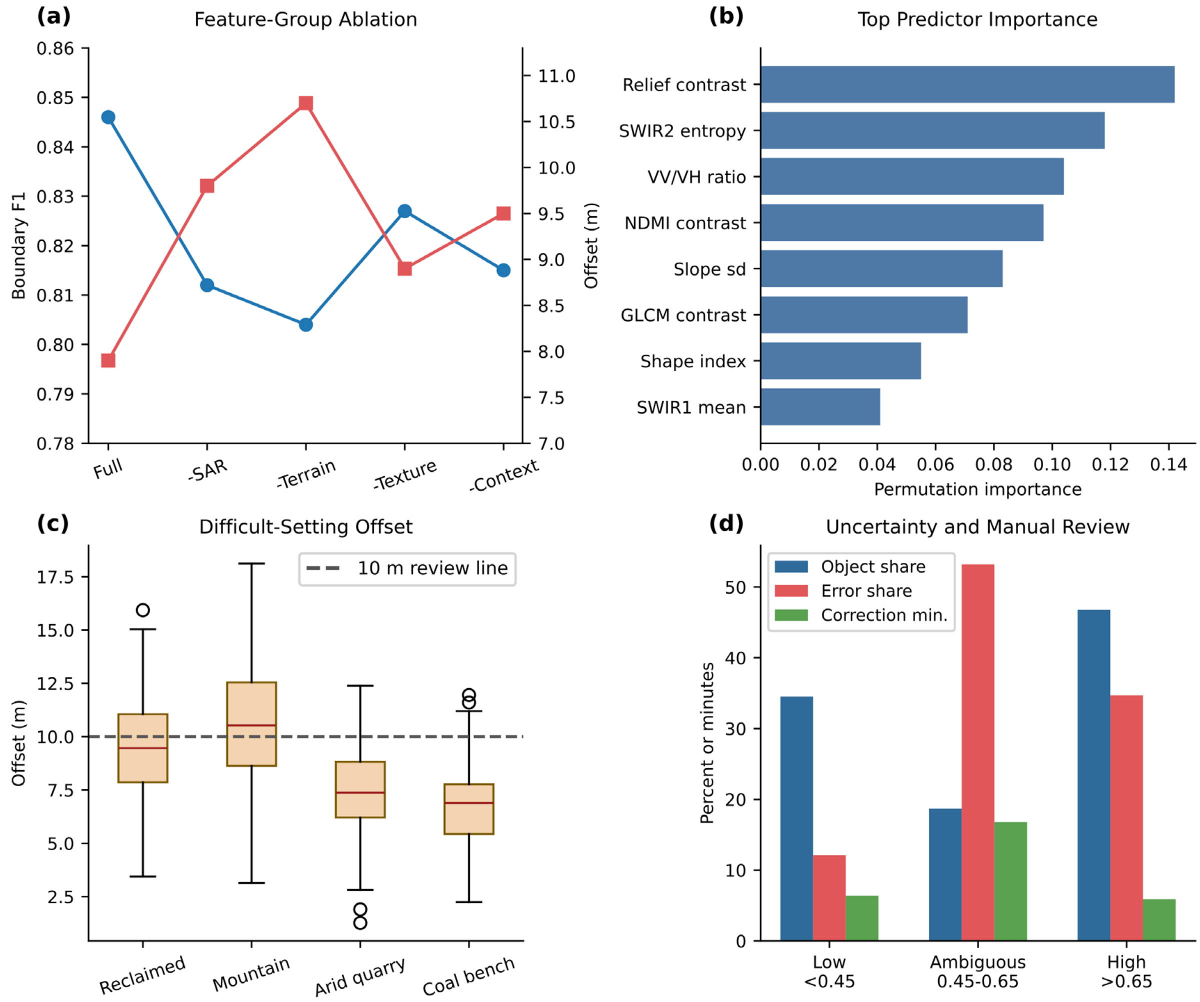

Figure 7. Ablation and engineering diagnostics: (a) feature-group ablation effects; (b) permutation importance of top predictors; (c) boundary-offset distributions in difficult settings; (d) uncertainty class versus manual correction effort.

Conclusion

This paper presents an AutoGluon-Tabular framework for open-pit mine boundary delineation from multi-source remote-sensing mosaics. The workflow converts optical, SAR, terrain, texture and contextual mosaic descriptors into object-level records, trains an automated stacked ensemble, calibrates boundary probabilities and reconstructs vector boundaries through probability-gradient filtering. The study demonstrates that tabular automated learning can address mine-boundary delineation as a measurable engineering task rather than only as a thematic land-cover classification problem.

Several limitations remain. The method depends on the quality of object segmentation, and boundary details may be lost where initial segments cross narrow benches or irregular pit walls. Reference labels also contain positional uncertainty, especially in reclaimed areas and in sites where the transition from active mine surfaces to surrounding land is gradual. Although calibration improves probability reliability, it cannot fully remove ambiguity caused by SAR layover, seasonal shadows, water-filled pits or newly disturbed ground absent from the training mosaics.

Future research should extend the framework toward temporal boundary tracking, active learning for uncertain segments and hybrid models that combine tabular ensembles with lightweight spatial encoders. More attention should also be given to cross-region transfer, automated feature-distribution monitoring and integration with mine rehabilitation records. These directions would make boundary delineation more robust for long-term monitoring and for operational mapping systems that must be updated as new remote-sensing mosaics become available.

Author Contributions

Kinga Pajor contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Bronisław Herdzik contributes to analysis, investigation, data collection, draft preparation. All authors have read and agreed with the manuscript before its submission and publication.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

References

[1] Wang, C., Chang, L., Zhao, L., & Niu, R. (2020). Automatic identification and dynamic monitoring of open-pit mines based on improved Mask R-CNN and transfer learning. Remote Sensing, 12(21), 3474. https://doi.org/10.3390/rs12213474

[2] Maxwell, A. E., Pourmohammadi, P., & Poyner, J. D. (2020). Mapping the topographic features of mining-related valley fills using Mask R-CNN deep learning and digital elevation data. Remote Sensing, 12(3), 547. https://doi.org/10.3390/rs12030547

[3] Nascimento, F. S., Gastauer, M., Souza-Filho, P. W. M., Nascimento, W. R., Jr., Santos, D. C., & Costa, M. F. (2020). Land cover changes in open-cast mining complexes based on high-resolution remote sensing data. Remote Sensing, 12(4), 611. https://doi.org/10.3390/rs12040611

[4] Balaniuk, R., Isupova, O., & Reece, S. (2020). Mining and tailings dam detection in satellite imagery using deep learning. Sensors, 20(23), 6936. https://doi.org/10.3390/s20236936

[5] Xie, H., Pan, Y., Luan, J., Yang, X., & Xi, Y. (2021). Open-pit mining area segmentation of remote sensing images based on DUSegNet. Journal of the Indian Society of Remote Sensing, 49(6), 1257-1270. https://doi.org/10.1007/s12524-021-01312-x

[6] Yu, X., Zhang, K., & Zhang, Y. (2022). Land use classification of open-pit mine based on multi-scale segmentation and random forest model. PLOS ONE, 17(2), e0263870. https://doi.org/10.1371/journal.pone.0263870

[7] Zhou, Y., Tian, S., Chen, J., Liu, Y., & Li, C. (2022). Research on classification of open-pit mineral exploiting information based on OOB RFE feature optimization. Sensors, 22(5), 1948. https://doi.org/10.3390/s22051948

[8] Du, S., Xing, J., Li, J., Du, S., Zhang, C., & Sun, Y. (2022). Open-pit mine extraction from very high-resolution remote sensing images using OM-DeepLab. Natural Resources Research, 31(6), 3173-3194. https://doi.org/10.1007/s11053-022-10114-y

[9] Yan, D., Zhang, H., Li, G., Li, X., Lei, H., Lu, K., Zhang, L., & Zhu, F. (2022). Improved method to detect the tailings ponds from multispectral remote sensing images based on Faster R-CNN and transfer learning. Remote Sensing, 14(1), 103. https://doi.org/10.3390/rs14010103

[10] Maxwell, A. E., & Warner, T. A. (2020). Thematic classification accuracy assessment with inherently uncertain boundaries: An argument for center-weighted accuracy assessment metrics. Remote Sensing, 12(12), 1905. https://doi.org/10.3390/rs12121905

[11] Fathololoumi, S., Karimi Firozjaei, M., Li, H., & Biswas, A. (2022). Surface biophysical features fusion in remote sensing for improving land crop/cover classification accuracy. Science of the Total Environment, 838, 156520. https://doi.org/10.1016/j.scitotenv.2022.156520

[12] Ren, B., Ma, S., Hou, B., Hong, D., Chanussot, J., Wang, J., & Jiao, L. (2022). A dual-stream high resolution network: Deep fusion of GF-2 and GF-3 data for land cover classification. International Journal of Applied Earth Observation and Geoinformation, 112, 102896. https://doi.org/10.1016/j.jag.2022.102896

[13] Judah, A., & Hu, B. (2022). An advanced data fusion method to improve wetland classification using multi-source remotely sensed data. Sensors, 22(22), 8942. https://doi.org/10.3390/s22228942

[14] Maxwell, A. E., & Shobe, C. M. (2022). Land-surface parameters for spatial predictive mapping and modeling. Earth-Science Reviews, 226, 103944. https://doi.org/10.1016/j.earscirev.2022.103944

[15] Cheng, G., Xie, X., Han, J., Guo, L., & Xia, G. S. (2020). Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 3735-3756. https://doi.org/10.1109/JSTARS.2020.3005403

[16] Rußwurm, M., & Körner, M. (2020). Self-attention for raw optical satellite time series classification. ISPRS Journal of Photogrammetry and Remote Sensing, 169, 421-435. https://doi.org/10.1016/j.isprsjprs.2020.06.006

[17] Campos-Taberner, M., García-Haro, F. J., Martínez, B., Izquierdo-Verdiguier, E., Atzberger, C., Camps-Valls, G., & Gilabert, M. A. (2020). Understanding deep learning in land use classification based on Sentinel-2 time series. Scientific Reports, 10, 17188. https://doi.org/10.1038/s41598-020-74215-5

[18] Zöller, M. A., & Huber, M. F. (2021). Benchmark and survey of automated machine learning frameworks. Journal of Artificial Intelligence Research, 70, 409-472. https://doi.org/10.1613/JAIR.1.11854

[19] He, X., Zhao, K., & Chu, X. (2021). AutoML: A survey of the state-of-the-art. Knowledge-Based Systems, 212, 106622. https://doi.org/10.1016/j.knosys.2020.106622

[20] Vaccaro, L., Sansonetti, G., & Micarelli, A. (2021). An empirical review of automated machine learning. Computers, 10(1), 11. https://doi.org/10.3390/computers10010011

[21] Heizmann, M., Braun, A., Glitzner, M., Günther, M., Hasna, G., Klüver, C., Krooß, J., Marquardt, E., Overdick, M., & Ulrich, M. (2022). Implementing machine learning: Chances and challenges. Automatisierungstechnik, 70(1), 90-101. https://doi.org/10.1515/auto-2021-0149

[22] Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U. R., Makarenkov, V., & Nahavandi, S. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. Information Fusion, 76, 243-297. https://doi.org/10.1016/j.inffus.2021.05.008

[23] Mena, J., Pujol, O., & Vitrià, J. (2021). A survey on uncertainty estimation in deep learning classification systems from a Bayesian perspective. ACM Computing Surveys, 54(9), 1-35. https://doi.org/10.1145/3477140

[24] Böken, B. (2021). On the appropriateness of Platt scaling in classifier calibration. Information Systems, 95, 101641. https://doi.org/10.1016/j.is.2020.101641

[25] Chen, P., Huang, H., & Shi, W. Z. (2022). Reference-free method for investigating classification uncertainty in large-scale land cover datasets. International Journal of Applied Earth Observation and Geoinformation, 107, 102673. https://doi.org/10.1016/j.jag.2022.102673

[26] Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence, 2, 56-67. https://doi.org/10.1038/s42256-019-0138-9

[27] Liu, B., Du, S., Du, S., & Zhang, X. (2020). Incorporating deep features into GEOBIA paradigm for remote sensing imagery classification: A patch-based approach. Remote Sensing, 12(18), 3007. https://doi.org/10.3390/rs12183007

[28] Li, X., Hu, T., Gong, P., Du, S., Chen, B., Li, X., Dai, Q., & Hack, J. (2021). Mapping essential urban land use categories in Beijing with a fast area of interest (AOI)-based method. Remote Sensing, 13(3), 477. https://doi.org/10.3390/rs13030477