Post-Fire Burned Area Mapping in Javanrud, Iran Using Sentinel-2 Imagery and Random Forest Classification

Hamed Hosein Amini [1*], Faraz Sarhadi [2]

[1] University of Tehran, School of Surveying and Geospatial Engineering, Tehran, Iran – amini79hamed@gmail.com
[2] University of Tehran, School of Surveying and Geospatial Engineering, Tehran, Iran – sarhadifaraz@gmail.com
* Corresponding author

Keywords: Fire; Sentinel-2; Burned area; dNBR; Machine learning; Change detection

Abstract
Timely delineation of burned areas is essential for post-fire assessment and recovery planning. In this study, we mapped wildfire impacts around Javanrud, Iran, using Sentinel-2 Level-2A imagery acquired before and after a summer 2022 event (pre-fire: 2022-08-16; post-fire: 2022-08-26). For each epoch, we computed the NDVI and NBR indices and derived their bi-temporal differences (dNDVI and dNBR). A Random Forest (RF) classifier was trained on a feature stack comprising three key VNIR/SWIR bands (B4, B8, B12) along with NDVI, NBR, dNDVI, and dNBR to distinguish burned from unburned pixels within an Area of Interest (AOI) defined as the burn polygon buffered by 5 km. Training labels were obtained from the provided burn shapefile (positive samples) and stratified random samples within the buffer zone (negative samples), with clouds and shadows masked out.
On held-out tiles, the RF model achieved an overall accuracy (OA) of 0.96, F1-score of 0.90, and AUC of 0.997, with a burned-area estimate of 13.07 ha. Permutation-based feature importance identified dNBR and SWIR–NIR features as dominant predictors, offering a physically consistent explanation of model behavior. The proposed workflow—combining simple spectral differencing with a lightweight supervised learner—provides a fast, reproducible, and cost-effective solution for post-fire mapping in heterogeneous, mountainous terrain.

1. INTRODUCTION

Wildfires have intensified across many regions over the past decade, amplifying ecological and socio-economic impacts and underscoring the need for rapid and reliable burned-area mapping to support response and recovery efforts. Satellite remote sensing—particularly Sentinel-2, with its 10 m VNIR and 20 m SWIR bands and frequent revisit capability—enables comprehensive burn-scar delineation and severity assessment at landscape scales.

Classical post-fire mapping workflows rely on vegetation- and moisture-sensitive indices such as the Normalized Burn Ratio (NBR) and Normalized Difference Vegetation Index (NDVI), along with their bi-temporal variants (dNBR and dNDVI) computed from pre- and post-fire imagery. While these indices are simple and effective, their performance can be limited by heterogeneous land cover, phenological variation, and topographic illumination effects, which may introduce commission and omission errors when fixed thresholds are used alone.

Supervised learning offers a practical approach to integrating spectral bands and change indices while accommodating nonlinear decision boundaries. In particular, the Random Forest (RF) algorithm provides high accuracy with minimal tuning, robustness to noisy features, and straightforward interpretability through variable-importance analysis—attributes that are especially valuable for time-critical crisis mapping. When combined with bi-temporal change features, RF can reduce fragmented false positives and refine fire perimeters compared with threshold-based methods.

This study focuses on a 2022 summer wildfire in the Javanrud region of Kermanshah Province, Iran. We developed a concise and operational pipeline that (i) selects cloud-free pre- and post-fire Sentinel-2 scenes; (ii) computes NDVI, NBR, dNDVI, and dNBR; (iii) trains an RF classifier on pre- and post-fire reflectance bands (B4, B8, B12) together with these indices using labels derived from a burn polygon and stratified unburned samples within a 118.36 ha area; and (iv) evaluates model performance on spatially disjoint tiles using metrics including overall accuracy (OA), Cohen's κ, precision, recall, F1-score, and ROC–AUC, alongside mapped burned area. The proposed workflow is streamlined, reproducible, and tailored to heterogeneous, mountainous terrain, with results confirming the dominant role of SWIR–NIR change features—particularly dNBR—in post-fire detection.

2. Study of Area

The study focuses on the surroundings of Javanrud in Kermanshah Province, western Iran, located on the western flank of the Zagros Mountains. The landscape is predominantly mountainous, with dissected valleys and foothills that create strong elevation, slope, and aspect gradients influencing both fire behavior and post-fire recovery. Land cover consists of a mosaic of Zagros oak woodlands (dominated by Quercus spp.), interspersed with patches of shrub–steppe, rangelands, and rain-fed agriculture and orchards situated in valley bottoms and gentler slopes. The region experiences a Mediterranean-type climate characterized by cool, wet winters and hot, dry summers. Prolonged summer drought and episodic winds create favorable conditions for wildfire spread. A notable wildfire occurred during the summer of 2022 within the Javanrud region, affecting woodland and rangeland areas across complex terrain. For this study, we delineated an Area of Interest (AOI) encompassing the burn scar and adjacent unburned reference zones to support training and validation.

The AOI spans approximately 1.09 km east–west and 1.08 km north–south (centered near 34.8236°N, 46.4945°E), covering about 118.36 ha. According to local situation reports and visual inspection of satellite imagery, the fire progressed generally from southwest to northeast, with the most severe impacts observed on south- and west-facing slopes and at mid-elevations where fuel continuity is higher.

To capture pre- and post-fire conditions, we used Sentinel-2 Level-2A imagery acquired before (2022-08-16) and after (2022-08-26) the event, both selected for minimal cloud contamination over the AOI. The post-fire scene corresponds closely to the burned-area stabilization period, allowing for reliable spectral differencing (dNBR, dNDVI). We also assembled ancillary datasets for contextual interpretation, including administrative boundaries (Kermanshah Province), a 10–30 m digital elevation model for terrain variables (elevation, slope, aspect), and recent

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W8-2025-69-2026 | © Author(s) 2026. CC BY 4.0 License.

land-cover data. Based on our mapping results (Section 4), the burned area within the AOI totals 13.07 ha, showing severity heterogeneity consistent with the topographic and fuel gradients of the Zagros foothills.

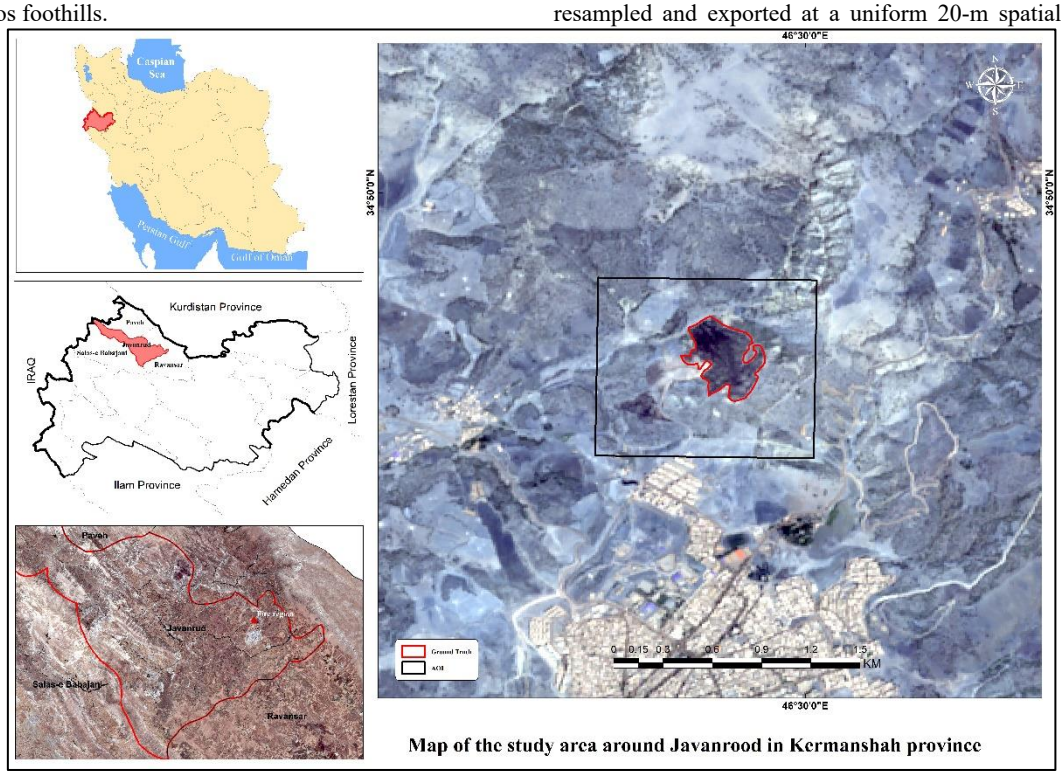

Figure 1- Map of the study area and the Javanrud fire perimeter in Kermanshah province.

3. Data and Materials

This section describes the satellite datasets used for burned-area detection in the Javanrud region, as well as the software platforms and analytical tools employed for data preprocessing, index computation, and machine-learning classification.

3.1 Study AOI and reference polygon

The Area of Interest (AOI) for this study was defined as a rectangular bounding box encompassing the full extent of the available burned-area reference polygon. The AOI was specifically designed to include the burn perimeter and its immediate surrounding landscape, enabling robust model training and validation while minimizing computational load.
This approach eliminates the need for a wide buffer zone and focuses the analysis on the most relevant area. The final AOI covers approximately 118.36 hectares (1.18 km²) and is delineated by a rectangular boundary surrounding the burn site. The center of the AOI is located near 34.8236° N, 46.4945° E (latitude, longitude). The available burned-area shapefile serves as the reference polygon, providing the training and validation data for supervised classification.

3.2 Satellite imagery (Sentinel-2 Level-2A)

We used Sentinel-2 Level-2A surface reflectance imagery (10–20 m) accessed and downloaded via Google Earth Engine (GEE). Two cloud-free (or low-cloud) epochs were selected:

- Pre-fire date: 2022-08-16
- Post-fire date: 2022-08-26

For each epoch, we extracted the following bands: B2 (Blue, 10 m), B3 (Green, 10 m), B4 (Red, 10 m), B8 (NIR, 10 m), B11 (SWIR1, 20 m), and B12 (SWIR2, 20 m). All bands were resampled and exported at a uniform 20-m spatial resolution, matching the native resolution of the SWIR channels to ensure consistency during training and analysis.
Cloud and shadow contamination were minimized using the Sentinel-2 Scene Classification Layer (SCL) available in GEE. A mask was applied to exclude SCL classes corresponding to clouds, cloud shadows, cirrus, and snow/ice (classes 3, 8, 9, 10, and 11), ensuring consistent pre- and post-fire conditions across the AOI.

3.3 Burn indices and derived variables

From the pre- and post-fire Sentinel-2 images, we computed vegetation- and moisture-sensitive spectral indices and their bi-temporal differences as follows:

$$NDVI = (B8 - B4) / (B8 + B4) \quad (1)$$

$$NBR = (B8 - B12) / (B8 + B12) \quad (2)$$

$$dNDVI = NDVIpre\_ - NDVIpost \quad (3)$$

$$dNBR = NBRpre - NBRpost \quad (4)$$

These indices constituted the core features for the initial burn-severity screening and guided the sampling strategy. The final machine-learning feature stack included multi-band reflectance (B2, B3, B4, B8, B11, B12) together with the derived indices (NDVI, NBR, dNDVI, dNBR). Terrain variables were optionally appended during sensitivity tests to assess topographic influence.

3.4 Reference data and sampling

The available burned-area shapefile (fireShape) served as the reference polygon for pixel labeling, where burned pixels were assigned a value of 1 and unburned pixels a value of 0. The

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W8-2025-69-2026 | © Author(s) 2026. CC BY 4.0 License.

sampling strategy aimed to create a balanced and spatially focused dataset within the rectangular AOI bounding box.

Sampling and Balancing

Sampling was conducted at a 20 m spatial scale within Google Earth Engine (GEE):

1. **Burned samples**: Pixels were randomly sampled within the fireShape polygon, with masking applied to exclude cloud and shadow contamination.
2. **Unburned samples**: A separate unburned region was defined using a geometric difference (AOI minus fireShape). From this region, 200 random pixels were sampled to balance approximately against 50 burned pixels.

Dataset Splitting

The final combined dataset (250 samples) was exported as a CSV file and processed in Python. It was divided into training (70%) and testing (30%) subsets using a stratified sampling approach to maintain class balance across both groups.

3.5 Implementation environment

All preprocessing steps—including date filtering, cloud masking, band and index computation, and data export—were performed in Google Earth Engine (GEE) using the Sentinel-2 Level-2A Surface Reflectance collection (COPERNICUS/S2_SR). The embedded Scene Classification Layer (SCL) was utilized for noise reduction and cloud/shadow filtering.

The exported multi-band raster files and reference vector data were subsequently transferred and analyzed locally within a Python environment. The workflow leveraged the pandas library for data handling, scikit-learn for Random Forest (RF) model training and evaluation, and rasterio for raster processing and final output generation.

4. METHODS

4.1 Workflow overview

We implemented a two-stage workflow specifically designed for burned-area mapping using Sentinel-2 imagery:

(i) Preprocessing and Feature Generation: This stage involved cloud masking and the computation of spectral change features, including dNBR and dNDVI, derived from the pre-fire and post-fire images.

(ii) Supervised Classification: A Random Forest (RF) model was trained on the generated features and applied across the AOI (rectangular bounding box) to produce a binary burned/unburned map.

4.2 Preprocessing in Google Earth Engine (GEE)

Two Sentinel-2 Level-2A scenes with minimal cloud cover were selected: pre-fire (2022-08-16) and post-fire (2022-08-26). The images were first filtered by date and spatially clipped to the rectangular AOI. Cloud and shadow contamination were minimized using the Scene Classification Layer (SCL), with pixels corresponding to clouds, cloud shadows, cirrus, and snow/ice masked out.

Bands B2, B3, B4, and B8 (10 m), as well as B11 and B12 (20 m), were retained. To ensure consistency during model training and data export, all bands were resampled to a uniform 20 m spatial resolution, matching the native resolution of the SWIR channels.

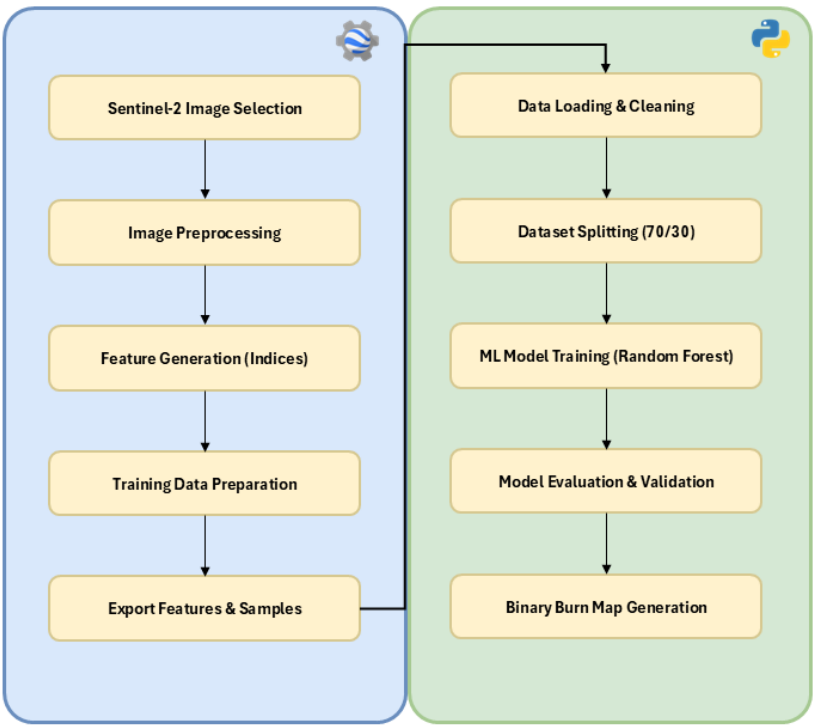

Figure 2- Burned Area Mapping Workflow

4.3 Spectral indices and change features

The Normalized Difference Vegetation Index (NDVI) and the Normalized Burn Ratio (NBR) were computed for both the pre- and post-fire epochs, and subsequently used to derive the bi-

temporal change features dNDVI and dNBR. The detailed formulas for these indices are provided in Section 3.3.
The final Random Forest (RF) feature set consisted of six reflectance bands (B2, B3, B4, B8, B11, and B12) along with four derived indices (NDVI, NBR, dNDVI, and dNBR). Terrain variables were generated for sensitivity testing but were not included in the core classification model.

4.4 Initial Sampling Design

An exploratory dNBR threshold was used solely to visualize the likely burn patterns and to guide sampling for the positive (burned) class. Training labels were derived from two sources:
(i) Positive Class (Burned): Samples were extracted from the designated burned-area shapefile.
(ii) Negative Class (Unburned): Samples were drawn outside the burn polygon but within the rectangular AOI bounding box.
During preprocessing in GEE, cloud and shadow pixels were excluded using the Scene Classification Layer (SCL) mask. The final sample set was balanced (50 burned and 200 unburned samples) to mitigate class imbalance. The combined dataset was subsequently divided into training (70%) and testing (30%) subsets using a stratified random sampling approach within the local Python environment.

4.5 Random Forest classifier

A Random Forest (RF) model was trained using the scikit-learn library in a Python environment. The model was fitted on the 70% training subset, utilizing the full spectral and derived index feature stack. Feature inputs were used without prior standardization or normalization.The key RF parameters were configured as follows:

- n_estimators = 100 : Number of trees in the forest.
- class_weight = "balanced": Used to mitigate residual class imbalance between the burned and unburned samples.
- random_state = 42: Ensured full reproducibility of the model results.

The trained model's performance was subsequently evaluated on the 30% held-out test subset.

4.6 Model interpretation

The relative importance of the input features was evaluated using the Gini Importance metric (also known as Mean Decrease in Impurity) embedded within the scikit-learn Random Forest model. As commonly observed in burned-area mapping studies, the dNBR index and the associated SWIR–NIR reflectance bands were expected to exhibit the highest predictive power. Features such as NDVI, dNDVI, and the remaining visible bands contributed to refining the classification boundaries, particularly in areas with subtle burn signals or complex topography.

4.7 Accuracy assessment and operating point

Model performance was rigorously evaluated using the 30% held-out test subset. The computed metrics included both standard per-class performance measures and overall diagnostic scores:

- Overall Metrics: Overall Accuracy (OA), Cohen's κ, and the Receiver Operating Characteristic – Area Under the Curve (ROC–AUC).
- Per-Class Metrics (derived from the confusion matrix): F1-score, Precision, Recall, False Positive Rate (FPR), and False Negative Rate (FNR).

The classification output from the Random Forest model was a binary burned/unburned map, generated using the model's default probability threshold of 0.5 to define class membership. The total estimated burned area (in hectares) obtained from the final prediction map is reported in the Results section.

4.8 Implementation details and reproducibility

All preprocessing steps—including filtering, cloud masking, index computation, and data export at a 20 m spatial scale—were performed in Google Earth Engine (GEE) over the defined AOI bounding box. Random Forest (RF) model training and evaluation were conducted in a Python environment using the scikit-learn package (version X.Y.Z) and the pandas library.
For full reproducibility, the random seed was fixed at 42 during both the sample balancing and the training/testing split procedures. The AOI definition, Sentinel-2 image IDs, sampling masks, and final training/testing datasets are archived and available for verification.

5. Results

The performance of the Random Forest (RF) classifier was rigorously evaluated using the 30% held-out test subset, and the computed metrics are summarized in Table 1.

The RF model demonstrated exceptionally high discriminatory power and classification accuracy. The overall accuracy (OA) reached 96.0%, and Cohen's Kappa (κ) coefficient was 0.872, indicating strong agreement beyond chance. Furthermore, the ROC–AUC score of 0.997 suggests that the model is highly effective at distinguishing between burned and unburned classes.

Table 1- Accuracy assessment of the Random Forest burned -area classification

Metric	Value
Overall Accuracy (%)	96.00
Kappa (κ)	0.872
Precision (%)	93.00
Recall (%)	87.00
F1-Score (%)	90.00
Balanced Accuracy (%)	92.50
False Alarm Rate (FPR, %)	1.67
Missed Detection Rate (FNR, %)	13.33
ROC–AUC	0.997

For the positive (burned) class, the model exhibited excellent performance, achieving a precision of 93.0%. This high precision is supported by a remarkably low false alarm rate (FPR) of only 1.67%, confirming the model's effectiveness in avoiding the misclassification of unburned areas as burned.
However, the recall for the burned class was 87.0%, corresponding to a missed detection rate (FNR) of 13.33%. This indicates that approximately 13% of the truly burned pixels were incorrectly classified as unburned. Such slight asymmetry in error types is commonly attributed to weaker burn signals along the fire perimeter or to residual effects of class imbalance in the reference data.

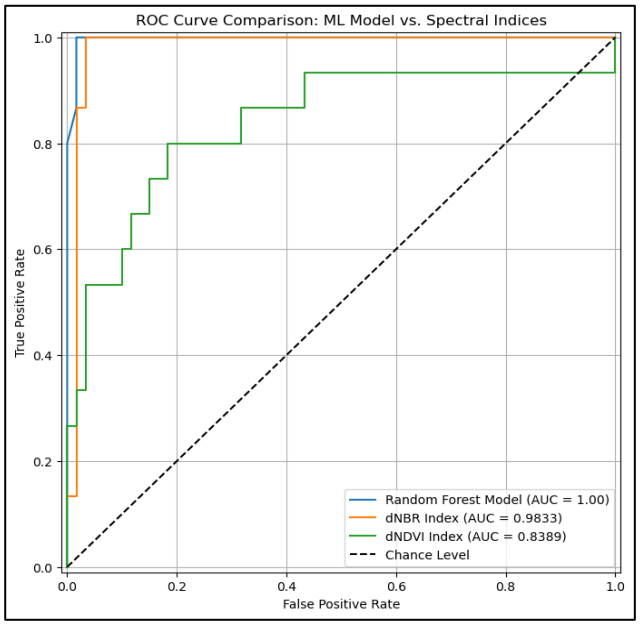

Figure 3- ROC Curve Comparison: ML Model vs Spectral Indices

yielding an AUC of 0.839. These results visually confirm that incorporating the full feature stack substantially enhances the model's discriminatory power beyond that of individual spectral indices such as dNBR.

Before training the RF model, the linear correlations among the input features were examined using the Pearson correlation matrix (Figure 4). Key observations regarding the relationships within the feature stack are as follows:

- **Strongest Correlations**: High positive correlation was observed between features conveying similar spectral information, such as nir_after and swir_after (r = 0.75). The strongest negative correlation occurred between red_before and NDVI_after (r = −0.78), reflecting the pronounced spectral contrast between pre-fire visible reflectance and post-fire vegetation response.
- **Relationship with dNBR**: The change index dNBR showed strong negative correlations with post-fire bands nir_after (r = −0.82) and swir_after (r = −0.78), confirming that the sharp post-fire decline in these reflectance bands is the primary driver of the positive dNBR signal.
- **Index Independence**: The correlation between dNBR and dNDVI was moderate (r = 0.47), suggesting that while related, each index contributes complementary information to the classification process.

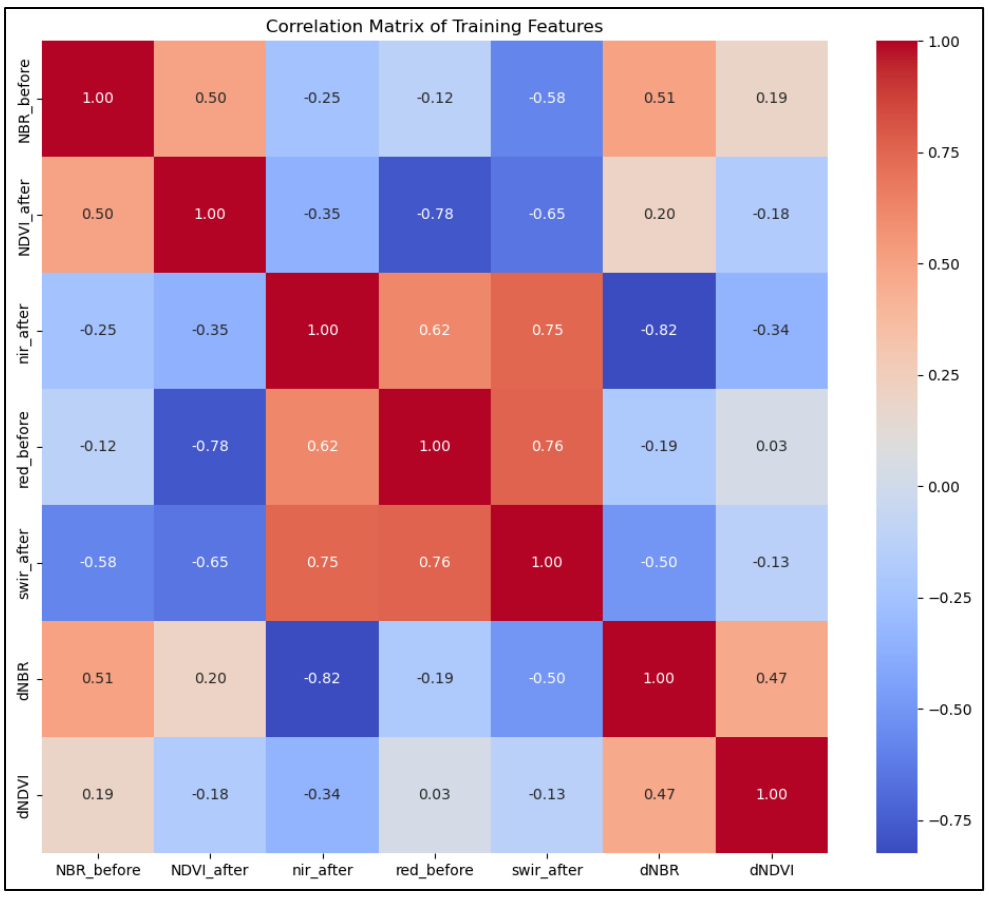

Figure 4- Correlation matrix of training features

Figure 3 presents the comparative ROC curves. The Random Forest (RF) model achieved a near-perfect AUC of 1.00 on the test dataset, confirming its superior classification capability relative to single-index methods. Among the individual indices, the dNBR—traditionally the benchmark for burn mapping—exhibited the highest standalone predictive power with an AUC of 0.983. The dNDVI index also showed strong potential,

Although several strong correlations were identified, the Random Forest classifier is known to be robust against multicollinearity, thereby justifying the inclusion of the full feature stack to maximize predictive performance.

The final binary classification map generated by the RF model over the AOI is shown in Figure 5, visually illustrating the

model's ability to accurately delineate the core burned patches from the surrounding unburned landscape.
Figure 5 also provides a quantitative comparison of the burned-area estimates:

- Ground Truth area: Derived from the reference shapefile, totaling 120,175 m^2.
- Predicted area: The Random Forest (RF) model estimated a total burned area of 130,700 m²..

The predicted burned area exceeded the reference by approximately 10,525 m², corresponding to an overestimation of about 8.75%. This slight overestimation is consistent with the model's high precision (93.0%) and very low false positive rate (FPR = 1.67%). It suggests that while the model effectively minimizes false alarms, the RF-predicted boundary tends to extend marginally beyond the human-digitized reference, successfully capturing faint burn signals along the fire perimeter.

6. discussion

The Random Forest (RF) model, trained on pre- and post-fire Sentinel-2 features—including six reflectance bands, NDVI, NBR, and their differenced forms (dNDVI and dNBR)—produced a coherent burned/unburned classification map for the Javanrud AOI. The RF approach significantly outperformed simple dNBR thresholding by reducing fragmented false positives and sharpening burned-area boundaries within heterogeneous topography.
Quantitatively, the model achieved an overall accuracy (OA) of 96.0%, an F1-score of 90.0%, and a ROC–AUC of 0.997 on the held-out test subset, using the default probability threshold of 0.5. Qualitatively, residual false alarms were concentrated over bright bare soils and recently tilled fields (FPR = 1.67%), while most missed detections occurred in patchy understory burns and shadowed north-facing slopes (FNR = 13.33%).
Bi-temporal change features proved pivotal: dNBR ranked highest in importance alongside SWIR bands (B12/B11), while NDVI/dNDVI and NIR/Red bands contributed to edge refinement—patterns that are physically consistent with the NIR–SWIR sensitivity to char and post-fire moisture loss. Although deep learning models could potentially yield smoother burn perimeters, the RF classifier provided a favorable speed–accuracy trade-off for time-critical mapping and performed effectively with limited labeled data derived from the burn polygon within the AOI.
Key limitations include potential label noise along the fire edge, phenological differences between the selected pre- and post-fire dates (2022-08-16 and 2022-08-26), and aspect-related illumination effects in mountainous terrain. Future work could incorporate illumination correction, expand multi-source reference data, and integrate Sentinel-1 SAR imagery to sustain performance under cloudy or smoky conditions. Overall, the RF-based, dNBR-centric pipeline delivers an accurate, interpretable, and operationally efficient solution for post-fire mapping in the Zagros foothills, yielding a final burned-area estimate of 13.07 ha suitable for management and recovery reporting.

7. Conclusions

This study presented a minimal yet effective pipeline for post-fire mapping in the Javanrud region, integrating two Sentinel-2 epochs with a Random Forest (RF) classifier trained on a compact set of physically meaningful features (spectral bands, NDVI, NBR, and their differenced forms: dNDVI and dNBR).
The core contribution of this work is operational: using freely available inputs from Google Earth Engine (GEE), lightweight preprocessing, and a single classical learner, the workflow successfully produced a clean binary burn map and a defensible burned-area estimate of 13.07 ha, without requiring extensive data or computational resources. The model demonstrated high discriminatory power, achieving a ROC–AUC of 0.997 and an overall accuracy of 96.0% on the held-out test subset.

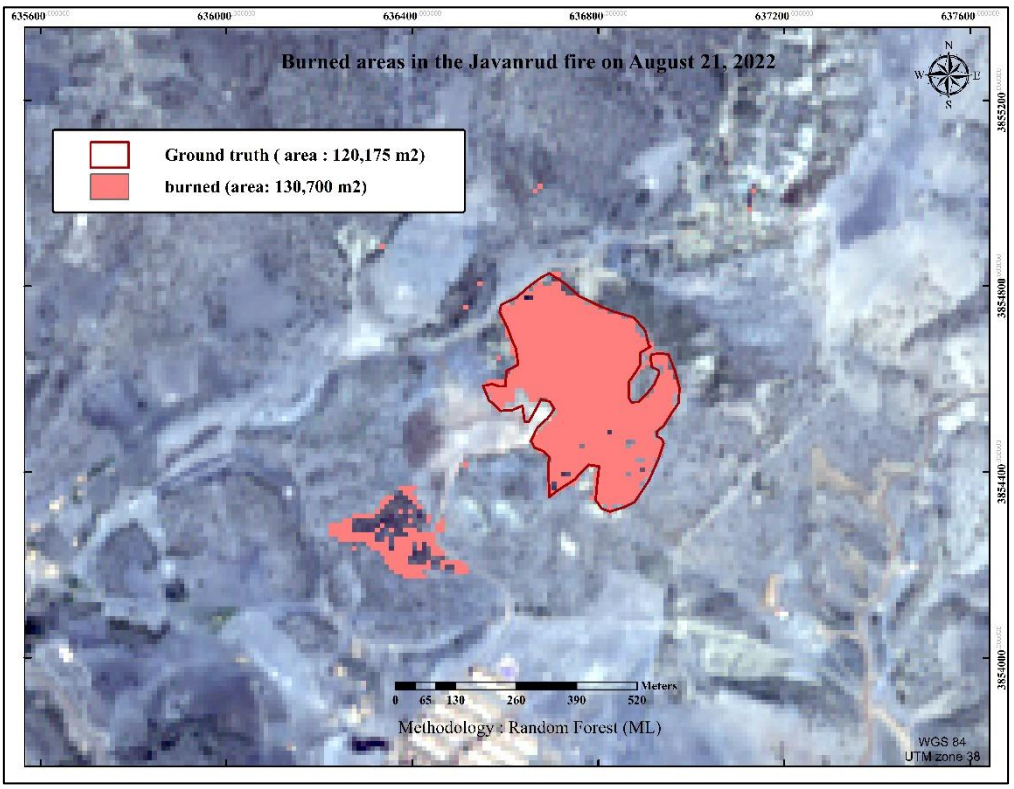

Figure 5- Burned area in JAVANRUD

While the Discussion section focused on detailed error analysis and performance interpretation, the emphasis here is on recommendations for future operational improvements:
(i) optimizing the decision threshold to align with specific mission objectives (e.g., prioritizing precision versus recall); (ii) enforcing spatial independence in training and testing splits to ensure fairer performance evaluation in regional-scale applications; and (iii) reporting accuracy metrics alongside concise uncertainty notes related to cloud/shadow masking and phenological variability.
The dNBR-centric RF approach is transferable to other wildfire events in the Zagros foothills with minimal retuning, as it relies on SWIR–NIR change signals that generalize well across landscapes. Under challenging conditions—such as heavy smoke, cloud cover, or steep illumination gradients—modest enhancements, including terrain and illumination normalization or Sentinel-1 data fusion, can further stabilize performance.
In summary, the proposed RF-based pipeline represents a robust operational baseline: fast to deploy, easy to audit, and readily integrable into regional post-fire assessment workflows.

8. References

Gündüz, H. İ., Torun, A. T., Gezgin, C., 2025. Post-Fire Burned Area Detection Using Machine Learning and Burn Severity Classification with Spectral Indices in İzmir: A SHAP-Driven XAI Approach. *Fire*, 8(4), 121. https://doi.org/10.3390/fire8040121

Lee, C., Park, S., Kim, T., Liu, S., Md Reba, M. N., Oh, J., Han, Y., 2022. Machine Learning-Based Forest Burned Area Detection with Various Input Variables: A Case Study of South Korea. *Applied Sciences*, 12(19), 10077. https://doi.org/10.3390/app121910077

Rege Cambrin, D., Colomba, L., Garza, P., 2025. Magnifier: A Multi-grained Neural Network-based Architecture for Burned Area Delineation. *arXiv preprint*. arXiv:2504.1958. https://doi.org/10.48550/arXiv.2504.1958

Zhao, Y., Ban, Y., 2024. RADARSAT Constellation Mission Compact Polarisation SAR Data for Burned Area Mapping with Deep Learning. *arXiv preprint*. arXiv:2412.11561. https://doi.org/10.48550/arXiv.2412.11561