Metric Learning Approach to Mineral Mapping Using Multi-Sensor Satellite Data

Katarzyna Jabłońska[1,2][0009−0006−6824−9437] and Maciej Zięba[1,3][0000−0003−4217−7712]

[1] Wrocław Univeristy od Science and Technology, Department of Artificial Intelligence Poland, Wrocław, Wybrzeże Wyspiańskiego 27
[2] Tooploox, Poland, Wrocław, Tęczowa 7
[3] TerraEye, Poland, Wrocław, Jana Długosza 60a

Abstract. We present a metric learning approach for mineral classification using a merging of hyperspectral and multispectral satellite data. Our method integrates spectral information from multiple providers, including EMIT, EnMAP, PRISMA, Sentinel-2 , and WorldView-3 in unified embeddings. To reflect realistic operational conditions, we introduce a training-time augmentation strategy in which a random subset of providers is masked, forcing the network to learn robust representations that generalize to incomplete sensor coverage. Trained on an augmented dataset derived from the USGS spectral library, the proposed network significantly outperforms classical one-class methods such as Isolation Forest and OCSVM in terms of F1-score, while maintaining competitive recall. Analysis of missing providers demonstrates that embeddings remain stable when only a subset of sensors is available, particularly when multispectral data are missing, whereas missing multiple hyperspectral providers has a substantial impact. The most robust embeddings are achieved when all hyperspectral providers are present, often in combination with WorldView-3. These results highlight the importance of hyperspectral data in discriminative mineral mapping. In general, our approach provides a flexible and effective framework for operational mineral classification under realistic multisensor conditions.

Keywords: Satellite data · Metric Networks · Data Fusion

1 Introduction

Spectral remote sensing, which encompasses both hyperspectral and multispectral imaging, enables a detailed characterization of surface materials based on their reflectance properties across multiple wavelengths. Multispectral and hyperspectral imaging systems offer complementary capabilities for Earth observation and material discrimination. Multispectral sensors, such as Sentinel-2 [1] and WorldView-3 [2], capture reflectance in a limited number of broad spectral bands, allowing large-scale monitoring of vegetation [3], water bodies [4] and land use [5] with high spatial coverage.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

Hyperspectral sensors, including EnMAP [6], PRISMA [7], and EMIT [8], record data in tens to hundreds of narrow contiguous bands, allowing detailed identification of surface materials based on their diagnostic spectral features. These combined modalities have found broad applications in agriculture [9], environmental monitoring [10], defense [11], and, most notably, in geological exploration and mineral mapping [12].

Despite significant progress, large-scale mineral detection using spectral satellite data remains challenging because of the heterogeneous nature of the available sensors. Hyperspectral missions provide rich spectral detail but limited spatial and temporal coverage, whereas multispectral systems offer frequent and wide-area observations but at coarser spectral resolution. As a result, no single provider can guarantee full data availability for a given area of interest, and models trained on one sensor type often fail to generalize to others.

Most existing machine learning methods are designed for a specific data provider, requiring sensor-specific preprocessing, retraining, or feature adaptation. This approach limits operational scalability and robustness, particularly in real-world scenarios where data from some providers are missing or incomplete. Furthermore, current fusion strategies typically rely on fixed sensor combinations and cannot adapt dynamically to partial multi-sensor coverage.

In this study, we address these limitations by introducing a unified metric learning framework that is capable of jointly leveraging data from multiple hyperspectral and multispectral providers. We concatenate spectral bands from EMIT, EnMAP, PRISMA, Sentinel-2, and WorldView-3 into a single feature vector, allowing the network to learn a shared spectral representation across heterogeneous sources. To simulate realistic observation conditions, we introduce a *random sensor masking* strategy in which a random number of providers are omitted during training. This forces the model to learn inter-provider relationships and remain robust when only partial data are available.

2 Related Works

2.1 Introduction to the Domain

Machine learning for mineral detection using remote sensing data has gained significant attention in recent years [13]. Hyperspectral data provide detailed spectral information that is crucial for identifying and mapping minerals. One of the key steps in this process is the extraction of meaningful features from raw spectral data, which greatly influences the performance of machine learning models [14].

2.2 Common Feature Extraction Methods

In most studies, hyperspectral data is obtained from a single sensor [15], and the analysis typically begins with atmospheric correction, which is often performed at the provider's site. After preprocessing, feature extraction is applied

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

to reduce dimensionality and highlight informative spectral patterns. The widely used feature extraction methods include spectral indices such as the Normalized Difference Vegetation Index (NDVI), the Soil Adjusted Vegetation Index (SAVI), and the Principal Endmember Analysis indices (PEAs) ([16], [17], [18], Component Analysis (PCA) ([19], [20], [21], [22]), Independent Component Analysis (ICA) ([23], [24]), Minimum Noise Fraction (MNF) combined with continuum removal [25], Sequential Maximum Angle Convex Cone (SMACC) [26] and Uniform Manifold Approximation and Projection (UMAP)([27], [28]).

In nearly all cases, the input to these algorithms is a vector of values derived from spectral bands of single sensors such as EMIT, EnMap or PRISMA.

2.3 Advanced and Data Fusion Approaches

Some studies have explored more advanced machine learning techniques and data fusion approaches [29]. These works often aim to combine multiple types of sensors or employ complex models to improve classification accuracy. For example, several studies have focused on merging LiDAR data with hyperspectral imagery ([30], [31], [32]). Other research has explored the fusion of optical and hyperspectral data. For example, Sentinel-2 multispectral imagery has been fused with PRISMA and EnMAP hyperspectral data to enhance the spatial resolution of hyperspectral products [33]. Transformer-based fusion approaches have also been proposed for combining Sentinel-2 and Sentinel-3 data using synthetic EnMAP-derived training samples [34]. These studies highlight the growing interest in heterogeneous multi-sensor data fusion but generally assume complete data availability or focus primarily on spatial–spectral resolution enhancement. In contrast, our method merges hyperspectral vectors from multiple providers and explicitly accounts for missing sensors by applying a random-provider dropout strategy during training, better reflecting realistic incomplete coverage scenarios.

2.4 Gap and Our Contribution

Although many studies rely on individual satellite datasets and relatively simple feature extraction methods, there is a lack of research that integrates multiple sources into a unified vector for machine learning-based mineral detection. In practice, hyperspectral coverage is often incomplete: some areas may only have observations from one or two providers due to cloud cover, acquisition schedules, or sensor limitations. Existing fusion approaches generally assume full coverage and cannot handle these partial observations effectively.

The present work addresses this gap by creating a merged feature vector from different providers, combined with a random-provider dropout training strategy. This enables robust and informative input even when only a subset of sensors is available in a given area. Importantly, our approach eliminates the need to train multiple models for different provider combinations or to design specialized fusion methodologies for each scenario; a single model is capable of handling any combination of available providers, simplifying deployment, and improving scalability.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

3 Methodology

3.1 Overview

The proposed approach learns a unified spectral embedding space that remains robust across multiple satellite sensors with differing spectral resolutions. For each sample, reflectance data from five sensors, EMIT, EnMAP, PRISMA, Sentinel-2, and WorldView–3, are concatenated into a single feature vector.

$$\boldsymbol{p} = [\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3, \boldsymbol{p}_4, \boldsymbol{p}_5],$$

where p_i denotes the sub-vector corresponding to the i-th provider. This unified representation enables learning of a sensor-invariant embedding function

$$f_\theta : R^{499} \to R^d,$$

which projects each spectrum into a latent space where samples of the same mineral are closer together than those of different minerals.

3.2 Random Sensor Masking During Training

To improve the model's ability to handle incomplete or heterogeneous observations from multiple providers, we introduce Random Sensor Masking as a stochastic augmentation step. For each input vector p, a random subset of sensors is masked to simulate missing data. Specifically, we first sample the number of missing sensors

$$n_{\text{missing}} \sim \text{Uniform}\{0, \text{max\_missing}\},$$

then randomly select the corresponding providers from the set of all five sensors. All spectral bands from the selected sensors are replaced with zeros:

$$\tilde{\boldsymbol{p}} = \boldsymbol{m} \odot \boldsymbol{p},$$

where the binary mask $m = [m_1, m_2, m_3, m_4, m_5]$ is defined as $m_i = 0$ if the i-th sensor is missing and 1 otherwise. This augmentation simulates realistic multi-sensor scenarios in which data from all providers are rarely available simultaneously.

3.3 Metric Learning Model Training

The model is trained using a learning objective metric based on triplet loss [35]. Each triplet consists of an anchor spectrum $\boldsymbol{x}_i$, a positive sample $\boldsymbol{x}_i^+$ from the same mineral class, and a negative sample $\boldsymbol{x}_i^-$ from a different class. After applying random sensor masking, their embeddings are calculated as $f_\theta(\tilde{\boldsymbol{x}}_i)$, $f_\theta(\tilde{\boldsymbol{x}}_i^+)$, and $f_\theta(\tilde{\boldsymbol{x}}_i^-)$, where $f_\theta : R^{499} \to R^p$ is the embedding function parameterized by the network weights $\boldsymbol{\theta}$.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

The pairwise Euclidean distances are defined as

$$D_{i,i^+} = \left\| f_\theta(\tilde{\boldsymbol{x}}_i) - f_\theta(\tilde{\boldsymbol{x}}_i^+) \right\|_2, \quad D_{i,i^-} = \left\| f_\theta(\tilde{\boldsymbol{x}}_i) - f_\theta(\tilde{\boldsymbol{x}}_i^-) \right\|_2.$$

The triplet loss function is expressed as

$$\mathcal{L}_{\text{triplet}} = \frac{1}{N_{\text{triplets}}} \sum_i \max\left(0, D_{i,i^+} - D_{i,i^-} + m\right),$$

where $m > 0$ is a margin that enforces a minimum separation between dissimilar samples in the embedding space.

3.4 Embedding-Based Classification

After training, the embedding function $f_\theta(\cdot)$ is used for mineral identification by comparing query embeddings to reference embeddings derived from full spectra. Triplet training is performed on complete spectral vectors, ensuring that class centroids represent full mineral signatures. During inference, a query spectrum $\boldsymbol{p}$ may include only a subset of sensors; missing providers are zero-filled using the same masking rule:

$$\tilde{\boldsymbol{p}} = \boldsymbol{m} \odot \boldsymbol{p}.$$

The embedding $f_\theta(\tilde{\boldsymbol{p}})$ is then compared to the class centroids computed from the full spectra:

$$\hat{y} = \arg\min_c D\left(f_\theta(\tilde{\boldsymbol{p}}), \boldsymbol{\mu}_c\right),$$

where $\boldsymbol{\mu}_c$ is the mean embedding of class c, and $D(\cdot, \cdot)$ denotes the Euclidean distance. This enables robust mineral identification even with partial multi-sensor coverage.

3.5 Data source

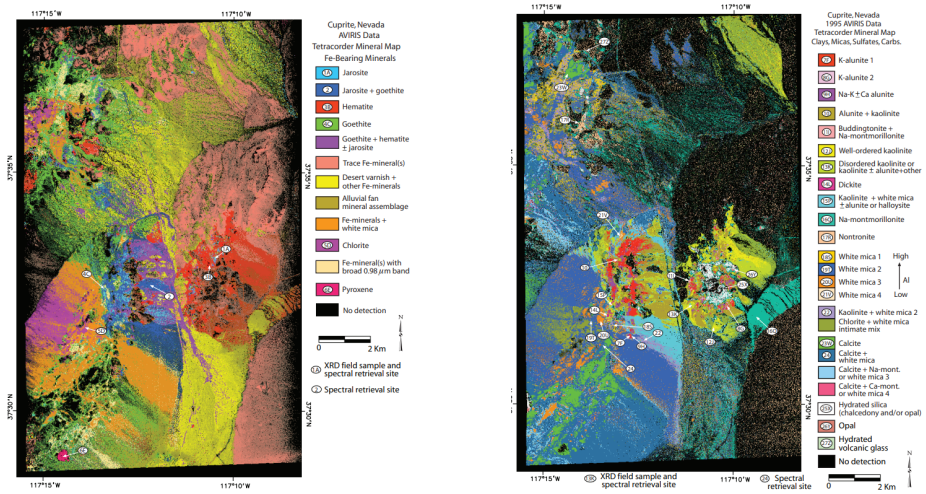
Satellite data We considered five satellite data providers commonly used for mineral detection: EMIT, EnMAP, PRISMA, Sentinel-2, and WorldView-3 [36]. EMIT, EnMAP, and PRISMA are hyperspectral missions covering the visible to short-wave infrared (VSWIR) range (approximately 400–2,500 nm), while Sentinel-2 and WorldView-3 provide multispectral observations.

EMIT, deployed on the International Space Station, acquires 285 spectral bands at 60 m spatial resolution. EnMAP and PRISMA provide hyperspectral imagery with 246 and 239 bands, respectively, both at 30 m resolution. Sentinel-2 offers 13 multispectral bands with spatial resolutions of 10–60 m, and WorldView-3 provides high-resolution multispectral and short-wave infrared data with spatial resolutions down to 1.24 m (multispectral) and 3.7 m (SWIR).

For validation, we used the Cuprite mining district (Nevada, USA), a well-studied hyperspectral benchmark with extensive hydrothermal alteration. Ground truth (GT) data were derived from published mineral maps [37] and vectorized

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

into geospatial polygons for evaluation. These maps are based on field observations, laboratory spectroscopy, and airborne imaging spectrometry, providing reliable reference labels. The GT polygons capture the spectral complexity of Cuprite, where multiple alteration minerals frequently co-occur within a single pixel, making them a realistic and challenging benchmark for algorithm assessment (Figures 1a and 1b).



(a) Map 1 (b) Map 2

Fig. 1: Ground-truth mineral maps of the Cuprite mining district (Nevada, USA), derived from [37].

USGS Spectral Library We used mineral reflectance spectra from the United States Geological Survey (USGS) spectral library [39] as the primary reference dataset. The library contains high-quality laboratory measurements of minerals, rocks, soils, vegetation, and man-made materials across the ultraviolet to thermal infrared range. Because of its standardized acquisition protocols and extensive spectral coverage, the USGS library serves as a reliable reference for hyperspectral analysis. In this study, ground-truth spectra are provided for training and evaluating the proposed embedding network, as well as for simulating sensor-specific responses in the multi-provider dataset.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

4 Experiments

4.1 Experimental Setup

The experiments aimed to evaluate the effectiveness of metric learning networks in embedding hyperspectral mineral spectra into a feature space where the spectra of the same mineral are closer together while being separated from different minerals. This is particularly important for mineral identification based on similarity, especially when labeled data are limited.

A synthetic data set was generated based on real spectral signatures of seven minerals commonly associated with porphyrical copper deposits (PCD): alunite, chlorite, calcite, goethite, hematite, jarosite, and kaolinite. The synthetic spectra incorporate realistic variations due to illumination, atmospheric effects, sensor noise, and mineralogical heterogeneity, enhancing the robustness of the learned representations.

All spectral data from different providers were resampled to a spatial resolution of 30 meters during the inference and validation steps, ensuring consistency in the EMIT, EnMAP, PRISMA, Sentinel-2 and WorldView-3 data. This allowed a fair comparison against the ground truth Cuprite maps.

Unless stated otherwise, the dataset construction procedure, data augmentation strategy, network architectures, loss functions, and training hyperparameters followed exactly the configuration described in [38].

4.2 Data preparation

We utilized spectra from the USGS Spectral Library to generate synthetic data, selecting pure mineral spectra rather than mixtures. These spectra were obtained using the Beckman 5270 high-resolution spectrometer, covering the 0.2-3 μm range [39]. As an initial step, we resampled the spectra to match the wavelength requirements of each satellite dataset, specifically the spectral ranges of EMIT, EnMAP, PRISMA, Sentinel-2, and WorldView-3. To mitigate the effects of atmospheric absorption and water characteristics ScienceDirect2025, Clark1999, we excluded bands in wavelength ranges 300–400, 910–1000, 1120–1190, 1340–1550, 1750-2050 and 2400–2500 nm.

The preprocessed spectra were concatenated into a single feature vector per sample, with the following band counts: EMIT 171, EnMAP 159, PRISMA 141, Sentinel-2 12, and WorldView-3 16, resulting in 499 bands per vector. Following our previous methodology, these vectors were enhanced using spectral transformations to simulate measurement variability and sensor noise, and then split into training, validation, and test sets. This augmented dataset provides diverse examples for robust model training while maintaining separate partitions for unbiased evaluation.

4.3 Random Sensor Masking

To simulate realistic multi-provider scenarios, we applied a training-time augmentation strategy in which a random subset of providers is selected for each

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

sample, and the bands corresponding to unselected providers are set to zero. This approach ensures that the network learns robust embeddings that can handle partial sensor coverage, mimicking real-world conditions where data from all sensors are rarely available simultaneously.

4.4 Network Configurations and Training

In our previous work, we systematically compared several shallow fully connected (FC) architectures for mineral spectra embedding and found that a two-layer network without dropout consistently achieved the best results across all evaluation metrics and test scenarios. Based on these findings, we adopt this configuration as the sole architecture in the present study.

The network consists of two linear layers with ReLU activations, mapping the input spectra (*input_bands*) to a fixed-dimensional latent representation (*embedding_size*). This simple yet effective design leverages the one-dimensional nature of spectral signals, where spatial information is absent, and global spectral relationships can be effectively captured without convolutional or sequence-based modeling.

Each FC block transforms the high-dimensional input spectra into a compact embedding space, where spectral similarity can be assessed directly in Euclidean distance. The resulting embeddings are optimized using Triplet loss, which enforces relative distance constraints among an anchor, a positive, and a negative sample. In our formulation, the anchor corresponds to a real spectrum from the USGS spectral library, the positive is a synthetically augmented version of the same spectrum, and the negative belongs to a different mineral class. This objective encourages spectra of the same mineral to cluster tightly in the latent space while pushing dissimilar classes apart, promoting robust and discriminative representations across sensors and acquisition conditions.

4.5 Evaluation Metrics

The performance of the model was assessed using the ground-truth mineral annotations described in Section 3.2. Since the metric learning networks output embeddings rather than class labels, we applied distance-based classification: for each pixel, the Euclidean distance between its embedding and the reference USGS mineral spectrum embedding was computed. Pixels with distances below a selected threshold were assigned to the corresponding mineral class. The threshold was chosen to maximize the F-score. We report standard classification metrics:

Recall Fraction of correctly identified mineral pixels among all true pixels; high values indicate that most true occurrences are detected.

Precision Fraction of correctly identified pixels among all predicted pixels; high values indicate few false positives.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

F-score Harmonic means of precision and recall, providing a balanced performance measure:

$$F = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}.$$

4.6 Missing-Provider Analysis

To assess robustness to incomplete data, we formed multisource spectral vectors by concatenating bands from all providers and systematically masked one or more providers by setting their corresponding bands to zero. Each masked spectrum was passed through the trained Siamese network and compared with the complete-spectrum embedding.

Similarity was quantified using cosine similarity ($\text{Sim}_{\cos}$), measuring angular consistency, and Euclidean distance (D_{Euc}), measuring the absolute embedding shift. This analysis was repeated for all combinations of missing providers to identify stable configurations with minimal embedding changes and critical providers whose absence caused substantial embedding drift.

4.7 Computational Complexity and Implementation Details

The proposed model is computationally efficient due to its simple fully connected architecture and was trained on a consumer-grade device with an Apple M4 chip. It consists of two fully connected layers operating on a 499-dimensional input, resulting in a low parameter count compared to deep architectures. Inference requires only a forward pass and a distance computation in the embedding space, scaling linearly with the number of reference embeddings. Although training is more demanding than classical methods (e.g., OCSVM, Isolation Forest) due to augmentation and triplet sampling, inference is significantly faster and does not require retraining for different sensor configurations.

5 Results

Table 1 presents the performance of the one-class mineral classification of the proposed metric learning network with random sensor masking, compared to the classical one-class approaches [36] and a previously proposed metric learning baseline without masking [38]. While the classical methods (OCSVM and Isolation Forest) were trained using five-fold augmented datasets, both metric learning models were trained on substantially larger, identically augmented datasets, allowing a direct assessment of the effect of sensor masking.

For Alunite, the masked FC_Baseline network achieved a recall of 0.664 and an F1-score of 0.474, substantially outperforming the classical one-class methods in terms of F1-score. Compared to the unmasked metric learning baseline, which achieved an F1-score of 0.353, the proposed masking strategy yields a notable improvement in classification balance. Although recall is moderately reduced, the higher F1-score indicates improved precision and more robust decision boundaries.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

A similar trend is observed for Jarosite. The masked FC_Baseline network achieves an F1-score of 0.437, outperforming both classical methods and the unmasked metric learning baseline (F1-score of 0.302). In contrast to Alunite, recall for Jarosite also increases with masking, suggesting enhanced generalization of the learned embeddings for this mineral class.

In general, these results demonstrate that the incorporation of random sensor masking into the metric learning framework consistently improves the performance of the one-class mineral classification, particularly in terms of balanced precision recall behavior, as reflected by higher F1-scores.

To further evaluate the robustness of the learned embeddings under realistic multi-sensor conditions, a missing-provider analysis was conducted for Jarosite and Alunite. Complete multisource spectral vectors were systematically masked by removing one or more satellite providers, and the resulting embeddings were compared to full-spectrum embeddings using cosine similarity and Euclidean distance (Table 2).

Table 2 summarizes the impact of missing providers on embedding stability. For clarity, we report the five most and least affected configurations on the basis of cosine similarity and Euclidean distance. Missing providers are listed explicitly, where combinations including hyperspectral sensors (EMIT, EnMAP, PRISMA) generally lead to larger deviations.

Table 1: Comparison of one-class mineral classification performance across classical one-class methods, prior metric learning network (MLN) approaches [36,38], and the proposed metric learning network with masking (this work).

Mineral	Method	Dataset Size	Recall	F1-Score
Alunite	OCSVM (5-aug) [36]	66,723	0.766	0.122
	Isolation Forest (5-aug) [36]	66,723	0.786	0.253
	MLN FC_Baseline (7-aug) [38]	648,000	0.750	0.353
	MLN FC_Baseline (masking, this work)	648,000	0.664	**0.474**
Jarosite	OCSVM (5-aug) [36]	58,748	0.871	0.212
	Isolation Forest (5-aug) [36]	58,748	0.884	0.228
	MLN FC_Baseline (7-aug) [38]	378,000	0.301	0.302
	MLN FC_Baseline (masking, this work)	378,000	0.448	**0.437**

The results in Table 2 show a clear asymmetry between multispectral and hyperspectral providers. The absence of multispectral sensors (Sentinel-2, WorldView-3) results in only minor changes in the embedding space, as indicated by consistently high cosine similarity (>0.999) and low Euclidean distance. In contrast, removing hyperspectral providers (EMIT, EnMAP, PRISMA) leads to significantly larger deviations, especially when multiple hyperspectral sources are missing simultaneously. This shows that hyperspectral data contribute the most discriminative information to the learned representations.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

Table 2: Effect of missing satellite providers on embedding similarity and distance for Jarosite and Alunite minerals

Jarosite			Alunite		
Metric Type	**Missing Providers**	**Value**	**Metric Type**	**Missing Providers**	**Value**
Top 5 Most similar Embedding by Cosine Similarity(CosSim)					
CosSim	Wv3	0.9999	CosSim	PRISMA	0.9997
CosSim	Sentinel-2	0.9998	CosSim	Sentinel-2	0.9996
CosSim	PRISMA, Wv3	0.9995	CosSim	EnMap, Wv3	0.9996
CosSim	EnMap	0.9995	CosSim	Wv3	0.9996
CosSim	EMIT, Sentinel-2	0.9995	CosSim	EnMap	0.9993
Top 5 Disimilar Embeddings by Cosine Similarity (CosSim)					
CosSim	EMIT, EnMap, PRISMA, Wv3	0.5671	CosSim	EMIT, EnMap, PRISMA, Wv3	0.8816
CosSim	EMIT, EnMap, PRISMA	0.8004	CosSim	EMIT, EnMap, PRISMA, Sentinel-2	0.9173
CosSim	EMIT, EnMap, PRISMA, Sentinel-2	0.9204	CosSim	EMIT, EnMap, PRISMA	0.9236
CosSim	EMIT, PRISMA	0.9923	CosSim	EnMap, PRISMA, Sentinel-2, Wv3	0.9926
CosSim	EnMap, PRISMA, Sentinel-2, Wv3	0.9936	CosSim	EMIT, EnMap	0.9937
Top 5 Most Disimilar Embeddings by Euclidean Distance (EucDist)					
EucDist	EMIT, EnMap, PRISMA, Wv3	0.9305	EucDist	EMIT, EnMap, PRISMA, Wv3	0.4865
EucDist	EMIT, EnMap, PRISMA	0.6318	EucDist	EMIT, EnMap, PRISMA, Sentinel-2	0.4068
EucDist	EMIT, EnMap, PRISMA, Sentinel-2	0.3991	EucDist	EMIT, EnMap, PRISMA	0.3908
EucDist	EMIT, PRISMA	0.1240	EucDist	EnMap, PRISMA, Sentinel-2, Wv3	0.1213
EucDist	EnMap, PRISMA, Sentinel-2, Wv3	0.1127	EucDist	EMIT, EnMap	0.1126
Top 5 Most Similar Embeddings by Euclidean Distance (EucDist)					
EucDist	Wv3	0.0137	EucDist	PRISMA	0.0234
EucDist	Sentinel-2	0.0220	EucDist	Sentinel-2	0.0275
EucDist	PRISMA, Wv3	0.0305	EucDist	EnMap, Wv3	0.0293
EucDist	EnMap	0.0306	EucDist	Wv3	0.0295
EucDist	EMIT, Sentinel-2	0.0331	EucDist	EnMap	0.0375

6 Conclusions

In this study, we investigated the use of a metric learning network for mineral classification based on multisource hyperspectral and multispectral data. Our results demonstrate that incorporating hyperspectral providers (EMIT, EnMAP, PRISMA) significantly improves classification performance, highlighting the value of hyperspectral data. The network remains functional even when only a single data provider is available, indicating robustness to incomplete multisensor coverage.

The absence of a single provider generally has a minor effect on embedding quality, particularly when the missing sensor is multispectral (WV3 or Sentinel-2). In contrast, the absence of multiple providers, especially hyperspectral ones, has a substantial impact on embeddings. For both Jarosite and Alunite, the five most affected cases always include at least two hyperspectral providers, while the most stable embeddings are achieved when all hyperspectral providers are present along with WV3.

These findings suggest that future work could explore models, such as autoencoders, capable of reconstructing missing sensor data based on the available providers, further enhancing robustness to incomplete datasets. In general, our metric learning approach effectively leverages multisource spectral data, produc-

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

ing embeddings that generalize across varying sensor availability and providing a foundation for operational mineral mapping under realistic data conditions.

Additionally, the proposed method demonstrates favorable scalability properties. Scalability of the proposed method depends on the dataset size and the number of reference embeddings. Due to the fixed-length input and shallow architecture, both training and inference scale linearly with the number of samples. For large-scale datasets, the method can be applied in a tile-based manner, with embeddings computed independently per pixel.

Acknowledgments. This work was financially supported by the Ministry of Education and Science. Katarzyna Jablonska reports that financial support for this research was provided by the Ministry of Education and Science.

Disclosure of Interests. The authors declare the following competing interests: Katarzyna Jablonska reports financial support was provided by the Ministry of Education and Science. The other author(s) declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. European Space Agency (ESA): Sentinel-2, Copernicus Earth Observation Program. Available at: `https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2`
2. Maxar Technologies: WorldView-3, commercial high-resolution satellite. Available at: `https://www.maxar.com/products/worldview-3`
3. M. F. Aslan, K. Sabanci, and B. Aslan, "Artificial Intelligence Techniques in Crop Yield Estimation Based on Sentinel-2 Data: A Comprehensive Survey," *Sustainability*, vol. 16, p. 8277, 2024, doi: 10.3390/su16188277. [Online]. Available: `https://doi.org/10.3390/su16188277`
4. J. Llodrà-Llabrés, J. Martínez-López, T. Postma, C. Pérez-Martínez, and D. Alcaraz-Segura, "Retrieving water chlorophyll-a concentration in inland waters from Sentinel-2 imagery: Review of operability, performance and ways forward," *International Journal of Applied Earth Observation and Geoinformation*, vol. 125, p. 103605, 2023, doi: 10.1016/j.jag.2023.103605. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1569843223004296`
5. A. Kumari and S. Karthikeyan, "Sentinel-2 Data for Land Use/Land Cover Mapping: A Meta-analysis and Review," *SN Computer Science*, vol. 4, no. 6, p. 815, 2023, doi: 10.1007/s42979-023-02214-0. [Online]. Available: `https://doi.org/10.1007/s42979-023-02214-0`
6. DLR/EOC: EnMap – Environmental Mapping and Analysis Program, German Aerospace Center (DLR), Earth Observation Center (EOC). Available at: `https://www.enmap.org`
7. ASI: PRISMA – PRecursore IperSpettrale della Missione Applicativa, Italian Space Agency (ASI), Earth Observation Program. Available at: `https://www.asi.it/en/prisma`
8. NASA/JPL-Caltech: EMIT – Earth Surface Mineral Dust Source Investigation, NASA Jet Propulsion Laboratory (JPL), International Space Station. Available at: `https://emit.earth`

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

9. C. Wang, B. Liu, L. Liu, Y. Zhu, J. Hou, P. Liu, and X. Li, "A review of deep learning used in the hyperspectral image analysis for agriculture," *Artificial Intelligence Review*, vol. 54, no. 7, pp. 5205–5253, 2021, doi: 10.1007/s10462-021-10018-y. [Online]. Available: `https://doi.org/10.1007/s10462-021-10018-y`
10. G. Tejasree and L. Agilandeeswari, "An extensive review of hyperspectral image classification and prediction: techniques and challenges," *Multimedia Tools and Applications*, vol. 83, no. 34, pp. 80941–81038, 2024, doi: 10.1007/s11042-024-18562-9. [Online]. Available: `https://doi.org/10.1007/s11042-024-18562-9`
11. M. Shimoni, R. Haelterman, and C. Perneel, "Hyperspectral Imaging for Military and Security Applications: Combining Myriad Processing and Sensing Techniques," *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, no. 2, pp. 101–117, June 2019, doi: 10.1109/MGRS.2019.2902525.
12. S. Peyghambari and Y. Zhang, "Hyperspectral remote sensing in lithological mapping, mineral exploration, and environmental geology: An updated review," *Journal of Applied Remote Sensing*, vol. 15, 2021, doi: 10.1117/1.JRS.15.031501. [Online]. Available: `https://doi.org/10.1117/1.JRS.15.031501`
13. E. Bedini, "The Use of Hyperspectral Remote Sensing for Mineral Exploration: A Review," *Journal of Hyperspectral Remote Sensing*, vol. 7, no. 4, pp. 189-211, 2017. Available: `https://www.researchgate.net/publication/363465665_The_use_of_hyperspectral_remote_sensing_for_mineral_exploration_a_review`
14. S. Thiruchittampalam, B. P. Banerjee, N. F. Glenn, and S. Raval, "A Systematic Review of Machine Learning-Based Remote Sensing Data Analysis for Geological and Mined Materials Characterisation," *European Journal of Remote Sensing*, vol. 58, no. 1, pp. 2524622, 2025. doi:10.1080/22797254.2025.2524622. Available: https://doi.org/10.1080/22797254.2025.2524622
15. Hajaj, S., El Harti, A., Beiranvand Pour, A., Jellouli, A., Adiri, Z., Hashim, M.: A review on hyperspectral imagery application for lithological mapping and mineral prospecting: Machine learning techniques and future prospects. *Remote Sensing Applications: Society and Environment* **35**, 101218 (2024).
16. Habashi, J., Mohammady Oskouei, M., Jamshid Moghadam, H., Beiranvand Pour, A.: Optimizing alteration mineral detection: A fusion of multispectral and hyperspectral remote sensing techniques in the Sar-e-Chah-e Shur, Iran. *Remote Sensing Applications: Society and Environment* **35**, 101249 (2024)
17. Kubo, T., Gonnokami, H., Hawu Hede, A.N., Koike, K.: Combining vegetation index with mineral identification for detection of high-geothermal-potential zones using hyperspectral satellite data. *Geothermics* **125**, 103194 (2025)
18. Shirkhani, M., Ghaderi, M., Rashidnejad-Omran, N.: Hyperspectral vs. multispectral remote sensing: ASTER and PRISMA performance in mapping IOA deposits in the Bafq mining district, Iran. *Ore Geology Reviews* **186**, 106923 (2025)
19. Abdelkader, M.A., Watanabe, Y., Shebl, A., El-Dokouny, H.A., Dawoud, M., Csámer, Á.: Effective delineation of rare metal-bearing granites from remote sensing data using machine learning methods: A case study from the Umm Naggat Area, Central Eastern Desert, Egypt. *Ore Geology Reviews* **150**, 105184 (2022)
20. Khedr, M.Z., Al Desouky, A.A., Kamh, S., Hauzenberger, C., Arai, S., Tamura, A., Whattam, S.A., Morishita, T., Lasheen, E.S., El-Awady, A.: Petrogenesis of Gerf Neoproterozoic carbonatized peridotites (Egypt): Evidence of convergent margin metasomatism of depleted sub-arc mantle. *Lithos* **450–451**, 107192 (2023)
21. Balázs, B., Bíró, T., Dyke, G., Singh, S.K., Szabó, S.: Extracting water-related features using reflectance data and principal component analysis of Landsat images. *Hydrological Sciences Journal* **63**(2), 269–284 (2018)

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

22. Abdelkader, M.A., Watanabe, Y., Shebl, A., El-Dokouny, H.A., Dawoud, M., Csámer, Á.: Effective delineation of rare metal-bearing granites from remote sensing data using machine learning methods: A case study from the Umm Naggat Area, Central Eastern Desert, Egypt. *Ore Geology Reviews* **150**, 105184 (2022)
23. Pour, A.B., Park, Y., Park, T.Y.S., Hong, J.K., Hashim, M., Woo, J., Ayoobi, I.: Evaluation of ICA and CEM algorithms with Landsat-8/ASTER data for geological mapping in inaccessible regions. *Geocarto International* **34**(7), 785–816 (2018)
24. Pour, A.B., Park, T.Y.S., Park, Y., Hong, J.K., Pradhan, B.: Fusion of DPCA and ICA algorithms for mineral detection using Landsat-8 spectral bands. In: *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 6067–6070 (2019)
25. Mondal, S., Guha, A., Pal, S.K.: Support vector machine-based integration of AVIRIS NG hyperspectral and ground geophysical data for identifying potential zones for chromite exploration – A study in Tamil Nadu, India. *Advances in Space Research* **73**(2), 1475–1490 (2024)
26. Shebl, A., Abdellatif, M., Abriha, D., Dawoud, M., Ali, M.A.H., Mahmoud, A.S., Kristály, F., Csámer, Á.: EnMap hyperspectral data in geological investigations: Evaluation for lithological and hydrothermal alteration mapping in Neoproterozoic rocks. *Gondwana Research* **143**, 91–124 (2025)
27. Myasnikov, E.: Using UMAP for Dimensionality Reduction of Hyperspectral Data. In: *2020 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon)*, pp. 1–5 (2020)
28. Hochstuhl, S., Weidner, U., Hammer, H., Thiele, A.: Evaluating uniform manifold approximation and projection for dimension reduction and visualization of PolInSAR features. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **V-1-2021**, 39–46 (2021)
29. He, J., et al. (2024). A Critical Review on Multi-Sensor and Multi-Platform Remote Sensing Data Fusion Approaches: Current Status and Prospects. *Remote Sensing*, 16(12), 2089. Available at: `https://www.mdpi.com/2072-4292/16/12/2089`
30. Wang, X., Dai, S., Wu, C., Iwahori, Y. (2023). Classification of Hyperspectral and LiDAR Data Using Multi-Modal Transformer Cascaded Fusion Net. *Remote Sensing*, 15(17), 4142. Available at: `https://www.mdpi.com/2072-4292/15/17/4142`
31. Wang, Y., et al. (2024). Joint Classification of Hyperspectral Images and LiDAR Data Based on Dual-Branch Transformer. *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1–14. doi:10.1109/TGRS.2024.xxxxxx
32. Wang, X., Dai, S., Wu, C., Iwahori, Y. (2024). Multimodal Semantic Collaborative Classification for Hyperspectral Images and LiDAR Data. *Remote Sensing*, 16(16), 3082. Available at: `https://www.mdpi.com/2072-4292/16/16/3082`
33. Alparone, L., Baronti, S., Garzelli, A., et al. (2024). Spatial Resolution Enhancement of Satellite Hyperspectral Data via Nested Hypersharpening with Sentinel-2 Multispectral Data. *Journal of Remote Sensing*, 16(8), 2456–2472. Available at: `https://flore.unifi.it/handle/2158/1379513`
34. Cristille, J., et al. (2024). Earth Observation Multi-Spectral Image Fusion with Transformers for Sentinel-2 and Sentinel-3 Using Synthetic EnMAP-Derived Training Data. *Remote Sensing*, 16(16), 3107. Available at: `https://www.mdpi.com/2072-4292/16/16/3107`
35. F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 815–823, IEEE Computer Society, 2015. doi: 10.1109/CVPR.2015.7298682. Available:

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2

https://openaccess.thecvf.com/content_cvpr_2015/html/Schroff_FaceNet_A_Unified_2015_CVPR_paper.html (accessed Sep. 8, 2025).
36. K. Jabłońska, M. Zięba, and M. Jakubowicz, "Satellite Data Based One-Class Classifier Models with Augmentation for Efficient Mineral Exploration," *Remote Sensing Applications: Society and Environment*, vol. XX, pp. XX–XX, 2025. doi:10.1016/j.rsase.2025.101316. Available: https://doi.org/10.1016/j.rsase.2025.101316
37. G. A. Swayze, R. N. Clark, A. F. H. Goetz, K. E. Livo, G. N. Breit, F. A. Kruse, S. J. Sutley, L. W. Snee, H. A. Lowers, J. L. Post, R. E. Stoffregen, and R. P. Ashley, "Mapping Advanced Argillic Alteration at Cuprite, Nevada, Using Imaging Spectroscopy," U.S. Geological Survey Publication, 2008. Available: https://pubs.usgs.gov/publication/70196084 (Last accessed 2025-01-02)
38. K. Jabłońska, M. Zięba, and D. Tanajewski, "Metric Learning Networks for Accurate Mineral Detection with Different Augmentation Techniques," *Expert Systems with Applications*, vol. 304, p. 130692, 2026. Available: https://www.sciencedirect.com/science/article/pii/S0957417425043076 (doi: 10.1016/j.eswa.2025.130692)
39. Kokaly, R. F., Clark, R. N., Swayze, G. A., Livo, K. E., Hoefen, T. M., Pearson, N. C., Wise, R. A., Benzel, W. M., Lowers, H. A., Driscoll, R. L., & Klein, A. J. (2017). *USGS Spectral Library Version 7*. U.S. Geological Survey, Data Series 1035. https://doi.org/10.3133/ds1035

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_2