Iteration-Based Feature Selection Method for Optimizing Feature Retention in PolSAR Image Classification

Ali Abdollahi, Tayebe Managhebi *, Mohammad SaadatSeresht

School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran – (aliabdollahi, t.managhebi, msaadat@ut.ac.ir)

KEY WORDS: PolSAR classification, feature selection, correlation-based method, median filtering, machine learning

ABSTRACT:

Polarimetric Synthetic Aperture Radar (PolSAR) provides rich scattering information that is highly valuable for land-cover classification. However, two major challenges remain: redundancy among polarimetric features and the presence of salt-and-pepper noise in classification maps. In this study, we propose a novel PolSAR classification framework that integrates an iterative correlation-based feature selection strategy with a rank-based post-processing approach. The iterative method progressively eliminates the most redundant features while retaining complementary and informative descriptors, thus preserving a larger and more discriminative feature space compared to conventional one-shot elimination. To evaluate its effectiveness, four machine learning algorithms—K-nearest neighbours (KNN), support vector machine (SVM), random forest (RF), and extreme gradient boosting (XGB)—were applied to a Gaofen-3 PolSAR dataset acquired over San Francisco. The results show that the proposed feature selection approach consistently improves classification performance, with accuracy gains of up to 4% across classifiers. Furthermore, applying a median filter as a post-processing step significantly enhances spatial coherence, achieving accuracies as high as 0.99 for the XGB classifier. These findings confirm that the proposed framework effectively addresses both feature redundancy and spatial noise, leading to more reliable and robust PolSAR classification outcomes.

1. INTRODUCTION

Image classification is a fundamental and long-standing task in the field of remote sensing, where the goal is to assign semantic labels to pixels or regions of an image based on their measured characteristics. It serves both as an end product in thematic mapping and as an intermediate step to support higher-level analyses, including change detection, object extraction, and environmental monitoring. In many cases, classification is used as a mask to isolate specific land-cover types or physical phenomena, thereby reducing the complexity and computational cost of subsequent processing steps. The quality of this initial classification stage can have a direct impact on the accuracy and reliability of downstream tasks [1].

Synthetic Aperture Radar (SAR) has emerged as a key imaging technology for classification purposes because it can operate in all-weather conditions, day and night, and is sensitive to a variety of physical properties, such as surface geometry, roughness, and dielectric constant. Unlike optical systems, SAR does not rely on sunlight and is therefore unaffected by cloud cover, making it invaluable in regions with frequent adverse weather. A particularly powerful variant of SAR is polarimetric SAR (PolSAR), which records the scattering of microwave signals in multiple polarization channels (HH, HV, VH, VV). This capability provides richer information about the scattering mechanisms present in each resolution cell, allowing for more detailed and physically meaningful feature extraction [2].

PolSAR data analysis often relies on polarimetric decomposition techniques to transform raw scattering information into interpretable parameters. Classical approaches include the Pauli decomposition, which provides a qualitative visualization of scattering types; the Cloude–Pottier entropy/anisotropy/alpha (H/A/α) decomposition, which quantifies scattering randomness and type; and model-based decompositions such as Freeman–Durden and Yamaguchi, which separate surface, double-bounce, and volume scattering components. More recent methods extend these frameworks to account for complex targets, multi-bounce effects, or additional scattering mechanisms [2, 3]. These decompositions yield a set of features that can be used individually or in combination as the input space for classification algorithms.

Over the past two decades, the literature has reported a wide variety of PolSAR classification strategies. Early studies were dominated by statistical approaches, most notably the complex Wishart maximum likelihood classifier, which operates directly on the polarimetric covariance or coherency matrices under the assumption of a complex Gaussian distribution. While computationally efficient, these methods often underperform in heterogeneous areas where class distributions deviate from gaussianity [4]. In response, machine learning methods such as support vector machines (SVM) [5], decision trees, random forests, and ensemble techniques have been adopted, often showing improved robustness to non-linear class boundaries. Deng et al. demonstrated that integrating polarimetric decomposition features with time–frequency domain features could significantly boost classification accuracy, particularly in complex urban environments [5]. Similarly, ensemble frameworks combining multiple classifiers have been explored to capitalize on the complementary strengths of different algorithms [6].

Another important dimension in this research area is feature selection. PolSAR imagery can yield a large number of candidate features, especially when multiple decomposition methods, frequency bands, and spatial texture measures are considered simultaneously. However, not all of these features contribute equally to classification performance; some may be

* Corresponding author

https://doi.org/10.5194/isprs-annals-X-4-W8-2025-25-2026

redundant or even detrimental. Traditional feature selection techniques include unsupervised dimensionality reduction methods such as principal component analysis (PCA), which project the data into a lower-dimensional subspace, and filter-based ranking methods like mutual information or correlation thresholding, which evaluate features individually against a criterion [7, 8]. More sophisticated strategies incorporate heuristic search algorithms—such as genetic algorithms, particle swarm optimization, or sequential forward selection—to identify optimal subsets [8, 9].

Despite these advances, conventional correlation-based feature elimination methods often remove all features with pairwise correlations above a certain threshold in a single step. While effective at reducing redundancy, this “one-shot” elimination can inadvertently discard features that, although correlated with others, still carry complementary discriminative information. Preserving a richer feature space can be especially beneficial when multiple classification algorithms are tested, as different classifiers may exploit different aspects of the feature set.

Post-processing is another critical stage in the classification pipeline. Even when classifier accuracy is high, classification maps may contain salt-and-pepper noise due to isolated misclassifications in homogeneous regions. A common remedy is to apply spatial filters such as majority voting within a local neighbourhood [10]. While such methods can effectively smooth noise, they may also erode sharp class boundaries, which is undesirable in applications requiring precise delineation [5]. Alternative approaches, such as rank-based smoothing filters, aim to improve spatial coherence without sacrificing boundary detail.

In this context, the present study proposes a PolSAR classification framework that addresses two key challenges: (1) retaining an informative and diverse feature set while reducing redundancy, and (2) improving the spatial consistency of classification outputs without over-smoothing critical boundaries. The first contribution is a correlation-based iterative feature selection strategy that progressively eliminates the most redundant features, as opposed to removing all correlated features in one pass. This approach maintains a higher-dimensional feature space, which in our experiments has led to improved classification accuracy compared to conventional threshold-based elimination. The second contribution is a median rank smoothing post-processing technique that replaces standard majority voting filters, producing classification maps that are both smoother and more faithful to fine-scale spatial patterns.

2. STUDY AREA AND DATASET

The Chinese Gaofen-3 (GF-3) satellite, launched in 2016, is the first C-band synthetic aperture radar (SAR) satellite with full polarimetric imaging capability in China. It provides high-resolution multi-polarization data with flexible imaging modes, making it highly suitable for applications in land cover mapping, environmental monitoring, and urban studies. In particular, its polarimetric SAR (PolSAR) data offer rich scattering information, which is crucial for improving the accuracy of classification tasks. The GF-3 dataset, was acquired in 2018 from San Francisco with a spatial resolution of 8 m is used in this research. The original image size is 5829 × 7173 pixels, and the cropped region used in our work corresponds to the coordinates (1144, 3464, 3448, 6376). Figure 1 presents a pseudo-colour image generated through Pauli RGB decomposition and the ground-truth map annotated by IPIU research team [11]. The reference data contain six land-cover categories: (1) Mountain, (2) Water, (3) Vegetation, (4) High-Density Urban, (5) Low-Density Urban, and (6) Developed area.

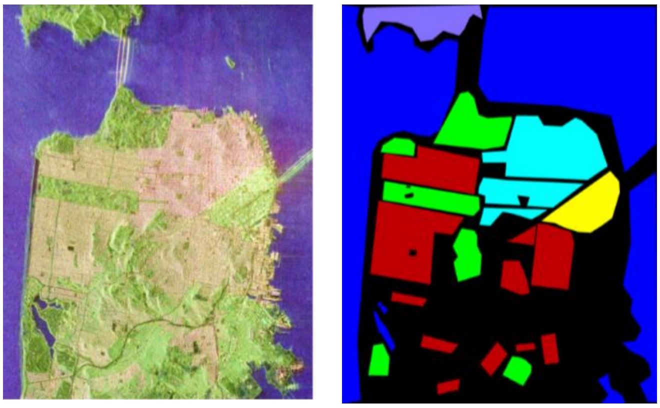


Figure 1. The study area and datasets

3. FUNDAMENTAL CONCEPTS

3.1 Polarimetric Descriptors

Polarimetric descriptors are mathematical representations of the scattering information contained in polarimetric synthetic aperture radar (PolSAR) data. For a fully polarimetric SAR system, the complete scattering matrix **S** is measured for each resolution cell, containing four complex elements corresponding to the horizontally and vertically transmitted and received waves: $S_{HH}, S_{HV}, S_{VH}, S_{VV}$. Under the assumption of reciprocity ($S_{HV} = S_{VH}$), this matrix provides the basis for deriving several alternative descriptor formats that are widely used in PolSAR data analysis [12].

The scattering matrix S is a 2×2 complex matrix containing the direct polarimetric measurements from the SAR system. Each element represents the complex scattering amplitude for a specific transmit–receive polarization combination. This primary descriptor is given as [13]

$$S = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \tag{1}$$

The lexicographic scattering vector is a 3×1 column vector obtained by rearranging the independent elements of the scattering matrix into a vector form. This representation is convenient for statistical analysis and is the basis for generating the covariance matrix. It is defined as [13]

$$L = \begin{bmatrix} S_{HH} & \sqrt{2}S_{HV} & S_{VV} \end{bmatrix}^t \tag{2}$$

where, t denotes the transpose operator.

The Pauli scattering vector is another 3×1 representation obtained by transforming the scattering matrix into the Pauli basis. This transformation expresses the scattering in terms of three canonical mechanisms: single-bounce (surface), double-bounce, and volume scattering. It is given as [13]

$$k = \frac{1}{\sqrt{2}}[S_{HH} + S_{VV} \quad S_{HH} - S_{VV} \quad 2S_{HV}]^t \tag{3}$$

The covariance matrix is a 3×3 Hermitian positive semidefinite matrix that captures the second-order statistical properties of the lexicographic scattering vector. This descriptor is widely used in statistical classification, speckle filtering, and model-based decomposition. It is computed as [13]

$$C = LL^* \tag{4}$$

where, * denotes the transpose conjugate operator.

The coherency matrix is also a 3×3 Hermitian positive semidefinite matrix, but it is computed from the Pauli scattering vector. It contains the second-order statistics expressed in the Pauli basis and is the standard input to many target decomposition algorithms. It is defined as [13]

$$T = KK^* \tag{5}$$

These polarimetric descriptors provide complementary views of the same scattering information and form the foundation for the polarimetric decomposition techniques that will be discussed in the next section.

3.2 Polarimetric decomposition

Polarimetric decomposition techniques aim to interpret the scattering mechanisms of targets observed in PolSAR data by expressing the measured scattering information in terms of a set of physically meaningful components. These methods can be broadly categorized into two groups: coherent and incoherent decompositions.

Coherent decompositions operate directly on the scattering matrix S and preserve the absolute phase information of the measured data [14]. They are suitable for situations where the resolution cell contains a single, well-defined scatterer (pure target), allowing for accurate retrieval of scattering characteristics at the individual pixel level.

In contrast, incoherent decompositions work on the covariance matrix C or the coherency matrix T, which are obtained by averaging the outer product of the scattering vectors over a neighborhood [15]. These matrices preserve only the relative phase information and represent the statistical properties of the scattering process. Incoherent methods are more appropriate for distributed targets or heterogeneous areas where each resolution cell contains a mixture of scattering mechanisms—such as urban regions, roads, vegetation, and water bodies—because they provide averaged parameters that are more robust to speckle noise.

In this study, due to the diversity of land-cover classes and the variety of scattering mechanisms present within the resolution cells of the image, we employ incoherent decompositions based on the covariance and coherency matrices. In the following, we briefly review several widely used incoherent decomposition methods.

Among incoherent decompositions, several model-based approaches are widely used. The Freeman–Durden decomposition expresses the total scattering as the sum of three independent components: surface scattering, double-bounce scattering, and volume scattering [16]. The Yamaguchi four-component decomposition extends this model by adding a helix scattering term, which improves the characterization of complex urban areas and man-made structures [17]. The Huynen decomposition focuses on parameterizing the target scattering in terms of physically interpretable parameters derived from the coherency matrix, often used in target recognition applications [18].

Another major class of incoherent decompositions relies on eigenvalue–eigenvector analysis of the coherency matrix. The most well-known is the Cloude–Pottier decomposition, which represents the scattering in terms of three parameters: entropy (H), anisotropy (A), and the mean alpha angle (α) [3]. Entropy measures the randomness of the scattering process, anisotropy describes the relative importance of the secondary scattering mechanisms, and the alpha angle provides information about the dominant scattering type. These parameters are widely used for unsupervised classification, clustering, and physical interpretation of PolSAR images.

In summary, coherent decompositions are most effective for pure targets with stable phase characteristics, while incoherent decompositions are preferred for heterogeneous environments where statistical averaging yields more reliable results. The choice between them depends on the nature of the study area and the desired balance between physical interpretability and robustness to speckle noise.

3.3 Feature selection

Feature selection is a crucial step in remote sensing image classification, particularly when dealing with high-dimensional datasets such as PolSAR imagery. The goal is to identify and retain the most informative and non-redundant features while discarding irrelevant or highly correlated ones. This process not only reduces computational complexity but can also significantly improve classification accuracy by mitigating the curse of dimensionality and overfitting [7, 19].

Feature selection methods can be broadly categorized into three main groups: filter methods, wrapper methods, and embedded methods [20].

Filter methods evaluate features based on statistical criteria, independent of any classifier. Common examples include Principal Component Analysis (PCA), which transforms the data into an orthogonal subspace that captures the maximum variance [21]; Minimum Noise Fraction (MNF), which aims to maximize the signal-to-noise ratio by first whitening the noise and then applying PCA [22]; and Fisher's Discriminant Ratio, which selects features that maximize between-class variance while minimizing within-class variance [23]. Another widely used filter approach is correlation-based feature selection, where features that exhibit high pairwise correlations beyond a certain threshold are eliminated to reduce redundancy [8].

Wrapper methods, on the other hand, evaluate subsets of features by training and testing a classifier, iteratively searching for the subset that yields the best performance. Examples include sequential forward selection and sequential backward elimination [24], as well as heuristic search algorithms such as genetic algorithms [25] and particle swarm optimization [25]. While potentially more accurate, wrapper methods are computationally expensive, especially for large feature spaces.

Embedded methods integrate feature selection into the classifier training process itself. Regularization-based models, such as LASSO [26] and decision tree–based algorithms like random forests [27], inherently perform feature ranking during model construction. These methods often strike a balance between the computational efficiency of filters and the accuracy of wrappers.

In the context of PolSAR data, feature selection plays a particularly important role because polarimetric decomposition methods can produce a large number of candidate features, many of which may be correlated or redundant. Careful selection ensures that the retained features capture complementary aspects of the scattering mechanisms, leading to more robust classification results. In this study, a correlation-based iterative selection strategy is employed to preserve a richer feature set while reducing redundancy, as will be described in the Proposed method section.

4. PROPOSED METHOD

The workflow of the proposed method is illustrated in Figure 2.

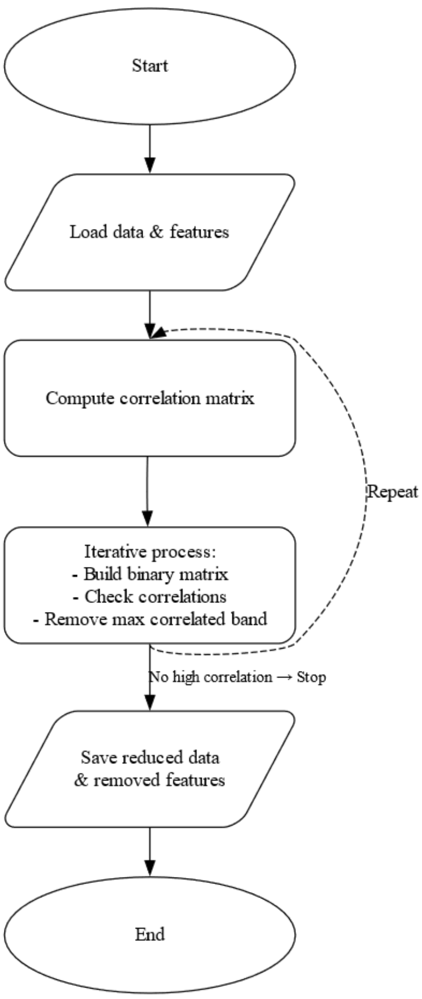
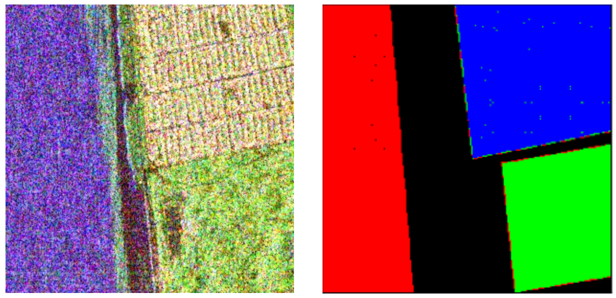


Figure 2. Workflow of the Proposed Method

4.1 Data Pre-processing

Considering the limitations of the computing hardware, a subregion of size 301 × 301 pixels was extracted from the cropped image for further processing. As illustrated in Figure. 3, this subregion contains three land-cover classes: urban area, water, and vegetation.



(a) (b)

Figure 3. Sub-region of study area used in this research (a: RGB Pauli, b: ground truth)

Due to the noisy nature of PolSAR data and the presence of relative phase information in the coherency matrix, the enhanced Lee filter with a kernel size of 5×5 was applied.

4.2 Feature Extraction

In this study, a comprehensive and diverse set of polarimetric features was extracted from the PolSAR data to fully leverage the rich scattering information contained within the covariance or coherency matrices. The feature vector encompasses a wide range of well-established decomposition theorems to characterize the scattering mechanisms of each target. Table 1 lists the name of each decomposition and the related features separately. It includes a total of 52 feature layers in the study area.

Name	Polarimetric features
Coherency matrix	T11, T22, T33, Span
Freeman 3	Vol, Odd, Dbl
H-A-Alpha	Entropy, Anisotropy, Alpha, Alpha1, Alpha2, Alpha3, Pedestal height, Shannon entropy, DERD, SERD, PA, PF, RVI
Cloude	T11, T22, T33
Holm 1	T11, T22, T33
Holm 2	T11, T22, T33
Huynen	T11, T22, T33
Scattering	Predominance, Polarization index, Conformity coefficient, Diversity, Degree of purity

Touzi	Alpha, Alpha1,2,3, Tau, Tau 1,2,3
VanZyl	Vol, Odd, Dbl
Yama 4	Vol, Odd, Dbl, Hlx

Table 1. Polarimetric decomposition used in this research

4.3 Feature Selection

In this study, the feature selection strategy represents the core innovation of our work. Increasing the number of features does not necessarily improve the performance of learning-based classification methods, unless the additional features contribute to the quality and discriminative power of the input space. Another critical issue is dimensionality reduction, which not only accelerates the training process but also improves model generalization. To address these challenges, we propose an iterative correlation-based approach for feature selection. The correlation analysis is conducted in a feature-to-feature manner to identify redundant descriptors, independent of class labels. Specifically, the normalized cross-correlation operator is first applied across all feature layers. A predefined correlation threshold is then used to identify highly correlated features. In each iteration, the feature with the maximum number of strong correlations is removed, and the procedure is repeated until all remaining features exhibit correlations below the threshold. This iterative mechanism ensures that the most independent and informative features are retained for the learning process. The key advantage of the proposed method lies in its iterative nature, which maximizes feature retention: removing a single feature in the early iterations can eliminate multiple redundant correlations, thereby preserving the largest possible set of independent features. Compared with conventional one-shot feature removal, the proposed iterative approach retains a larger number of independent features, which—as demonstrated in the next section—positively impacts the classification accuracy.

4.4 Classification

K nearest neighbours (KNN), support vector machine (SVM), Random Forest (RF), and extreme gradient boosting (XGB) were employed to evaluate the effectiveness of the proposed feature selection strategy. These algorithms were chosen due to their proven capability in handling high-dimensional remote sensing data and their complementary strengths in terms of bias-variance trade-offs. To ensure fair and optimized performance, the hyperparameters of each model were systematically tuned using a grid search procedure, tailored specifically to the characteristics of the employed PolSAR dataset.

4.5 Post-processing

After obtaining the initial classification maps, several post-processing strategies can be applied to enhance their accuracy. Common approaches include majority voting (mode filtering), morphological operations, and spatial smoothing filters, all of which aim to reduce salt-and-pepper noise and improve the spatial consistency of classification results. In this study, we specifically investigate the effect of applying rank-based image processing filters, such as the median filter, as a post-processing step on the classification map. The performance of the median filter will also be compared with other post-classification refinement techniques to evaluate its effectiveness in improving the final accuracy. The median filter is a non-linear, rank-order filter that replaces the value of each pixel with the median of its neighbourhood. This property makes it particularly effective in suppressing impulsive noise while preserving important structural boundaries, which is crucial for maintaining the spatial integrity of classified maps.

5. RESULTS AND DISCUSSION

In the first stage of the results, the proposed iterative feature selection method was evaluated by comparing it with the conventional one-shot feature elimination approach. For this comparison, the set of remaining features obtained from each method was used to train and classify the data with all four machine learning algorithms (KNN, SVM, random forest, and XGB). This allowed a direct assessment of the impact of feature selection strategy on classification performance. It is important to note that a threshold of 0.9 (90%) was employed to define high correlation. also, due to the limited size of the dataset, a training-to-testing ratio of 75:25 was adopted in all experiments to ensure a fair and consistent evaluation.

From the initial set of 52 features, the one-shot correlation-based elimination approach retained only 13 features with correlation values below the 90% threshold. In contrast, the proposed iterative correlation-based method preserved 28 features that satisfied the same criterion. This outcome clearly demonstrates, as an initial step, the superiority of the proposed method in maintaining a larger pool of uncorrelated and informative features, thereby confirming its effectiveness in comparison with the conventional one-step approach. the list of removed features under each method is summarized in Table 2**.**

One-shot correlation-based eliminated features	Iterative correlation-based eliminated features
Yamaguchi4_Y4O_Vol	Conformity
Yamaguchi4_Y4O_Odd	Yamaguchi4_Y4O_Odd
Yamaguchi4_Y4O_Dbl	Alpha
VanZyl3_Vol	VanZyl3_Odd
VanZyl3_Odd	Yamaguchi4_Y4O_Dbl
VanZyl3_Dbl	scatt_predominance
TSVM_alpha_s1	T11.huynen
TSVM_alpha_s	VanZyl3_Vol
Scatt_predominance	Scatt_diversity
Scatt_diversity	T33.holm2
Depolarisation_index	T11.holm2
Degree_purity	T22.holm1
Conformity	Serd
T33.huynen	TSVM_alpha_s1
T23.huynen	Depolarisation_index
T11.huynen	T33.holm1
T33.holm2	T11.holm1
T22.holm2	RVI
T11.holm2	Degree_purity
T33.holm1	T22.holm2
T22.holm1	Polarisation_fraction
T11.holm1	Derd
Serd	Alpha2
RVI	Freeman_Odd
Polarisation_fraction	
Pedestal	
Entropy	
Derd	
Anisotropy	

Alpha2	
Alpha1	
Alpha	
Freeman_Vol	
Freeman_Odd	
Freeman_Dbl	
T33	
T33.cloude	
T22.cloude	
T33.cloude	

Table 2. Eliminated features

The comparative results of the one-shot correlation elimination and the proposed iterative correlation-based feature selection method are summarized in Table 3. As can be observed, the one-shot approach yields classification accuracies and related metrics (Precision, Recall, and F1-score) in the range of 0.80–0.83 across the tested classifiers (KNN, SVM, RF, and XGB). In contrast, the proposed iterative method consistently improves the performance of all models, achieving values between 0.84 and 0.87 for all evaluation measures. Although the numerical improvements appear modest (approximately 0.03–0.04), they are systematic across all classifiers and metrics, which demonstrates the robustness of the proposed approach. Importantly, the higher F1-scores indicate that the iterative method not only increases overall accuracy but also provides a better balance between Precision and Recall. These results confirm that preserving a larger set of non-redundant features through the iterative elimination process leads to enhanced classification accuracy and generalization capability compared to the conventional one-shot feature selection strategy.

Recall	Precision	F1-score	Accuracy	model	Feature selection mode
0.80	0.79	0.79	0.80	KNN	One-shot correlation elimination
0.85	0.83	0.83	0.83	SVC	
0.83	0.82	0.82	0.83	RF	
0.83	0.82	0.82	0.83	XGB	
0.84	0.83	0.84	0.84	KNN	Iterative correlation elimination (Proposed method)
0.86	0.85	0.85	0.85	SVC	
0.87	0.86	0.85	0.86	RF	
0.87	0.86	0.86	0.86	XGB	

Table 3. Comparative results of the one-shot correlation elimination and iterative feature selection method

Considering that the XGB classifier achieved the highest performance among the evaluated models, its classification outcomes are further illustrated in Figure 4.

To improve the quality of spatial classification maps and reduce noise, two post-processing approaches—majority voting and a 3×3 median filter—were evaluated. The results in Table 4 show that the median filter consistently enhanced performance metrics across all models (KNN, SVC, RF, and XGB), including overall accuracy, precision, recall, and F1 score. In particular, for stronger models such as RF and XGB, metrics reached approximately 0.98–0.99, highlighting the filter's effectiveness in eliminating scattered errors and enhancing spatial coherence. In comparison, majority voting provided moderate improvements for weaker models (e.g., KNN and SVC) but generally underperformed relative to the median filter, with metrics around 0.93 based on the XGB model. This indicates that majority voting is more suitable for aggregating outputs from multiple diverse classifiers, whereas rank-based filters are more effective for single-model spatial data. Overall, the 3×3 median filter proves to be a superior post-processing strategy, improving quantitative accuracy while also producing smoother maps and reducing spurious classification artifacts.

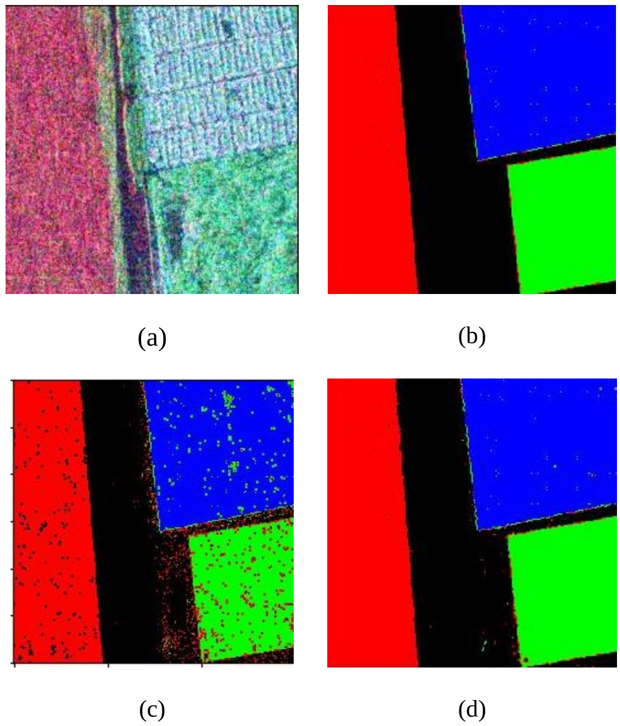


Figure 4. XGB results (a: Pauli false color, b: ground truth, c: classification map, d: median filtered classification map)

Strategy	Index	KNN	SVC	RF	XGB
Median filter 3*3	Accuracy	0.92	0.91	0.98	0.99
	F1-score	0.92	0.91	0.98	0.99
	Precision	0.92	0.92	0.98	0.99
	Recall	0.93	0.91	0.99	0.99
Majority voting	Accuracy	0.93			
	F1-score	0.93			
	Precision	0.93			
	Recall	0.94			

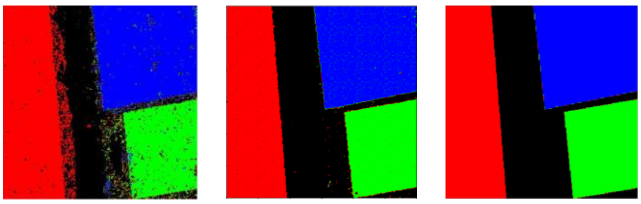
Table 4. Median filter effect on the classification results across all models

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-W8-2025-25-2026 |

It is worth noteing that, the very high accuracy (0.99) is influenced by the limited scene extent and class separability, not overfitting, as confirmed via test separation and cross-model consistency.

The visual results of applying these two post-processing methods, along with the ground truth, can be examined in the Figure 5.

The experimental results highlight the effectiveness of the proposed framework in addressing two long-standing challenges in PolSAR classification: feature redundancy and spatial noise. The iterative correlation-based feature selection preserved a larger pool of uncorrelated and complementary features compared with conventional one-shot elimination, which systematically improved the accuracy and generalization of all tested classifiers.



(a) (b) (c)

Figure 5. Classification map post processing (a: majority voting, b: 3*3 median filter on XGB, c: ground truth)

Among all tested classifier, XGB consistently demonstrated the best performance, indicating its suitability for handling high-dimensional PolSAR feature spaces. Furthermore, the application of the median filter as a post-processing step proved superior to majority voting by significantly enhancing spatial coherence while preserving boundary details. These improvements confirm that combining advanced feature selection with rank-based spatial refinement can substantially increase both quantitative accuracy and the visual quality of classification maps. The outcomes suggest that the proposed framework can be extended to larger datasets and more complex land-cover scenarios, offering a robust solution for operational PolSAR image analysis.

REFERENCES

1. Ramzi, P., F. Samadzadegan, and P. Reinartz, Classification of hyperspectral data using an AdaBoostSVM technique applied on band clusters. IEEE journal of selected topics in applied earth observations and remote sensing, 2013. 7(6): p. 2066-2079.
2. Karachristos, K., G. Koukiou, and V. Anastassopoulos, A review on PolSAR decompositions for feature extraction. Journal of Imaging, 2024. 10(4): p. 75.
3. Cloude, S.R. and E. Pottier, An entropy based classification scheme for land applications of polarimetric SAR. IEEE transactions on geoscience and remote sensing, 2002. 35(1): p. 68-78.
4. Lee, J.-S., M.R. Grunes, and G. De Grandi, Polarimetric SAR speckle filtering and its implication for classification. IEEE Transactions on Geoscience and remote sensing, 1999. 37(5): p. 2363-2373.
5. Deng, L., Y.-n. Yan, and C. Wang, Improved POLSAR image classification by the use of multi-feature combination. Remote Sensing, 2015. 7(4): p. 4157-4177.
6. Camps-Valls, G. and L. Bruzzone, Kernel-based methods for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing, 2005. 43(6): p. 1351-1362.
7. Guyon, I. and A. Elisseeff, An introduction to variable and feature selection. Journal of machine learning research, 2003. 3(Mar): p. 1157-1182.
8. Farahani, G., Feature Selection Based on Cross-Correlation for the Intrusion Detection System. Security and Communication Networks, 2020. 2020(1): p. 8875404.
9. Khan, A.H., et al., A genetic algorithm based feature selection approach for microstructural image classification. Experimental Techniques, 2022. 46(2): p. 335-347.
10. Hajizadeh, R., A. Aghagolzadeh, and M. Ezoji, Mutual neighborhood and modified majority voting based KNN classifier for multi-categories classification. Pattern Analysis and Applications, 2022. 25(4): p. 773-793.
11. Liu, X., et al. PolSF: PolSAR image datasets on san Francisco. in International Conference on Intelligence Science. 2022. Springer.
12. Cloude, S.R. and E. Pottier, A review of target decomposition theorems in radar polarimetry. IEEE transactions on geoscience and remote sensing, 2002. 34(2): p. 498-518.
13. Van Zyl, J.J. and H.A. Zebker, Imaging radar polarimetry. Progress in Electromagnetics Research, 1990. 3: p. 277-326.
14. Ulaby, F.T. and C. Elachi, Radar polaritnetry for geoscience applications. 1990.
15. Lee, J.-S. and E. Pottier, Polarimetric radar imaging: from basics to applications. 2017: CRC press.
16. Freeman, A. and S.L. Durden, A three-component scattering model for polarimetric SAR data. IEEE transactions on geoscience and remote sensing, 2002. 36(3): p. 963-973.
17. Yamaguchi, Y., et al., Four-component scattering model for polarimetric SAR image decomposition. IEEE Transactions on geoscience and remote sensing, 2005. 43(8): p. 1699-1706.
18. Huynen, J.R., Phenomenological theory of radar target. Ph. D. Dissertation, Technology University Delft, 1970.
19. Bhanu, B. and Y. Lin, Genetic algorithm based feature selection for target detection in SAR images. Image and Vision Computing, 2003. 21(7): p. 591-608.
20. Liu, H. and H. Motoda, Feature selection for knowledge discovery and data mining. Vol. 454. 2012: Springer science & business media.
21. Wold, S., K. Esbensen, and P. Geladi, Principal component analysis. Chemometrics and intelligent laboratory systems, 1987. 2(1-3): p. 37-52.
22. Green, A.A., et al., A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on geoscience and remote sensing, 1988. 26(1): p. 65-74.

23. Fisher, R.A., The use of multiple measurements in taxonomic problems. Annals of eugenics, 1936. 7(2): p. 179-188.
24. Pudil, P., J. Novovičová, and J. Kittler, Floating search methods in feature selection. Pattern recognition letters, 1994. 15(11): p. 1119-1125.
25. Xue, B., M. Zhang, and W.N. Browne, Particle swarm optimisation for feature selection in classification: Novel initialisation and updating mechanisms. Applied soft computing, 2014. 18: p. 261-276.
26. Tibshirani, R., Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B: Statistical Methodology, 1996. 58(1): p. 267-288.
27. Breiman, L., Random forests. Machine learning, 2001. 45(1): p. 5-32.

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-W8-2025-25-2026 | © Author(s) 2026. CC BY 4.0 License.