DOI: 10.34749/3061-1466.2026.28

Explainable AI for Efficient Hyperspectral Band Selection in Textile Recycling: A Score-CAM Approach

Achraf Guenounou
University of Udine
guenounou.achraf@spes.uniud.it

Jean-Philippe Andreu, Harald Ganster
JOANNEUM RESEARCH
{first-name}.{last-name}@joanneum.at

Abstract

To enable the transition to cost-effective, real-time multispectral sensors, this study introduces a novel Explainable AI (XAI) framework for spectral band selection by adapting Score-CAM—typically used for 2D images—to 1D hyperspectral signals. This XAI-driven approach is rigorously evaluated against established chemometric and machine learning baselines, including Weighted Regression Coefficients (WRC), Variable Importance in Projection (VIP), and the Successive Projections Algorithm (SPA). This method reduces data volume by over 90% while matching full-spectrum baseline performance. Ultimately, this research validates XAI as an interpretable, robust tool for designing efficient, low-cost optical sorting systems for the circular economy.

1 Introduction

The topic of sustainability is increasingly impacting our lives. We are facing various challenges to ensure our social and economic models are in harmony with the natural cycle of life. In our case, we have focused on one specific area: the classification of textile waste. With the advent of fast fashion, the economic impact of this sector has become impossible to ignore. The textile recycling market is projected to reach $6.42 billion by 2025 and is expected to grow to $9.23 billion between 2030 and 2035 [1].

It is essential to have technologies capable of precise detection of the fabrics in question, while also respecting production parameters that allow for case-specific detection and reduce production costs. Here, a key technology emerges: hyperspectral cameras. The idea behind this is that to analyze a material, we can observe the behavior of its molecules, such as C-H and O-H bonds.

By capturing the vibrations of these components, we can distinguish waves that become identifiable for each individual material. This curve is also known as a spectral fingerprint because it is unique. Specifically, our detection targets a specific range, the Short-Wave Infrared (SWIR) [2]. In our study this is covered by two distinct hyperspectral cameras, operatin in SWIR ranges: one spanning from 950 nm to 1700 nm, and the other from 1000 nm to 2500 nm, both widely utilized in the recycling sector for precise classification.

However, the fingerprint, or the material's spectrum, presents a significant challenge: being a continuous spectrum, the number of values captured for each individual pixel varies from 100 to 300 in the hyperspectral case. This introduces a major bottleneck as the space required to store and process these images increases enormously, slowing down image processing and model training times.

Indeed, such a high information density introduces various problems, including excessively correlated bands, which carry nearly similar information, and bands with background noise. These factors risk

The Third Austrian Symposium on AI, Robotics, and Vision (AIROV26).

misleading the classification and reducing the overall accuracy of the model. A good practice in this area is therefore to perform a dimensionality reduction of the spectral hypercube, allowing these redundant bands to be removed by focusing on a set of limited and localized features.

Crucially, identifying a restricted subset of highly discriminative wavebands offers a profound industrial advantage: it paves the way for replacing expensive, full-spectrum hyperspectral cameras with low-cost, application-specific multispectral optical sensors, dramatically reducing hardware costs in recycling facilities.

In the current state of the art, various techniques exist to address this issue. Typically, solutions that implement statistical methodologies are relied upon, as they are less computationally expensive and produce good results. The solution we present, however, shifts the focus to the interpretability of artificial intelligence models by implementing Score-CAM, a method originally developed for the analysis of classic 2D RGB images, adapting it to the hyperspectral case as a 1D vector ($1 \times$ number of bands).

2 State of the Art and Related Work

The high dimensionality of Hyperspectral Imaging (HSI) presents significant challenges in terms of data redundancy and computational overhead, a phenomenon widely known as the curse of dimensionality. To mitigate these issues, current literature generally divides dimensionality reduction techniques into two main categories: feature extraction and feature selection.

Feature extraction methods, most notably Principal Component Analysis (PCA) [3] and its variants, project the original data into a new, low-dimensional embedding space. However, this mathematical transformation inherently destroys the physical meaning of the specific spectral bands, rendering the resulting features difficult to interpret from a chemical perspective. Consequently, for applications requiring physical transparency, feature selection methods—which aim to determine a highly informative subset of the original, unaltered HSI bands—are strictly preferred. In this regard, Tunny et al. [4] conducted a detailed study of various techniques for performing feature selection with corresponding results in the hyperspectral field.

To make a fair comparison and provide a comprehensive overview, we focused on methodologies that are conceptually similar to our presented approach, namely those that determine the importance of a band by assigning it a relevance score.

2.1 Traditional Statistical Methods

In the current state of the art, methods relying on statistical principles are widely used due to their computational efficiency and ease of implementation. Variable Importance in Projection (VIP) [5] selects optimal variables by calculating a score derived from the variable's weight and measured variance. Similarly, the Weighted Regression Coefficient (WRC) [5] measures the relationship between standardized predictors and response variables using the regression coefficient β. Alternatively, the Successive Projections Algorithm (SPA) [6] iteratively selects a subset of variables to minimize redundant information and collinearity.

Despite their computational efficiency, these linear methods share a critical limitation: they struggle to capture the complex, non-linear relationships between distant spectral bands, necessitating more advanced approaches.

2.2 Neural Network-Based Approaches

To intrinsically model these non-linear relationships, deep learning architectures are increasingly adopted. Autoencoders, such as the SRL-SOA architecture [7], learn efficient data representations by compressing inputs into a lower-dimensional latent space and minimizing reconstruction error. Attention Mechanisms [8] offer another powerful alternative, dynamically assigning weights to spectral bands to indicate their relevance. Finally, Reinforcement Learning (RL) [9] frames band selection as a sequential decision-making process, though these approaches are complex to implement and demand substantial training data.

While Artificial Neural Networks (ANNs) excel at capturing complex, higher-order interdependencies among hyperspectral features, they inherently act as black boxes. This opacity prevents a clear understanding of the model's decision-making process and the physical relevance of the selected features.

3 Foundation of the Score-CAM Framework

Early visual explanation techniques, such as Class Activation Mapping (CAM) [10] and Grad-CAM [11], highlighted discriminative regions using network activations and gradients. However, gradient-based methods suffer from critical vulnerabilities: gradient saturation, noisy visual outputs, and the "False Confidence" phenomenon—where highly weighted features paradoxically contribute less to the actual prediction.

To address the limitations of gradient-based variations, Score-CAM [12] was introduced as a novel, gradient-free alternative. Instead of relying on local gradient sensitivity, Score-CAM encodes the importance of activation maps through the global contribution of the corresponding input features. Specifically, it evaluates the weight of each activation map through its forward passing score on the target class.

Mathematically, for a given convolutional layer l and a class of interest c, the Score-CAM saliency map is defined as a linear combination of activation maps passed through a ReLU function:

$$L^c_{Score-CAM} = \text{ReLU}\left(\sum_k \alpha^c_k A^k_l\right) \tag{1}$$

Where A^k_l is the k-th activation map of layer l, and α^c_k is its corresponding weight. Unlike Grad-CAM, Score-CAM defines this weight as the Channel-wise Increase of Confidence (CIC):

$$\alpha^c_k = C(A^k_l) = f(X \circ H^k_l) - f(X_b) \tag{2}$$

Here, X_b is a known baseline input, $f(\cdot)$ represents the network's output, and $\circ$ denotes the Hadamard Product. H^k_l acts as a mask generated from the upsampled activation map.

3.1 Synthesis and Research Objectives

While Score-CAM has proven highly effective at generating interpretable saliency maps for 2D spatial images, its potential for analyzing 1D continuous spectral signals remains unexplored. This study addresses this gap by adapting the Score-CAM framework to evaluate 1D hyperspectral data. We systematically compare this novel approach against established chemometric methods (such as WRC, VIP, and SPA). Ultimately, this analysis aims to establish a robust methodology for transitioning from complex hyperspectral data cubes to efficient, task-specific multispectral sensor architectures.

4 Methodology

4.1 Data Preparation and Preprocessing

The proposed XAI framework was evaluated using two distinct hyperspectral datasets. Both datasets operate within the Short-Wave Infrared (SWIR) spectrum, ensuring that the fundamental chemical absorption features of the textiles are properly captured.

- **Dataset A - Polyester and Cotton (512 bands):** This dataset comprises two chemically distinct classes: Cotton (a natural cellulosic fiber) and Polyester (a synthetic polymer). Due to their chemical differences, their spectral signatures exhibit high contrast. This dataset serves as a foundational proof-of-concept to demonstrate that macroscopic chemical disparities can be resolved with extremely low dimensionality; as our preliminary tests showed, utilizing a severely restricted subset of only two specific spectral bands is sufficient to achieve a classification accuracy of 98%.

- **Dataset B - Linen and Cotton (288 bands):** To stress-test the feature selection framework, a much more complex dataset was utilized, consisting of 16 fine-grained classes (8 variations of Cotton and 8 variations of Linen). Because both cotton and linen are cellulose-based natural fibers, their hyperspectral signatures are remarkably similar. This dataset challenges the model to identify subtle spectral characterizations that differentiate sub-variants of the same chemical family.

The discrepancy in the total number of spectral bands between the two datasets is due to the use of two distinct hyperspectral cameras during the acquisition campaigns. Specifically, Dataset A was acquired using a sensor operating in the 950–1700 nm range, producing 512 spectral bands, while Dataset B was captured using a broader SWIR camera covering the 1000–2500 nm range, yielding 288 bands.

To preserve local spatial information and support the overall classification process 3×3 pixels were adopted. The data was systematically partitioned into a split of 1800 samples for training, 600 for validation, and 600 for testing.

However, to rigorously isolate the chemical signature of the materials and prevent spatial texture from contaminating the feature extraction, the first phase of the methodology utilized only the central pixel of each patch.

4.2 1D-CNN and XAI Band Selection

A 1D Convolutional Neural Network (1D-CNN) was trained on a single-pixel classification task. Once trained, this 1D-CNN served as the baseline to decode its internal reasoning. For each class through our 1D implementation of Score-CAM average saliency profiles were computed, from which the top-K bands ($K \in \{1, 4\}$) for each class were extracted. The final reduced feature subset (S_{final}) was defined as the mathematical union of these class-specific peaks ($S_{final} = \bigcup_c S_c$).

4.3 Downstream Classification on Sparse Data

To evaluate the discriminative power of the selected feature subset (S_{final}), the original continuous hyperspectral data was masked, retaining only the intensity values at the top-K selected wavelengths. These newly generated sparse feature vectors were then tested across three distinct classification architectures.

First, the original 1D-CNN was retrained on the reduced dataset to assess how convolutional filters perform when the continuous spectral morphology is disrupted. Subsequently, the classification paradigm was expanded to include non-linear tabular approaches, specifically utilizing a Multi-Layer Perceptron (MLP).

5 Results and Discussion

5.1 Dimensionality Reduction via Score-CAM

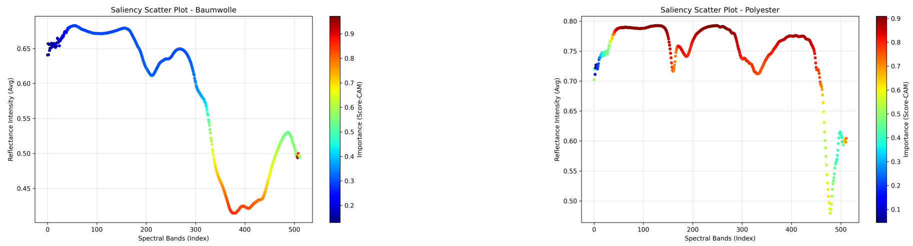
The first step in the process was to train a 1D-CNN on the full spectrum to perform the classification task. This step is crucial, as the feature extraction phase performed by the convolutional section of the network will allow us to identify and select the most discriminating bands in our dataset.

Once the network is trained, it is reused by Score-CAM to perform a classification analysis and identify the most impactful bands. In our adaptation, we ensured that it could generate a heatmap plot on the hyperspectral signal highlighting the most significant bands. The results for the various identified classes are shown in Figure 1.

5.2 1D-CNN Classification

The bands with the highest saliency scores are saved, separated for each class, in a single file, allowing us to compute the top K bands as follows:

- **Extreme Reduction ($K = 1$):** Extracting only the single most activated band per class.



(a) Cotton

(b) Polyester

Figure 1: Average heatmap result of Score-CAM from samples of Dataset A. Image represents the average saliency score for each band. The red regions indicate the most important bands for the classification of each class, while the blue regions indicate the least important bands.

- **Optimal Reduction** ($K = 4$)**:** Extracting the top 4 most activated bands per class. Empirical testing indicated a clear performance plateau: reducing the subset below four bands ($K \leq 3$) caused a severe degradation in classification accuracy due to the loss of critical secondary markers, whereas expanding the subset beyond four bands ($K \geq 5$) yielded negligible accuracy improvements. Therefore, $K = 4$ was empirically identified as the optimal inflection point.

Using the reduced sets, we proceeded to train the 1D-CNN for the respective datasets. Results are detailed in Table 1 and Table 2.

5.3 MLP for Classification

While the 1D-CNN was essential for processing the full hyperspectral signature and guiding the Score-CAM feature extraction, applying a convolutional architecture to a highly sparse and discrete set of selected bands is structurally suboptimal.

Once the optimal wavelengths are isolated, the continuous sequential correlation between adjacent bands is broken. Consequently, for the final classification stage on the reduced dataset, we transition from the 1D-CNN to a lightweight Multi-Layer Perceptron (MLP).

Unlike convolutional networks, an MLP is inherently designed to process discrete, non-sequential feature vectors, making it the mathematically sound choice for this task. Furthermore, the MLP significantly reduces computational overhead and inference time, fully aligning with the industrial objective of deploying fast, real-time multispectral sorting systems for textile recycling.

To properly format the input for this new architecture, the sparse selected bands extracted from a 3×3 spatial patch were flattened into a single input vector. This spatial-spectral approach allows the network to achive a better understanding of the local neighborhood around the selected spectral features, which is crucial for distinguishing subtle differences in the chemical composition of the materials.

Model Performances

The contrast between Dataset A and Dataset B perfectly illustrates the challenges of high-dimensional spectral reduction and the phenomenon of topological disruption. For the highly contrasted Dataset A, macroscopic chemical differences allowed the 1D-CNN to perform exceptionally well on the reduced set, even slightly exceeding its full-spectrum baseline at $K = 4$ (99.42% vs. 99.33%) by effectively using the Score-CAM selection as a noise filter. This demonstrates a crucial advantage of the proposed framework: when distinguishing materials with macroscopic chemical differences, isolating a small subset of primary and secondary chemical peaks effectively removes background spectral noise, effectively using the Score-CAM selection as a noise filter and intrinsically enhancing the model's discriminative power.

However, the limitations of applying convolutions to sparse data became glaringly evident in the highly complex Dataset B. When the 1D-CNN was retrained on the reduced bands, its accuracy dropped from the 96.78% baseline down to 95.49% ($K = 4$) and collapsed to 81.17% ($K = 1$). Because the feature selection process discards contiguous wavelengths, the convolutional filters lose the continuous morphological shape of the spectral peaks, rendering the sequential processing of the CNN highly ineffective.

Conversely, the MLP architecture completely mitigated this topological disruption. At $K = 4$, the MLP achieved an outstanding 98.56% accuracy on Dataset B, actually surpassing the full-spectrum baseline while processing only a fraction of the original data.

Furthermore, the MLP exhibited significantly shorter training times compared to the 1D-CNN. This confirms that combining XAI-guided spectral reduction with a lightweight classifier delivers a highly optimal solution, maintaining high accuracy with minimal computational overhead.

To visually demonstrate the effectiveness of this approach, Figure 2a, Figure 2b and Figure 2c compare the spatial classification maps generated by the MLP at extreme ($K = 1$) and optimal ($K = 4$) reduction levels against the expected Ground Truth. Finally, a comprehensive comparison of the classification accuracies achieved by both the 1D-CNN and the MLP models across all spectral ranges (full-spectrum, $K = 1$, and $K = 4$) for Dataset A and Dataset B is detailed in Tables 1 and 2.

Table 1: Classification Accuracy on Dataset A (2 Classes: Cotton-Polyester).

Model Architecture	**Full Spectrum (512 Bands)**	**Extreme Reduction** ($K = 1$)	**Optimal Reduction** ($K = 4$)
1D-CNN	99.33%	74.58%	99.42%
MLP (with 3×3 patch)	-	**99.42%**	99.25%

Table 2: Classification Accuracy on Dataset B (16 Classes: Fine-grained Cotton-Linen).

Model Architecture	**Full Spectrum (288 Bands)**	**Extreme Reduction** ($K = 1$)	**Optimal Reduction** ($K = 4$)
1D-CNN	96.78%	81.17%	95.49%
MLP (with 3×3 patch)	-	89.90%	**98.56%**

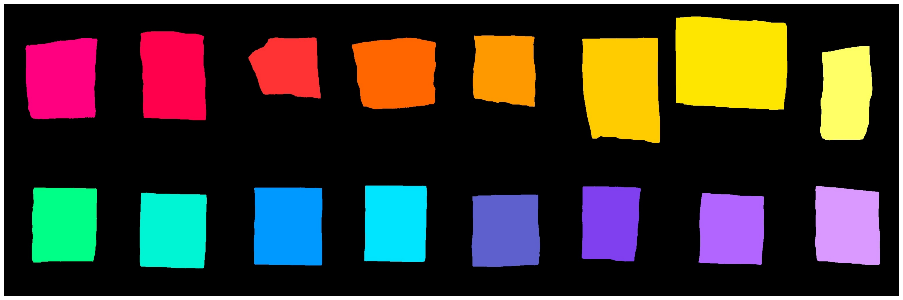
5.4 Comparison with Classical Statistical Feature Selection

To fairly evaluate the proposed Score-CAM framework against traditional statistical methodologies, we compared their classification performance using highly restrictive feature subsets. For Dataset A, Score-CAM under extreme reduction ($K = 1$) yielded 2 bands, and the statistical algorithms (PLS-VIP, SPA, and WRC) were tuned to select an identical subset size of 2 bands. For Dataset B, Score-CAM at the optimal reduction threshold ($K = 4$) yielded 43 bands, while the statistical methods were tuned to select comparable subsets ranging from 33 to 40 bands.

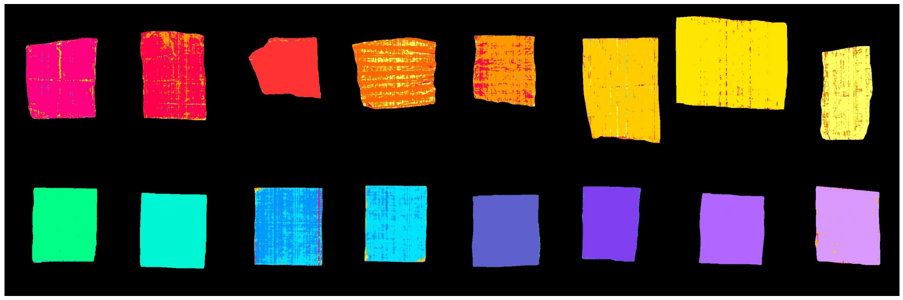
The empirical results of this comparison, evaluated using the MLP classifier, are detailed in Table 3 and Table 4. As the data shows, the Score-CAM framework exhibits highly competitive performance. On Dataset A, Score-CAM perfectly matches the 99.42% accuracy of the best-performing statistical methods (PLS-VIP and SPA) while outperforming WRC. On Dataset B, Score-CAM performs on par with the traditional methods, slightly outperforming SPA and matching PLS-VIP, while marginally trailing WRC, utilizing a comparable number of spectral bands.

This robust outcome for statistical methods is mathematically consistent. Algorithms such as SPA, WRC, and PLS-VIP are explicitly designed to optimize statistical variance, minimize collinearity, or directly reduce the classification error, allowing them to naturally achieve peak quantitative performance with highly compressed subsets. However, they operate as opaque mathematical constructs, lacking any physical or visual interpretability.

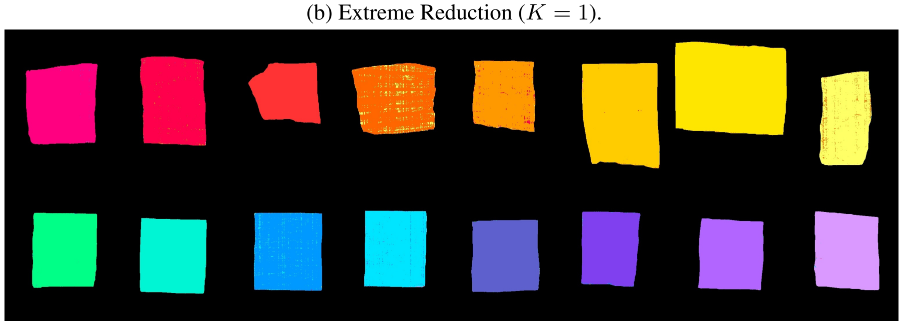
Any slight trade-off in dimensionality or marginal accuracy differences exhibited by our proposed method is fundamentally offset by the advantages of Explainable AI. Score-CAM provides a direct, interpretable link between the neural network's decision-making process and the actual chemical



(a) Ground Truth.



(b) Extreme Reduction ($K = 1$).



(c) Optimal Reduction ($K = 4$).

Figure 2: Spatial classification maps generated by the MLP for Dataset B under different reduction scenarios, compared against the Ground Truth.

absorption peaks of the textile materials. In real-world industrial recycling applications, deploying a transparent model that visually highlights *why* a classification is made—grounding the AI's behavior in physical spectroscopy—provides a critical layer of trust and diagnostic capability that purely statistical methods cannot offer.

Table 3: Performance Comparison on Dataset A (2 Classes: Cotton-Polyester) using MLP classifier.

Feature Selection Method	**Classifier**	**Selected Bands**	**Accuracy**
Variable Importance (PLS-VIP)	MLP	2	99.42%
Weighted Regression (WRC)	MLP	2	90.42%
Successive Projections (SPA)	MLP	2	99.42%
Score-CAM ($K = 1$)	**MLP**	**2**	**99.42%**

6 Conclusions and Future Work

The transition toward a circular economy requires fast, accurate, and scalable sorting of recyclable textiles. While Short-Wave Infrared (SWIR) Hyperspectral Imaging provides the necessary chemical

Table 4: Performance Comparison on Dataset B (16 Classes: 8 of Cotton and 8 of Linen) using MLP classifier.

Feature Selection Method	Classifier	Selected Bands	Accuracy
Variable Importance (PLS-VIP)	MLP	33	98.57%
Weighted Regression (WRC)	**MLP**	**40**	**98.95%**
Successive Projections (SPA)	MLP	39	98.39%
Score-CAM ($K = 4$)	MLP	43	98.56%

precision to distinguish complex blends, its high dimensionality poses a severe computational bottleneck for real-time industrial deployment. This study successfully addressed this challenge by proposing a novel, Explainable AI (XAI) driven framework for extreme spectral dimensionality reduction, adapting Score-CAM to 1D continuous hyperspectral signals.

By utilizing a 1D-CNN as a baseline "oracle" and decoding its internal reasoning, the proposed methodology successfully isolated the most critical chemical absorption peaks required to differentiate highly correlated materials, such as 16 fine-grained variations of cellulosic fibers.

The empirical evaluation of downstream classifiers on the reduced feature subsets yielded several critical methodological discoveries:

Topological Disruption and Spatial Compensation: The study demonstrated that deep convolutional networks struggle with sparse, non-contiguous spectral vectors due to the loss of continuous peak morphology. However, transitioning to a Multi-Layer Perceptron (MLP) architecture that incorporates a flattened 3×3 local spatial patch successfully mitigated this disruption. The spatial-spectral MLP achieved an outstanding 98.56% accuracy on the highly complex 16-class dataset (Dataset B) using an optimal reduction subset of 43 bands ($K = 4$), actually surpassing the full-spectrum baseline while discarding the vast majority of the data volume.

Physics-Awareness vs. Mathematical Optimization: When evaluated against classical chemometric feature selection methods (such as PLS-VIP, SPA, and WRC), the statistical algorithms achieved highly comparable results (e.g., 98.39%–98.95% vs. 98.56% on Dataset B) using slightly fewer bands (33–40 vs. 43). Furthermore, on Dataset A, Score-CAM perfectly matched the peak accuracy of the best statistical methods (99.42%) using an identical subset of just 2 bands. While classical algorithms achieve peak quantitative performance with highly compressed subsets by acting as opaque mathematical operators designed strictly to minimize collinearity, Score-CAM acts as a transparent, physics-aware selector. Any marginal trade-off in dimensionality is fundamentally offset by its ability to isolate explainable wavebands that directly correspond to the true physical and chemical absorption peaks of the textiles.

Future Work

Future research should focus on expanding the dataset to include a wider variety of synthetic polymers, organic blends, and contaminated post-consumer garments. Furthermore, the theoretical class-specific band subsets identified by the 1D Score-CAM algorithm could be physically validated by engineering a prototype multispectral optical sensor array. Finally, future work could explore the integration of this 1D spectral feature selection with 2D spatial segmentation models, creating a comprehensive, real-time spatio-spectral pipeline for robotic sorting architectures in textile recycling facilities.

7 Acknowledgments

This research was partially funded by Austrian Research Promotion Agency FFG through project StraTex (project number FO999911974).

Bibliography

[1] Fortune Business Insights. *Textile Recycling Market Size, Share & Industry Report, 2034*. https://www.fortunebusinessinsights.com/textile-recycling-market-109273. Ultimo accesso: 25 marzo 2026. 2026.

[2] Jordi-Roger Riba, Rosa Cantero, and Rita Puig. "Classification of Textile Samples Using Data Fusion Combining Near- and Mid-Infrared Spectral Information". In: *Polymers* 14.15 (2022). ISSN: 2073-4360. DOI: 10.3390/polym14153073. URL: https://www.mdpi.com/2073-4360/14/15/3073.

[3] Andrzej Maćkiewicz and Waldemar Ratajczak. "Principal components analysis (PCA)". In: *Computers & Geosciences* 19.3 (1993), pp. 303–342. DOI: https://doi.org/10.1016/0098-3004(93)90090-R. URL: https://www.sciencedirect.com/science/article/pii/009830049390090R.

[4] Salma Sultana Tunny, Hanim Z. Amanah, Mohammad Akbar Faqeerzada, Collins Wakholi, Moon S. Kim, Insuck Baek, and Byoung-Kwan Cho. "Multispectral Wavebands Selection for the Detection of Potential Foreign Materials in Fresh-Cut Vegetables". In: *Sensors* 22.5 (2022). ISSN: 1424-8220. DOI: 10.3390/s22051775. URL: https://www.mdpi.com/1424-8220/22/5/1775.

[5] Svante Wold, Michael Sjöström, and Lennart Eriksson. "PLS-regression: a basic tool of chemometrics". In: *Chemometrics and intelligent laboratory systems* 58.2 (2001), pp. 109–130. DOI: 10.1016/S0169-7439(01)00155-1.

[6] Mário CU Araújo, Tereza CB Saldanha, Roberto KH Galvao, Takashi Yoneyama, HC Chame, and V Visani. "The successive projections algorithm for variable selection in spectroscopic multicomponent analysis". In: *Chemometrics and Intelligent Laboratory Systems* 57.2 (2001), pp. 65–73. DOI: 10.1016/S0169-7439(01)00119-8.

[7] Mete Ahishali, Serkan Kiranyaz, Iftikhar Ahmad, and Moncef Gabbouj. "SRL-SOA: Self-Representation Learning with Sparse 1D-Operational Autoencoder for Hyperspectral Image Band Selection". In: *2022 IEEE International Conference on Image Processing (ICIP)*. 2022, pp. 2296–2300. DOI: 10.1109/ICIP46576.2022.9897863.

[8] Pablo Ribalta Lorenzo, Lukasz Tulczyjew, Michal Marcinkiewicz, and Jakub Nalepa. "Hyperspectral Band Selection Using Attention-Based Convolutional Neural Networks". In: *IEEE Access* 8 (2020), pp. 42384–42403. DOI: 10.1109/ACCESS.2020.2977454.

[9] Lichao Mou, Sudipan Saha, Yuansheng Hua, Francesca Bovolo, Lorenzo Bruzzone, and Xiao Xiang Zhu. "Deep Reinforcement Learning for Band Selection in Hyperspectral Image Classification". In: *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022), pp. 1–14. DOI: 10.1109/TGRS.2021.3067096.

[10] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. "Learning deep features for discriminative localization". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 2921–2929. DOI: 10.1109/CVPR.2016.319. URL: https://arxiv.org/abs/1512.04150.

[11] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. "Grad-cam: Visual explanations from deep networks via gradient-based localization". In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 618–626. DOI: 10.1109/ICCV.2017.74. URL: https://arxiv.org/abs/1610.02391.

[12] Haofan Wang, Zifan Wang, Mengnan Du, Fan Yang, Zijian Zhang, Sirui Ding, Piotr Mardziel, and Xia Hu. "Score-CAM: Score-weighted visual explanations for convolutional neural networks". In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 2020, pp. 24–25. DOI: 10.1109/CVPRW50498.2020.00020. URL: https://arxiv.org/abs/1910.01279.