TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Chang, C., Peng, S., Wang, X., Zong, S., Hu, W., Cheng, H., & Maio, F. D. (2026). A representation-first framework for laser-induced breakdown spectroscopy (LIBS) based quality assurance in recycled aggregate processing. *Results in Engineering*, *30*, Article 110902. https://doi.org/10.1016/j.rineng.2026.110902

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Results in Engineering 30 (2026) 110902

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Research paper

A representation-first framework for laser-induced breakdown spectroscopy (LIBS) based quality assurance in recycled aggregate processing

Cheng Chang [a,b,*], Siwei Peng [c], Xiaorong Wang [d], Shuai Zong [e], Wei Hu [f,g,**], Hao Cheng [h], Francesco Di Maio [a]

[a] Resource & Recycling, Department of Engineering Structures, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN, Delft, the Netherlands
[b] Faculty of Innovation Engineering, Macau University of Science and Technology, Avenida Wai Long, Taipa, 999078, Macao SAR, China
[c] School of Architecture, Southeast University, 210096, Nanjing, China
[d] Nanjing Engineering Design Branch, China 17th Metallurgical Group Co., Ltd., 210000, Nanjing, China
[e] School of Civil Engineering, Wuhan University, 430072, Wuhan, China
[f] School of Artificial Intelligence (School of Future Technology), Nanjing University of Information Science and Technology, 210044, Nanjing, China
[g] School of Mechanical and Aerospace Engineering, Nanyang Technological University, 639798, Singapore
[h] School of Civil and Environmental Engineering, Nanyang Technological University, 639798, Singapore

ARTICLE INFO

Keywords:
Recycled aggregates
Construction and demolition waste
Laser-induced breakdown spectroscopy
Spectral representation
Quality control
Conveyor-belt sorting
Circular construction

ABSTRACT

Ensuring consistent quality of recycled aggregates is essential for their wider use in circular construction. Laser-induced breakdown spectroscopy (LIBS) enables rapid elemental inspection, yet its performance in automated sorting systems is strongly shaped by how raw spectra are represented. This study adopts a representation-first benchmarking perspective and evaluates four representative feature families, namely variance-driven Principal Component Analysis (PCA), manifold learning-based Isomap, label-driven Partial Least Squares Discriminant Analysis (PLS-DA), and cepstral envelope-line separation, together with a raw-spectrum logistic-regression baseline and a histogram gradient boosting (HGB) reference model. The benchmark uses 24,000 single-shot spectra collected from ten material classes under conveyor-belt conditions, and repeated stratified random-split evaluation is used to assess the robustness of the comparative results across data splits. Across the evaluated models, cepstral features deliver the strongest overall performance, while PCA and the raw-spectrum logistic-regression baseline remain closely competitive. Isomap broadens the comparison toward non-linear manifold-based embeddings but does not improve performance in the present dataset, and PLS-DA shows the weakest stability under strong channel collinearity and class overlap. The results indicate that explicit feature extraction is not uniformly beneficial across all methods, but that spectral representation remains a major source of performance variation under controlled conveyor-like LIBS acquisition. In particular, cepstral features provide the most favourable balance among classification performance, robustness to baseline variation, and compactness, whereas PCA remains attractive when interpretability is prioritised. These findings provide a controlled benchmark and practical guidance for designing reliable and explainable LIBS-based quality-assurance pipelines for recycled-aggregate processing.

1. Introduction

The large-scale recycling of construction and demolition waste has become a key component of circular construction and sustainable resource management [1–3]. Among these waste streams, end-of-life concrete represents the dominant fraction, and its upcycling into recycled coarse and fine aggregates (RCA and RFA) offers a promising route to reduce reliance on virgin raw materials [4–6]. However, the quality of these secondary aggregates is often compromised by contaminants such as brick, foam, glass, gypsum, mineral fibres, plastics, and wood [7–10].

* Corresponding author at: Resource & Recycling, Department of Engineering Structures, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN, Delft, the Netherlands.
** Corresponding author at: School of Mechanical and Aerospace Engineering, Nanyang Technological University, 639798, Singapore.
E-mail addresses: chang-cheng@outlook.com, C.Chang-1@tudelft.nl (C. Chang), huwei@nuist.edu.cn (W. Hu).

https://doi.org/10.1016/j.rineng.2026.110902
Received 3 December 2025; Received in revised form 13 April 2026; Accepted 5 May 2026
Available online 5 May 2026
2590-1230/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Even at low concentrations, these impurities can significantly deteriorate the mechanical and durability properties of recycled concrete, limiting its structural performance and raising safety concerns. Accurate and efficient contaminant identification is therefore a prerequisite for ensuring that recycled aggregates meet the quality requirements of high-grade applications. In practice, inline quality assurance must meet industrial throughput requirements, where latency and computational footprint can be as critical as accuracy.

Laser-Induced Breakdown Spectroscopy (LIBS) has become a widely recognised analytical technique for this purpose due to its ability to perform rapid, in-situ elemental analysis across a wide variety of materials. A high-energy laser pulse focused on the sample surface generates a transient microplasma whose emission spectrum reflects the elemental composition of the ablated material [11–14]. The emitted light is dispersed and recorded by a spectrometer, producing high-dimensional spectra that capture both major and trace elements. LIBS has already been applied in environmental monitoring, geological exploration, and industrial quality assurance because of its speed, versatility, and minimal sample preparation [15–17]. In the context of recycled aggregates, LIBS offers particular advantages: it can be integrated into conveyor-belt sorting lines, enabling the single-shot detection of contaminants during high-throughput processing.

Nevertheless, LIBS data pose unique challenges for industrial deployment. Each single-shot spectrum contains thousands of wavelength channels, with substantial redundancy, noise, and baseline drift caused by varying plasma conditions or environmental factors. Moreover, many contaminants produce overlapping emission signatures, particularly when RCA particles retain cementitious residues [18]. Direct classification on raw spectra is therefore suboptimal. Effective feature extraction is required to reduce dimensionality, suppress irrelevant variation, and highlight the discriminative information needed for reliable classification.

A range of feature extraction methods has been adopted in LIBS research. Common LIBS representations include unsupervised compression such as Principal Component Analysis (PCA) [19–22] and supervised latent-variable representations such as Partial Least Squares Discriminant Analysis (PLS-DA) [23–27], which can reduce sensitivity to baseline variations. More recently, Cepstrum analysis, traditionally developed for speech signal processing, has been applied to spectroscopic data to separate slow-varying spectral envelopes from fine emission structures, thereby improving robustness to baseline fluctuations. Despite their frequent use, their comparative behaviour in multi-contaminant recycling streams, including robustness and computational cost under throughput constraints, remains insufficiently documented. This motivates a focused benchmark of representative feature spaces under conveyor-like acquisition.

From both an industrial quality assurance and a data-processing viewpoint, this gap is critical. Feature extraction is not a neutral step: the way spectral information is represented determines the stability, interpretability, and efficiency of subsequent classification. Yet most prior work has focused either on proposing new classifiers or on incremental variations in preprocessing [28–31]. In particular, whether representation choice matters more than classifier complexity for stability under baseline drift remains underexplored.

This study introduces a representation-first framework for deploying LIBS-based quality assurance systems in recycling plants, and validates the framework through case studies. In the revised benchmark, we compare four representative spectral representations—PCA, Isomap, PLS-DA, and Cepstrum—using the same lightweight classifier to isolate the effect of representation under conveyor-belt LIBS acquisition. We further include a raw-spectrum logistic-regression baseline and an end-to-end histogram gradient boosting (HGB) [32–35] reference model to contextualise the contribution of explicit feature extraction. Evaluation covers accuracy-oriented metrics, error structure, and repeated stratified random-split performance, together with indicators relevant to real-time deployment. Here, the term “representation-first” is used in a controlled benchmarking sense rather than as an exhaustive survey of all possible spectral embeddings. The selected methods are intended to span several practically relevant feature families, including linear variance-based, non-linear manifold-based, supervised latent-variable, and transform-based representations.

The main contributions of this work are threefold. First, we propose a representation-first framework that explicitly treats spectral representation as a design choice under robustness and deployment constraints. Second, we conduct a controlled benchmark of four representative feature families on a ten-class conveyor-belt LIBS dataset using a consistent classifier and evaluation protocol, and further contextualise these results with a raw-spectrum logistic-regression baseline and an HGB reference model. Third, we summarise practical considerations for selecting representations in inline recycled-aggregate quality assurance while clarifying the scope and limits of the present benchmark.

The remainder of this paper is organised as follows. Section 2 reviews related work and situates the present study within existing LIBS and feature extraction research. Section 3 describes the dataset and experimental setup. Section 4 presents the feature extraction methods under study, while Section 5 outlines the evaluation protocol. Section 6 reports the comparative results and discusses their implications for both signal processing methodology and industrial practice. Section 7 concludes the paper and outlines directions for future research.

2. Related work

LIBS has been widely applied for rapid material characterisation, and numerous studies have demonstrated its potential for recycling and resource recovery applications. Early work on LIBS in construction and demolition waste mainly focused on demonstrating feasibility, for instance by distinguishing concrete particles from contaminants such as glass or brick under controlled laboratory conditions [36]. More recent efforts have advanced towards inline sorting systems, integrating LIBS with conveyor belts and automated data processing pipelines [4]. These studies consistently confirm that while LIBS provides rich elemental information, the spectral data are high-dimensional and subject to strong variability, requiring dedicated processing strategies to extract reliable discriminative features [29,37].

2.1. Feature extraction in LIBS studies

Feature extraction has been a central concern in LIBS research. PCA is the most common unsupervised approach, employed for dimensionality reduction, data visualisation, and clustering of spectral signatures. Numerous studies report that PCA can separate major classes of materials such as metals, plastics, and mineral-based aggregates [38]. However, PCA is variance-driven and not optimised for class separation, which limits its effectiveness in cases where inter-class differences are subtle or confounded by background variation.

Non-linear manifold-learning methods provide another relevant family of spectral representations [39,40]. Approaches such as Isomap seek a low-dimensional embedding that preserves geodesic neighbourhood structure in the original high-dimensional space, and may therefore capture non-linear class geometry that is not well represented by linear projections [39]. Although such methods are less common in deployment-oriented LIBS pipelines, they offer a useful reference for assessing whether non-linear structure provides discriminative value beyond linear and transform-based representations [40,41].

PLS-DA introduces supervision by constructing latent variables that maximise covariance between spectra and class labels. Several studies in LIBS have employed PLS-DA for contaminant identification, particularly in binary or few-class problems [36]. PLS-DA can outperform PCA when class labels correlate strongly with systematic variance, but it is also sensitive to noise and collinearity, which are prevalent in LIBS spectra containing thousands of highly correlated channels [30].

More recently, Cepstrum analysis has been proposed as a feature

extraction tool for spectral data. Originally developed in speech processing and widely used in vibration analysis, the Cepstrum operates on the log-spectrum to separate slowly varying envelopes from rapidly oscillating structures. When applied to LIBS, this transformation can attenuate baseline drift and enhance discriminative emission peaks. While promising, its application to LIBS remains relatively unexplored, and few systematic comparisons with PCA or PLS-DA exist.

2.2. Classification and machine learning approaches

Beyond feature extraction, numerous classifiers have been applied to LIBS spectra, ranging from traditional linear discriminants to modern machine learning. Support Vector Machines (SVM), Random Forests, and Gradient Boosted Trees have shown competitive performance in various LIBS datasets [18,42]. More recently, convolutional neural networks (CNNs) have been explored for end-to-end spectral classification [43–45]. These methods highlight the feasibility of direct classification from raw or minimally processed spectra. However, such approaches often trade interpretability and robustness for accuracy, and their performance is highly dependent on dataset size, hyperparameter optimisation, and regularisation strategies.

2.3. Limitations in current literature

Despite these advances, several limitations persist. First, many LIBS studies report results on relatively simple datasets (e.g., binary classification of one contaminant vs. clean aggregate), which do not reflect the multi-class, noisy conditions encountered in industrial recycling. Second, comparisons are often made between different classifiers rather than between feature extraction strategies [46]. As a result, it remains unclear whether performance differences arise from the choice of classifier or from the underlying representation of the spectra. Third, systematic evaluation across multiple contaminants with rigorous metrics (e.g., macro-F1, ROC, and error structures) is rare, although such evaluation is essential to assess robustness in realistic settings [4].

2.4. Positioning of the present study

The present work builds on this literature by focusing specifically on the role of feature extraction in LIBS-based contaminant identification. Instead of emphasising classifier novelty, we benchmark four conceptually distinct feature extraction approaches—variance-driven (PCA), manifold-based (Isomap), label-driven (PLS-DA), and envelope/fine-structure separation (Cepstrum)—within a unified experimental framework. To ensure fairness, all feature-based pipelines are evaluated with a common classifier (Logistic Regression), while a raw-spectrum logistic-regression baseline and an HGB reference model provide additional points of comparison. This design enables us to isolate the impact of representation and to identify the mechanisms that govern robustness, interpretability, and efficiency in LIBS classification.

3. Dataset and experimental setup

3.1. Material samples

The experimental dataset was derived from end-of-life (EoL) concrete and demolition waste collected at multiple dismantling sites in the Netherlands. To minimise cross-contamination, selective demolition procedures were adopted, ensuring that concrete was removed with minimal admixture of foreign materials. Additional contaminants, including brick, foam, glass, gypsum, mineral fibres, plastics, cement paste, and wood, were manually separated from the demolition streams. Together with the RCA and RFA, these materials constituted the ten target classes used in this study.

For each class, materials collected from multiple sites were pooled to form a representative reference batch. Samples were prepared to reflect typical particle sizes encountered in industrial recycling operations. Material geometry therefore varies by class, ranging from granular aggregates to irregular fragments or fibrous pieces, and discrete physical specimens are not defined in this pooled design. During acquisition, all materials were measured in an air-dried state and spread on the conveyor to a bed thickness of approximately 4 cm. This dataset design reflects the compositional heterogeneity of demolition waste and supports controlled benchmarking under realistic multi-class conditions.

3.2. Experimental setup

The experimental setup (Fig. 1) was designed to emulate inline quality inspection on conveyor belts, as encountered in industrial recycling facilities. It comprised three main components: a laboratory-scale conveyor belt, a Q-switched Nd:YAG laser source, and an Echelle spectrometer for plasma emission capture.

The Q-switched Nd:YAG laser (TRLi DPSS Series) operated at the fundamental wavelength of 1064 nm, with a pulse energy of 170 mJ, a repetition rate of 100 Hz, and a pulse width of 8–10 ns. The pulse energy was chosen to ensure reliable plasma formation and adequate signal strength under in-air, conveyor-like acquisition on heterogeneous particle surfaces; we note that such conditions can increase shot-to-shot variability, which is reflected in the single-shot benchmark reported in this study. The beam was focused vertically onto the sample stream using a 300 mm focal length lens. Each pulse ablated a small volume of material, producing a transient plasma.

Plasma emission was collected through a collection lens module and transmitted via optical fibre to an Iris series Echelle spectrometer (SPECTRAL Industries, the Netherlands), covering a spectral range of approximately 179–1200 nm. To reduce the influence of the early plasma continuum, a detection delay of 1.5 μs was introduced, controlled by a digital delay generator (Quantum Composers).

The conveyor belt was operated at a constant speed of 20 cm/s. Combined with the 100 Hz pulse repetition, this configuration yielded one laser shot every ~2 mm along the belt, ensuring dense spatial sampling of moving particles. All measurements were performed under standard atmospheric conditions, replicating the non-vacuum environment typical of industrial practice.

The optical system operated in a fixed-focus configuration without automatic focus adjustment. During acquisition, the material layer was levelled to an approximately consistent height, while residual small height variations were treated as part of the single-shot variability of the benchmark.

3.3. Data acquisition

Each laser shot generated a single LIBS spectrum corresponding to a localised spot on the particle surface. For each material type, 2400 single-shot spectra were acquired during a single conveyor pass (Fig. 2), resulting in a balanced dataset across all ten classes. Materials were not recirculated and no repeated passes were performed; shots were taken at spatially separated locations along the belt to avoid repeated firing on the same spot. The 2400 spectra per class correspond to all acquired single-shot spectra used in this study, and no additional post hoc filtering or discarding of unstable events was performed. The raw spectra [47], consisting of 11,790 intensity channels per shot, served as the basis for subsequent feature extraction and classification analyses.

4. Feature extraction methods

LIBS spectra contain thousands of wavelength channels per measurement, with strong collinearity and significant noise originating from plasma fluctuations, baseline drift, and surface heterogeneity. Directly applying classifiers to such high-dimensional raw data often leads to overfitting, computational inefficiency, and reduced robustness. Feature extraction therefore plays a critical role: by transforming spectra into

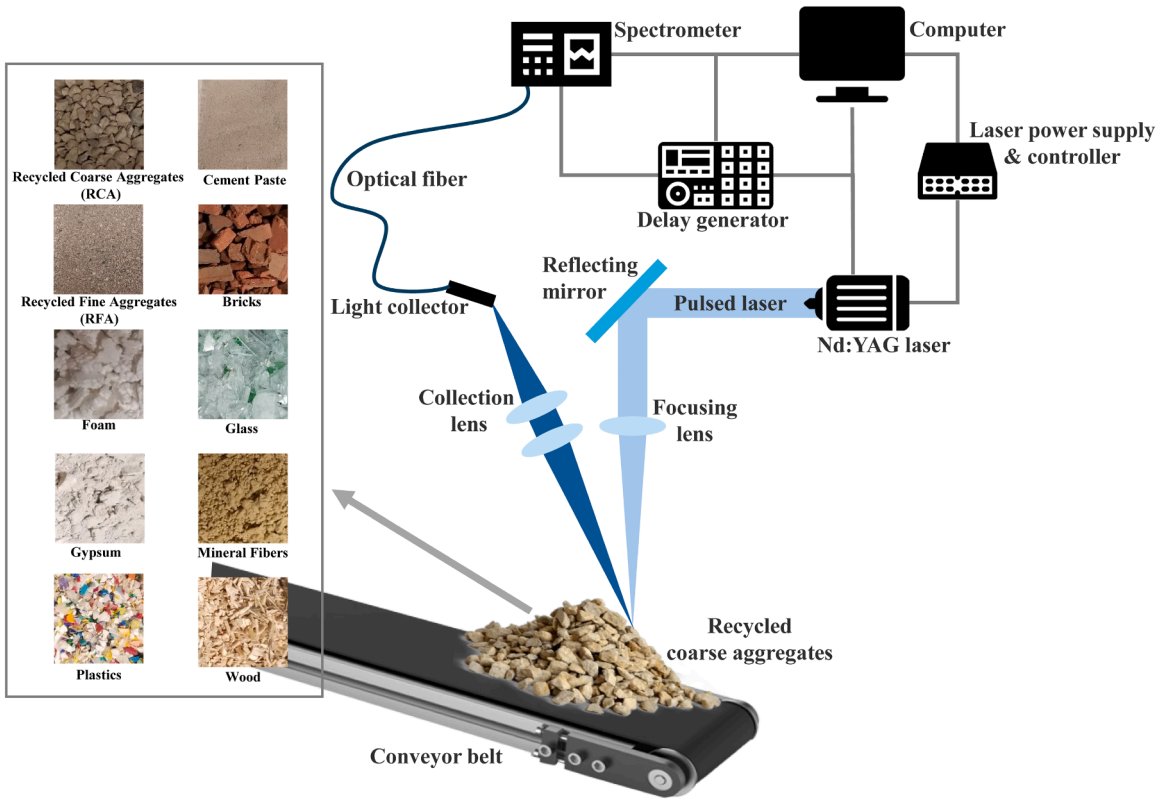


Fig. 1. Experimental setup.

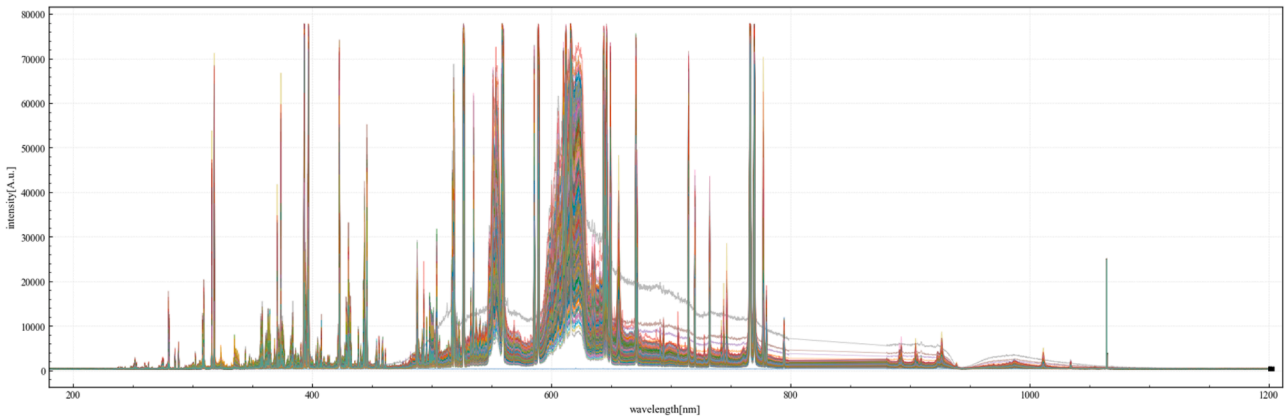


Fig. 2. Representative LIBS spectra of cement paste samples (multi-shot measurements).

compact and discriminative representations, it reduces dimensionality, suppresses irrelevant variation, and emphasises spectral information that aligns with class boundaries.

This study considers four feature extraction strategies, PCA, Isomap, PLS-DA, and Cepstrum analysis, together with a raw-spectrum logistic-regression baseline and an HGB reference model. The methods represent four conceptually distinct approaches: variance-driven, manifold-based, label-driven, and envelope–fine-structure separation. They are chosen as representative case studies for a deployment-oriented LIBS benchmark, rather than as an exhaustive survey of all spectral representation methods. A complete summary of the parameter settings used in the benchmark is provided in Supplementary Table S1.

4.1. Raw-spectrum logistic-regression (LR) baseline

To isolate the contribution of explicit feature extraction, we additionally trained a logistic-regression (LR) baseline directly on the raw spectra. The same train/test protocol and optimisation framework were used as for the feature-based logistic-regression pipelines, so that performance differences can be attributed primarily to the input representation rather than to the classifier family.

4.2. PCA

PCA is a widely used unsupervised dimensionality reduction technique that projects high-dimensional data into a lower-dimensional subspace while retaining the maximum possible variance [48,49]. In the context of LIBS, where each single-shot spectrum consists of thousands of wavelength–intensity pairs, PCA helps to eliminate redundancy, reduce noise, and reveal latent spectral structures that are not directly observable in the raw data.

Formally, given a dataset of n single-shot LIBS spectra, arranged in a

matrix:

$$X \in \mathbb{R}^{n \times p} \tag{1}$$

where p is the number of recorded wavelengths (11,790 in our dataset), PCA proceeds by centering the spectra and computing the covariance matrix:

$$C = \frac{1}{n-1} X^{\top} X \tag{2}$$

The eigendecomposition of C yields:

$$Cv_j = \lambda_j v_j, j = 1, 2, \ldots, p \tag{3}$$

where v_j are the eigenvectors (principal axes) and λ_j are the eigenvalues, ordered such that $\lambda 1 \geq \lambda 2 \geq \cdots \geq \lambda p$. The eigenvectors define orthogonal directions in spectral space, and their associated eigenvalues indicate the amount of variance explained.

The projection of spectra onto the first k principal components is given by:

$$T = XV_k, V_k = [v_1, v_2, \ldots, v_k] \tag{4}$$

where $T \in \mathbb{R}^{n \times k}$ contains the principal component scores, representing compressed and noise-reduced features. In practice, only the first few components (typically explaining >95 % of the variance) are retained for subsequent classification.

In this study, each LIBS spectrum was treated as a vector in $\mathbb{R}^p$. PCA was applied globally across all training samples, ensuring that the resulting projection directions captured the dominant spectral variations across the ten material classes. The reduced feature vectors T were then used as input to a LR classifier. This setup allows a fair comparison between PCA and other feature extraction methods, isolating the effect of the representation from that of the classifier.

The interpretability of PCA is also valuable in LIBS analysis. Principal component loadings often highlight specific wavelength regions corresponding to elemental emission lines (e.g., Ca, Si, Fe). This facilitates physical interpretation of classification boundaries, linking algorithmic decisions to underlying material properties. However, PCA remains variance-driven and label-agnostic: the dominant directions may capture global spectral variation (e.g., baseline shifts or common peaks) rather than features most relevant for class discrimination. This explains why PCA achieves competitive but not always optimal performance in distinguishing contaminants with overlapping emission patterns, such as foam and glass. For construction quality assurance, PCA preserves interpretability by linking decisions to elemental emission lines, supporting standards compliance and operator diagnosis.

4.3. Isomap

Isomap is a non-linear manifold-learning method that seeks a low-dimensional embedding while preserving geodesic distances between neighbouring samples in the original high-dimensional space. Unlike PCA, which is restricted to linear projections, Isomap can capture curved low-dimensional structure when spectral variation is not well approximated by a single linear subspace.

In the context of LIBS, Isomap is relevant as a reference non-linear representation because heterogeneous materials may exhibit class structure that is shaped not only by linear intensity variation but also by non-linear relations induced by overlapping emission features, baseline effects, and matrix-dependent responses. Including Isomap in the present benchmark therefore helps assess whether a manifold-based embedding provides discriminative value beyond linear or transform-based representations.

In this study, Isomap was fitted on the training spectra only, and the resulting low-dimensional embedding was used as input to the same logistic-regression classifier adopted for the other feature-based pipelines. This preserves comparability and isolates the contribution of the representation from that of the classifier. The neighbourhood size and embedding dimensionality were selected empirically to balance local structure preservation and classification stability.

4.4. PLS-DA

PLS-DA is a supervised dimensionality reduction method that constructs latent variables by maximising the covariance between the predictor matrix X (spectral intensities) and the response matrix Y (class labels). Unlike PCA, which identifies variance-maximising directions without considering class information, PLS-DA seeks projections that are directly predictive of the categorical response.

Given a spectral dataset $X \in \mathbb{R}^{n \times p}$ with n LIBS spectra and p wavelength channels, and a response indicator matrix $Y \in \mathbb{R}^{n \times c}$ encoding class membership across c material classes, PLS-DA decomposes the data into:

$$X = TP^{\top} + E \tag{5}$$

$$Y = UQ^{\top} + F \tag{6}$$

where $T \in \mathbb{R}^{n \times k}$ and $U \in \mathbb{R}^{n \times k}$ are the score matrices, $P \in \mathbb{R}^{p \times k}$ and $Q \in \mathbb{R}^{c \times k}$ are loadings, and E,F are residuals. The latent variables T are obtained such that the covariance between T and U is maximised. This can be expressed as an optimisation problem:

$$\max_{w,q} \; cov(Xw, Yq) \tag{7}$$

subject to $\| w \| = 1$, $\| q \| = 1$.

The resulting weight vectors w define directions in the spectral space that best align with the discriminant structure encoded in Y. Once computed, the latent projections T serve as compressed, label-driven features that can be used for classification.

PLS-DA can be viewed as a supervised projection method. Starting from the training spectra X, we encode class membership as a one-hot response matrix Y, and extract k latent components that capture the class-discriminative structure in the data. The resulting score matrix T can be written as $T = XW$, where W collects the PLS weight vectors w obtained from the optimisation in Eq. (7). These latent scores provide a low-dimensional, label-informed representation and are used as inputs to the logistic-regression classifier, analogous to PCA scores or truncated cepstral coefficients.

In this study, PLS-DA was applied to the ten-class LIBS dataset, and LR was subsequently trained on the extracted latent scores to ensure comparability with PCA and cepstral features.

The potential advantage of PLS-DA in LIBS analysis lies in its ability to incorporate label information directly into feature construction, theoretically leading to enhanced separability between material classes. This is particularly relevant when inter-class differences are subtle but correlated with specific wavelength bands (e.g., Ca or Si emission peaks).

However, PLS-DA is also highly sensitive to the statistical structure of LIBS data. The thousands of correlated wavelength channels in each spectrum introduce strong collinearity, which destabilises the estimation of latent variables. Moreover, when class boundaries overlap significantly—as in the case of recycled aggregates and cement paste—PLS-DA tends to overfit training data and exhibits poor generalisation. In our experiments, these effects manifested as a substantial reduction in macro-F1 compared with PCA and cepstral methods, accompanied by pronounced off-diagonal entries in the confusion matrix. Here, the confusion matrix is reported with rows denoting the true classes and columns denoting the predicted classes; diagonal entries indicate correct classifications, whereas off-diagonal entries indicate misclassifications (Fig. 6).

In summary, PLS-DA provides a supervised framework for feature extraction, but its effectiveness on LIBS spectra is limited by data

collinearity and class overlap. While it can outperform PCA in controlled or low-dimensional chemometric settings, its instability in high-dimensional, noisy, multi-class environments makes it less reliable for inline contaminant identification in recycled aggregates. Under multi-contaminant mixtures and collinearity, PLS-DA can trade stability for label alignment, which may limit reliability under plant conditions.

4.5. Cepstrum analysis

Cepstrum analysis is a signal processing technique originally developed in speech analysis and vibration diagnostics, designed to separate slowly varying spectral envelopes from rapidly oscillating fine structures. Here, we extract cepstral features from the Cepstrum, defined as the inverse Fourier transform of the log-magnitude spectrum. In simple terms, this transform separates the slowly varying spectral background from the finer peak pattern, which reduces sensitivity to baseline drift. When applied to LIBS spectra, the Cepstrum offers a compact representation that suppresses baseline drift while preserving discriminative emission features.

Given a single-shot LIBS spectrum $x[n]$ sampled across p wavelengths, the Cepstrum is obtained by first computing the discrete Fourier transform (DFT):

$$X[k] = \sum_{n=0}^{p-1} x[n] e^{-\frac{j2\pi kn}{p}}, k = 0, 1, \ldots, p-1 \tag{8}$$

The log-magnitude spectrum is then taken:

$$L[k] = \log(|X[k]| + \epsilon) \tag{9}$$

where ϵ is a small stabilisation constant to avoid singularities. Finally, the cepstral coefficients are given by the inverse DFT of $L[k]$:

$$c[m] = \frac{1}{p} \sum_{k=0}^{p-1} L[k] e^{\frac{j2\pi km}{p}}, m = 0, 1, \ldots, p-1 \tag{10}$$

Low-order coefficients $c[m]$ describe the slowly varying spectral envelope, which captures baseline drift and broad continuum variations. Higher-order coefficients encode the fine oscillatory structures, corresponding to sharp atomic emission lines. By truncating the Cepstrum to a limited number of coefficients, one obtains a compact feature vector that retains discriminative information while attenuating noise and drift.

In this study, cepstral coefficients were computed for each LIBS spectrum and truncated to retain only the first few dozen coefficients, corresponding to the most informative envelope and line-structure components. These reduced-dimension feature vectors were then used as inputs to LR classifiers, allowing direct comparison with PCA- and PLS-DA-based pipelines.

The advantage of Cepstrum analysis in LIBS is twofold. First, it provides robustness against baseline drift caused by variations in plasma conditions, dust scattering, or optical transmission, which are common in conveyor-belt environments. Second, it enhances separation of classes characterised by strong emission lines (e.g., gypsum and mineral fibres), yielding nearly diagonal confusion matrices in our experiments.

Nevertheless, Cepstrum analysis also exhibits limitations. Because truncation discards higher-order coefficients, it may lose sensitivity to very narrow or weak spectral lines. This explains why the method underperformed on wood, where subtle organic emission features were critical for discrimination. Moreover, Cepstrum coefficients are less directly interpretable than PCA loadings, as their relation to individual emission lines is indirect.

In summary, Cepstrum analysis transfers a well-established technique from speech and vibration domains into LIBS spectroscopy, where it achieves high robustness and accuracy. Its strong performance across most material classes in our dataset demonstrates the value of adopting cross-domain signal processing tools for complex, noisy spectral data. Low-quefrency cepstral coefficients suppress baseline drift while retaining discriminative lines, improving robustness under variable industrial environments.

4.6. End-to-end baseline: HGB

To provide a reference point for evaluating the value of explicit feature extraction, an end-to-end classifier based on HGB was included in this study. HGB is a tree ensemble method that combines gradient boosting with histogram-based binning of features, allowing efficient training on high-dimensional datasets. Unlike PCA, PLS-DA, or cepstral analysis, HGB operates directly on the raw spectra, learning discriminative decision rules without intermediate feature transformations.

Formally, HGB builds an additive model of the form:

$$f(x) = \sum_{m=1}^{M} \gamma_m h_m(x) \tag{11}$$

where each $h_m(x)$ is a decision tree fitted to the residuals of the previous stage, and γ_m are weights determined by gradient descent on a loss function (e.g., cross-entropy for classification). The histogram approximation accelerates training by discretising continuous spectral intensities into a fixed number of bins, reducing memory and computational cost.

For a multi-class problem with classes $y \in \{1, \ldots, C\}$, HGB minimises the multinomial deviance:

$$L = -\sum_{i=1}^{n} \sum_{c=1}^{C} 1(y_i = c) \log p_{ic} \tag{12}$$

where p_{ic} is the predicted probability of sample i belonging to class c, and $1(\cdot)$ is the indicator function.

In the context of LIBS, HGB has two attractive properties. First, it can capture non-linear interactions between spectral channels, which may be overlooked by linear projection methods. Second, it avoids manual feature design, learning discriminative rules directly from the raw intensity values. In our dataset, HGB achieved competitive ROC-AUC values, confirming its ability to exploit the high-dimensional structure of LIBS spectra.

However, HGB also has notable limitations. Its performance depends heavily on hyperparameter tuning (tree depth, learning rate, bin size), and the resulting decision rules are not readily interpretable in terms of physical emission lines. For industrial LIBS sorting, this lack of transparency is problematic: operators cannot easily trace a misclassification back to specific spectral features, making diagnostics and maintenance more difficult. Moreover, HGB showed reduced stability at decision boundaries compared with feature-based methods, particularly for overlapping classes such as RCA and cement paste. While competitive, end-to-end models can increase computational footprint and reduce explainability, which is disadvantageous for embedded, auditable quality assurance pipelines.

In this study, HGB is therefore treated as a baseline reference rather than a deployable solution. It provides an important point of comparison, illustrating the relative advantage of principled feature extraction methods in terms of robustness, interpretability, and consistency across contaminant classes.

5. Evaluation protocol

The evaluation protocol was designed to provide a fair and consistent comparison of six methods under study: raw-spectrum LR, PCA+LR, Isomap+LR, PLS-DA, Cepstrum+LR, and the HGB reference model. All methods were applied to the same dataset of ten material classes (brick, cement paste, foam, glass, gypsum, mineral fibres, plastics, RCA, RFA, and wood), with 2400 single-shot LIBS spectra per class, yielding a total of 24,000 spectra. Each spectrum consisted of 11,790 intensity channels across the wavelength range 179–1200 nm.

5.1. Data partitioning

The dataset was randomly divided into training and test subsets using a stratified split to preserve class balance. The split was performed at the shot level by randomly partitioning the 2400 single-shot spectra per class into training and test sets. The training set was used for feature extraction and classifier training, while the test set was used to evaluate shot-level classification performance on held-out spectra. Stratification ensured that class proportions were maintained across the splits, preventing performance estimates from being biased by class imbalance. Accordingly, the reported metrics should be interpreted as a controlled shot-level benchmark; physical-sample-level generalisation is not claimed here and is discussed as a limitation.

5.2. Feature extraction and model training

For each pipeline, feature extraction was carried out only on the training data, and the resulting projection parameters were subsequently applied to the test data. This procedure avoided information leakage between train and test sets. LR was employed as a common classifier for PCA, Isomap, PLS-DA, and cepstral features, ensuring comparability across feature representations. An additional logistic-regression baseline was trained directly on the raw spectra. HGB served as a separate reference model on the raw spectra.

To assess the stability of the comparative results with respect to data partitioning, we additionally performed repeated stratified random-split evaluation and report mean values together with standard deviations for the main performance metrics. This analysis is intended to reduce sensitivity to a single random split and to provide a more reliable basis for comparing methods whose performance differences are modest.

5.3. Performance metrics

Performance was assessed using multiple complementary metrics:

Accuracy: the overall proportion of correctly classified spectra.

Precision: the proportion of spectra predicted as a given class that truly belong to that class, reflecting the tendency toward false positives.

Recall: the proportion of spectra from a given class that are correctly identified, reflecting the tendency toward false negatives.

F1 score: the harmonic mean of precision and recall, providing a balanced summary of classification performance when both false positives and false negatives are relevant.

Macro-averaged metrics: precision, recall, and F1 averaged with equal weight across classes, so that each class contributes equally regardless of its sample count.

Weighted metrics: precision, recall, and F1 averaged using class support as weights, so that classes with more test spectra contribute proportionally more to the final score.

Receiver Operating Characteristic (ROC) and Area Under the Curve (AUC): evaluated using a one-vs-rest (OvR) scheme, in which each class is treated in turn as the positive class against all remaining classes. Macro-averaged OvR ROC curves were plotted to assess separability across all thresholds.

Support: the number of test spectra belonging to a given class.

These metrics were reported both in tabular form and as graphical outputs (confusion matrices, ROC curves, and class-wise performance bar charts). The combined use of accuracy, macro-F1, and ROC-AUC ensured that both overall performance and class-level behaviour were captured.

5.4. Visualisation and comparative analysis

To enable direct interpretation of classifier behaviour, confusion matrices were generated for all six methods, highlighting systematic misclassifications such as the overlap between RCA and cement paste, or between wood and foam. ROC curves were compared across methods to illustrate separability. The results of all six methods were also compiled in comparative figures, providing a consistent overview of their relative strengths and weaknesses.

5.5. Rationale for evaluation design

The choice of evaluation metrics and visualisations reflects the specific challenges of LIBS spectral classification. While overall accuracy provides a convenient summary, it can obscure systematic misclassifications in classes with overlapping spectral features. For this reason, the macro-F1 score was emphasised, ensuring that all material classes contributed equally to the performance evaluation, regardless of their intrinsic separability.

Precision and recall were used to disentangle false positive and false negative behaviours, which are particularly relevant in the recycling context: misclassifying a contaminant as aggregate (false negative) directly compromises material quality, whereas misclassifying an aggregate as contaminant (false positive) primarily reduces yield.

The inclusion of ROC curves and AUC values enabled a threshold-independent assessment of separability. This is important in practice, since decision thresholds may vary depending on whether quality assurance prioritises minimising false positives or false negatives. The shape of the ROC curves further reveals whether certain feature extraction methods provide consistent discriminability across all classes or whether their advantage is concentrated in specific regions of the decision space.

Finally, confusion matrices and class-wise bar charts were incorporated to visualise error structures explicitly. These tools highlight recurring misclassification patterns such as the overlap between RCA and cement paste, or the difficulty of distinguishing wood from other organic contaminants. By exposing such patterns, the evaluation goes beyond reporting high-level metrics and provides insight into the operational reliability of each method under realistic multi-class conditions.

Together, this evaluation protocol ensures that the comparative analysis captures not only headline accuracy but also the robustness, interpretability, and operational trade-offs of different feature extraction strategies in LIBS.

6. Results and discussion

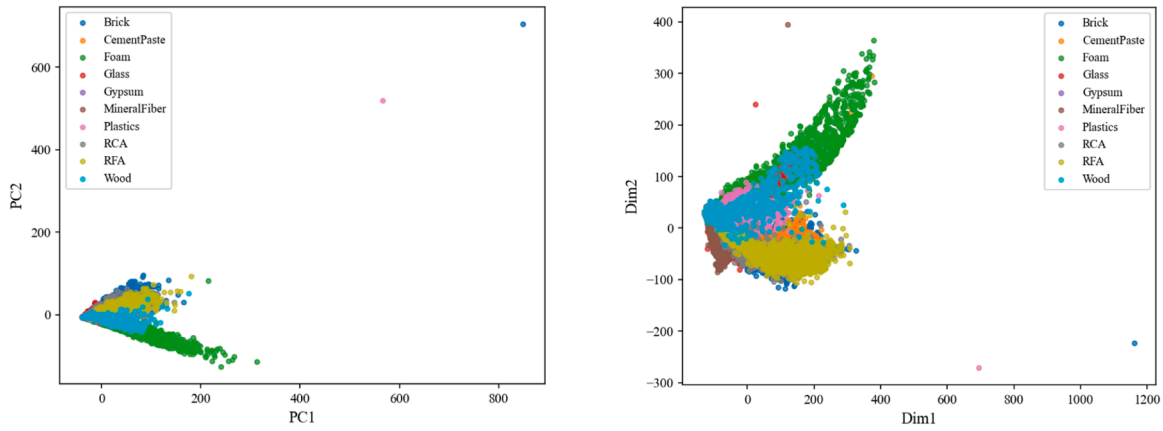
6.1. Feature-space visualisation and representation interpretability

To complement the quantitative benchmark, Fig. 3 visualises the geometry induced by two representative embeddings. Fig. 3(a) shows the two-dimensional PCA embedding of all spectra. Several classes already separate along the dominant variance directions, but substantial overlap remains for challenging pairs such as Foam and Glass, and for cement-related classes such as RCA, RFA, and CementPaste. This is consistent with PCA providing compact and interpretable linear structure, while still being influenced by shared global variance.

Fig. 3(b) shows the corresponding two-dimensional Isomap embedding. The manifold-based projection unfolds part of the non-linear geometry of the dataset, but local overlap and several dispersed points remain. This visual behaviour is consistent with the lower classification stability of Isomap in the benchmark and suggests that non-linear structure is present, but not sufficiently regular to outperform the strongest compact linear and transform-based representations in this dataset.

6.2. Overall classification performance

The comparative analysis shows that the choice of spectral representation has a substantial impact on classification behaviour under conveyor-like LIBS acquisition. Among the evaluated methods, Cepstrum achieves the strongest overall performance. PCA and the raw-



(a)PCA embedding (2D) *(b)Isomap embedding (2D)*

Fig. 3. Feature-space visualisation used to support the interpretation of the benchmark.

spectrum logistic-regression baseline remain closely competitive, indicating that explicit feature extraction is not uniformly beneficial in this dataset. The addition of Isomap broadens the benchmark toward a nonlinear manifold-based embedding, but its performance is clearly below that of the strongest compact linear and transform-based representations. PLS-DA remains the weakest feature-based method, whereas HGB provides a useful reference model with competitive ranking ability but lower overall stability than the best-performing pipelines.

The gap between feature-driven pipelines and the end-to-end baseline is significant: a 2–3 % improvement in macro-F1, when averaged over ten classes, tends to translate into additional contaminants being correctly classified in every thousand samples processed. This margin is meaningful in high-throughput recycling plants, where misclassifications can compromise downstream concrete quality. Figs. 4 and 5 compile these findings. For plant operations, such gains translate into fewer mis-sorted particles at high throughput and therefore higher usable yield of recycled aggregates.

To assess the sensitivity of the benchmark to data partitioning, Fig. 5 summarises performance across repeated stratified random-split runs. The repeated-split results confirm the main ranking observed in the illustrative split shown in Fig. 4: cepstral features remain the strongest compact representation, PCA and the raw-spectrum logistic-regression baseline remain closely competitive, Isomap remains clearly below the strongest compact representations, and PLS-DA remains the least stable method. This repeated-split analysis therefore supports the robustness of the comparative ordering and reduces the risk of over-interpreting a

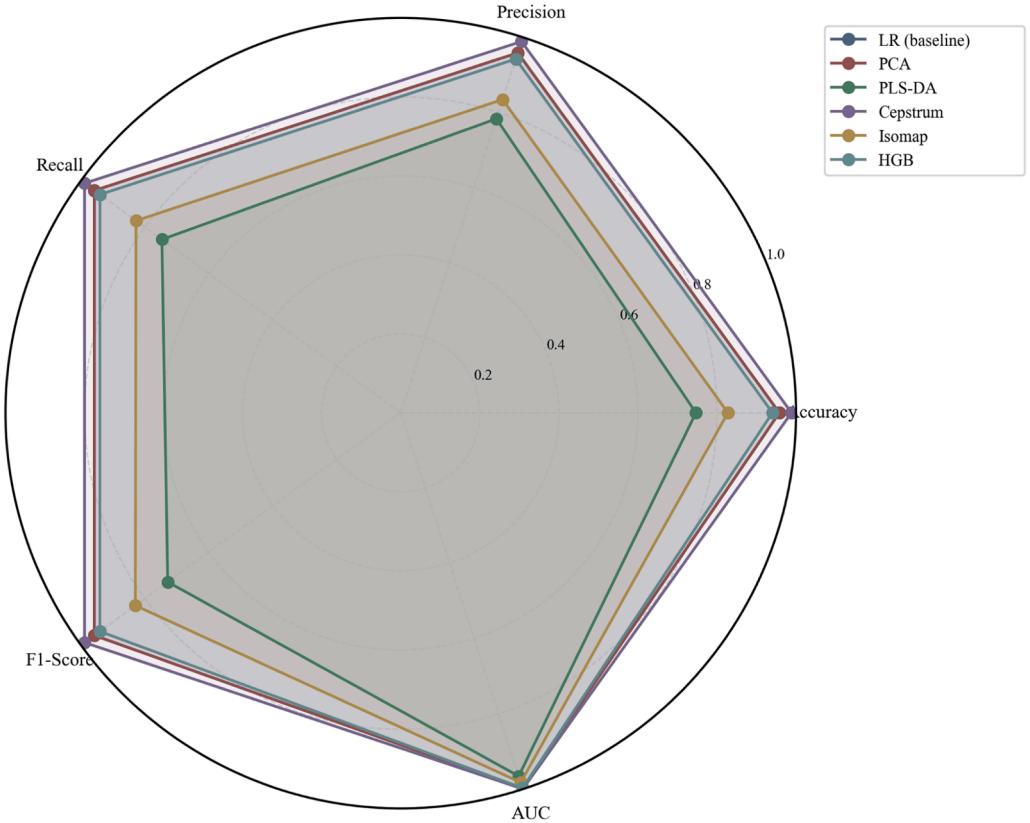


Fig. 4. Comparative performance across six methods in the illustrative stratified train/test split.

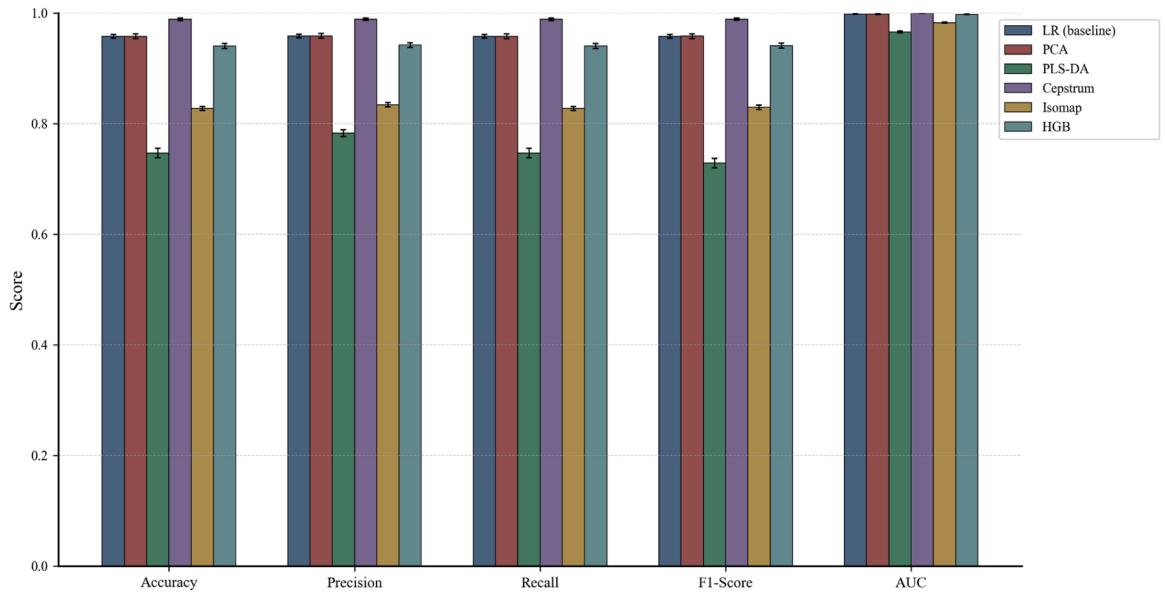


Fig. 5. Comparative performance across repeated stratified random-split evaluation, reported as mean ± std over 10 runs.

single partition.

6.3. Class-wise behaviour

At the class level, some materials are consistently well separated across all methods, whereas others expose the limitations of specific feature spaces (Table 1).

Gypsum, CementPaste, Plastics, and MineralFiber are robustly detected (F1 $>$ 0.970) by both Cepstrum and PCA, indicating that their LIBS spectra contain strong, distinct signatures. For construction quality assurance, such separability supports consistent compliance with material acceptance criteria.

Wood emerges as the most challenging contaminant. Cepstrum yields a recall of 0.806, suggesting that cepstral coefficients may underrepresent some narrow spectral cues. PCA recovers higher recall (0.965) but at the expense of precision, revealing a trade-off between robustness to baseline drift and sensitivity to fine-line details. In practice, this motivates augmenting low-quefrency cepstral vectors with higher-quefrency terms or PCA scores and applying a reject option for borderline spectra.

Foam and Glass show mutual confusion under PCA, with off-diagonal entries in the 3–5 % range, consistent with partially overlapping mineral-related emission signatures and shot-to-shot variability (Fig. 6). A plausible physical explanation is that construction glass is dominated by silicate and alkali/alkaline-earth constituents, so its spectra tend to emphasise Si-rich and alkali-related regions together with Ca-bearing features. Polymeric foam, in contrast, has a much weaker mineral signature and its spectra are more sensitive to surface residues, fillers, or attached fines. Under conveyor-like single-shot acquisition, this combination can yield spectra in which mineral-related regions are comparatively weak or partially masked, leading to overlap with the lower-intensity or more variable glass shots in the PCA space. This interpretation is consistent with the observed bidirectional confusion and highlights that the separation is not driven by a single isolated line, but by relative intensity patterns across mineral-dominated regions. Operationally, this argues for class-specific thresholds or cost-sensitive tuning to protect downstream concrete quality where glass mis-sorts are critical.

RCA vs CementPaste is the most problematic pair in PLS-DA (RCA recall = 0.171), while Cepstrum and PCA both preserve recall $>$0.950. The misclassification reflects a real physical ambiguity: RCA particles often retain surface cement paste, resulting in near-indistinguishable spectra. From a materials perspective, both classes are calcium- and silicate-bearing, and their LIBS spectra share strong cementitious emission regions. Separation therefore depends less on the presence or absence of an element and more on relative contributions. Cement paste is dominated by cement hydration products and typically shows more pronounced cement-related regions, whereas RCA contains aggregate phases together with varying amounts of adhered paste, which can increase the relative contribution of aggregate-related regions. When adhered paste coverage is high, single-shot spectra on RCA surfaces can closely resemble cement paste spectra, which naturally increases confusion. This also explains why supervised projections can become unstable for this pair under strong collinearity, while representations that better control baseline variation remain more robust in this benchmark.

These observations highlight that cepstral features, while generally superior, may attenuate subtle class-specific cues, whereas PCA preserves variance that sometimes aligns better with weaker contaminants.

6.4. Receiver-operating characteristics and error structure

ROC curves (Fig. 7) provide further insight into threshold-independent separability. Cepstrum achieves the highest macro-AUC, while PCA, the raw-spectrum logistic-regression baseline, and HGB remain closely competitive. Isomap shows a clear reduction in separability, and PLS-DA remains the weakest method. These patterns are consistent with the macro-F1 results and indicate that the strongest methods retain their advantage across a broad range of operating thresholds.

Confusion matrices (Fig. 6) reveal compact diagonals for Cepstrum and PCA, with misclassifications concentrated in predictable pairs (Wood↔Foam/Glass, RCA↔CementPaste/RFA). HGB exhibits broader off-diagonals for RCA and Wood, pointing to reduced robustness in borderline classes. PLS-DA, in contrast, produces widespread off-diagonal entries across cementitious and organic classes, reflecting instability of its latent space under spectral collinearity. These patterns justify class-aware decision thresholds and targeted recalibration for RCA–cement paste and wood-adjacent confusions. These error patterns can be related back to compositional similarities between the corresponding materials. For example, gypsum and cement paste exhibit strong mineral signatures and are therefore separated reliably by both PCA and cepstral features, whereas pairs that share overlapping mineral regions or that differ mainly by subtle relative-intensity cues are more

Table 1
Classification Report.

(a) LR

Class	Precision	Recall	F1-score	Support
Brick	0.973	0.960	0.966	480
CementPaste	0.983	0.977	0.980	480
Foam	0.962	0.908	0.935	480
Glass	0.907	0.912	0.910	480
Gypsum	0.988	0.994	0.991	480
MineralFiber	0.976	0.948	0.962	480
Plastics	0.989	0.969	0.979	480
RCA	0.919	0.948	0.933	480
RFA	0.972	0.944	0.958	480
Wood	0.867	0.963	0.912	480
accuracy		0.952		4800
macro avg	0.954	0.952	0.953	4800
weighted avg	0.954	0.952	0.953	4800

(b) PCA

Class	Precision	Recall	F1-score	Support
Brick	0.979	0.958	0.968	480
CementPaste	0.985	0.979	0.982	480
Foam	0.958	0.910	0.934	480
Glass	0.902	0.925	0.914	480
Gypsum	0.994	0.996	0.995	480
MineralFiber	0.980	0.938	0.958	480
Plastics	0.992	0.973	0.982	480
RCA	0.933	0.958	0.946	480
RFA	0.977	0.956	0.966	480
Wood	0.872	0.965	0.916	480
Accuracy	—	—	0.956	4800
Macro avg	0.957	0.956	0.956	4800
Weighted avg	0.957	0.956	0.956	4800

(c) Isomap

Class	Precision	Recall	F1-score	Support
Brick	0.960	0.894	0.926	480
CementPaste	0.899	0.906	0.902	480
Foam	0.853	0.846	0.849	480
Glass	0.741	0.585	0.654	480
Gypsum	0.948	0.956	0.952	480
MineralFiber	0.864	0.821	0.842	480
Plastics	0.875	0.923	0.899	480
RCA	0.663	0.677	0.670	480
RFA	0.738	0.735	0.737	480
Wood	0.558	0.698	0.620	480
accuracy		0.804		4800
macro avg	0.810	0.804	0.805	4800
weighted avg	0.810	0.804	0.805	4800

(d) PLS-DA

Class	Precision	Recall	F1-score	Support
Brick	0.857	0.875	0.866	480
CementPaste	0.858	0.779	0.817	480
Foam	0.935	0.810	0.868	480
Glass	0.574	0.627	0.600	480
Gypsum	0.794	0.990	0.881	480
MineralFiber	0.867	0.854	0.860	480
Plastics	0.676	0.942	0.787	480
RCA	0.845	0.171	0.284	480
RFA	0.603	0.762	0.673	480
Wood	0.286	0.292	0.289	480
Accuracy	—	—	0.710	4800
Macro avg	0.730	0.710	0.693	4800
Weighted avg	0.730	0.710	0.693	4800

(e) Cepstrum

Class	Precision	Recall	F1-score	Support
Brick	0.965	0.977	0.971	480
CementPaste	0.996	0.994	0.995	480
Foam	0.858	0.946	0.900	480
Glass	0.936	0.969	0.952	480
Gypsum	0.998	1.000	0.999	480
MineralFiber	0.998	0.996	0.997	480
Plastics	0.970	0.994	0.981	480

(*continued on next page*)

Table 1 (*continued*)

(e) Cepstrum

Class	Precision	Recall	F1-score	Support
RCA	0.955	0.973	0.964	480
RFA	0.979	0.969	0.974	480
Wood	0.985	0.806	0.887	480
Accuracy	—	—	0.962	4800
Macro avg	0.964	0.962	0.962	4800
Weighted avg	0.964	0.962	0.962	4800

(f) HGB

Class	Precision	Recall	F1-score	Support
Brick	0.974	0.946	0.960	480
CementPaste	0.989	0.969	0.979	480
Foam	0.899	0.906	0.902	480
Glass	0.897	0.927	0.912	480
Gypsum	0.987	0.979	0.983	480
MineralFiber	0.985	0.960	0.973	480
Plastics	0.969	0.967	0.968	480
RCA	0.842	0.923	0.881	480
RFA	0.962	0.894	0.927	480
Wood	0.877	0.894	0.885	480
Accuracy	—	—	0.936	4800
Macro avg	0.938	0.936	0.937	4800
Weighted avg	0.938	0.936	0.937	4800

sensitive to the chosen representation. In this sense, interpretability in our framework refers to connecting dominant regions captured by the representation to plausible material constituents and to explaining recurrent confusions using these compositional overlaps.

6.5. Mechanistic interpretation of representation behaviour

The evaluated methods differ in how they represent or exploit spectral variation, which helps explain the error patterns observed in this benchmark.

Cepstrum is computed from the log-magnitude spectrum and maps spectra into the quefrency domain, which separates slowly varying background trends from finer peak structure. Low-order terms mainly reflect broad baseline variation, whereas higher-order terms capture sharper spectral detail. This separation helps reduce sensitivity to baseline variation and instrument drift in conveyor-like LIBS acquisition. It can be less effective for classes where discrimination relies on subtle narrow-band cues, such as Wood in our dataset.

PCA compresses spectra by projecting onto directions of maximal variance without using class labels. This yields stable dimensionality reduction and competitive separation for most classes. However, variance that is not class-specific can dominate when classes share similar emission patterns. The Foam and Glass confusion is consistent with this limitation, where shared variance can reduce separability in the PCA space.

Isomap seeks a low-dimensional embedding that preserves neighbourhood structure through geodesic distances on a sample graph. In principle, this allows non-linear class geometry to be unfolded relative to linear projections such as PCA. In the present dataset, however, the manifold embedding remains more sensitive to local sampling structure and spectral overlap, which likely contributes to its lower overall stability and weaker classification performance. This suggests that non-linear structure is present but not sufficiently regular to outperform the more compact and stable linear or transform-based representations in this benchmark.

PLS-DA introduces supervision by constructing latent variables that align spectral variation with class labels. In high-dimensional LIBS data with strong channel collinearity, these latent directions can become sensitive to sampling and overlap between classes. In our benchmark, this behaviour is reflected in poorer separation for overlapping material pairs such as RCA and CementPaste, and in reduced macro-F1 relative to PCA and Cepstrum.

HGB learns decision rules directly from the raw spectra and can achieve competitive ranking performance as reflected by AUC. In this dataset, the lower macro-F1 suggests that threshold-based decisions can be less stable for borderline spectra. In addition, the lack of a direct mapping from model rules to emission features limits diagnostic interpretation compared with representations such as PCA loadings or cepstral coefficients.

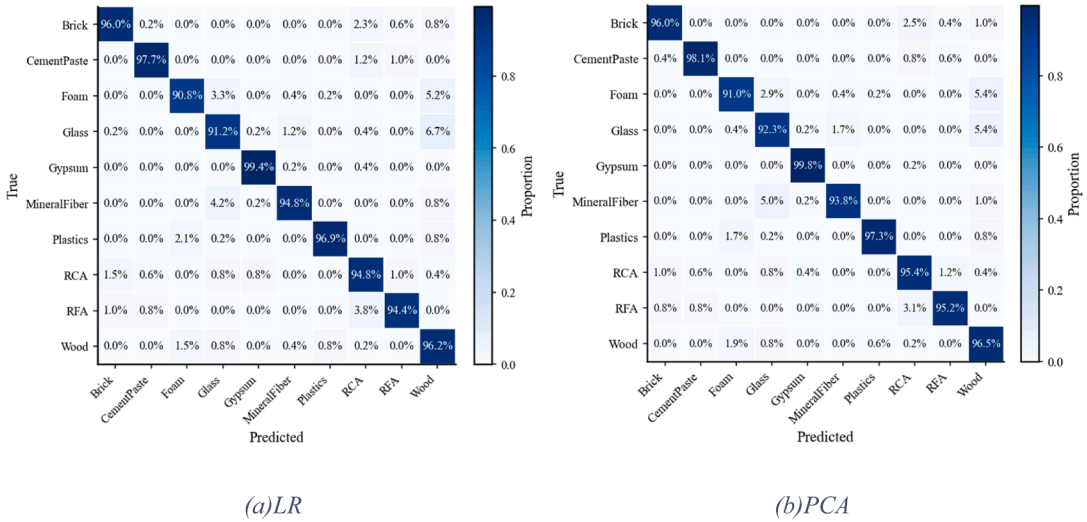
6.6. Implications for industrial LIBS sorting, scope, and limitations

The discussion below is framed by the evaluation design. Results are obtained under a controlled, class-balanced, shot-level benchmark, and the train/test split is performed at the shot level. The metrics therefore characterise classifier behaviour on held-out spectra under these controlled conditions. System-level performance under continuous mixed-material conveyor operation is not evaluated in this study.

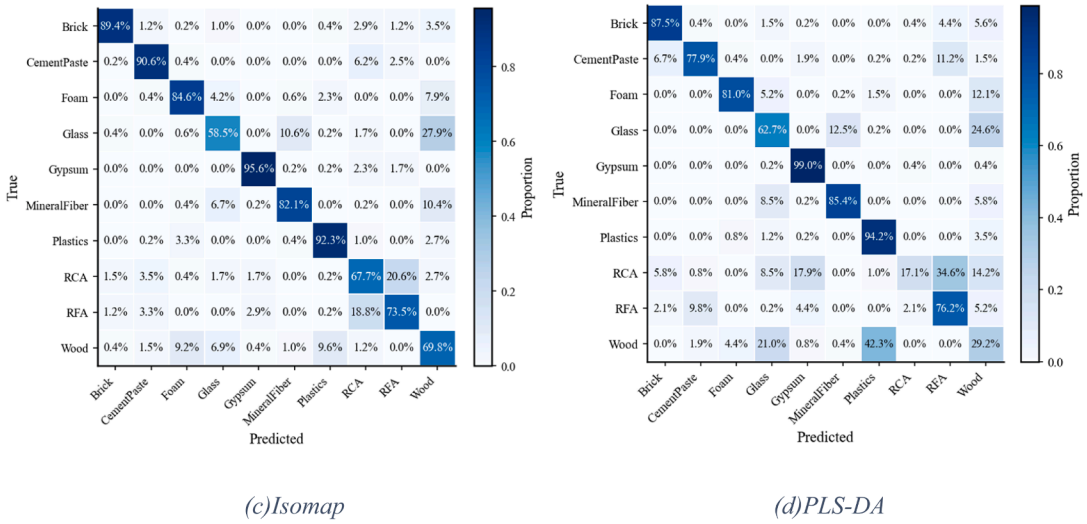
Within this benchmark, representation choice has a larger impact on macro-F1 and class-wise error structure than switching between the evaluated classifiers. This supports treating spectral representation as a primary design variable when benchmarking robustness to baseline variation and redundancy in LIBS spectra. It also provides practical guidance for selecting compact features under compute constraints.

For deployment-oriented design, cepstral features provide the strongest compact representation when baseline variation is a dominant source of instability. PCA offers a close alternative, with the added benefit that loadings can be linked to elemental emission regions for diagnostic interpretation. The raw-spectrum logistic-regression baseline shows that acceptable classification performance can already be obtained without explicit dimensionality reduction in this dataset, but this comes at the cost of reduced compactness and weaker interpretability of the input representation. Isomap provides a useful non-linear reference, yet its lower stability and stronger dependence on neighbourhood structure make it less immediately attractive for transparent deployment than PCA or cepstral features. PLS-DA remains less suitable in this setting due to instability under collinear channels and class overlap, while HGB serves as a useful reference model but with more limited interpretability.

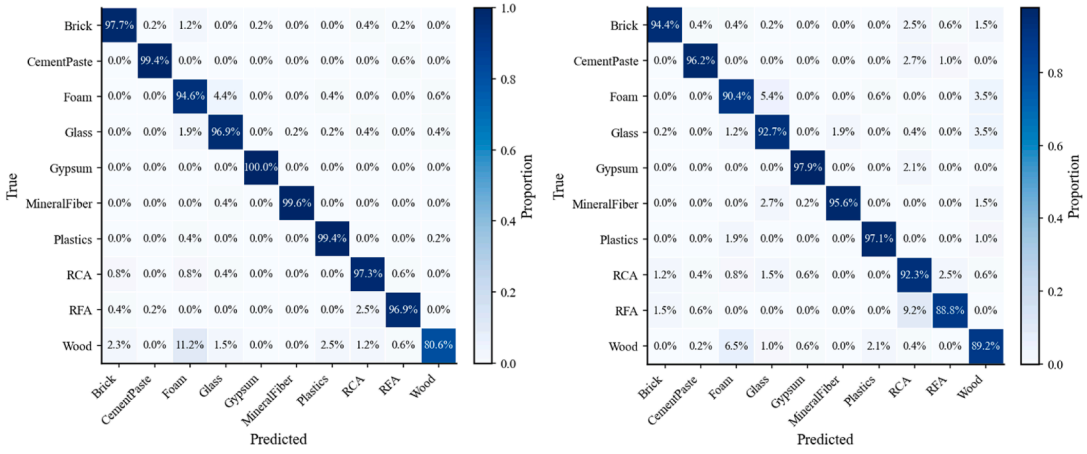
From a computational perspective, compact representations reduce the inference burden by compressing spectra from about 12k channels to a few dozen features. This supports predictable, low-latency inference with linear classifiers. In contrast, tree-based models can introduce sharper threshold effects, which may increase decision variability for



(a)LR (b)PCA



(c)Isomap (d)PLS-DA



(e)Cepstrum (f)HGB

Fig. 6. Confusion matrices for six methods.

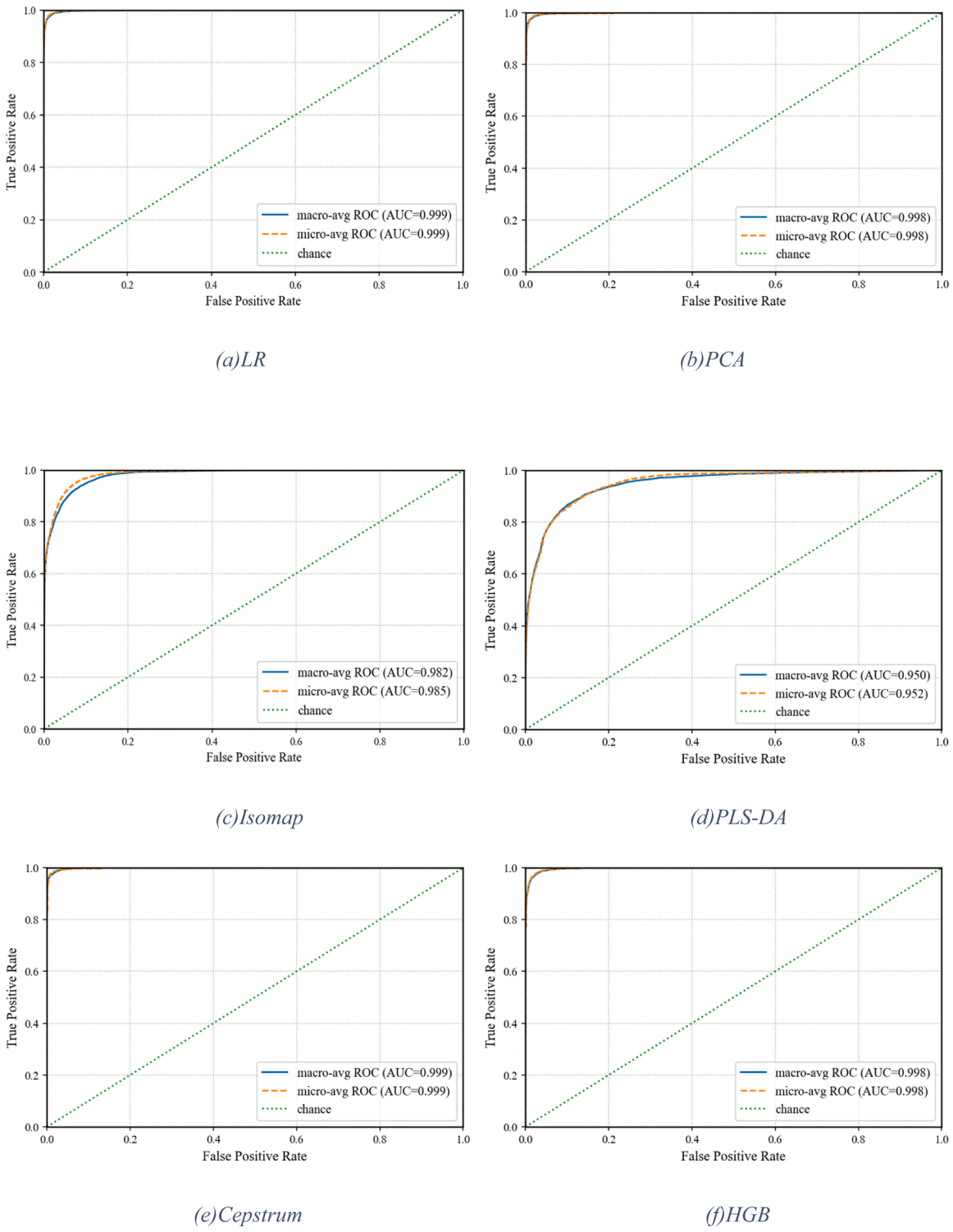


Fig. 7. ROC curves for six methods.

spectra near class boundaries.

In practical deployment terms, the benchmark suggests two complementary design routes. When baseline variation and background fluctuations dominate, cepstral features offer a compact representation that reduces sensitivity to these effects while keeping inference lightweight. When interpretability and traceability to elemental emission regions are priorities, PCA provides competitive performance together with physically meaningful loadings that support diagnostics and audit. By contrast, PLS-DA should be treated cautiously in multi-class LIBS settings with strong collinearity and class overlap, and end-to-end models such as HGB are best used as performance references rather than as diagnostic tools, given the limited transparency of their decision logic.

Two aspects of the current evaluation constrain how far the results can be generalised beyond this controlled benchmark. First, the split is performed at the shot level on spectra drawn from pooled reference batches, rather than at the level of distinct physical specimens. As a result, the test set is independent at the level of held-out shots, but it

does not guarantee independence with respect to specimen-specific factors such as surface condition, composition micro-heterogeneity, or local contamination patterns. This can inflate apparent generalisation relative to a setting where the model must transfer to unseen specimens, and it particularly affects conclusions about robustness for classes that exhibit strong within-material variability. Second, the benchmark is class-balanced and does not evaluate continuous mixed-material conveyor operation in which classes appear sequentially or interleaved, with changing class priors and transition regions between materials. In such streams, system-level performance depends not only on per-shot classification accuracy, but also on temporal consistency, the handling of borderline spectra at material boundaries, and the cost of false positives and false negatives under realistic class frequencies. Therefore, the reported metrics should be read as evidence of comparative classifier behaviour under controlled conditions. Demonstrating deployment-level performance will require validation with physical-sample-level separation and stream-level testing under mixed-material conveyor scenarios.

7. Conclusion and outlook

This study introduces and validates a representation-first framework for LIBS-based quality assessment of recycled aggregates. Using principal components, Isomap embeddings, cepstral coefficients, and partial least-squares discriminant analysis as case studies, together with a raw-spectrum logistic-regression baseline and an HGB reference model, this work supports a representation-focused perspective on LIBS classification design within the present controlled benchmark, while also showing that the benefit of explicit feature extraction depends on the specific representation chosen.

In our experiments, the combination of cepstral features with LR achieved the best balance of accuracy, robustness, and efficiency, consistent with the value of separating baseline envelopes from fine spectral lines for noisy, high-dimensional spectra. The combination of principal components with LR follows closely, offering lightweight computation and clear interpretability, with slightly higher confusion where narrow-line classes overlap. Partial least-squares discriminant analysis, while label-driven, can be sensitive to collinearity and overlapping classes without careful regularisation. The end-to-end tree-ensemble baseline, although competitive in receiver-operating-characteristic behaviour, offers less transparent decision logic and showed lower boundary stability than the strongest feature-based pipelines in this benchmark, which may limit its usefulness in auditable operations.

Methodologically, the results illustrate that tools originally developed in acoustics can be informative for spectroscopic data analysis. This cross-domain transfer highlights the continued value of principled feature extraction for high-dimensional, drift-prone spectra and related sensing tasks, especially when interpretability and robustness matter alongside nominal accuracy.

From a deployment-oriented perspective, the present benchmark suggests that pipelines based on cepstral features or principal components are promising candidates for inline contaminant detection under latency and resource constraints. However, these implications should be interpreted cautiously, since the current study evaluates controlled held-out spectra rather than continuous mixed-material conveyor operation. Accordingly, any expected reduction in mis-sorting, yield improvement, or downstream quality benefit remains to be validated at the system level.

Within the present benchmark, our contribution is twofold: first, we show that cepstral analysis is a strong candidate for spectra in which envelope-line separation helps mitigate baseline variation and fine-structure overlap; second, by framing inline sorting as a constrained systems problem, we clarify why smooth, interpretable, and compact feature spaces can be preferable to more complex models when robustness, auditability, and computational practicality must be considered together.

Future work will: (a) quantify sensitivity to cepstral order and principal-component dimensionality under explicit latency and computational-footprint budgets, reporting per-spectrum transform time and model size; (b) integrate adaptive drift compensation by monitoring low-quefrency coefficients or principal-component loadings as triggers, with documented recalibration intervals; (c) develop hybrid designs that couple physics-informed representations with modern classifiers while preserving traceability for quality audits; (d) conduct physical-sample-level and cross-plant validation on mixed streams and higher conveyor speeds, with systematic reporting of inference latency, batch-to-batch stability, and yield impact.

In summary, this work provides a controlled representation-focused benchmark for LIBS-based quality assessment of recycled aggregates. Within this benchmark, spectral representation remains a major source of performance variation, but the benefit of explicit feature extraction depends on the specific representation chosen and should not be treated as universal across all methods. Compact and interpretable representations such as cepstral features and principal components remain especially attractive for deployment-oriented LIBS design, while broader system-level validation under mixed-material conveyor operation remains an essential next step.

CRediT authorship contribution statement

Cheng Chang: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Siwei Peng:** Visualization, Investigation, Formal analysis. **Xiaorong Wang:** Writing – original draft, Visualization. **Shuai Zong:** Validation, Data curation. **Wei Hu:** Visualization, Software, Conceptualization. **Hao Cheng:** Conceptualization. **Francesco Di Maio:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was partially supported by the European Union's Horizon 2020-funded Project "Innovative Circular Economy Based solutions demonstrating the Efficient Recovery of valuable material resources from the Generation of representative end-of-life building materials" (ICEBERG, grant agreement No. 869336).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.rineng.2026.110902.

Data availability

Data will be made available on request.

References

[1] A. Abushanab, W. Alnahhal, Life cycle cost analysis of sustainable reinforced concrete buildings with treated wastewater, recycled concrete aggregates, and fly ash, Results Eng. 20 (2023) 101565, https://doi.org/10.1016/J.RINENG.2023.101565.

[2] M. Casale, C. Oggeri, P. Rossi, G.A. Dino, Performance assessment of recycled aggregates from C&DW plants for road embankments: influence of composition and treatment methods, Results Eng. 28 (2025) 107340, https://doi.org/10.1016/J.RINENG.2025.107340.

[3] C. Chang, F. Di Maio, R. Bheemireddy, P. Posthoorn, A.T. Gebremariam, P. Rem, Intelligent optimization of particle size distribution in unscreened recycled coarse aggregates using 3D surface analysis, J. Ind. Inf. Integr. 46 (2025) 100864, https://doi.org/10.1016/J.JII.2025.100864.

[4] C. Chang, F. Di Maio, R. Bheemireddy, P. Posthoorn, A.T. Gebremariam, P. Rem, Rapid quality control for recycled coarse aggregates (RCA) streams: multi-sensor integration for advanced contaminant detection, Comput. Ind. 164 (2025) 104196, https://doi.org/10.1016/J.COMPIND.2024.104196.

[5] Z.M. Bakr, D.K.I. Jaf, Performance enhancement of recycled aggregate concrete (RAC) incorporating pre-soaked recycled concrete aggregate (RCA) and silica-based additives, Results Eng. 28 (2025) 108318, https://doi.org/10.1016/J.RINENG.2025.108318.

[6] J. Li, X. Zhang, X. Xu, H. Wu, Fully recycled aggregate concrete composite functional additives: proportioning tests and verification of engineering adaptability, Results Eng. 27 (2025) 106910, https://doi.org/10.1016/J.RINENG.2025.106910.

[7] E. Harrison, A. Berenjian, M. Seifan, Recycling of waste glass as aggregate in cement-based materials, Env. Sci. Ecotechnol. 4 (2020) 100064, https://doi.org/10.1016/J.ESE.2020.100064.

[8] S. Hameed, S.A.A. Gillani, M. Tahir, R. Hameed, S. Abbas, M.L. Nehdi, M. Ahmad, Investigating lightweight recycled brick aggregate concrete incorporating EPS beads: application to masonry units, Results Eng. 25 (2025) 104019, https://doi.org/10.1016/J.RINENG.2025.104019.

[9] T.P. Huynh, T.M. Le, N.C. Ngan, An experimental evaluation of the performance of concrete reinforced with recycled fibers made from waste plastic bottles, Results Eng. 18 (2023) 101205, https://doi.org/10.1016/J.RINENG.2023.101205.

[10] F.W. Al-Awabdeh, M.J. Al-Kheetan, Y.S. Jweihan, H. Al-Hamaiedeh, S.H. Ghaffar, Comprehensive investigation of recycled waste glass in concrete using silane treatment for performance improvement, Results Eng. 16 (2022) 100790, https://doi.org/10.1016/J.RINENG.2022.100790.

[11] F.J. Fortes, J. Moros, P. Lucena, L.M. Cabalín, J.J. Laserna, Laser-induced breakdown spectroscopy, Anal. Chem. 85 (2013) 640–669, https://doi.org/10.1021/AC303220R/ASSET/IMAGES/LARGE/AC-2012-03220R_0007.JPEG.

[12] V. Majidi, M.R. Joseph, Spectroscopic applications of laser-induced plasmas, Crit. Rev. Anal. Chem. 23 (1992) 143–162, https://doi.org/10.1080/10408349208050852;REQUESTEDJOURNAL:JOURNAL:BATC20;WGROUP:STRING:PUBLICATION.

[13] D.A. Cremers, L.J. Radziemski, Handbook of Laser-Induced Breakdown Spectroscopy: Second Edition, 2nd Edition, Handbook of Laser-Induced Breakdown Spectroscopy, 2013, pp. 1–423, https://doi.org/10.1002/9781118567371.

[14] D. Fernandes Andrade, E.R. Pereira-Filho, D. Amarasiriwardena, Current trends in laser-induced breakdown spectroscopy: a tutorial review, Appl. Spectrosc. Rev. 56 (2021) 98–114, https://doi.org/10.1080/05704928.2020.1739063.

[15] G.S. Senesi, Laser-induced breakdown spectroscopy (LIBS) applied to terrestrial and extraterrestrial analogue geomaterials with emphasis to minerals and rocks, Earth. Sci. Rev. 139 (2014) 231–267, https://doi.org/10.1016/J.EARSCIREV.2014.09.008.

[16] R.S. Harmon, R.E. Russo, R.R. Hark, Applications of laser-induced breakdown spectroscopy for geochemical and environmental analysis: a comprehensive review, Spectrochim. Acta B At. Spectrosc. 87 (2013) 11–26, https://doi.org/10.1016/J.SAB.2013.05.017.

[17] T. Chen, T. Zhang, H. Li, Applications of laser-induced breakdown spectroscopy (LIBS) combined with machine learning in geochemical and environmental resources exploration, TrAC Trends Anal. Chem. 133 (2020) 116113, https://doi.org/10.1016/J.TRAC.2020.116113.

[18] S. Moncayo, S. Manzoor, F. Navarro-Villoslada, J.O. Caceres, Evaluation of supervised chemometric methods for sample classification by laser induced breakdown spectroscopy, Chemom. Intell. Lab. Syst. 146 (2015) 354–364, https://doi.org/10.1016/J.CHEMOLAB.2015.06.004.

[19] C. Chang, F. Di Maio, P. Rem, A.T. Gebremariam, F. Mehari, H. Xia, Cluster-based identification algorithm for in-line recycled concrete aggregates characterization using laser-induced breakdown spectroscopy (LIBS), Resour. Conserv. Recycl. 185 (2022) 106507, https://doi.org/10.1016/J.RESCONREC.2022.106507.

[20] C. Chang, F. Di Maio, R. Bheemireddy, P. Posthoorn, A.T. Gebremariam, P. Rem, Enhancing quality inspection efficiency and reliability of unscreened recycled coarse aggregates (RCA) streams using innovative mobile sensor-based technology, Dev. Built. Environ. 21 (2025) 100611, https://doi.org/10.1016/J.DIBE.2025.100611.

[21] C. Chang, F. Di Maio, P. Rem, Real-time sensor-based characterization of recycled coarse aggregates (RCA): advancing sustainable construction through automated quality assessment, in: Proceedings of The Sixth International Conference, 2025, pp. 656–667.

[22] C. Chang, Quality assessment & monitoring of recycled coarse aggregates, 2025. https://doi.org/10.4233/uuid:ca80614c-747f-4ef8-95aa-78c0b193d261.

[23] B.K. Alsberg, D.B. Kell, R. Goodacre, Variable selection in discriminant partial least-squares analysis, Anal. Chem. 70 (1998) 4126–4133, https://doi.org/10.1021/AC980506O/ASSET/IMAGES/LARGE/AC980506OF00005.JPEG.

[24] I. Marquetti, J.V. Link, A.L.G. Lemes, M.B. dos S. Scholz, P. Valderrama, E. Bona, Partial least square with discriminant analysis and near infrared spectroscopy for evaluation of geographic and genotypic origin of arabica coffee, Comput. Electron. Agric. 121 (2016) 313–319, https://doi.org/10.1016/J.COMPAG.2015.12.018.

[25] K.Y. Peerbhay, O. Mutanga, R. Ismail, Commercial tree species discrimination using airborne AISA Eagle hyperspectral imagery and partial least squares discriminant analysis (PLS-DA) in KwaZulu–Natal, South Africa, ISPRS. J. Photogramm. Remote Sens. 79 (2013) 19–28, https://doi.org/10.1016/J.ISPRSJPRS.2013.01.013.

[26] L. Lenhardt, R. Bro, I. Zeković, T. Dramićanin, M.D. Dramićanin, Fluorescence spectroscopy coupled with PARAFAC and PLS DA for characterization and classification of honey, Food Chem. 175 (2015) 284–291, https://doi.org/10.1016/J.FOODCHEM.2014.11.162.

[27] A.M. Jiménez-Carvelo, S. Martín-Torres, F. Ortega-Gavilán, J. Camacho, PLS-DA vs sparse PLS-DA in food traceability. A case study: authentication of avocado samples, Talanta 224 (2021) 121904, https://doi.org/10.1016/J.TALANTA.2020.121904.

[28] L. Jolivet, M. Leprince, S. Moncayo, L. Sorbier, C.P. Lienemann, V. Motto-Ros, Review of the recent advances and applications of LIBS-based imaging, Spectrochim. Acta B At. Spectrosc. 151 (2019) 41–53, https://doi.org/10.1016/J.SAB.2018.11.008.

[29] L. Brunnbauer, Z. Gajarska, H. Lohninger, A. Limbeck, A critical review of recent trends in sample classification using laser-induced breakdown spectroscopy (LIBS), TrAC Trends Anal. Chem. 159 (2023) 116859, https://doi.org/10.1016/J.TRAC.2022.116859.

[30] G. Galbács, A critical review of recent progress in analytical laser-induced breakdown spectroscopy, Anal. Bioanal. Chem. 407 (25) (2015) 7537–7562, https://doi.org/10.1007/S00216-015-8855-3, 2015407.

[31] F. Boué-Bigne, Laser-induced breakdown spectroscopy and multivariate statistics for the rapid identification of oxide inclusions in steel products, Spectrochim. Acta B At. Spectrosc. 119 (2016) 25–35, https://doi.org/10.1016/J.SAB.2016.02.018.

[32] S. Chowdhury, A.K. Saha, D.K. Das, R2 Score estimation with various time series models incorporating 1000 epochs for optimization of accuracy of renewable energy future trends, in: Proceedings of the 4th International Conference on Ubiquitous Computing and Intelligent Information Systems, ICUIS 2024, 2024, pp. 1906–1911, https://doi.org/10.1109/ICUIS64676.2024.10867081.

[33] A. Rösner, A. Gegov, A. Hopgood, O. Lamptey, D. Ouelhadj, S. Da Deppo, Machine learning-based classification of extremism using explainable artificial intelligence, in: International IEEE Conference Proceedings, IS, 2024, https://doi.org/10.1109/IS61756.2024.10705241.

[34] Z. Haider Jaffari, H. Jeong, J. Shin, J. Kwak, C. Son, Y.G. Lee, S. Kim, K. Chon, K. Hwa Cho, Machine-learning-based prediction and optimization of emerging contaminants' adsorption capacity on biochar materials, Chem. Eng. J. 466 (2023) 143073, https://doi.org/10.1016/J.CEJ.2023.143073.

[35] Y. Ahmed, K.R. Dutta, S.N.C. Nepu, M. Prima, H. AlMohamadi, P. Akhtar, Optimizing photocatalytic dye degradation: a machine learning and metaheuristic approach for predicting methylene blue in contaminated water, Results Eng. 25 (2025) 103538, https://doi.org/10.1016/J.RINENG.2024.103538.

[36] H. Xia, M.C.M. Bakker, Reliable classification of moving waste materials with LIBS in concrete recycling, Talanta 120 (2014) 239–247, https://doi.org/10.1016/J.TALANTA.2013.11.082.

[37] P. Pořízka, J. Klus, E. Képeš, D. Prochazka, D.W. Hahn, J. Kaiser, On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review, Spectrochim. Acta B At. Spectrosc. 148 (2018) 65–82, https://doi.org/10.1016/J.SAB.2018.05.030.

[38] V.C. Costa, F.W.B. Aquino, C.M. Paranhos, E.R. Pereira-Filho, Identification and classification of polymer e-waste using laser-induced breakdown spectroscopy (LIBS) and chemometric tools, Polym. Test. 59 (2017) 390–395, https://doi.org/10.1016/J.POLYMERTESTING.2017.02.017.

[39] J.B. Tenenbaum, V. De Silva, J.C. Langford, A global geometric framework for nonlinear dimensionality reduction, Science 290 (1979) (2000) 2319–2323, https://doi.org/10.1126/SCIENCE.290.5500.2319;ISSUE:ISSUE:DOI.

[40] D. Lunga, S. Prasad, M.M. Crawford, O. Ersoy, Manifold-learning-based feature extraction for classification of hyperspectral data: a review of advances in manifold learning, IEEE Signal. Process. Mag. 31 (2014) 55–66, https://doi.org/10.1109/MSP.2013.2279894.

[41] Y. Chen, M.M. Crawford, J. Ghosh, Applying nonlinear manifold learning to hyperspectral data for land cover classification, in: International Geoscience and Remote Sensing Symposium (IGARSS), 2005, pp. 4311–4314, https://doi.org/10.1109/IGARSS.2005.1525872, 6.

[42] X. Li, S. Yang, R. Fan, X. Yu, D. Chen, Discrimination of soft tissues using laser-induced breakdown spectroscopy in combination with k nearest neighbors (kNN) and support vector machine (SVM) classifiers, Opt. Laser. Technol. 102 (2018) 233–239, https://doi.org/10.1016/J.OPTLASTEC.2018.01.028.

[43] Y. Yu, M. Yao, When convolutional neural networks meet laser-induced breakdown spectroscopy: end-to-end quantitative analysis modeling of ChemCam spectral data for major elements based on ensemble convolutional neural networks, Remote Sens. 15 (2023) 3422, https://doi.org/10.3390/RS15133422, 2023Page 3422 15.

[44] J. Castorena, D. Oyen, A. Ollila, C. Legget, N. Lanza, Deep spectral CNN for laser induced breakdown spectroscopy, Spectrochim. Acta B At Spectrosc. 178 (2021) 106125, https://doi.org/10.1016/J.SAB.2021.106125.

[45] X. Xu, F. Ma, J. Zhou, C. Du, Applying convolutional neural networks (CNN) for end-to-end soil analysis based on laser-induced breakdown spectroscopy (LIBS) with less spectral preprocessing, Comput. Electron. Agric. 199 (2022) 107171, https://doi.org/10.1016/J.COMPAG.2022.107171.

[46] M. Grandini, E. Bagli, G. Visani, Metrics for multi-class classification: an overview, (2020). https://arxiv.org/pdf/2008.05756 (accessed August 28, 2025).

[47] Laser-induced breakdown spectroscopy (LIBS) scanning data of recycled building materials (dataset), (n.d.). https://data.4tu.nl/datasets/f8155313-9b68-46cc-ba4d-8f473cfdc190/1 (accessed September 2, 2025).
[48] J.T. Vogelstein, E.W. Bridgeford, M. Tang, D. Zheng, C. Douville, R. Burns, M. Maggioni, Supervised dimensionality reduction for big data, Nat. Commun. 12 (2021) 1–9, https://doi.org/10.1038/S41467-021-23102-2/MEDIAOBJECTS/41467_2021_23102_MOESM2_ESM.PDF.
[49] P. Ray, S.S. Reddy, T. Banerjee, Various dimension reduction techniques for high dimensional data analysis: a review, Artif. Intell. Rev. 54 (2021) 3473–3515, https://doi.org/10.1007/S10462-020-09928-0/FIGURES/27.