EPiC Series in Built Environment

Volume 7, 2026, Pages 595–604

Proceedings of Associated Schools of Construction 62nd Annual International Conference

Real-Time Hyperspectral Classification of Visually Similar Materials: A Pilot Study

Rana Muhammad Irfan Anwar [1], Eric M. Wetzel[1]

[1]Auburn University

Visual inspection and Red, Green, Blue (RGB) cameras remain the dominant method for assessing surface finish and quality in construction materials. However, this approach is inherently subjective and constrained by human color perception, particularly with visually similar coatings. Similarly, conventional RGB imaging also fails to capture the subtle spectral cues that distinguish surface finishes sharing similar color and texture, leading to potential inaccuracies when verifying completed work. Reliable differentiation of surface coatings is essential for automated progress tracking applications. To address these limitations, this pilot study investigates the use of hyperspectral imaging (HSI) for the non-destructive differentiation of coated and uncoated gypsum surfaces under controlled lighting. Experiments were conducted in a laboratory under controlled illumination as a pilot study, using a Kelvin Play LED source. The lighting temperature varied from 6000 K to 8000 K to improve spectral separability. Data was collected as hyperspectral files in JSON format and applied Min-max normalization. Two classifiers, Support Vector Machine (SVM) and Random Forest Classifier (RFC), were trained and evaluated using spectra. The RFC achieved over 95% accuracy in real-time classification under 8000 K illumination. The live system, implemented via the Living Optics SDK (v1.9.0), predicted surface types directly on a grayscale camera feed with OpenCV overlays. The results confirm that hyperspectral sensing, coupled with optimized lighting and machine learning, can enable reliable, real-time differentiation of construction surface materials. The findings establish a strong foundation for extending hyperspectral inspection to mobile robotic platforms for autonomous, on-site progress tracking.

Keywords: Hyperspectral Imaging, Progress Tracking, Machine Learning

Introduction

Accurate progress tracking is essential to construction project management (Zou et al., 2017). Project schedules, billing cycles, and productivity analytics all rely on knowing precisely which materials or finishing stages have been completed on-site (Kim et al., 2013). For example, a wall that has been primed but not yet coated with its final semi-gloss layer may appear, under bright or inconsistent lighting, nearly identical to one that is fully finished. When progress assessments depend on visual documentation or RGB-camera imagery, these subtle differences can be easily overlooked, leading to reporting errors, miscommunication, or premature sign-off of incomplete work.

Traditional visual inspection methods include manual walkthroughs, image-based tracking, and 360-degree video cameras (Ekanayake et al., 2021). Modern market solutions such as OpenSpace and Track3D employ these image data streams to algorithmically estimate percentage completion across trades and building areas (OpenSpace, 2025; Track3D, 2025). While these platforms show strong potential for automating site documentation, they remain influenced by human subjectivity and the spectral limitations inherent to conventional RGB cameras (Musarat et al., 2024). RGB imaging and video ingest only three bands (red, green, and blue), providing limited sensitivity to material composition. As a result, visually similar surfaces, such as plain gypsum, primer-coated, and painted boards, produce nearly indistinguishable pixel values in standard photographs—particularly under variable jobsite lighting. In many real-world conditions, this makes it challenging for automated systems to reliably differentiate between early, intermediate, and final finishing stages. Although vendors continue to improve their algorithms and integrate additional sensing modalities (e.g., depth or LiDAR), current RGB-based approaches still face inherent constraints when fine-grained material distinctions are required.

Limitations of RGB and Multispectral Approaches in Construction Monitoring

The construction industry has increasingly adopted visual analytics and computer-vision tools to automate progress tracking, site documentation, and quality assurance (Sami Ur Rehman et al., 2022). Most of these systems rely on standard RGB cameras, drones, or mobile devices to capture periodic site images for assessing percent complete or generating as-built models. While cost-effective and widely accessible, RGB imaging captures only three broad spectral bands, red, green, and blue, each encompassing wide wavelength intervals ($\approx$ 100 nm). This narrow spectral sampling provides color information but very little insight into a material's chemical or physical composition.

Consequently, when surfaces differ primarily in coating type or finish rather than color, such as unpainted gypsum, primed wallboard, and painted finishes, the pixel values recorded by RGB sensors are nearly indistinguishable. Algorithms trained on RGB data often interpret all these surfaces as visually identical, classifying them under a single category. Such misclassification directly affects automated progress estimation by overstating completion levels and can mislead project stakeholders during remote inspections.

Multispectral imaging (MSI) offers a partial improvement by adding several discrete spectral bands beyond RGB, typically in the near-infrared region (Tretter et al., 2005). However, MSI systems collect non-contiguous and limited bands, often fewer than twenty, providing only coarse spectral resolution. This is insufficient to capture the subtle reflectance shifts caused by light absorption on visually similar materials. In addition, both RGB and MSI techniques remain highly sensitive to illumination variability, shadows, and camera calibration drift, all of which are common in construction environments with mixed lighting conditions. As a result, these methods cannot reliably determine whether a wall is in the priming, intermediate, or final coating stage, a distinction crucial for accurate progress tracking and quality control.

Hyperspectral Imaging as a Material Diagnostic and Progress-Tracking Tool

Hyperspectral imaging (HSI) overcomes these limitations by capturing a contiguous spectrum, often hundreds of narrow bands, each only 5 - 10 nm wide, for every pixel in an image (Bhargava et al., 2024). Unlike RGB or MSI, HSI provides both spatial and spectral information, producing a "hypercube" that characterizes each surface based on its unique reflectance signature. Because these signatures are determined by the intrinsic optical and chemical properties of materials rather than surface color alone, HSI can differentiate coatings, pigments, and binders that appear visually identical.

HSI's strength lies in combining imaging and spectroscopy, allowing per-pixel spectral analysis. A study by Ptacek et al (2021) demonstrated that Near Infrared Region (NIR) spectra distinguished curing regimes and ages of young concrete. Clustering reliably differentiated well-cured vs poorly cured samples. Demonstrated a non-destructive method to evaluate curing quality (Ptacek et al., 2021). Another study by Vitek collected spectral data from common construction and demolition (C&D) materials. When a Multilayer Perceptron (MLP) model was applied, results showed that adding only two near-infrared bands (approximately 700 nm and 900 nm) to the standard RGB channels significantly increased classification accuracy. These additional bands greatly enhanced material discrimination, demonstrating that real-time sorting can be achieved with only a few carefully selected spectral channels. (Vítek et al., 2025).

In construction applications, this capability translates into objective material-stage identification. For instance, the calcium-sulfate base of raw gypsum exhibits a broad, diffuse reflectance across the visible spectrum. Primer coatings alter the visible reflectance spectrum through the addition of pigments, binders, and polymeric films that reduce diffuse scattering. Semi-gloss paints, in contrast, introduce stronger specular reflection components, producing more pronounced angle-dependent highlights across visible and near-infrared wavelengths. Capturing these features allows HSI to determine precisely which stage of wall finishing has been achieved. When integrated into automated workflows, these spectral insights can enhance progress-tracking systems, enabling accurate percent-complete estimation and minimizing traditional camera errors.

Furthermore, recent advances in snapshot hyperspectral cameras, such as the Living Optics (LO) camera used in this study, have made real-time imaging feasible. Combined with controlled illumination and lightweight machine-learning pipelines, HSI can deliver on-site, instantaneous classification of construction materials. This creates new possibilities for robotic inspection, BIM-linked progress analytics, and data-driven quality assurance, moving beyond traditional visual and camera-based inspection toward fully autonomous construction monitoring.

Research Objectives

The objectives were:

1. To design and execute a controlled hyperspectral experiment comparing plain gypsum, primer-coated, and semi-gloss-coated surfaces (off the shelf paint coatings).

2. To develop and test a real-time classification pipeline integrating the Living Optics camera, preprocessing, and machine learning.

3. To evaluate classifier performance under varying illumination (6000 K vs. 8000 K).

4. To establish a foundation for future robotic HSI systems for autonomous material identification.

Experimental Setup

All experiments were performed inside the Construction Automation, Robotics and Visualization (CARV) lab at Auburn university. To eliminate interference, trials were conducted at night with all ambient lights off. Both the camera and the Kelvin Play light were mounted on tripods and positioned vertically downward toward the material samples placed on the floor, as shown in Figure 1. The equipment and software environment are presented in the following Table 1.

Table 1. Equipment and Software

Component	Specification
Hyperspectral Camera	Living Optics (LO) Snapshot HSI Camera
Spectral Range	440–900 nm (VNIR)
Spectral Resolution	≈ 8 nm FWHM
Lighting Source	Kelvin Play LED, Tunable 2000-20000K, CRI > 95
Samples	2 × 2 ft Plain gypsum board, primer-coated board, and semi-gloss painted board
Software Environment	Python 3.8, LO SDK v1.9.0, scikit-learn 1.4.1, OpenCV 4.9

The camera and light were separated by approximately 0.5 m, and the field of view was adjusted to fully encompass each board. The lighting temperature was initially 6000 K but later increased to 8000 K to maximize spectral separability between primer and semi-gloss coatings.

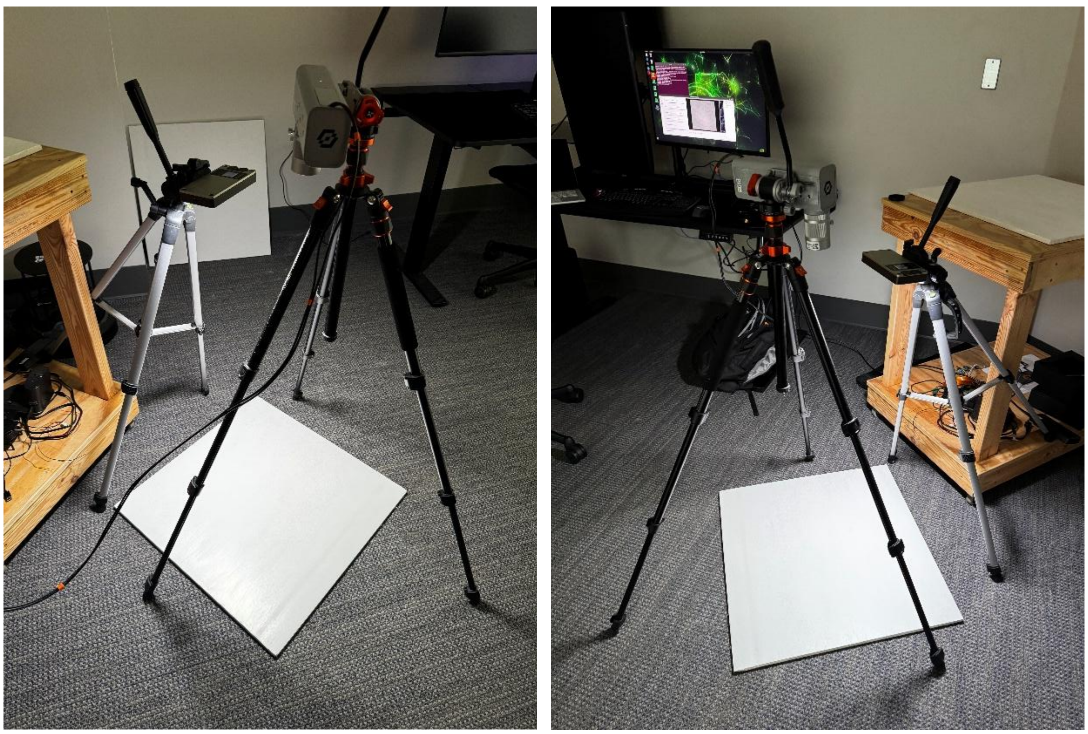

Figure 1. Experimental Setup

Data Collection and Preprocessing

Initial Spectral Datasets

The first acquisition phase produced three JSON datasets as “gyp-6000k-45.json”, “primer-6000k-45.json”, and “semigloss-6000k-45.json”. Each contains 46 reflectance spectra (r0-r45) captured under 6000 K lighting. These static samples established baseline separability among materials.

Live Spectral Acquisition

To enable real-time classification, a Python script (log_live_spectrum.py) was written to capture live spectra directly from the camera. Each frame was decoded using the LO SDK's "`SpectralDecoder`", averaged across all pixels, labeled, and saved as individual JSON files in a live-samples directory. Approximately 20 labeled samples per class (gypsum, primer, semi-gloss) were collected under 8000 K lighting. This iterative sampling addressed early misclassifications of the primer due to insufficient data diversity.

Spectral Normalization

Each spectrum contained 96 bands. To account for variations in illumination intensity, data was normalized using "`MinMaxScaler`" from scikit-learn, scaling linearly all reflectance values to the [0, 1] range. This ensured uniform feature scaling across all input spectra before model training.

Data Quality Checks

During preprocessing, several issues were identified and corrected:

- Black preview output: This issue occurred when the preview window displayed a completely black image instead of the expected spectral visualization. It was caused by the incorrect handling of the data tuple returned from the hyperspectral decoding function. The problem was resolved by referencing the second element of the decoded tuple, which contains the processed reflectance data rather than the raw, uninitialized array.

- Shape mismatch: The "shape" refers to the dimensional structure of the spectral data array, specifically its height (H), width (W), and number of spectral bands (B). A mismatch in shape indicates that the data dimensions did not align properly for visualization or analysis. This was corrected by reshaping the spectral array into the standardized format (H, W, B) using a custom tiling function to ensure consistent spatial-spectral alignment.

- Lighting inconsistencies: Addressed by using an 8000 K illumination setting and maintaining a fixed geometry between the camera and light source, improving uniformity across all captured datasets.

Machine Learning Models

Selecting appropriate classification algorithms was a critical step in developing the hyperspectral material identification pipeline. Because the dataset consisted of a limited number of high-dimensional spectra (96 spectral bands and approximately 45 total labeled samples), the chosen algorithms needed to:

1. Perform well with small sample sizes

2. Handle nonlinear decision boundaries inherent in spectral data

3. Require minimal parameter tuning to maintain real-time feasibility.

Two algorithms were therefore selected for evaluation: a Support Vector Machine (SVM) with a radial basis function (RBF) kernel, and a Random Forest Classifier (RFC).

Support Vector Machine (SVM)

The SVM was selected as a strong baseline classifier because it performs well with high-dimensional data (such as hyperspectral bands) even when the number of training samples is relatively small. This makes it suitable for pilot-scale hyperspectral datasets where spectral features greatly outnumber the available samples. The model was implemented as `SVC(kernel='rbf', gamma='scale')`, where the "`kernel='rbf'`" allows the classifier to handle non-linear relationships between spectral features, and the `gamma='scale'` parameter automatically adjusts the kernel's sensitivity based on the variance of the input data. This configuration ensures stable performance without overfitting in small yet complex datasets.

Random Forest Classifier (RFC)

To improve robustness and interpretability, an RF classifier was also implemented as `RandomForestClassifier(n_estimators=100, max_depth=None)`. Random forests operate by constructing an ensemble of multiple decision trees, where each tree is trained on a random subset of the training data and spectral bands. The final classification is determined through a majority voting process across all trees, which enhances stability and reduces the influence of noise or outliers. The parameter "`n_estimators=100`" defines the number of trees in the ensemble, providing an effective balance between computational efficiency and predictive accuracy, while "`max_depth=None`" allows each tree to grow until all data patterns are fully explored. This ensemble-based approach enables the model to capture nonlinear relationships between spectral features and coating categories without extensive parameter tuning, making it both robust and interpretable for small-scale hyperspectral datasets.

Comparison and Final Selection

The results were obtained from training and evaluating the two classification models using the preprocessed hyperspectral dataset of coated and uncoated gypsum samples. Although the SVM achieved a satisfactory training accuracy of approximately 93%, the RFC demonstrated superior performance (>95%) and more consistent class separation when tested under 8000 K illumination. In addition, the trained RFC model required less computational time during real-time inference and showed greater robustness to lighting variations and spectral noise, indicating its suitability for deployment in on-site inspection scenarios.

For these reasons, the RFC was selected as the final model for real-time classification within the Living Optics hyperspectral pipeline. Both the trained model "rf_live_classifier.joblib" and scaler "rf_live_scaler.joblib" were stored for integration into the real-time pipeline.

Feature analysis

Feature importance analysis indicated that spectral bands between 520–640 nm and 760–880 nm contributed most to classification, correlating with pigment absorption and surface gloss transitions.

Real-Time Classification Pipeline

The real-time inference system was implemented using the LO SDK and OpenCV

1. Frame Capture: Each frame was decoded with a spectral decoder.

2. Spectral Averaging: Mean spectra were computed across either the full frame or small tiled regions.

3. Normalization: Using the pre-trained "`MinMaxScaler`".

4. Prediction: The Random Forest model predicted the most probable material class.

5. Visualization: Predicted class name was overlaid on the grayscale live feed using OpenCV text and color-coded labels (blue = semi-gloss, green = primer, red = gypsum).

This pipeline provided real-time feedback at ~10 fps on an Ubuntu workstation (Nvidia's Jetson AGX Orin) capable of delivering 275 Tera Operations Per Second (TOPS) of Artificial Intelligence (AI) performance.

Results and Analysis

Spectral Distinction

Distinct reflectance profiles were observed:

- Plain Gypsum: Broad, relatively flat curve with minor absorption around 450 nm.
- Primer-Coated Gypsum: Reduced reflectance in the blue-green region (~480–550 nm).
- Semi-Gloss: Higher reflectance beyond 700 nm due to specular behavior from polymer resins.

These results matched the initial static datasets and confirmed improved separation under 8000 K lighting.

Quantitative Metrics

Table 2 provides the quantitative metrics of the performed experimental study.

Table 2. Quantitative Metrics

Model	Lighting	Accuracy	Key Misclassifications
SVM (RBF)	6000 K	93.0 %	Primer ↔ Semi-Gloss
Random Forest	6000 K	95.6 %	Rare Gypsum ↔ Primer confusion

The increased color temperature to 8000 K yielded higher inter-class centroid distances and reduced variance within classes, confirming the sensitivity of spectral separability to illumination quality.

Euclidean Distance Analysis

The progressive distance values corroborate the model's decision boundaries and physical coating differences. To verify separation, Euclidean distances between mean class centroids were computed:

- Gypsum–Primer: 0.142
- Primer–Semi-Gloss: 0.187
- Gypsum–Semi-Gloss: 0.201

Real-Time Performance

The optimized pipeline maintained stable throughput without dropped frames. Predicted class labels updated continuously with minimal latency (< 0.2 s). The system demonstrated resilience to slight orientation and distance changes between the camera and surface.

Discussion

Influence of Lighting Conditions

Illumination was a critical factor in classification accuracy. Increasing the color temperature from 6000K to 8000K enhanced contrast in the blue NIR transition region, where semi-gloss coatings exhibit distinctive specular features. This demonstrates that controlled lighting calibration is as important as model selection for HSI-based classification.

Model Interoperability

The Random Forest model provided better interpretability than SVM through feature importance metrics, allowing physical correlation between bands and coating behavior. These findings support future data-driven optimization of illumination wavelengths and sensor selection.

Comparison with Conventional Methods

A qualitative comparison was conducted using RGB images of the same gypsum samples under identical lighting conditions. The RGB camera provided only visual color differences, which appeared nearly indistinguishable between the primer and semi-gloss coatings. No quantitative image-based classification was performed; however, visual inspection confirmed that conventional RGB imaging lacked sufficient spectral sensitivity to distinguish these coatings. Only the hyperspectral system captured the spectral nuances induced by primer pigments and gloss resins. This underscores the superior discriminatory capability of HSI for subtle material differences.

Challenges Encountered

- Data Quantity: A limited number of samples restricted deep learning exploration.
- Spectral Noise: Minor band-edge noise near 440 nm and 900 nm.
- Illumination Stability: Even small shifts in lamp distance altered reflectance intensity; consistent geometry was critical.
- Primer Misclassification: Resolved through additional samples and lighting adjustment.

Despite these challenges, the final pipeline remained stable and reproducible.

Future Improvements

The next iteration will include:

- Dimensionality reduction using Principal Component Analysis (PCA) or t-SNE (t-Distributed Stochastic Neighbor Embedding) to visualize spectral clusters.

- Confidence thresholding to filter uncertain predictions.
- Region-based classification, enabling per-pixel material mapping rather than frame-averaged labeling.
- Expanded dataset incorporating other visually similar construction materials.
- Adaptive illumination calibration, allowing model retraining for site-specific lighting.

These improvements will transition the system from controlled laboratory validation to robust, field-deployable robotic inspection.

Conclusions

This pilot study demonstrated that a LO hyperspectral camera, paired with controlled Kelvin Play illumination and an RF classifier, can accurately differentiate between plain, primer-coated, and semi-gloss-painted gypsum boards in real time. Key technical insights include:

1. Lighting color temperature substantially influences spectral separability.
2. Random Forest models outperform SVM for this dataset.
3. Real-time classification is achievable using the LO SDK and OpenCV integration.
4. The approach provides a foundation for autonomous robotic inspection.

The successful implementation inside the laboratory verifies hyperspectral imaging as a powerful non-destructive technique for distinguishing visually similar construction materials and paves the way for on-site robotic deployment.

Acknowledgement

The authors acknowledge the use of an Artificial Intelligence (AI) tool, specifically OpenAI's ChatGPT, to assist in improving the clarity, grammar, and overall readability of this manuscript. All intellectual content and interpretations are the authors' original work.

References & Appendices

Bhargava, A., Sachdeva, A., Sharma, K., Alsharif, M. H., Uthansakul, P., & Uthansakul, M. (2024). Hyperspectral imaging and its applications: A review. *Heliyon*, *10*(12), e33208. https://doi.org/10.1016/J.HELIYON.2024.E33208

Ekanayake, B., Wong, J. K. W., Fini, A. A. F., & Smith, P. (2021). Computer vision-based interior construction progress monitoring: A literature review and future research directions. *Automation in Construction*, *127*, 103705. https://doi.org/10.1016/J.AUTCON.2021.103705

Kim, C., Kim, C., & Son, H. (2013). Automated construction progress measurement using a 4D building information model and 3D data. *Automation in Construction*, *31*, 75–82. https://doi.org/10.1016/J.AUTCON.2012.11.041

Musarat, M. A., Khan, A. M., Alaloul, W. S., Blas, N., & Ayub, S. (2024). Automated monitoring innovations for efficient and safe construction practices. *Results in Engineering*, *22*, 102057. https://doi.org/10.1016/J.RINENG.2024.102057

OpenSpace. (2025). *OpenSpace: The Visual Intelligence Platform for Builders*. https://www.openspace.ai/

Ptacek, L., Strauss, A., Hinterstoisser, B., & Zitek, A. (2021). Curing Assessment of Concrete with Hyperspectral Imaging. *Materials 2021, Vol. 14, Page 3848*, *14*(14), 3848. https://doi.org/10.3390/MA14143848

Sami Ur Rehman, M., Shafiq, M. T., & Ullah, F. (2022). Automated Computer Vision-Based Construction Progress Monitoring: A Systematic Review. *Buildings 2022, Vol. 12, Page 1037*, *12*(7), 1037. https://doi.org/10.3390/BUILDINGS12071037

Track3D. (2025). *Track3D - Reality Intelligence Platform*. https://track3d.ai/

Tretter, D., Memon, N., & Bouman, C. A. (2005). Multispectral Image Coding. *Handbook of Image and Video Processing*, 747–760. https://doi.org/10.1016/B978-012119792-6/50107-8

Vítek, S., Zbíral, T., & Nežerka, V. (2025). Using hyperspectral imaging to identify optimal narrowband filter parameters for construction and demolition waste classification. *Resources, Conservation and Recycling*, *215*, 108123. https://doi.org/10.1016/J.RESCONREC.2025.108123

Zou, Y., Kiviniemi, A., & Jones, S. W. (2017). A review of risk management through BIM and BIM-related technologies. *Safety Science*, *97*, 88–98. https://doi.org/10.1016/J.SSCI.2015.12.027