Data-Driven Color Classification Using the Spectrometer Module of the SENSIPATCH Wearable System

Hamza Mustafa[1,2], Simone Contardi[2], Luigi Ferrigno[1], Filippo Milano[1], Mario Molinara[1]
[1]*Dept. of Electrical and Information Engineering*, *University of Cassino and Southern Lazio*, 03043 Cassino, Italy
European University of Technology EUt+
European Union
[2]*Sensichips srl*, 04011 Aprilia, Italy
{hamza.mustafa,ferrigno, filippo.milano, m.molinara }@unicas.it
simone.contardi@sensichips.com

***Abstract*—Accurate and portable color detection technologies are increasingly relevant in smart city contexts, where infrastructure maintenance, asset monitoring, and compliance assessment require reliable, in-situ data collection. This paper presents a wearable system for color classification based on the SENSIPATCH platform, which integrates a multi-wavelength spectrometer and machine learning-based processing. To simulate a diffusive and homogeneous optical background that mimics real-world conditions, such as urban materials placed on reflective or scattering surfaces, we placed multiple sheets of white paper behind the PANTONE samples during acquisition. This setup does not interfere directly with the light path between the sensor and the sample, but it affects the overall backscattering environment. The spectral responses obtained from six discrete LEDs were used as input features for four supervised machine learning models. Among them, the Support Vector Machine (SVM) achieved the highest classification accuracy of 93.13%. The results demonstrate the system's feasibility as a low-cost, portable solution for automated visual inspection tasks within smart city applications.**

***Index Terms*—Color classification, Spectrometer, Wearable sensing, Machine learning, SENSIPATCH**

I. Introduction

Color monitoring in urban infrastructure, such as road signs, street furniture, and painted surfaces is essential for maintenance, safety, and compliance in smart cities. Conventional visual inspections are subjective and inefficient, while high-end spectrometers, though accurate, are often bulky and costly. Recent advances in low-cost optical sensors and machine learning enable portable, data-driven color classification systems. These have shown promise in agriculture [1], biomedical analysis [2], fashion [3], and forensics [4]. More complex methods like hyperspectral imaging [5], embedded multispectral sensors [6], and IoT-based tools such as Skinly [7] offer higher precision but often at the expense of cost or wearability [8]. This study investigates the use of the SENSIPATCH platform [9], a compact, wearable spectroscopic system developed by Sensichips [10]. It integrates the SENSIPLUS microsensor [11] and six LEDs covering visible to near-infrared wavelengths. PANTONE [12] color cards were evaluated under diffusive conditions by overlaying 15 layers of white paper to simulate real-world interference such as skin or fabric.

This work aims to:

- Evaluate the SENSIPATCH's accuracy under scattering conditions.
- Train machine learning models using raw spectral signals from the device.

II. Methodology

A. Experimental Setup

The experimental setup involved covering official PANTONE color samples with fifteen layers of standard white A4 paper to mimic real-world diffusive interference. The SENSIPATCH device was placed beneath the stack, ensuring contact with the bottom of the color sample. This design simulated conditions where color detection occurs through scattering surfaces like skin or fabric.

B. Data Acquisition

The SENSIPATCH device includes six LEDs at specific wavelengths: Blue (468 nm), Green (523 nm), Yellow (593 nm), Red (645 nm), and Infrared (850 nm and 950 nm). Each LED was activated in sequence while the central photodiode measured reflected light. A lock-in amplifier extracted the in-phase current response from the sensor under each wavelength.

For each color sample:

- Five independent measurements were collected.
- Each measurement consisted of 1,151 samples.
- The first 150 samples captured ambient conditions (no card).
- From sample 151 onward, a color card with 15 paper layers was added.
- A baseline correction was applied by subtracting the mean signal of the ambient samples from the rest.

Ten conditions were measured: nine PANTONE colors and plain white paper, plus an empty baseline (AIR). Each data point consisted of six features one for each LED wavelength.

III. Machine Learning Models

Four supervised models were trained on the baseline-corrected spectral data: Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbors (KNN), and Multi-Layer Perceptron (MLP). Five-fold cross-validation ensured robustness. The six spectral responses served as direct inputs.

Model hyperparameters included:

- SVM: linear kernel, probability estimation enabled
- RF: 200 estimators
- KNN: $k = 150$
- MLP: two hidden layers with 100 and 50 neurons

IV. Results

The performance of four supervised machine learning models SVM, MLP, KNN, and RF was evaluated using five-fold cross-validation. Table I presents the average accuracy, precision, recall, and F1-score across folds. SVM achieved the highest mean accuracy of 93.13%, outperforming the other classifiers across all evaluation metrics. Its performance was consistent across folds, with a low standard deviation of 3.97%, indicating robustness against sampling variation. MLP followed closely with 92.01% accuracy but exhibited higher variability across folds, suggesting it may be more sensitive to training conditions.

KNN performed reasonably well (88.99%) but was more prone to misclassification, particularly in spectrally similar classes. Random Forest yielded the lowest accuracy (79.16%), potentially due to overfitting or limited generalization capacity when using raw spectral input without feature engineering.

TABLE I
Model Performance Metrics (Mean of 5 Folds)

Model	Accuracy	Precision	Recall	F1
SVM	**93.13%**	92.79	93.13	91.13
MLP	92.01%	91.19	92.01	90.32
KNN	88.99%	85.21	89.01	85.83
RF	79.16%	73.16	73.00	72.87

To further assess model behavior at the class level, the mean confusion matrix for the SVM model is presented in Figure 1. The matrix highlights that several classes, such as `PURPLE_5125`, `RED_1797`, and `YELLOW_7548`, were classified with 100% accuracy. Minor misclassifications were observed primarily among spectrally similar classes, for example, `BLACK_19-4305` was confused with `Blue_18_4525` and `GREEN_7726`, and `LIGHTESTSKY_4804` was misclassified as `BROWN_16-1439`.

V. Conclusion

This study demonstrates the viability of wearable color classification under scattering conditions using the SENSIPATCH spectrometer. The system achieves high accuracy without extensive preprocessing, confirming its potential for smart city maintenance tasks. Future efforts will expand the dataset, refine the hardware setup, and explore hybrid models for enhanced generalization.

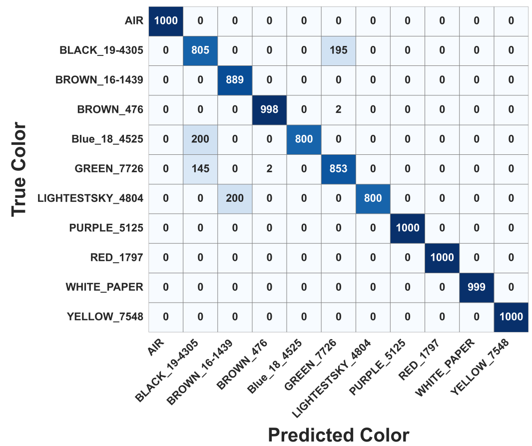

Fig. 1. Mean Confusion Matrix for SVM

Acknowledgment

This work was co-funded by the European Union Horizon Europe program with the project TARA (GA 101057524).

References

[1] S. Ueno and O. Sakai, "Data driven calibration of color-sensitive optical sensor by supervised learning for botanical application," *IEEE Sensors Journal*, vol. 22, no. 12, pp. 11915–11927, 2022.

[2] L.-C. Hsu, S. Hsu, T.-H. Tan, C.-H. Cheng, and C.-C. Chang, "Developing low-cost mobile device and apps for accurate skin spectrum measurement via low-cost spectrum sensors and deep neural network technology," *Sensors*, vol. 22, no. 22, p. 8844, 2022.

[3] Z. Li, G. Golwala, S. Sundaram, and J. Allebach, "A data-driven approach for garment color classification in on-line fashion images," *Electronic Imaging*, vol. 31, pp. 1–5, 2019.

[4] M. García García, M. A. González Arrieta, S. Rodríguez González, S. Márquez-Sánchez, and C. F. Da Silva Ramos, "Graffiti identification system using low-cost sensors," 2024.

[5] W. Zhu, P. Sang, and Y. He, "Facial skin colour classification using machine learning and hyperspectral imaging data," *IET Image Processing*, vol. 16, no. 2, pp. 509–520, 2022.

[6] D. M. Crafa, S. Di Giacomo, C. Fiorini, and M. Carminati, "Neural networks embedded in wearable devices: a preliminary digital vs. analog comparison," in *2023 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*, pp. 587–592, IEEE, 2023.

[7] M. d. P. B. Tobar, S. Clemann, R. Hagens, S. Pagel-Wolff, S. Hoppe, P. Behm, F. Engelhard, M. Langhals, S. Gallinat, A. Zhavoronkov, *et al.*, "Skinly: A novel handheld iot device for validating biophysical skin characteristics," *Skin Research and Technology*, vol. 30, no. 3, p. e13613, 2024.

[8] Y. Tang, S. Song, S. Gui, W. Chao, C. Cheng, and R. Qin, "Active and low-cost hyperspectral imaging for the spectral analysis of a low-light environment," *Sensors*, vol. 23, no. 3, p. 1437, 2023.

[9] SENSIPATCH, "https://sensichips.com/design-kit/wearables-kit/," 2025. Accessed: 2025-05-31.

[10] Sensichips, "https://sensichips.com/," 2025. Accessed: 2025-05-31.

[11] H. Mustafa, M. Molinara, L. Ferrigno, and M. Vitelli, "A deep learning system for water pollutant detection based on the sensiplus microsensor," in *International Conference on Pattern Recognition*, pp. 192–203, Springer, 2024.

[12] Pantone, "https://www.pantone.com/," 2025. Accessed: 2025-05-31.