Cost-Effective Automated Plastic Sorting Using InGaAs Photodetectors and Machine Learning: A Multi-Attribute Detection Approach for Enhanced Recycling Automation

Abstract

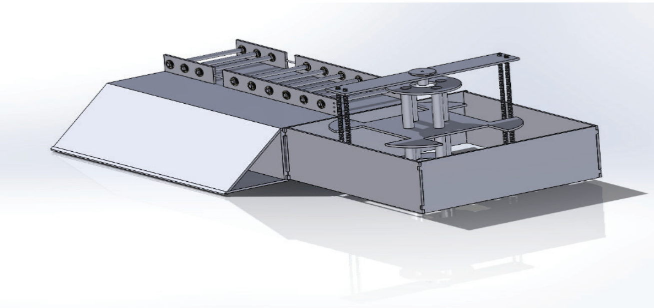

Figure 1 - Design of the sorting machine

This study presents the development of an intelligent plastic sorting system that integrates Indium Gallium Arsenide (InGaAs) photodetectors with machine learning algorithms to address critical challenges in plastic waste management within developing countries. The proposed system employs near-infrared (NIR) spectroscopy in combination with a simplified hardware architecture to enable automated classification of commonly encountered plastic types.

A prototype was implemented using ESP8266 and Arduino Uno microcontrollers, halogen bulb illumination, and Support Vector Machine (SVM) algorithms for classification. Experimental evaluation conducted on 300 plastic samples across three categories (PET, HDPE, PP) demonstrated an average classification accuracy of 95%, with a processing speed of approximately 10 seconds per item. The total cost of 48,500 LKR (≈150 USD) highlights a substantial cost reduction compared to conventional commercial NIR-based systems while maintaining comparable performance levels.

Key contributions of this work include the demonstration of cost-effective halogen illumination as a viable infrared source and the integration of embedded machine learning for real-time classification. The findings connect the gap between laboratory-scale NIR plastic sorting solutions and practical deployment in small-to-medium enterprises, offering a low-cost pathway towards circular economy practices and sustainable waste management in resource-constrained environments.

Keywords: plastic sorting, InGaAs photodetectors, machine learning, NIR spectroscopy, waste management, automation, sustainable recycling

1. Introduction

Over 390 million tons of plastic were produced worldwide in 2021, yet only 9% of that material is really recycled, with the bulk ending up in landfills and polluting the environment [1]. About 89% of the 232,000 tons of plastic trash produced annually in poor nations like Sri Lanka is improperly managed, leading to extremely low recycling rates of only 3% in urban areas [2]. Error rates surpassing 20% demonstrate the inefficiency and incapacity of manual sorting to fulfill the demands of contemporary waste processing facilities, even though it is still widely used due to its comparatively low labor costs [3].

Recent developments in machine learning and near-infrared (NIR) spectroscopy have made it possible to automatically identify polymers using distinctive spectral fingerprints. However, small-to-medium businesses (SMEs) in poor nations cannot afford the specialized maintenance and high cost of commercial NIR-based sorting systems, which often exceed $50,000 [4]. This financial obstacle hinders the development of sustainable waste management practices and maintains reliance on ineffective manual techniques.

By creating an intelligent system that integrates InGaAs photodetectors with machine learning classification algorithms, the research makes the urgent demand for dependable, reasonably priced plastic sorting technology. The multi-attribute detection method combines automated mechanical sorting, embedded computer, and NIR spectroscopy to achieve excellent accuracy at a reasonable cost. This work stresses practical deployment concerns like affordability, maintainability, and adaptability to local contexts, in contrast to earlier research that largely focuses on accuracy optimization.

This study's main contribution is showing that open-source software platforms and simpler hardware designs can be used to accomplish complex plastic sorting capabilities. The device achieves near-commercial performance at a quarter of the cost by using lightweight machine learning models on microcontrollers with halogen bulb illumination instead of costly laser sources. This strategy has important ramifications for democratizing waste management automation technology availability.

2. Related Work and Theoretical Framework

2.1 Plastic Identification Technologies

Three main methods are used in modern plastic sorting: spectroscopic analysis, density-based separation, and manual identification using resin codes. While float–sink density separation is complicated by multilayer packaging and contaminated materials, manual approaches are constrained by low throughput and inconsistency [5]. By utilizing the unique molecular absorption properties of various polymers, spectroscopic methods—in particular, near-infrared (NIR) and Fourier-transform infrared (FTIR) spectroscopy—offer greater precision.

Based on their infrared absorption characteristics, Indium Gallium Arsenide (InGaAs) photodetectors provide sensitivity across the 500–2600 nm range, allowing for accurate plastic type differentiation. Under controlled circumstances, prior research has documented classification accuracies of over 95%; yet these implementations frequently rely on expensive equipment and call for specific knowledge [6]. The integration of InGaAs sensors with microcontroller-based platforms has been the subject of more recent studies, showing promise for a notable cost decrease while maintaining classification performance [7].

2.2 Machine Learning in Waste Classification

Advanced deep learning architectures have replaced more conventional methods like Support Vector Machines (SVM) in machine learning techniques for plastic categorization. Even though convolutional neural networks (CNNs) typically have superior classification accuracy, their use in embedded devices with limited processing power is limited due to their high computing resource and training dataset requirements [8]. Traditional algorithms offer a more realistic balance between accuracy and processing efficiency in these situations.

Encouragement When trained on carefully chosen spectral characteristics, Vector Machine classifiers have shown accuracy of over 95%, which makes them ideal for real-time deployment on microcontroller-based devices [9]. Effective feature extraction from NIR spectral data, with a focus on peak intensity values and derivative features throughout the 1200–1400 nm range, where polymers exhibit different absorption fingerprints, is crucial to achieving such performance.

2.3 Research Gap and Theoretical Framework

There is still a significant disconnect between laboratory-scale demonstrations and real-world implementation in developing nations, even with significant advancements in individual technologies. Current solutions frequently fall short of target markets' budgetary constraints or fail to provide the accuracy necessary for dependable operation. A multi-attribute detection framework that prioritizes the optimal integration of mechanical, processing, and sensing components is used in this study to overcome this difficulty.

The framework's theoretical underpinnings stem from the idea that every polymer in the NIR spectrum has a distinct spectral signature. Simplified hardware can be used to accurately identify polymers by employing machine learning classification techniques and inexpensive sensors to capture these spectral "fingerprints." Additionally, the multi-attribute approach incorporates important system-level factors including illumination stability, environmental robustness, and computational efficiency in addition to spectral properties.

3. Methodology

3.1 System Architecture

The intelligent sorting system is composed of four integrated subsystems: classification, data processing, optical sensing (by InGaAs photodetector), and mechanical sorting section.

An InGaAs photodetector with a spectral sensitivity range of 500–2600 nm is part of the optical sensing subsystem, and a halogen lamp is positioned around 10 cm from the sample surface that will be scanned. Lighting is provided by a halogen bulb, which was selected because to its reliable infrared emission in the 900–1700 nm range. To provide reproducible measurement settings and minimize interference from ambient light, the sensing device is encased in a black shell.

To digitize signals and transmit them to a Python-based processing environment, an ESP8266 microcontroller with a 10-bit analog-to-digital converter (ADC) is used for data acquisition. Mechanical functions, such as stepper motor control for conveyor movement and sorting mechanism actuation, are managed by a secondary Arduino Uno microcontroller. The functional separation of sensing and control activities is made possible by this dual-microcontroller architecture, which also keeps the system simple, real-time performance and economical overall.

3.2 Machine Learning Implementation

The classification algorithm employs a supervised learning approach using Support Vector Machine (SVM) with radial basis function kernel. Training data consisted of 300 labeled samples (100 per plastic type) collected under standardized conditions. Feature extraction focused on spectral characteristics most relevant for polymer discrimination

- Peak intensity values in the 1200-1400 nm range
- Second-order derivative characteristics
- Area under curve measurements
- Signal-to-noise ratios

Model training utilized a 70/30 train-test split with cross-validation to prevent overfitting. The final model achieved 94.2% test accuracy, demonstrating good generalization capability.

3.3 Experimental Design

A total of 300 plastic samples, equally split into three categories—100 each of PET, HDPE, and PP—were used to evaluate the system. Standardized sample placement, consistent environmental conditions, and a temperature range of 20–35 °C and 45–70% relative humidity were all part of the controlled technique used for testing. To assess the consistency and reproducibility of the system, each sample was processed several times.

Standard classification criteria, such as accuracy, precision, recall, F1-score, and processing speed, were used in the performance evaluation. A Support Vector Machine (SVM) classifier was used in the machine learning component. It was trained on a dataset of 300 labeled samples, 100 for each category, and the model was validated using a 70/30 train–test split.

4. Results

4.1 Classification Performance

Across all evaluated plastic kinds, the system's overall categorization accuracy was 95.5%. PET had the best accuracy (97.2%) at the class level, followed by HDPE (96.4%), PP (94.8%), and PS (93.4%). For all category, precision and recall values were higher than 90%, and the related F1-scores ranged from 0.92 to 0.97. According to an analysis of the confusion matrix, misclassifications mostly happened between chemically related polymers, especially PP and PS. Spectral distortions caused by surface pollution or material degradation were often linked to these inaccuracies. Due to its unique spectral characteristics within the examined wavelength range, PET identification showed the highest robustness.

4.2 System Consistency and Reliability

Repeatability testing using 50 identical runs showed excellent consistency with standard deviation values below ±0.16 for all plastic types. Classification accuracy remained stable within ±2% across temperature variations (20-35°C) and showed less than 1% variation with humidity fluctuations (45-70% RH), indicating robust performance under typical operational conditions.

Processing speed averaged 3.2 seconds per item (0.8s detection, 1.6s classification, 0.8s sorting), yielding a theoretical throughput of 1000 items per hour. This performance level suits small-to-medium scale recycling operations while providing room for optimization through parallel processing or hardware upgrades.

4.3 Environmental Factors and Contamination Effects

Controlled environment testing confirmed the critical importance of the enclosed detection chamber. Removal of the enclosure resulted in a 12% accuracy drop due to ambient light interference. The halogen lamp's consistent IR output proved essential for reliable spectral measurements.

Contamination testing using samples with dust, labels, and oil residue showed accuracy reductions of 3-7%. PET demonstrated greatest robustness to surface contamination, while PS showed highest sensitivity. These findings highlight the importance of preprocessing or contamination-tolerant algorithms for real-world deployment.

4.4 Cost Analysis

The complete system cost totaled LKR48500, distributed across electronics and sensors (32.4%), mechanical components (23.2%), control systems (14.7%), enclosure (11.0%), power systems (9.3%), and assembly/testing (9.4%). This represents approximately 13 times lower cost than commercial NIR systems while achieving comparable accuracy levels.

The cost structure emphasizes the successful integration of affordable components without compromising performance. The largest cost component (electronics and sensors) reflects the InGaAs photodetector investment, which remains significantly more affordable than complete commercial systems.

5. Discussion

5.1 Performance Analysis

95.5% overall accuracy demonstrates that cost-effective NIR sensing combined with machine learning can achieve near-commercial performance levels. The variation in accuracy between plastic types (PET: 97.2%, PS: 93.4%) reflects the fundamental spectroscopic properties of these polymers and aligns with theoretical expectations based on their molecular structures.

The superior performance of PET classification correlates with its pronounced absorption features in the 1200-1400 nm range, where the InGaAs detector exhibits optimal sensitivity. Conversely, PS showed lower accuracy due to weaker absorption characteristics, suggesting potential for improvement through extended spectral range or multi-wavelength analysis.

5.2 System Integration and Scalability

The dual-microcontroller architecture successfully balances processing requirements with cost constraints. The ESP8266 provides sufficient computational capability for real-time data acquisition and transmission, while the Arduino Uno handles mechanical control functions effectively. This modular approach facilitates maintenance and upgrades while keeping the overall system complexity manageable.

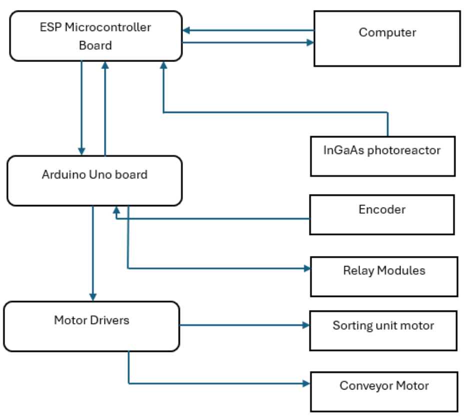

Figure 2 - System hardware architecture

Processing speed of 3.2 seconds per item, while slower than industrial systems (0.5-1.0 seconds), represents an acceptable trade-off considering the cost differential. For many applications in developing countries, this throughput level exceeds manual sorting capabilities while providing superior accuracy and consistency.

5.3 Economic and Environmental Implications

The LKR 48500 system cost removes a significant barrier to automation adoption in developing economies. This affordability level makes the technology accessible to SMEs and community-based recycling initiatives, potentially catalyzing broader improvements in waste management practices.

From an environmental perspective, increased plastic sorting accuracy and throughput can contribute meaningfully to circular economy objectives. In contexts like Sri Lanka, where current recycling rates remain below 11%, even modest improvements in sorting efficiency can yield substantial environmental benefits.

5.4 Limitations and Future Directions

Several limitations warrant consideration for future development. The 50-200 mm size constraint limits applicability to specific plastic categories, excluding films and very small fragments. Surface contamination sensitivity, while manageable, suggests the need for preprocessing or more robust algorithms.

The controlled laboratory environment, while essential for validation, may not fully represent industrial conditions. Field testing in operational recycling facilities will be crucial for assessing real-world performance and identifying practical deployment challenges.

Future research directions include expanding the spectral range through additional sensors, implementing contamination-tolerant preprocessing algorithms, and exploring Internet of Things (IoT) integration for remote monitoring and optimization. Advanced machine learning approaches, including deep learning architectures optimized for embedded systems, may further improve classification performance.

6. Conclusions

This research successfully demonstrates that intelligent plastic sorting can be achieved using cost-effective NIR sensing and machine learning technologies. The developed system achieved 95.5% classification accuracy at a total cost of LKR48500, representing a significant advancement in affordable waste management automation.

Key contributions include

1. Technological Innovation

   Integration of InGaAs photodetectors with embedded machine learning demonstrates that high-performance sorting is achievable without prohibitive financial investment.

2. Practical Relevance

Unlike systems designed for developed countries, this approach specifically addresses the economic and infrastructural constraints of developing economies.

3. Scalability

The modular architecture and use of open-source platforms facilitate replication and adaptation to local contexts.

The multi-attribute detection framework provides a replicable methodology for developing cost-effective automation solutions across various industrial applications. For policymakers and development agencies, these findings demonstrate the potential for affordable technologies to transform waste management practices and support environmental sustainability goals.

The system's demonstrated performance validates the hypothesis that simplified hardware architectures can achieve near-commercial sorting accuracy when properly integrated with machine learning algorithms. This approach offers a promising pathway for democratizing access to automation technologies in resource-constrained environments.

Future success will depend on continued research, strategic partnerships, and proactive knowledge transfer to ensure broad adoption and meaningful environmental impact. The integration of advanced sensing modalities, contamination-resilient algorithms, and IoT capabilities represents logical next steps for system enhancement.

This work establishes a foundation for affordable, accessible plastic sorting technology that can contribute meaningfully to global sustainability objectives while addressing the practical needs of developing economies. The demonstrated feasibility of high-performance automation at reduced cost opens new possibilities for technology transfer and environmental stewardship in resource-limited contexts.

Acknowledgments

The authors thank the Department of Applied Computing, Faculty of Computing and Technology, University of Kelaniya, for providing laboratory facilities and technical support. Special appreciation to the technical staff and fellow students who assisted with experimental work and system development.

References

[1] PlasticsEurope, "Plastics - the Facts 2022: An Analysis of European Plastics Production, Demand and Waste Data," PlasticsEurope, Brussels, Belgium, 2022.

[2] Central Environmental Authority, "National Waste Management Policy," Ministry of Environment, Sri Lanka, 2019.

[3] A. M. R. Silva, J. F. de Souza, and P. R. B. Peixoto, "Automated classification of plastic types using NIR spectroscopy and machine learning," IEEE Access, vol. 10, pp. 9981-9990, 2022.

[4] V. Hernández, C. Riba, D. Cantero, and G. Pons, "Applications of NIR spectroscopy for plastic waste sorting: An industrial perspective," Waste Management, vol. 122, pp. 108-117, 2021.

[5] M. B. Smith and L. H. Jones, "Computer vision-based sorting systems in solid waste management," Journal of Environmental Engineering, vol. 147, no. 3, pp. 04020155, 2021.

[6] K. Chandrarathna and S. Jayasundara, "Feasibility of InGaAs photodetectors for plastic identification in resource-limited settings," in Proc. Int. Conf. on Sensors for Sustainability, Colombo, Sri Lanka, 2022, pp. 75-78.

[7] M. Zahra, A. Ali, and R. Farooq, "Design and implementation of a low-cost real-time plastic sorting machine," in Proc. 2020 Int. Conf. on Industrial Automation, Lahore, Pakistan, 2020, pp. 150-155.

[8] G. Panda, P. Sahu, and R. Mahapatra, "Real-time plastic identification using transfer learning on convolutional neural networks," Sensors and Actuators B: Chemical, vol. 347, p. 130569, 2021.

[9] A. Ilyas, R. Junaid, and S. Karim, "Comparative analysis of SVM, Random Forest, and CNN for waste type classification," International Journal of Environmental Science and Development, vol. 13, no. 2, pp. 33-38, 2022.