A MACHINE LEARNING DRIVEN ULTRAFAST MID-INFRARED SPECTROSCOPIC SYSTEM FOR MIXED PLASTIC CHARACTERIZATION

Saar Bar-Ziv – Michigan Technological University
Amir Hussain Idrisi – Michigan Technological University
Adeyinka Adekunle – Michigan Technological University
Ezra Bar-Ziv – Michigan Technological University
Fei Long* – Michigan Technological University
* Correspondence: Fei Long, flong@mtu.edu

Abstract:
The U.S. Environmental Protection Agency estimates that over 35.6 million tons of plastic waste are generated annually in the United States, yet only 8.7% is recycled, with approximately 75% ending up in landfills. Among the landfilled plastic waste, 15%—equivalent to more than 2.5 million tons annually—comprise black and dark-colored plastics. This high disposal rate is primarily due to the inability of existing optical sensors, such as near-infrared (NIR) and camera systems used in Material Recovery Facilities, to detect black plastics due to their low reflectivity in the near-visible wavelength range. Mid-infrared (MIR) technology has the potential to overcome this limitation by providing accurate compositional information independent of plastic color. However, traditional MIR systems suffer from slow measurement speeds, limiting their applicability in industrial settings. To address this challenge, we developed a ultra high-speed MIR characterization system capable of measuring of up 400-10,000 infrared spectra per second, integrated with a belt conveyor operating at 20-600 feet per minute. This innovation bridges the gap between laboratory-grade MIR analysis and the operational requirements of industrial recycling facilities. To ensure reliable performance despite the low signal-to-noise ratio inherent to high-speed MIR measurements, we implemented robust wavelet data preprocessing techniques. Additionally, we developed a lightweight machine learning (ML) model that achieves over 95%-99% accuracy in characterizing mixed plastic waste, including black plastics. Our ML model is designed for efficiency and operation without reliance on large-scale data centers, enabling seamless integration into industrial control systems. By facilitating the accurate identification and recovery of black plastics, our MIR characterization system offers a practical solution to increase recycling rates and reduce landfill waste. This innovation supports the development of a sustainable and circular economy in plastic recycling, addressing critical environmental challenges and unlocking new opportunities for resource recovery.

1. Introduction and Motivation
The global plastic waste crisis has reached alarming levels, with more than 35.6 million tons of plastic wastes were generated, within which, only 8.7% was recycled, while about 75% was disposed of in landfills.[1] Addressing this challenge requires the development of an efficient and advanced system for plastic waste recycling. The recycling process follows a pipeline that involves several interconnected stages, including collection, transportation, sorting, and recovery. One of the most significant hurdles in this chain, particularly before the recovery stage, is the accurate sorting and classification of plastics. Out of the 75% landfilled municipal plastic waste (MPW), 15% are black and dark-colored plastics,[2] corresponding to more than 2.5 million tons annually. This high landfill rate is largely due to the limitations of current optical sensors, such as Near-Infrared (NIR) and cameras commonly employed in MRFs, are unable to detect black and dark-colored plastics. Recent developments on Mid-Infrared (MIR) spectroscopy provided an attractive solution because of its capabilities of the accurate chemical composition characterization including black and dark-colored plastics, and its comparable measurement speed to the current optical sensors used in the recycling facilities. Enabled by ML, we have developed an ultrafast MIR system that can measure mixed plastic waste (MPW) on a belt conveyor in real-time. The schematic of our system is shown in Figure 1. This system has the potential to be

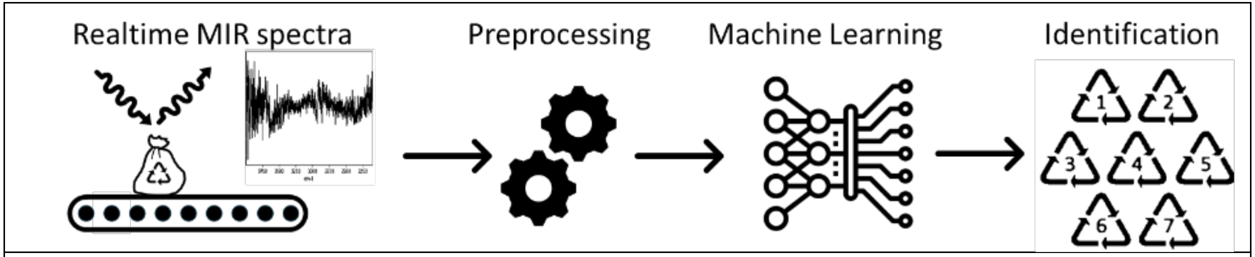


Figure 1. A machine learning driven ultrafast mid-infrared spectroscopic system for mixed plastic characterization.

combined with the current optical sensors in industry, and provide comprehensive information of MPW including composition, size, shape, color, and even brand, with unprecedented accuracy.

2. Review of Related Work

Many waste facilities, including municipal recycling centers and commercial recycling companies, have adopted NIR spectroscopy as a standard practice in their operations. NIR is a non-destructive technique that utilize reflectance spectroscopy in the wavelength of 0.8–2.5 μm to identify plastics, based on the stretching vibration modes of covalent bonds. One of NIR's key advantages is its rapid data acquisition, capable of analyzing up to hundreds of spectra per second. However, a significant limitation of NIR is its inability to detect black and dark-colored plastics, due to the low reflectance in the NIR range, resulting in insufficient signal-to-noise ratios.[3-5] Black plastics, constituting approximately 15% of MPW, are consequently often sent to landfills or incineration. An alternative to NIR is mid-infrared (MIR) spectroscopy, which operates in the 2.5–5 μm wavelength range and can reliably detect black and dark-colored plastics. Historically, the low measurement speed of MIR limited its industrial applications. However, recent advancements have enabled high-speed MIR characterization through photon up-conversion technology. This technique uses sum-frequency generation in a $\chi(2)$-nonlinear lithium niobate ($LiNbO_3$) crystal to upconvert MIR light (2.0–5.0 μm) to the near-visible range (695–877 nm).[6-10] The upconversion minimizes thermal noise and allows the use of silicon-based CMOS-array detectors, which exhibit much higher sensitivity than traditional MIR detectors such as HgCdTe or PbSe.[11] Leveraging this fast MIR technology, we have developed a machine learning (ML)-driven characterization system, which is capable of measuring 400 – 10,000 spectra per second, enabling the analysis of plastic waste moving at 20 – 600 ft/min. Our MIR system not only combines the high accuracy of the mid-infrared range with the speed comparable to NIR, but also addresses the critical limitation of detecting black plastics, thereby offering substantial improvements to overall recycling efficiency.

For industrial applications, NIR and MIR spectroscopic data often have low signal-to-noise ratios (SNR) due to fast measurement speeds and additional noise introduced by conveyor belt vibrations. To address this, ML models using convolutional neural network (CNN) have been developed to classify noisy spectra effectively.[12-14] CNNs, a type of deep learning method, excel in automatic feature extraction using convolutional and pooling layers.[15] However, the nonlinear nature of CNNs, coupled with noise and variability in spectra, can hinder meaningful data interpretation. To improve the quality and interpretability of IR spectra, we have developed comprehensive preprocessing techniques such as wavelet analysis.[16] Unlike Fourier Transform or smoothing methods, which use a single cut-off frequency and risk either retaining noise or losing critical information, wavelet analysis decomposes signals into multiple frequency components, localizing both wavelength and frequency information. This enables targeted noise reduction and enhanced feature extraction, particularly for spectra with complex peaks located across different regions, thereby significantly enhancing the fidelity of spectroscopic data for industrial recycling applications.

3. Technology Approach

The experimental set up is shown in Figure 2. The IR source (a 1000 °C silicon nitride, Si_3N_4 light source, 4.5 mm in diameter and 17 mm long, heated by 70 W electric power, Hawkeye Technologies model IR-Si311) was placed at the focus of an aluminum elliptical reflector which focused the light at 200 mm from the front surface of the reflector, projecting the light on a gold diffuser, generating a circular area with around 10 mm diameter. Some of the reflected light from the plastic surface was collected by a 1 in. parabolic gold-coated aluminum mirror (with a focal length of 200 mm) that collimated light from the diffuser, after which a 40 mm CaF_2 lens focused the light into a 200 μm core indium fluoride (InF_3) fiber that was connected to an MIR spectrometer (NLIR S2050, Denmark). The spectrometer is based on sum-frequency generation in a $\chi(2)$-nonlinear $LiNbO_3$ crystal that upconverts MIR light from the band 2.0−5.0 μm to the near-visible region 695−877 nm. The upconversion spectrometer offers several advantages, including the exclusion of most thermal noise and the utilization of silicon-based CMOS array detectors that exhibit superior detectivities compared to traditional MIR detectors such as HgCdTe (MCT) or PbSe array detectors. The CMOS detector comprises 2048 pixels, and the spectral resolution is $<6\ cm^{-1}$. The plastic samples are placed on a belt conveyor moving at speeds from 20 – 600 ft/min. To establish the ML training dataset, we selected 5 commercially available plastic materials commonly found in mixed plastic waste (MPW): black polystyrene (PS), black polyethylene (PE), deep blue polypropylene (PP), clear Polyethylene terephthalate (PET) and black polycarbonate (PC). These plastic samples were 1 mm thick sheets obtained from ePlastics (San Diego, California, United States). In addition, we also considered the black urethane, which is the conveyer belt material, referred to as the background (BK) in the following context. The samples were placed on the belt conveyor, and 10,000 spectra of each sample were collected at various measurement rates. We have developed a comprehensive data preprocessing method to reduce the noise and enhance the signature peaks of MIR spectra. The raw data (2.0 – 5.5 μm) measured at 200 spectra per second are shown in Figure 3 (a1-f1). We selected the wavelength range of 2.4 – 4.8 μm which is the range of the signature peaks of polymers, as shown in Figure 3 (a2-f2). Butterworth filter was applied to remove the high frequency noise, followed by multiresolution analysis (MRA) based on wavelet transformation. The corresponding spectra of each polymer after preprocessing are shown in Figure 3 (a3-f3). The SNR of the signature peaks in the selected wavelength range were significantly enhanced by our preprocessing method, which helps the CNN feature extraction and improves the accuracy and robustness of the ML model.

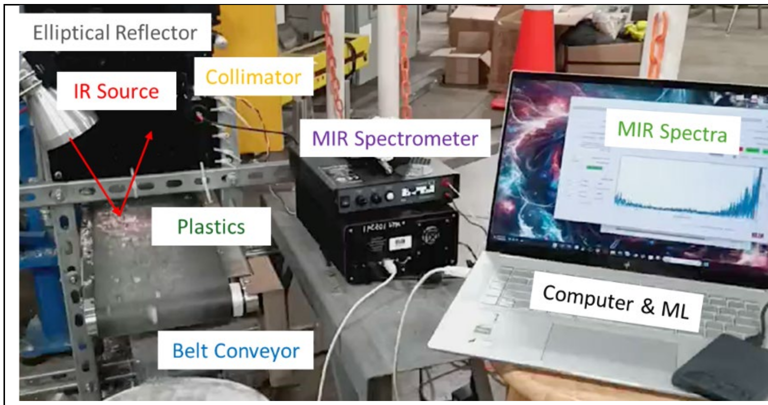


Figure 2. Experimental setup of the ML-MIR system for MPW characterization.

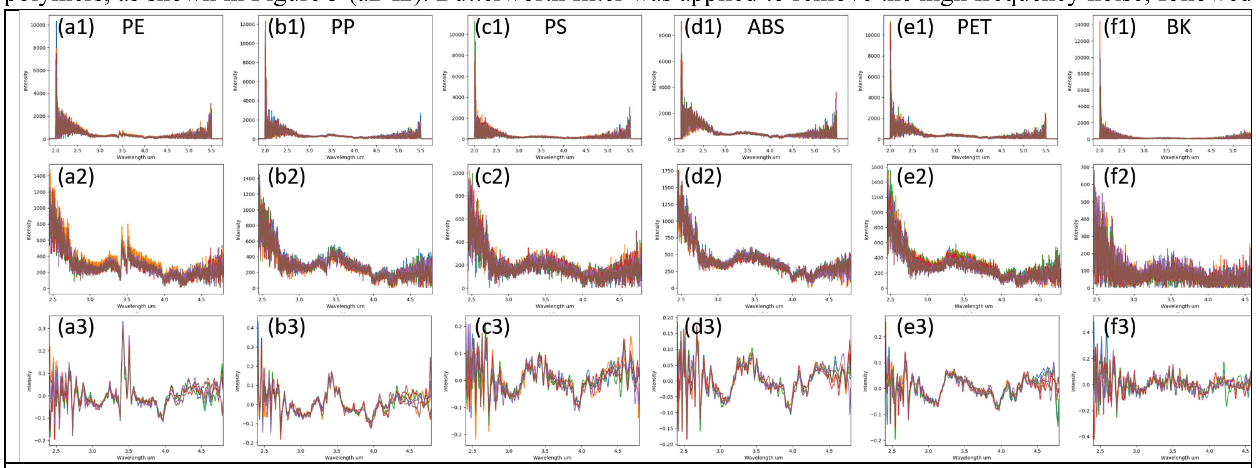


Figure 3. (a-f) Plots of 50 randomly selected MIR spectra of each polymer. (g-l) The corresponding spectra after preprocessing including low-pass filter and MRA based on wavelet transformation.

4. Discussion

We have developed a ML model using 1-D CNN to classify the polymers. The architecture is shown in Figure 4 (a). Four convolutional layers, three max-pooling layers, and three fully connected layers were included in the model. Each convolutional layer contained 64 filters of size 3, whereas the max-pooling layer consists of filter of size 2. Each fully connected layer contained 64 nodes. We used rectified linear units (ReLUs) between layers as activation functions. We added a dropout ratio of 0.2 between each of two fully connected layers to prevent overfitting. The activation function of the output layer was SoftMax, and the loss function was categorical cross-entropy. We leveraged a 5-fold cross-validation strategy to ensure the reliability of our CNN model. Initially, we evenly split the dataset into five random subsets. During each cross-validation iteration, we dedicated one subset (20% of the total data) to testing, and the remaining four subsets (80%) to model training. This five-fold procedure assured each subset was used as test data exactly once, enabling a comprehensive model evaluation across diverse data sections. To ensure homogeneity in each subset before sampling, we applied stratification, maintaining an equivalent proportion (16.7%) of each plastic type in the training and testing sets. The training confusion matrix is shown in Figure 4 (b) with near 100% accuracy.

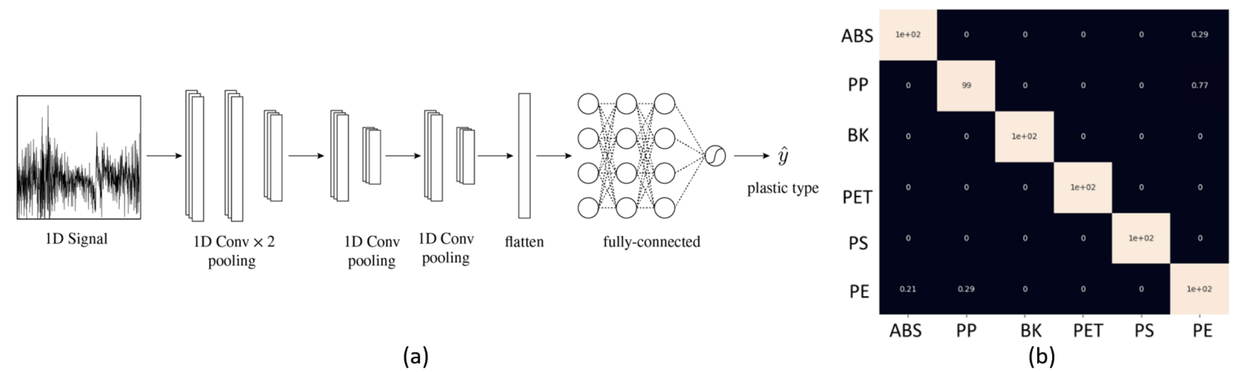
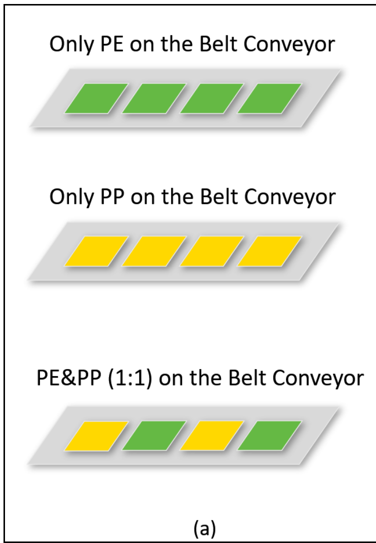


Figure 4. (a) The architecture of the 1D CNN model. (b) Confusion matrix showing near 100% training accuracy.

To validate the ML model prediction accuracy, we designed control experiments to simulate MPW on a belt conveyor, as shown in Figure 5 (a). Since the MIR is based on surface reflectance, the resulting compositional information should be proportional to the surface area of each polymer. PE and PP sheets were cut into ~10 mm^2 flakes. We collected six validation datasets: (i) two datasets with only PE flakes on the belt conveyor, and the ML model should predicts only PE and BK; (ii) two datasets with only PP flakes on the belt conveyor, and the ML model should predicts only PP and BK; and (iii) two datasets with both PE and PP with the ratio of 1:1 on the belt conveyor, and the ML model should predicts PE, PE and BK, with PE:PP approximately 1:1. We collected 500 spectra for each dataset, and the results predicted by the ML model are shown in Figure 5 (b). For the two PE datasets, 3.7% and 4.9% of the spectra were misclassified, shown as "Other". For the two PP datasets, 2.3% and 1.2% were misclassified. The results of the PE and PP mixture are encouraging, with only 2.2% and 1.7% were misclassified. More importantly, the ratios of PE and PP are almost identical, which are consistent with the designed control experiments. The validation results showed that our ML model can predict mix plastics on a belt conveyor with and overall >95% accuracy.



PE dataset 1	
BK	81.9%
PE	14.4%
Other	3.7%

PE dataset 2	
BK	82.1%
PE	13.0%
Other	4.9%

PP dataset 1	
BK	81.0%
PE	16.7%
Other	2.3%

PP dataset 2	
BK	78.0%
PE	20.8%
Other	1.2%

PEPP dataset 1	
BK	72.7%
PE	12.7%
PP	12.4%
Other	2.2%

PEPP dataset 2	
BK	78.7%
PE	9.9%
PP	9.7%
Other	1.7%

(b)

Figure 5. (a) Schematic of the validation experiments. (b)ML prediction results showing >95% accuracy.

The trained model was then tested by plastic waste sourced from a MRF in the United States. PP and PE flakes were pre-sorted and shredded to 5~10 mm flakes. 10 flakes of each polymer were used, and 100 spectra on each flake were collected, totaling 500 PE and 500 PP spectra, respectively. The raw spectra are shown in Figure 6 (a-b). Significant variations in signal intensity were observed. The variations were due to the (1) different surface roughness of the plastic flakes, resulting in different surface reflectivity; (2) Irregular flake shapes; (3) Contaminants on the surface. After preprocessing, the signal variations and the noise were significantly reduced, and the characteristic features of PE and PP spectra were clearly enhanced. The proposed ML model predicted >99% accuracy on both polymers, demonstrating the potential of characterizing real plastic waste. More experiments will be conducted to test the model with a wide range of waste plastics.

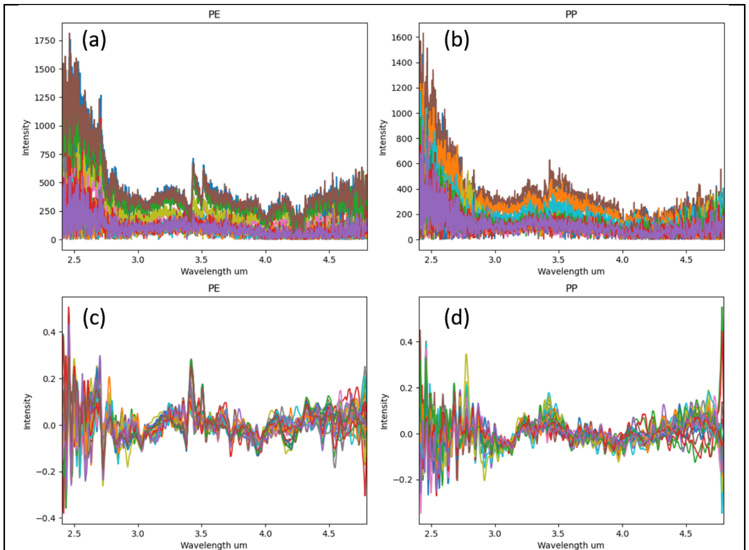


Figure 6. (a-b) MIR spectra of PE and PP waste samples sourced from an MRF. (c-d) MIR spectra after pre-processing.

5. Conclusions & Recommendations

We have developed a machine learning driven ultrafast mid-infrared spectroscopic system for mixed plastic characterization. The system can measure 400 – 10,000 MIR spectra per second on a belt conveyor moving at 20 – 600 ft per minute, compared to the NIR sensors currently used in the recycling facilities. The main advantage of our system is the ability to measure black and dark-colored plastics, which consist of 15% total plastic waste. An ML model was developed using 1D CNN to characterize the measured spectra in real time. To improve the ML accuracy, we have developed comprehensive data preprocessing method including filtration and MRA based on wavelet transformation. The model achieved near 100% training accuracy, and >95% prediction accuracy on mixed plastics moving on a belt conveyor. Future work will focus on expanding the ML database to include more polymer types.

The limitation of MIR is the limited penetration depth, which makes it difficult to accurately characterize plastics covered with labels, such as shrink-sleeve bottles commonly seen in MPW. We will explore the integration of the MIR system with a camera system to provide comprehensive information of MPW, including composition, color, size, shape, and even brand. The integrated characterization system will provide accurate MPW information for the downstream processes, such as automatic sorters, to improve the recycling rates.

6. Acknowledgements

The authors acknowledge support from: National Science Foundation (NSF-PFI-RP Award Number 2234450), and U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Bioenergy Technologies Office under Award Number DE-EE0010294.

7. References

1. Advancing Sustainable Materials Management: 2018 Tables and Figures, USEPA, https://www.epa.gov/, 2018.
2. Andrew Turner, Black plastics: Linear and circular economies, hazardous additives and marine pollution, Environment International, 117, 308, 2018.
3. Ragaert, K., L. Delva, and K. Van Geem, Mechanical and chemical recycling of solid plastic waste. Waste Management, 69, 24, 2017.
4. Faraca, G. and T. Astrup, Plastic waste from recycling centres: Characterisation and evaluation of plastic recyclability. Waste Management, 95, 388, 2019.
5. Offer Rozenstein, Eldon Puckrin, Jan Adamowski, Development of a new approach based on midwave infrared spectroscopy for post-consumer black plastic waste sorting in the recycling industry, Waste Management, 68, 38, 2019.
6. Becker, W., K. Sachsenheimer, and M. Klemenz, Detection of Black Plastics in the Middle Infrared Spectrum (MIR) Using Photon Up-Conversion Technique for Polymer Recycling Purposes. Polymers, 9, 435, 2017.
7. Ajanta Barh, Christian Pedersen, and Peter Tidemand-Lichtenberg, Ultra-broadband mid-wave-IR upconversion detection, Optics Letters, 42, 1504, 2017.
8. Søren M. M. Friis and Lasse Høgstedt, Upconversion-based mid-infrared spectrometer using intra-cavity LiNbO3 crystals with chirped poling structure, Optics Letters, 44, 4231, 2019.
9. Khalil Eslami Jahromi, Qing Pan, Lasse Høgstedt, Søren M. M. Friis, Amir Khodabakhsh, Peter Morten Moselund, and Frans J. M. Harren, Mid-infrared supercontinuum-based upconversion detection for trace gas sensing, Optics Express, 27, 24469, 2019.

10. Lichun Meng, Lasse Høgstedt, Peter Tidemand-Lichtenberg, Christian Pedersen, and Peter John Rodrigo, Enhancing the detectivity of an upconversion single-photon detector by spatial filtering of upconverted parametric fluorescence, Optics Express, 26, 24712, 2018.
11. Rasmus L. Pedersen, Dina Hot, and Zongshan Li, Comparison of an InSb Detector and Upconversion Detector for Infrared Polarization Spectroscopy, Applied Spectroscopy, 72, 793, 2018.
12. Shengli Jiang, Zhuo Xu, Medhavi Kamran, Stas Zinchik, Sidike Paheding, Armando G. McDonald, Ezra Bar-Ziv, Victor M. Zavala, Using ATR-FTIR spectra and convolutional neural networks for characterizing mixed plastic waste. Computers & Chemical Engineering, 155, 107547, 2021.
13. Stas Zinchik, Shengli Jiang, Søren Friis, Fei Long, Lasse Høgstedt, Victor M. Zavala, Ezra Bar-Ziv, Accurate Characterization of Mixed Plastic Waste Using Machine Learning and Fast Infrared Spectroscopy, ACS Sustainable Chem. Eng. 9, 14143, 2021.
14. Fei Long, Shengli Jiang, Adeyinka Gbenga Adekunle, Victor M Zavala, Ezra Bar-Ziv, Online Characterization of Mixed Plastic Waste Using Machine Learning and Mid-Infrared Spectroscopy, ACS Sustainable Chem. Eng., 10, 16064, 2022.
15. Liu, Y., H. Pu, and D.-W. Sun, Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices. Trends in Food Science & Technology, 113, 193-204, 2021.
16. Fei Long, Shengli Jiang, Ezra Bar-Ziv, Victor M. Zavala, Robust plastic waste classification using wavelet transform multi-resolution analysis and convolutional neural networks, Comput. Chem. Eng. 181, 108516, 2024.

8. About the Author(s)

Saar Bar-Ziv is an undergraduate student of Mechanical Engineering at Michigan Technological University. His research focuses on material characterization and machine learning.

Adeyinka Adekunle holds a M.S. degree from University of Lagos, Nigeria and is a research engineer at Michigan Technological University. His research focuses on material processing and system development.

Amir Hussain Idrisi is an Assistant Teaching Professor of Mechanical Engineering at Michigan Technological University. He is actively involved in the research and education of advanced manufacturing with years of experience in machine learning.

Fei Long (Presenter) is an Assistant Professor of Mechanical Engineering at Michigan Technological University. His current major research interests are material characterization, machine learning, and sustainable manufacturing.

Ezra Bar-Ziv is a Professor of Mechanical Engineering at Michigan Technological University. His current major research activity is wastes valorization, including material handling, thermal processes, power generation, clean combustion, and solvent-based technology for extraction of pure resins and biogenic materials.