A Novel Approach for Sensor Fusion Object Detection in Waste Sorting: The Case of WEEE

Ali Nazeri[1], Christiane Plociennik [2], Malte Vogelgesang [3], Chanchan Li [4], Martin Ruskowski [5]

Abstract: This paper investigates the application of AI-based methods for characterizing waste materials in sorting processes. With the increasing use of sensors in waste sorting systems, there is an opportunity to integrate data and improve accuracy. AI methods, such as deep object detection models, have the potential to optimize waste management processes and promote sustainability. This research examines the utilization of Sensor Fusion Object Detection in a multi-sensor sorting system, focusing on two different data fusion methods: concatenation and image mirroring. In the first approach, image data is concatenated with data from a hyperspectral near-infrared camera (NIR) and an inductive sensor, where dimensionality reduction techniques are applied to the data from both sensors. The second approach relies on a specific combination of NIR and inductive sensor data to simulate the format of image data. A Siamese Object Detection architecture is developed to train the model. The real-world testing results show that both approaches improve waste characterization accuracy and reliability by augmenting the models' mean average precision (mAP). These findings demonstrate the potential for AI-based methods to transform the waste separation and management process, leading to more sustainable practices and resource efficiency.

Keywords: Deep Neural Networks, Sensor-Based Sorting, E-Waste, Siamese Networks

1 Introduction

The concept of sustainable development refers to the circular economy approach, which aims to minimize the use of materials at different stages of the product life cycle, including production, distribution, and consumption [Pl22]. This approach focuses on increasing the efficiency of material use through the adoption of practices such as recycling and reuse. In this way, achieving significant reductions in material waste and environmental impact while promoting economic and social growth is feasible. The design and ingredients of products are subject to increasingly rapid changes, and it is increasingly heterogeneous even within "identical" product classes. Hence, characterizing discarded products and channeling them into the best available recycling process has become a big challenge. On the other hand, more sensors are now being employed in the waste sorting process. As a result, it is

[1] DFKI, IFS Department, 67663 Kaiserslautern, Germany Ali.Nazeri@dfki.de
[2] DFKI, IFS Department, 67663 Kaiserslautern, Germany Christiane.Plociennik@dfki.de
[3] Fraunhofer IWKS, Digitalization of Resources, 63755 Alzenau, Germany malte.vogelgesang@iwks.fraunhofer.de
[4] Fraunhofer IWKS, Digitalization of Resources, 63755 Alzenau, Germany Chanchan.Li@iwks.fraunhofer.de
[5] DFKI, IFS Department, 67663 Kaiserslautern, Germany Martin.Ruskowski@dfki.de

indeed possible to use more data and integrate it to recognize products more accurately. The implementation of waste sorting systems has a significant impact on promoting sustainability by enabling the separation of recyclable materials from non-recyclable waste [Bi18]. In the past, waste sorting systems relied on manual methods. However, the integration of intelligent automated methods is becoming more common. AI methods can greatly improve the accuracy and efficiency of waste sorting, enabling real-time monitoring with the ability to adapt to changes in the waste stream. This has the potential to transform the waste separation and management process [AW19]. Furthermore, AI-based methods can optimize waste management processes to improve resource efficiency and reduce the environmental impact of waste disposal. For example, dynamic routing algorithms can optimize waste collection routes and reduce the time, cost, and fuel consumption associated with waste transportation [YTZ22], or AI-based waste-to-energy conversion methods can extract energy from the waste while minimizing greenhouse gas emissions [BR15].
In this research, we are looking to investigate how to use AI-based methods to detect materials and objects in the waste sorting process. The concept was first described in [Pl22]. In this study, we use a multi-sensor sorting system which is presented in Figure 1. The goal is to fuse the available sensory data in this system to improve the accuracy and reliability of characterizing waste objects. In this paper, two different approaches are considered for data fusion: concatenating the data and image mirroring. Both of these approaches have been implemented with a deep object detection model (SSD: Single Shot MultiBox Detector [Li16]) and have been examined in a real-life environment. In the following subsections, the sorting system is described in detail and an overview of the problem is given. Section 2 discusses related work to object detection and sensor fusion in waste sorting systems. Section 3 describes the data exploration. In Section 4, the proposed approaches for sensor fusion and the model architectures are introduced. In Section 5, the specification of the models and the results of the training process are presented. Section 6 analyzes the results and concludes the paper.

1.1 Problem Definition

The present work aims to address a crucial issue in the recycling process of waste electrical and electronic equipment (WEEE), where characterizing discarded objects is mostly done by hand. The large variety of different product types and the differing respective recycling treatments, regarding legislative and economic constraints, lead to a labor-intensive mainly manual procedure. The perspective of this work is to support the automation of this sorting step with the development of a precise model that can characterize objects accurately by combining information from different sensors. The main goal of this study is to explore different approaches for integrating multiple sensors into an object detection model. Hence, this study contributes to the advancement of waste recycling technology by developing a more accurate and efficient object detection model for multi-sensor sorting systems. The sorting system utilized in this study is equipped with a tri-sensor configuration to accurately determine the properties of objects. There are three different types of sensors, line-scan cameras, a near-infrared camera, and an electromagnetic inductive sensor. which can capture the different attributes of objects. Before this study, the sorting system worked with an AI-Based module that provides image classification with high accuracy (>80%). That module used only image data to classify the objects, but the system was not able to detect the position of the objects, and also other sensors were being used separately. The goal is to integrate the data obtained from these three sensors to achieve maximum accuracy in the object detection task during the sorting process. The substantial variation in sensor specifications significantly amplifies the complexity of this task. Therefore, innovative approaches for integrating these sensor data must be developed.

1.2 The Waste Sorting System

The waste sorting system that is being used in this work is located at the Fraunhofer Research Institution for Materials Recycling and Resource Strategies IWKS at its Alzenau site, as shown in Figure 1. It consists of several modules on the scale of a pilot plant that can be configured and operated independently to simulate industrial sorting procedures. A zig-zag air separator and a flip-flow screen allow for classification, whereas magnetic and eddy-current separators can be utilized to separate metals. The most versatile module is a sensor-based sorting machine (Varisort Compact) by Sesotec, which is the focus of this article. A conveyor belt accelerates the objects to be sorted. They pass different sensors and are released from the conveyor belt at its end, leaving in a horizontal trajectory. Compressed air is then used mid-air to eject particles meeting the sorting criteria from the rest of the stream. The passing and the ejected fractions then fall into separate chutes and onto further conveyor belts for subsequent use.

Abb. 1: Waste Sorting System at Fraunhofer IWKS

2 State of the Art

As mentioned, the sorting machine contains three types of sensors, a line-scan camera, a hyperspectral near-infrared camera, and an electromagnetic inductive sensor. A line-scan camera can capture RGB information in the visible range and deliver data with a predefined principle. In this case, two RGB line-scan cameras are placed above the conveyor belt to further increase the resolution for color, shape, and object recognition. The hyperspectral near-infrared (NIR) camera can discriminate between various materials, such as plastic polymers, paper, and wood, based on their molecular characteristics by collecting chemometric information on particles. Finally, the inductive sensor is utilized to distinguish between conductive materials, such as metals. After sensor data acquisition and analysis, an array of valves is activated to release compressed air on the individual particles.

2.1 State-of-the-art Sorting Systems and WEEE recycling

Current state-of-the-art waste sorting systems utilize direct and indirect sorting mechanisms [GHT17]. Direct sorting utilizes differences in physical properties like density between objects to separate a waste stream into different fractions. Indirect sorting, on the other hand, uses sensors to detect object properties, before a separate mechanism is activated to dislocate them from the waste stream when sorting criteria are met. A large share of established sensor-based sorting machines are equipped with sensors in different ranges of the electromagnetic spectrum, such as near-infrared, the visible spectrum, ultra-violet, x-ray, or magnetic induction. Even though several publications could be identified by Kroell et al. about the use of machine learning on sensor data in waste recycling processes, only a few research has yet been published about the use of multiple sensors and AI to approach waste sorting [Kr22]. WEEE in the format of electrical and electronic devices is currently mainly being sorted by hand. The objective is to channel the devices into the corresponding treatment pathways, depending on legal demands and expected economic outcomes. The main route in WEEE recycling is mechanical treatment. Here, the electrical and electronic equipment is automatically comminuted via a shredder, crusher, mill, or similar machines due to the complexity of its components. Separating different materials from one another and reducing the size for easier sorting is necessary. Magnetic and eddy-current separation are then used to extract ferrous and non-ferrous metals, such as copper (e.g. in circuit boards) or aluminum, from the mix. Due to the intricate components of e-waste, the sorting processes are usually ineffective and recycling rates are low. So sensor-based sorting is utilized to improve the performance [Fo20]. Although these sensors can be employed either individually or in conjunction to tackle intricate sorting tasks, achieving accurate sorting of objects with multiple components, such as differentiating batteries from other electronic objects, remains a challenge.

2.2 Sensor Fusion in Object Detection

Object detection is a highly relevant topic in computer vision. Advances in image processing technology combining affordable computing platforms with novel methods, particularly those relying on deep learning, are revolutionizing the computer vision field and providing new opportunities for research with larger and more varied data sets.
Sensor fusion object detection refers to the integration of data from multiple sensors in order to enhance the accuracy and robustness of object detection. This approach is particularly useful in complex environments, such as those encountered in the waste recycling process, where multiple types of sensory inputs are required to fully characterize the objects to be detected. The use of multiple sensors, such as RGB cameras, near-infrared sensors, and inductive sensors, allows for a more complete representation of the objects. The integration of these sensory inputs poses significant challenges, including the fusion of data from sensors with different modalities, resolutions, and measurement characteristics. To overcome these challenges, various sensor fusion techniques have been developed, including feature-level fusion, decision-level fusion, and model-level fusion [Ca19].
One of the key challenges in sensor fusion object detection is the development of an appropriate architecture that can effectively integrate the sensory inputs and produce an accurate and robust object detector. The architecture should be designed to handle the high-dimensional data from the various sensors and to effectively fuse the data in a way that maximizes the information gained from each sensor. The selection of the appropriate architecture will depend on several factors, including the number and types of sensors used, the complexity of the environment, and the computational resources available [Ye21].
The majority of research on the subject of sensor fusion for object detection focuses on fusing different vision sensors, like RADAR, LiDAR, and RGB cameras. Different approaches to 3D object detection were proposed by [Li19] and [MCH+19]. Another aspect of this research area is different sensor fusion architectures, which have been investigated in several works. The main question regarding sensor fusion architecture that has been considered by [PD18] is how and when the sensory data should be combined or concatenated to achieve the best performance in the object detection process – in other words, finding out what the best strategy for data fusion is. This question leads us to different approaches as follows: Early Fusion, Middle Fusion, and Late Fusion [ZW19].

3 Data Exploration

As noted in Section 1.1, the multi-sensor sorting system used in this study consists of three sensors: line-scan camera, NIR, and inductive sensor. These sensors capture the data of the waste pieces on the moving conveyor belt. The NIR sensor captures 96 values for each pixel and inductive sensor data consists of 4 values for each pixel. The line-scan camera offers RGB values. The entire procedure entails data alignment followed by the application of a dimensionality reduction technique to preprocess the ultimate dataset for the data fusion strategy. For the collection of data, small discarded electrical and electronic devices of the collection category 5 in Germany were obtained from local recycling companies. In this study, four typical categories of e-waste are collected: routers, digital cameras, and two types of mobile phones (smartphones and key-operated phones). To test the sorting of pollutants in the waste stream, AAA batteries were added to the mix. To replicate an authentic sorting process, the devices were intentionally intermixed and subjected to multiple passes through the sorting machine to obtain a larger dataset comprising diverse and representative observations (see Table 1 for details).

Tab. 1: Overview of Objects and Datasets

	Camera	Router	Smart Phones	Key-operated Phones	AAA-Battery
Items (Devices)	70	65	93	100	120
Data Samples	165	225	402	228	225

4 Model Architecture

The selection of an appropriate model architecture for a sensor fusion object detection task is a crucial aspect in the development of such systems. Several factors must be taken into account in making this decision, including the type of object detection task being performed, the type of sensors used, available computational resources, and desired performance characteristics [AZ13]. In the case of sensor fusion object detection, a model architecture that is capable of processing multiple types of sensor data and integrating them is desired [ZW19]. The type of sensors used also plays a significant role in the selection of a model architecture. Different sensors provide different types of information about objects and this information may need to be combined and processed differently. For instance, a color camera provides visual information about an object's appearance while a LiDAR sensor provides information about the object's shape and location. Hence, a model architecture must be selected that can effectively process and integrate information from these different sensors. Computational resources must also be considered. Object detection models can be computationally intensive, particularly deep learning-based models, and the available computational resources must be taken into account when selecting a model architecture. This may involve trade-offs between performance and computational efficiency to ensure real-time operation. Desired performance characteristics, such as accuracy, speed, robustness, and generalization ability, must be considered in selecting a model architecture.

4.1 Base Model

Object detection models are used to do tasks such as object recognition, localization, and categorization. These models typically employ deep neural networks to learn and make predictions about objects in an image or a video stream. One such architecture is the Single Shot MultiBox Detector (SSD) model [Li16]. The SSD model architecture is designed to perform object detection in a fast and efficient manner. It is a single-shot model, meaning it makes predictions in a single forward pass through the network. This contrasts with two-shot models, which require two separate forward passes through the network to make predictions. The SSD model employs a base network, typically a pre-trained deep convolutional neural network (DCNN), to extract feature maps from the input image. These feature maps are then processed by several prediction layers, which produce bounding box proposals and class predictions. The predictions are then post-processed to obtain the final object detections [Li16]. One key aspect of the SSD model architecture is the use of anchor boxes, or priors, which are pre-defined bounding boxes with specific aspect ratios and scales. The prediction layers are designed to predict the locations of these anchor boxes relative to the objects in the image, as well as their class probabilities. This allows the SSD model to make predictions for objects of various sizes and shapes in a single forward pass through the network. In addition, the SSD model employs a multi-scale prediction strategy, which enables the model to detect objects at multiple scales in a single forward pass. This is accomplished by using multiple feature maps of different resolutions as input to the prediction layers, allowing the model to learn representations at multiple scales [Li16]. Overall, the SSD model architecture offers several advantages, including fast detection speed, efficient computation, and robust performance on a variety of object detection benchmarks. These qualities make the SSD model a popular choice for real-world object detection applications, such as autonomous vehicles, security systems, and industrial automation.

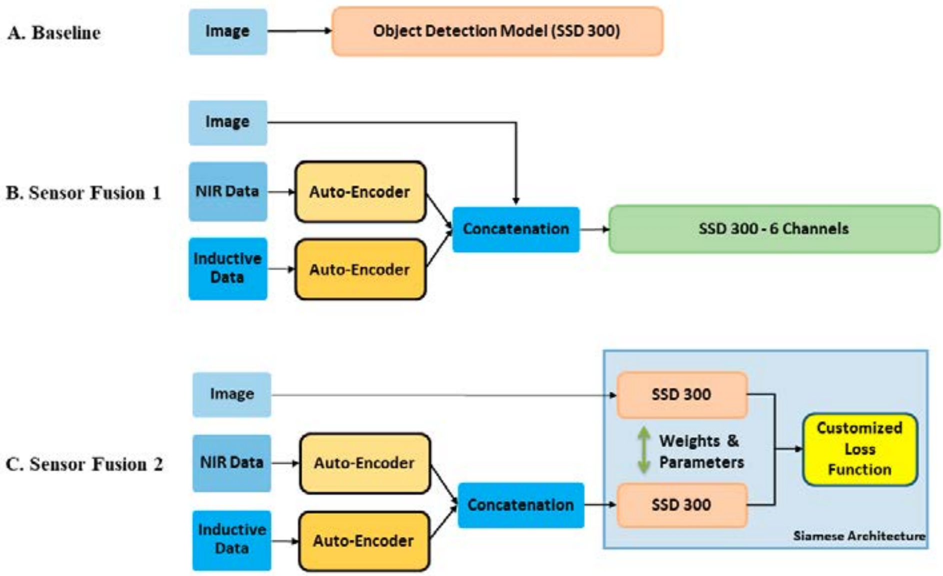

Abb. 2: Model Architectures

4.2 Sensor Fusion: Approach 1

In the first approach of data fusion, the image data is concatenated with information from the near-infrared (NIR) and inductive sensors. As previously mentioned, the NIR data consists of 96 values, while the inductive sensor data has 4 values per pixel. To ensure a suitable feature space, a dimensionality reduction method must be used that provides good performance on big data like the NIR dataset as well. Hence for each dataset, a deep autoencoder is trained to generate a latent space and extract feature channels. The underlying assumption is that since the main model used is object detection, which typically relies on 3 channels of image data, we concatenate 2 channels from the NIR data and 1 channel from the inductive sensor with the image data. Consequently, the resulting data encompasses 6 channels. Since the SSD model's architecture is initially designed for 3-channel input, the initial layers of the model are modified and adapted to adjust the 6-channel input. This adaptation allows for the effective integration of the combined data into the object detection framework. Figure 2.B represents the pipeline of the model architecture in this approach.

4.3 Sensor Fusion: Approach 2

In the alternative approach, which is called image mirroring, the fusion of NIR data and inductive sensor data is achieved through concatenation after applying a dimensionality reduction technique. This results in the creation of a 3-channel input data structure, mirroring the format of image data. These three channels consist of two channels derived from NIR data and one channel from inductive sensor data. The proposed model architecture follows a Siamese structure, comprising two identical models that are trained together. Within the Siamese model, input samples are mapped into a shared feature space. Each input sample is individually passed through one of the twin networks. Notably, these twin networks are trained simultaneously with shared weights, enabling them to collectively learn and represent the data effectively. By adopting this approach, the model benefits from training with both images and the combined information from NIR and inductive sensor data. This integration of multiple data sources aims to enhance the reliability and accuracy of the model. The inclusion of image data allows the model to leverage visual patterns and cues, while the NIR and inductive sensor data provide complementary information regarding material composition and electromagnetic properties. By utilizing all available information, the model can make more informed and precise predictions. Figure 2 illustrates the base model (2.A), the first sensor fusion model (2.B), and the Siamese model for the second sensor fusion approach (2.C).

5 Training and Results

Following the completion of data cleaning and matching procedures, a data set containing 1245 images along with their respective near-infrared (NIR) and inductive sensor data have been used for training (80%) and evaluating (20%) the baseline, sensor fusion 1, and 2 models. The results of training these models (refer to Table 2) show that both sensor fusion approaches provide a higher mean average precision (mAP) in comparison to the baseline model for waste object detection. Furthermore, the first sensor fusion model shows slightly better performance, at least with the current dataset. It shows that using additional sensors and their associated data in the waste sorting process delivers an improvement in the accuracy of detecting waste objects.

Tab. 2: mAP of the Models (overall and per class)

		Camera	Router	Smart Phone	Key-operated Phone	AAA Battery	Overall
Baseline:	SSD300	0.85	0.85	0.74	0.76	0.68	**0.777**
Sensor Fusion 1:	SSD300 (6 Channels)	0.76	0.95	0.92	0.80	0.86	**0.859**
Sensor Fusion 2:	SSD300 Siamese Model	0.79	0.89	0.91	0.83	0.83	**0.841**

6 Conclusion

The findings of this study demonstrate the effectiveness of AI-based sensor fusion techniques, specifically concatenation and image mirroring, in enhancing the accuracy of object detection within a waste sorting system. Integration of multi-sensor data led to a considerable improvement in mean average precision (mAP) for the object detection models. These findings show the potential of AI methods and machine learning models to enhance the waste sorting process. It is important to note that this paper presents a partial overview of the ongoing research, and certain details linked to data matching, model architecture, and training procedures have been omitted due to the page limit for "Work in Progress" papers. Further steps in this research include an in-depth analysis of the proposed customized loss function for the Siamese architecture. Additionally, exploring various sensory data sources and evaluating the effectiveness of each sensor within the waste sorting system present interesting points for future research.

Literatur

[AW19] Adedeji, O.; Wang, Z.: Intelligent waste classification system using deep learning convolutional neural network. Procedia Manufacturing/, 2019.

[AZ13] Angelova, A.; Zhu, S.: Efficient object detection and segmentation for fine-grained recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2013.

[Bi18] Bircanoğlu, C.; Atay, M.; Beşer, F.; Genç, Ö.; Kızrak, M. A.: RecycleNet: Intelligent waste sorting using deep neural networks. In: 2018 Innovations in intelligent systems and applications (INISTA). IEEE, 2018.

[BR15] Brunner, P. H.; Rechberger, H.: Waste to energy–key element for sustainable waste management. Waste management/, 2015.

[Ca19] Cai, J.; Meng, Z.; Khan, A. S.; Li, Z. et al.: Feature-level and model-level audiovisual fusion for emotion recognition in the wild. In: Conference on Multimedia Information Processing and Retrieval. IEEE, 2019.

[Fo20] Forti, V.; Balde, C. P.; Kuehr, R.; Bel, G.: The Global E-waste Monitor 2020: Quantities, flows and the circular economy potential./, 2020.

[GHT17] Gundupalli, S. P.; Hait, S.; Thakur, A.: A review on automated sorting of source-separated solid waste for recycling. Waste management/, 2017.

[Kr22] Kroell, N.; Chen, X.; Greiff, K.; Feil, A.: Optical sensors and machine learning algorithms in sensor-based material flow characterization for mechanical recycling processes: A literature review. Waste Management/, 2022.

[Li16] Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A. C.: Ssd: Single shot multibox detector. In: Computer Vision–ECCV: 14th European Conference, Amsterdam, The Netherland, Proceedings. 2016.

[Li19] Liang, M.; Yang, B.; Chen, Y.; Hu, R.; Urtasun, R.: Multi-task multi-sensor fusion for 3d object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.

[MCH+19] Meyer, G.; Charland, J.; Hegde, D. et al.: Sensor fusion for joint 3d object detection and semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2019.

[PD18] Pfeuffer, A.; Dietmayer, K.: Optimal sensor data fusion architecture for object detection in adverse weather conditions. In: 2018 21st International Conference on Information Fusion (FUSION). IEEE, 2018.

[Pl22] Plociennik, C.; Pourjafarian, M.; Nazeri, A.; Windholz, W.; Knetsch, S.; Ciroth, A.; Lopes, A.; Hagedorn, T.; Vogelgesang, M. et al.: Towards a Digital Lifecycle Passport for the Circular Economy. Procedia CIRP/, 2022.

[Ye21] Yeong, D. J.; Velasco-Hernandez, G.; Barry, J.; Walsh, J.: Sensor and sensor fusion technology in autonomous vehicles: A review. Sensors/, 2021.

[YTZ22] Yang, J.; Tao, F.; Zhong, Y.: Dynamic routing for waste collection and transportation with multi-compartment electric vehicle using smart waste bins. Waste Management & Research/8, 2022.

[ZW19] Zhang, H.; Wang, J.: Towards adversarially robust object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision. 2019.