Sorting of Smartphone Components for Recycling Through Convolutional Neural Networks

Álvaro G. Becker∗, Marcelo P. Cenci†, Thiago L. T. da Silveira∗, Hugo M. Veit†
∗Institute of Informatics, Federal University of Rio Grande do Sul
†LACOR - Department of Materials, Federal University of Rio Grande do Sul
Emails: {agbecker, tltsilveira}@inf.ufrgs.br, {marcelo.cenci, hugo.veit}@ufrgs.br

Abstract—The recycling of waste electrical and electronic equipment is an essential tool in allowing for a circular economy, presenting the potential for significant environmental and economic gain. However, traditional material separation techniques, based on physical and chemical processes, require substantial investment and do not apply to all cases. In this work, we investigate using an image classification neural network as a potential means to control an automated material separation process in treating smartphone waste, acting as a more efficient, less costly, and more widely applicable alternative to existing tools. We produced a dataset with 1,127 images of pyrolyzed smartphone components, which was then used to train and assess a VGG-16 image classification model. The model achieved 83.33% accuracy, lending credence to the viability of using such a neural network in material separation.

I. INTRODUCTION

In a report released by the United Nations University (UNU) in 2020, the global generation of waste electrical and electronic equipment (WEEE) was estimated at 53.6 million tons annually, or 7.3 kg per capita, with WEEE being the fastest-growing solid waste stream in recent years (from 9.2 million tons in 2014 to a projected 74.7 million tons annually by 2030) [1]. The context of WEEE generation also includes a high degree of informality in end-of-life management, with only 17.4% being properly documented and disposed of through formal means, primarily due to technological challenges in collection and recycling faced by the actors involved in this process [1]. From this scenario, the report emphasizes that recycling is a fundamental strategy for minimizing the environmental and societal impacts of the WEEE generation, as it is an essential component of the 2030 Agenda for Sustainable Development under the following United Nations Sustainable Development Goals: Goal 3 (Good Health and Well-being), Goal 6 (Clean Water and Sanitation), Goal 8 (Decent Work and Economic Growth), Goal 11 (Sustainable Cities and Communities), Goal 12 (Responsible Consumption and Production), and Goal 14 (Life Below Water).

Over the past decade, there has been a concentration of scientific efforts to find recycling solutions for WEEE. Typically, methods established in the metallurgical industry are adapted for WEEE processing. It is the case of the company Umicore, considered a global benchmark in the field, which has its processes based on copper and lead metallurgy, adding only 15% of WEEE to the primary ores and recovering only the most precious metals, such as gold and silver [2], [3]. At the base of the recycling chain, the collection and handling of WEEE are still carried out using inefficient methods, with a predominance of manual labor and massive wastage of valuable components [4], [5].

The ongoing work presented here stems from research efforts to increase the efficiency of processes and valorization of WEEE (in this case, applied to smartphones) in the early stages of recycling, where waste is managed by small and medium-sized recyclers lacking appropriate technology for these tasks. One of the main objectives is to minimize the need for manual handling of WEEE by automating the processes. By improving the efficiency of the early beneficiation stages (collection, sorting, and pretreatments), the downstream metallurgy processes are expected to be positively impacted in terms of material recovery.

The project begins with processing whole smartphones in pyrolysis furnaces (degradation of polymers in the absence of oxygen, generating byproducts with high energy value) to open the devices and release the components that will be subsequently separated. This approach considers all the components in the recycling chain, and no material is wasted. In Fig. 1, degraded smartphones resulting from pyrolysis are shown. Following the pyrolysis process, a critical step is separating batteries from other electronic components (screens, printed circuit boards (PCBs), and metal parts) because the
methods for recovering materials from batteries differ from those used for other components. Batteries of smartphones contain high concentrations of lithium, which is considered a critical and strategic material for many nations and companies due to the limited availability of primary ores and limited international supply [6]. The recycling of batteries emerged as a strategy to mitigate this problem, but specific methods of recycling are needed. If the batteries are not separated from the other WEEE, the lithium is lost in the slag of the typical pyrometallurgical processes, or diluted among other elements, hindering hydrometallurgical approaches [7] [8].

At this point, the challenge of separating batteries was addressed through a machine learning strategy for image-based component separation, which has the potential to achieve efficient separation without hindering the subsequent recycling processes through metallurgical separation processes. The general idea is to allow the industry to perform these activities automatically, in which a detector and a mechanical sorting device could be coupled to a conveyor belt to separate components.

The rest of this paper is organized as follows. Section II reviews the few related works that aim to sort waste using automated learning-based solutions. The proposed methodology is described in Section III. The results are presented and discussed in Section IV. Section V concludes this paper and indicates future works.

II. RELATED WORKS

Image-based waste separation has gained traction recently [9]–[13], with predominant applications in separating urban waste such as paper, plastic, glass, and metals.

Lu and Chen [9], in a literature review from 2022, found eight studies applying artificial neural networks for waste separation. Traditional backbones are trained and applied to the waste separation task in most works found. Bobulski and Kubanek [10] developed a convolutional neural network (CNN) to segregate different types of plastics for recycling, achieving an accuracy of over 99%. Zhang et al. [11] used a DenseNet169 [14] to segregate different kinds of household waste (glass, paper, textiles, metals, and plastics) and achieved 82% accuracy in their tests.

For WEEE applications, some studies specific to the field were published between 2022 and 2023. Yang et al. [12] applied the YOLOv4 network [15] to classify WEEE that potentially have internal batteries from those that do not, such as laptops and printers, achieving an accuracy of 90.1%. Lu et al. [13] used the YOLOv3 network [16] to detect and separate previously disconnected electronic components from PCBs, such as capacitors and transistors, with accuracies exceeding 90%.

To the best of our knowledge, the present work is the first to use CNNs to classify WEEE dismantled components focusing on recycling. We note that sorting pyrolyzed components turns out to be a challenging task since the individual components lack characteristic texture and shape. Separating these components (screens, PCBs, metals, and batteries) is a critical step in the recycling routes for most WEEE, and automating this step, besides being an operational and academic innovation, can generate significant increases in profitability and material recovery efficiency.

Specifically, we aim to achieve high accuracy in battery separation, creating a material stream concentrated in batteries. In other words, to prevent other components from being misclassified as batteries and to prevent batteries from being misclassified as other components.

III. METHODOLOGY

To carry out the project associated with this paper, 123 smartphones were gathered through collecting campaigns and the support of partner companies and research projects. In addition, 27 detached smartphone batteries were also gathered with support of partners. No restrictions were imposed on any characteristics of the devices, except for the requirement that they adhere to the smartphone-style. The following sections provide a detailed overview of the methodological steps applied here, which can be observed in Fig. 2.

A. Smartphone pyrolysis

The smartphones were processed in a batch electric resistance furnace. The process conditions were as follows: nitrogen atmosphere, temperature of 600°C, ambient pressure, heating rate of 300°C/h, and residence time of 1 hour. Fig. 1 provides examples of the materials that result from the process. The material was also submitted to a screening with a 2cm opening to remove small particles. These small particles, after the battery separation, can be sent back to the flow of other components (PCBs, glass, and screens).
The technique of pyrolysis, applied to WEEE, generates various benefits to the downstream recycling chain. By the degradation of polymers, encapsulated metals are liberated, and the mass concentration of valuable materials increases, facilitating the posterior chemical attacks. In addition, magnets are demagnetized, the total mass to be processed decreases, metals are kept in their reduced form (preventing oxidation), and energy valuable liquids and gases are generated. The technique of pyrolysis is a hotspot in recent research, being applied to many WEEE as the first step of recycling routes. Pyrolysis furnaces are commonly used in the treatment of wastes, especially for organic wastes, and are expected to conquer more space in the WEEE recycling industry in the middle term.

B. Image capture

The image capture took place with the material randomly selected. The decision was made to position various components in a single image to facilitate future study of the image-based detection task. The present article, however, tackles the multiclass classification problem.

In total, 300 3072 × 3072 colored images were captured. Each image contains at least one battery, one piece of metal or PCB, and another random piece. The pieces were randomly placed on a background prepared in gray, black, or white (in equal proportions) under variable ambient lighting (no artificial illumination is used). The images are captured perpendicularly, roughly 50 cm from the background. Fig. 3 shows an example image captured on a white background with one battery, one PCB, a small piece of glass, and another small metal piece.

After capturing the images, the components were flipped to display the reverse side, their positions were swapped, and another shot was captured. Thus, for each component, two images are captured, one for each face, without duplication. After capturing both images, the components were stored separately to prevent repetition.

C. Image annotation and pre-processing

The images were annotated using the Roboflow platform [17]. All components in each image were manually outlined with a polygonal drawing tool. The components were then assigned one of four classes: Metal Piece, Battery, Glass, or Printed Circuit Board (PCB). These four classes are omnipresent in every smartphone. Indeed, after pyrolysis, they are visually the only ones to remain in gross granulometry along with the batteries.

Very small components in some pictures were not annotated, as they were considered to have little or negative contribution to learning. They were instead left as part of the background, acting as visual noise elements that will also be present in real-life applications.

The annotated images were exported in an oriented bounding box (OBB) format [18]. Then, for each component, a square, horizontal bounding box was circumscribed about the component’s original OBB. The contents of these square BBs were then exported and resized to 500 × 500 pixels, thus composing a dataset of individual component images. Fig. 4 illustrates this process for the components shown in Fig. 3.

In total, the dataset contained 1,127 images. The images were split in a 70:20:10 ratio between training, validation, and test sets, respectively, with the split happening after the cropping process, given the uneven distribution of different components within the original images. The distribution of these images among the different classes and subsets is shown in Table I. Note that, due to the small size of the dataset, we avoided discarding any component images, despite the uneven quantities of different materials present in the images. Because of this, the entire dataset is slightly unbalanced, presenting 19.25%, 21.83%, 26.62%, and 32.30% of metal piece, PCB, battery, and glass instances, respectively.

D. Neural network training

For the desired image classification model, the VGG-16 [19] architecture was used as a backbone, being chosen due to its shallow depth being appropriate for the small dataset used [20]. We only adjust the last VGG-16 layer to the desired number of classes.

The network we used was pre-trained on the ImageNet dataset [21]. Our ablation study conducted the impact of pre-training, which significantly improves results, as discussed later. To account for the small dataset, data augmentation was
used. A random combination of transformations was applied to each image for each training epoch. More precisely, we considered:

(i) Rotation, within \( \pm 45^\circ \);
(ii) Shear, within \( \pm 5^\circ \);
(iii) Zoom, up to 20%;
(iv) Channel shifts, within \( \pm 10 \);
(v) Horizontal flips;
(vi) Vertical flips.

The shear, zoom, and channel shift transformations were constrained within narrow value ranges so as not to deform the images to the point of impairing learning.

The training was conducted on an NVIDIA 4070 Ti GPU. The training was set to run for a maximum of 100 epochs, with early stopping enabled with a patience value of 10 epochs. This condition monitored the validation accuracy so as to export the model from the epoch where it achieved its highest value in that time frame. Batch size 32 was used, with categorical cross entropy as the loss function and the Adam optimizer with a learning rate \( 10^{-3} \).

IV. RESULTS AND DISCUSSION

A. Training results

In this paper, we consider the overall accuracy, and per-class precision and recall as figures of merit. The training accuracy is calculated as

\[
A = \frac{N_{True}}{N_{Total}},
\]

where \( N_{True} \) is the number of correct predictions by the model and \( N_{Total} \) is the total number of predictions made by the model.

The precision of a given class is calculated as

\[
P = \frac{N_{TP}}{N_{TP} + N_{FP}}
\]

and recall as

\[
R = \frac{N_{TP}}{N_{TP} + N_{FN}},
\]

where \( N_{TP} \) is the number of true positives guessed by the model (that is, elements of the class in question which were correctly labeled as such), \( N_{FP} \) is the number of false positives (elements incorrectly labeled as belonging to that class), and \( N_{FN} \) is the number of false negatives (elements belonging to that class which were labeled as something else).

The resulting model was trained for 20 epochs (due to the early stopping), with epoch 10 yielding the best results. It achieved a training accuracy of 82.49% and a loss of 1.3541. Meanwhile, the validation accuracy reached 79.28%, with a validation loss of 1.9309. Fig. 5 shows the graphs for Accuracy and Loss over time for training and accuracy metrics.

Running an inference test with the model on the test dataset resulted in a mean precision of 77.29% and a mean recall of 77.35%. The confusion matrix of this test is given in Table II, while the precision and recall metrics for each class are shown in Table III.

These results show that the network did gain the ability to generalize what it learned, and show promise for even more robust results from this classification network if trained on a larger dataset, given the small size of the one used. It should also be noted that the high contrast in precision and recall values between classes is likely attributable to the unbalanced nature of the dataset, seeing as the lowest values are associated with the two least represented classes (Metal Piece and PCB). Expanding the dataset in subsequent works would allow for eventual discarding of images so as to test training the model with a balanced dataset.

Regarding the materials flow for recycling, the model’s performance was surprisingly successful, even more so considering that many opportunities for improvements and alternative tests are still to be considered. It is essential to highlight that the study’s primary goal is to separate batteries. Thus, the results of precision (90.32%) and recall (93.33%) for the class of batteries determine the study’s success.

Of the thirty tested batteries, twenty-eight were correctly identified as batteries, and two were considered PCBs (93.33% of recall). As stated, batteries must be separated to recover lithium through specific metallurgical processes [8]. Batteries falling into the flow of PCBs preclude the recovery of this element. In the flow of the class of batteries, there were 28 batteries, two metal pieces, and one PCB (precision of 90.32%). The two metal pieces in the flow of batteries do not hinder the recyclability of any of the two materials. Indeed, batteries contain a metal casing [22], easily treated in the typical recycling route of batteries. The one PCB in the flow
of batteries means that the valuable materials present at the PCBs (gold, silver, and others) would not be recovered in their typical recycling route. Finally, the model efficiently removed the glass from other flows of materials (83.33% of recall). This is an important outcome, considering that glass is not a valuable material and only increases the mass of the materials needing treatment [23].

B. Ablation studies

1) Pre-training: Because the background of the images in our dataset is quite simple (nearly smooth white, gray, or black surface) and the pyrolyzed materials do not present color variation, we questioned the necessity of pre-training. It is reasonably assumed that pre-training the model on large datasets such as ImageNet [21] might be excessive for this setup problem.

As such, to verify the validity of using a pre-trained network, a version of the model was trained with randomly initialized weights to compare its performance with the pre-trained model. All the training conditions were identical between models other than the initial weights.

The randomly initialized model trained for 32 epochs, achieving its best results at epoch 22, with a training accuracy of 46.98% and loss of 1.2000, and a validation accuracy of 45.95% and loss of 1.2115. The training graphs are shown in Fig. 6, with the accuracy graph showing its learning process was inconsistent.

In the inference evaluation with the test dataset, the model had a mean precision of 38.71% and recall of 40.49%. In fact, it did not label a single metal piece component correctly, as can be seen in Table IV. The precision and recall for each class is shown in Table V.

Thus, the results with the randomly initialized network are drastically worse than the pre-trained network, justifying the choice to use a pre-trained model.

2) Binary classification: Because the work mainly emphasizes separating batteries from other materials, we also questioned if we could not optimize learning for only detecting what is and isn’t a battery, especially considering that the battery class has the most visually consistent appearance. For that reason, we trained a version of the model with only two classes: Battery and Other (including PCBs, glasses, and metal pieces). This model was also pre-trained on ImageNet, and all training conditions except for the classes were equal to those of the first model. The resulting model trained for a total 20 epochs, with the best one being epoch 10. It had a training accuracy of 94.71% and loss of 0.2481, with a validation accuracy of 91.89% and loss of 0.2816. The confusion matrix is shown in Table VI, and the class-wise precision and recall are shown in Table VII. The training graphs are shown in Fig. 7.

The drawback of the binary approach is that it is not possible to understand which component is contaminating the flow of batteries and therefore improve the performance of the model based on this information. As discussed in Section IV-A, if the only class being incorrectly sorted alongside batteries is metal...
TABLE IV
CONFUSION MATRIX OF THE RANDOMLY INITIALIZED MODEL’S TEST PREDICTIONS

<table>
<thead>
<tr>
<th></th>
<th>Metal Piece</th>
<th>Battery</th>
<th>PCB</th>
<th>Glass</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Actual</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metal Piece</td>
<td>0</td>
<td>7</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Battery</td>
<td>1</td>
<td>14</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>PCB</td>
<td>3</td>
<td>3</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Glass</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>31</td>
</tr>
</tbody>
</table>

TABLE V
PRECISION AND RECALL FOR INFERENCE IN EACH CLASS FOR THE RANDOMLY INITIALIZED MODEL

<table>
<thead>
<tr>
<th>Class</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metal Piece</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Battery</td>
<td>56.00%</td>
<td>46.67%</td>
</tr>
<tr>
<td>PCB</td>
<td>36.84%</td>
<td>29.17%</td>
</tr>
<tr>
<td>Glass</td>
<td>50%</td>
<td>86.11%</td>
</tr>
</tbody>
</table>

Table VI
CONFUSION MATRIX OF THE BINARY MODEL’S TEST PREDICTIONS

<table>
<thead>
<tr>
<th></th>
<th>Battery</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Actual</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Battery</td>
<td>26</td>
<td>4</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>81</td>
</tr>
</tbody>
</table>

Table VII
PRECISION AND RECALL FOR INFERENCE IN EACH CLASS FOR THE BINARY MODEL

<table>
<thead>
<tr>
<th>Class</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Battery</td>
<td>100.00%</td>
<td>86.67%</td>
</tr>
<tr>
<td>Other</td>
<td>95.29%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

Fig. 7. Accuracy graphs (upper) and losses (bottom) of training and validation along the training.

In this work, we present the possibility of using an image classification neural network as a cheaper and more efficient alternative in separating materials for recycling WEEE. We created an image dataset with 300 pictures of assorted components from pyrolyzed smartphones and an adapted version of this dataset for image classification with 1,127 individual images of these components. We then trained an image classification model on this dataset to differentiate between metal pieces, battery, PCB, and glass components, achieving an overall accuracy of 82.49%.

The experiment results show promise for using a neural network to separate WEEEs, especially considering the small dataset used. Future works could expand the goal of this study to a detection problem, creating a model capable of both classifying and locating multiple components at once in real-time, which could even further allow for a compact and efficient automated separation system. The relatively high accuracies obtained with a lightweight network could also feasibly allow for low-cost yet accurate embedded systems to be developed for this purpose. Future works could look into testing other models—both alternative architectures and possibly detection models—as well as investigating the impact of the unbalanced dataset and ways to mitigate it, such as with class-specific augmentations or by using weighted loss functions.

Even considering the study’s limitations, the model used achieved its purpose (separation of batteries) with high accuracies, with significant potential to be implemented on recycling routes and increase their effectiveness in the valorization of the residues. This approach can be implemented in a recycling line coupled with a mechanical sorting system. A camera pointing perpendicularly on a conveyor belt could feed a detection system that informs the mechanical devices to separate the target components. It is important to highlight that most electronics usually contain the same components as smartphones (metal pieces, glass, PCBs, and batteries). Thus, this approach has the potential to be applied to other types of devices, such as laptops and tablets. In addition to the simple separation of batteries, more studies can be carried out to refine the separation of the other classes of materials with high accuracy.

Finally, the study created a dataset that could serve as a basis for the future development of models for similar cases. The fact that the dataset was annotated with polygonal masks also allows the annotations to be exported in different formats for different purposes, such as fine-grained semantic segmentation. We also intend to add more images to the dataset and, ultimately, make it public.

ACKNOWLEDGMENTS

We acknowledge the financial support from the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul
References


