Waste Electrical and Electronics Equipment Classification using Deep Neural Network Sensor Fusion

Lasse Mølgaard

Abstract— **The recycling of Waste Electrical and Electronic Equipment is an issue of growing importance. Classification and sorting of these products in recycling plants can be a promising approach to leverage the full value of these waste streams.**

We present a sensor fusion approach employing convolutional neural networks to classify WEEE products using an RGB camera with NIR capability in a waste sorting scenario. We demonstrate that convolutional neural networks trained on natural images can be used as feature extractors before a final output neural network layer that adapts to this domain of image classification. This approach means that adaptation to this domain can be done with a relatively small amount of training data.

The paper presents a dataset containing 10 classes of WEEE objects, such as cellphones, blender sticks, power drills and electrical kettles, recorded in a realistic sorting scenario.

The resulting sensor fusion network produces a classification accuracy of 73% for images with overlapping objects from the 10 classes.

I. Introduction

The recycling of Waste Electrical and Electronic Equipment(WEEE) is an increasingly important issue. Growing volumes of WEEEand stronger legislation on the handling of discarded electrical devices makes efficient sorting of these items an interesting topic.

Application of sensor based sorting to waste streams is an increasing industry. Many different sorting technologies have been employed to build automated sorting systems to be able to sort commonly occurring waste fractions such as glass, plastics, paper, and metals. Successful sensor technologies include color cameras, near-infrared (NIR) spectroscopy, and X-ray fluorescence. These technologies perform sorting based on physical properties of the materials, e.g. to separate clear from colored plastics or to detect of polyethylene or polypropylene samples.

Current recycling plants for handling of WEEE predominantly apply a process where the recyclable products are crushed or shredded to split the products into fragments that can be sorted based on simple characteristics such as color for printed circuit boards, magnetization for ferrous metals, etc. Before shredding, objects must be sorted by hand to remove hazardous objects from the waste stream, such as batteries or objects that have high value. This manual sorting process could be automated by using robot-based sorting at an early stage of the recycling plant.

We propose to detect complete WEEE objects based on imaging technologies. The classification of natural objects has been made possible with the recent developments in deep neural networks. The application of convolutional neural networks (CNN) to big datasets to classify objects in natural images[3] has resulted in very successful network structures [4], [6] that can detect objects from thousands of categories.

The research has been performed as a part of the INNOSORT project (http://innosort.teknologisk.dk/), which is funded by the Danish Agency for Science, Technology and Innovation.

This paper presents a novel dataset of WEEE objects obtained using the INNOSORT sensor platform[7]. The combination of color images and NIR are used to train a fusion network that employs both imaging modalities to achieve increased object classification performance.

II. Method

Applying CNNs for waste sorting poses a number of issues. 1) Dataset size - the number of images needed to train a CNN is very large. Typically training requires thousands of annotated images which may be very costly to obtain for a waste sorting task. 2) Sensor fusion - most of the efforts in developing CNNs has focused on RGB images as these are readily available databases of annotated images. Extending the modality with NIR[5], depth[2] is a less investigated area of research.

The former issue of obtaining enough data has been addressed using a combination of two techniques; transfer learning and data augmentation. Transfer learning is a technique tries to leverage knowledge learned in one problem and to transfer this knowledge to a related problem. In practice this can be accomplished by training a CNN on a large dataset (such as Imagenet for natural objects) and then fine-tune the trained network on the target problem. This has shown to be a viable way to successfully apply convolutional networks to problems with small training sets[1]. Data augmentation is used to increase the effective number of images in a small dataset by performing a number of different transformations of the images. In the context of waste sorting the objects in the waste streams can have any orientation and can be obscured in a number of ways when they move around on a conveyor belt. Therefore the training data can be augmented by implementing a number of transformations that can emulate these sorts of alterations to the images.

A. *Sensor fusion*

Bringing the Scanlab array of sensors into play despite the small number of samples available is a key challenge in the Innosort project. At present, we have tested the the use of the RGB and the NIR imaging modality. The two modalities have been used as input to an Inception network[6] which is one of the most recent CNNs proposed

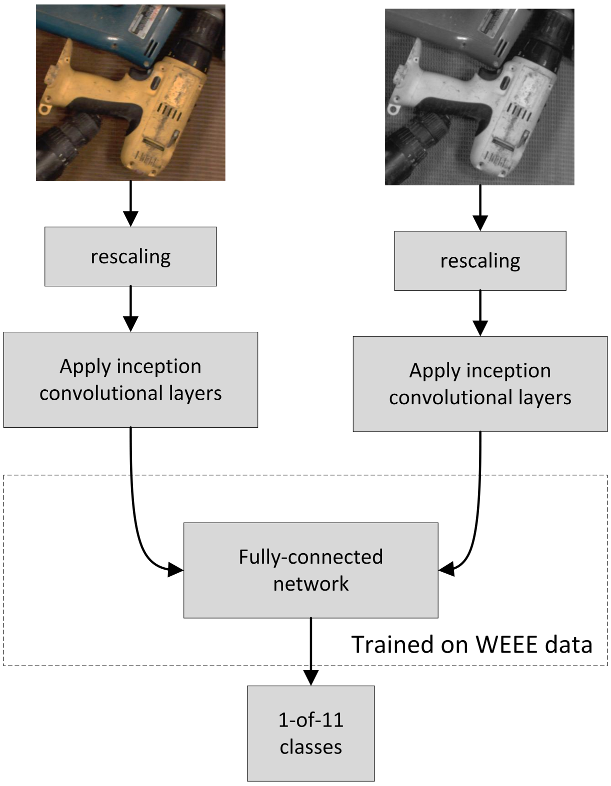


Fig. 1: Neural network structure for late sensor fusion of RGB and NIR images.

for image classification. The architecture can be seen in figure

Using the Inception network for classification of NIR images is accomplished by interpreting NIR image as a gray scale image, which is then fed into the network using three channels. This approach is somewhat redundant but saves the effort of retraining a network from scratch.

A straightforward way to combine the information from the two image modalities is to let each network produce predictions on which class is in the image and then do an aggregation of these predictions. This can be done using a voting scheme or more complex ensembling methods.

The approach taken here is to combine the information using the intermediate representation of the Inception network after the convolutional layer outlined with the red box in figure **??**, and then retraining a new fully connected layer to combine this information. This network structure will leverage some of the lower level information extracted from the images in a useful way.

III. WEEE SENSOR FUSION DATASET

We present a novel dataset for WEEE classification and segmentation. The dataset was produced using the Scanlab demonstrator system at the Danish Technological Institute [7]. The Scanlab demonstrator was fitted with one CCD camera with NIR capability. Alignment of RGB and NIR images was therefore automatically aligned and did not require further processing.

TABLE I: Summary of images and objects for each waste type in the recorded WEEE dataset

Type	Recorded images	Annotated images	No. of annotated objects
Blendersticks	66	46	83
Car stereos	60	29	36
Cellphones	162	41	174
Hair dryers	79	38	53
Harddrives	67	22	67
Irons	126	51	78
Electrical Kettles	95	40	92
Laptop computers	124	63	86
Electrical mixers	74	36	115
Powerdrill	114	39	70
Mixed	449	-	352

A. Samples

The WEEEdataset samples were gathered from a simulated WEEE fraction. A collection of objects were borrowed from a municipal recycling station, classified into ten groups of objects, as listed in table I

The set of items are inspired by objects that normally appear in a mixture of WEEE collected at a recycling station. The images were obtained by placing the objects on a conveyor belt moving the objects under the camera. from each group being loaded onto the belt. This has produced a collection of images where only one type of objects is in each image. Each image has then manually been annotated, by placing a bounding box around each object in the image. Table I summarizes data in the dataset. Examples of the sample images are shown in figure 7.

In addition to the data generated with each WEEE type isolated in the Scanlab a number of sessions were recorded where a mixture of the objects fed onto the conveyor. This mixed fraction simulates what we would see in a pre-sorting scenario in a WEEE sorting plant. The objects, therefore, overlap in random ways and may also be rotated in different ways, as can be seen in figure 2.

IV. CLASSIFICATION RESULTS

The current solution is based on using the Tensorflow framework. The choice of using Tensorflow for this task is based on the flexibility of the framework and that it is used by a broad developer base. This means that many pre-trained models are made available by different developers that can be adapted for new applications.

One of these pre-trained networks is the Inception network[6] which has been trained on the Imagenet dataset using a farm of high-end GPUs.

Using the pre-trained Inception network for the WEEE dataset can be exploited by applying transfer learning. Transfer learning is practically used in this application by retraining the final layer of the Inception network from scratch while leaving all the others untouched. This approach was proposed in [1].

Although this approach does not on par with training a network fully from scratch, it has proven effective for many applications, and the adaptation was possible to perform on a laptop without the use of a GPU.

Fig. 2: Example mixed WEEE fractions.

The WEEE dataset has been used to retrain the Inception network by using the images recorded for each class of objects. The retrained network was consequently trained to classify an input image as one of the 10 object classes or as an empty image.

The performance of the neural network was then evaluated on the images recorded with mixed objects. This split of training the network on clean fractions and then evaluating on a mixture of objects aligns with the envisioned practical use of robotic sorting of WEEE.

A. *Object classification performance*

Initial retraining of the network using the basic dataset of 854 annotated object images resulted in a correct classification rate of 64% using the RGB modality.

To improve on this performance we implemented a number of random transformations of the images, such as rotation, flipping, and cropping which made it possible to augment the dataset in a meaningful way. Using these transformations a dataset of 5,000 images per object class was created, i.e. such that 50,000 images were used to retrain the network. Training the network exclusively using the augmented set of RGB images improved the classification rate to 72.4% on the test set of held out mixed fraction of WEEE objects. The errors are summarized in a confusion matrix, shown in figure 3. It is evident that the car stereos and cellphones are the most difficult to detect. The car stereos are most often confused with laptops, which seems explainable given that they are mostly seen as rectangular gray objects in the images. Similarly using the NIR images resulted in a classification rate of 71.5% on the test set. The confusion matrix, shown in figure 4, shows that the car stereos are also difficult to recognize using this modality. On the other hand, cellphones are better recognized as the classification rate rises to 59.7% compared to 49.3% compared to the RGB-based classification. This difference in classification performance for the single classes indicates that a fusion of the two modalities is a viable option.

1) Sensor fusion: The results of using the fusion network gives a small improvement over the single modality networks. The detection of cell phones is better than for RGB alone, while the other object classes are recognized at the same level.

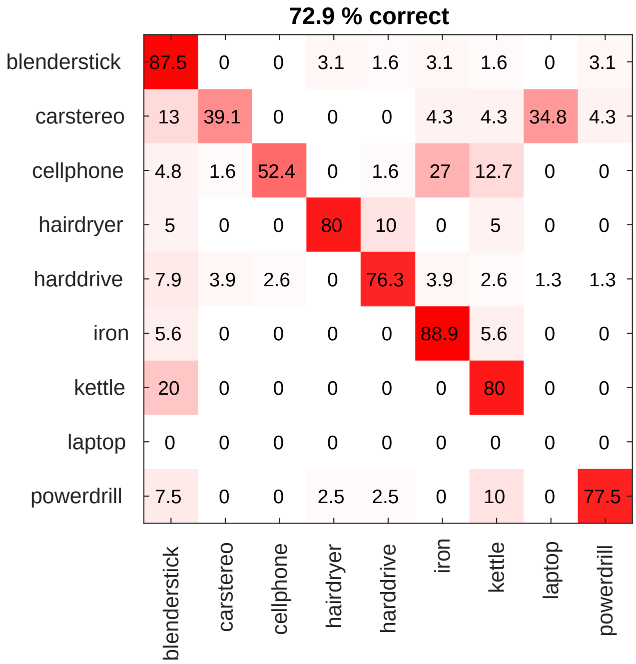


Fig. 3: Confusion matrix for classification on images from mixed WEEE fraction using only RGB images.

A number of wrongly classified images are shown figure 6 to highlight what kind of errors are being made by the network. In many cases, the images in the mixed fraction include objects from multiple classes, which means that the network might identify another object type than the one that was labeled manually. This could also sometimes mean that the object at the bottom would be unpickable by a robot and therefore missing this object is not that crucial. In other cases, some objects do not have visible features for the CNN to use, for instance, whether a gray rectangular object is a laptop or a car stereo. These observations would imply that a practical use of object based detection would require that objects are in some way spread out on the conveyor belt such that the overlap between objects is minimized. Furthermore,

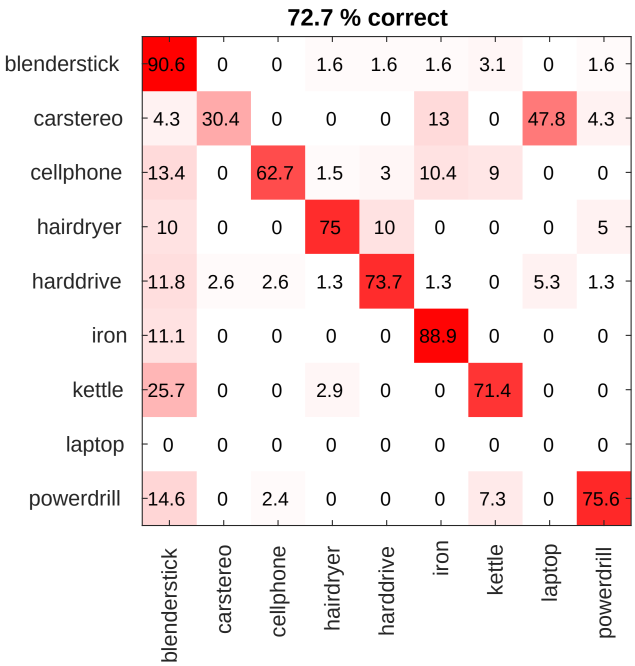


Fig. 4: Confusion matrix for classification on images from mixed WEEE fraction using only NIR images.

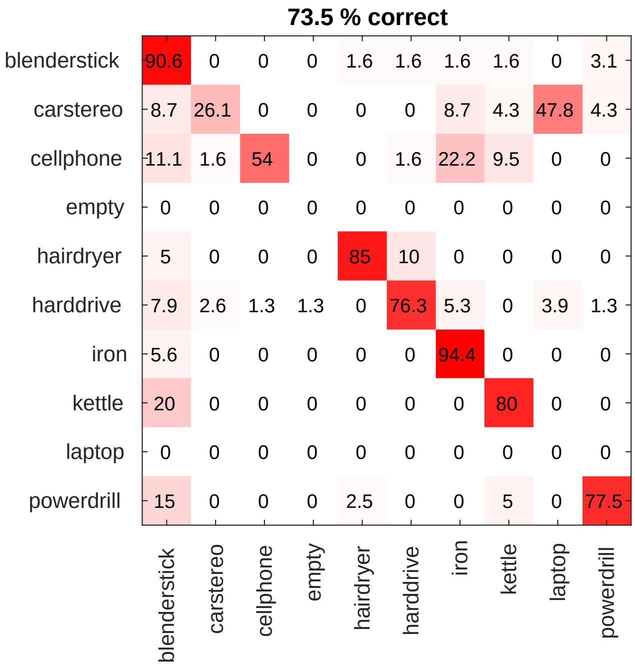
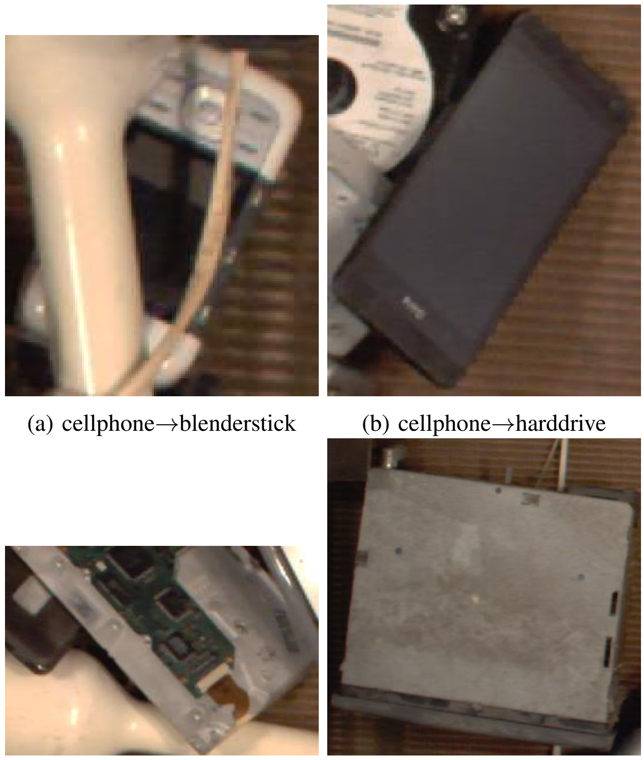


Fig. 5: Confusion matrix for classification on images from mixed WEEE fraction using fusion network.



(a) cellphone→blenderstick (b) cellphone→harddrive

(c) harddrive→blenderstick (d) car stereo→laptop

Fig. 6: Images of objects wrongly classified by the sensor fusion network.

we could add auxiliary input variables in addition to the images to the CNN, e.g. to measure the size of objects or naturally add the other available sensor modalities.

V. Discussion and Conclusion

The experiments have shown that it is possible to use a pre-trained convolutional neural network to classify 10 different types of WEEE objects. The implemented system has not directly addressed the problem of detecting where objects are placed on the conveyor belt. Fortunately, this extension of CNNs can be implemented by using a sliding window approach to classifying all areas in an image. Given the stationary background of the images; the number of subimages that must be classified can be filtered using standard foreground detection tools from computer vision.

The proposed training procedure suggests a practical approach to learning new classes. Each class should be fed into Scanlab and using a foreground detector will automatically suggest the parts of training images that can be used as inputs for training the CNN.

Future experiments could add other imaging modalities such as thermography and stereo imagery. The stereography modality would anyway have to be added for integration with a robot to estimate where to grab objects. Including depth information has also been shown to help in training CNNs for object detection even when depth information is not available at test time[2].

The accuracy of the classifier reached about 73% on the challenging task of classifying objects mixed in a realistic waste stream. The challenges include that the objects are placed on top of each other and are oriented in very differently from what is seen in the training phase.

The insights gained for this WEEE object pre-sorting scenario could be used as a basis for practical setup of sorting systems where a known set of objects are valuable to remove from a waste stream.

ACKNOWLEDGMENT

REFERENCES

[1] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. In *ICML*, pages 647–655, 2014.

[2] Judy Hoffman, Saurabh Gupta, Jian Leong, Sergio Guadarrama, and Trevor Darrell. Cross-modal adaptation for rgb-d detection. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5032–5039. IEEE, 2016.

[3] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[4] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252, 2015.

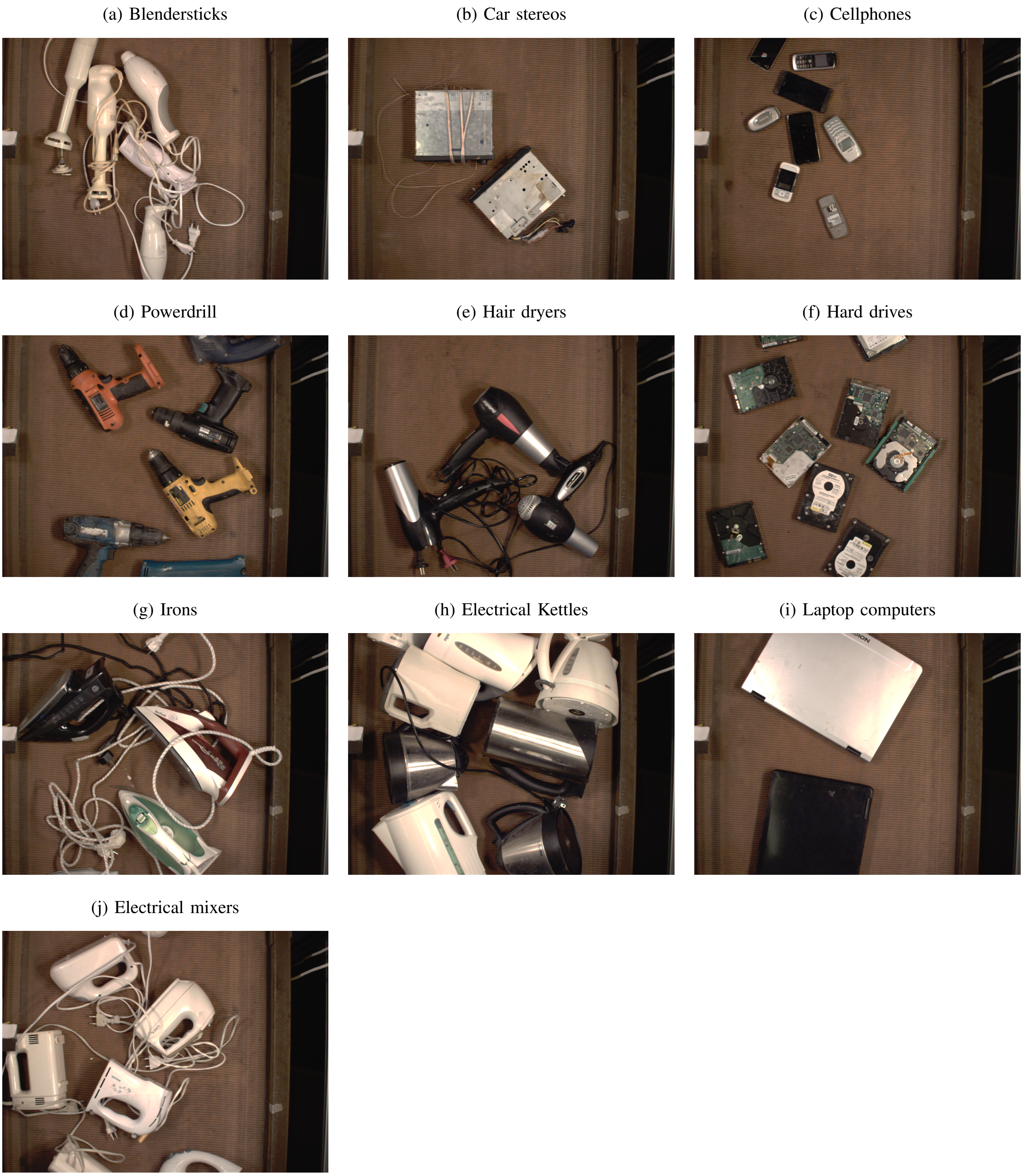
[5] Inkyu Sa, Zongyuan Ge, Feras Dayoub, Ben Upcroft, Tristan Perez, and Chris McCool. Deepfruits: A fruit detection system using deep neural networks. *Sensors*, 16(8), 2016.

[6] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. *CoRR*, abs/1512.00567, 2015.

[7] Mikkel Viager. ScanLab - Forsøgsopstilling for analyse af affaldsstrømme. Technical report, Teknologisk Institut, Forskerparken 10F, 5230 Odense, Robotteknologi, 2014.

APPENDIX

Figure 7 shows sample images of the types of WEEE photographed.



(a) Blendersticks (b) Car stereos (c) Cellphones
(d) Powerdrill (e) Hair dryers (f) Hard drives
(g) Irons (h) Electrical Kettles (i) Laptop computers
(j) Electrical mixers

Fig. 7: Example RGB images WEEE objects used for experiments.