Development of a method of detection and classification of waste objects on a conveyor for a robotic sorting system

A V Seredkin¹,², M P Tokarev¹, I A Plohih¹, O A Gobyzov¹ and D M Markovich¹

¹Kutateladze Institute of Thermophysics, Siberian Branch of RAS, 1 Lavrentyev Avenue, Novosibirsk, 630090, Russia
²Novosibirsk State University, 2 Pirogova Street, Novosibirsk, 630090, Russia

E-mail: sasha7-93@mail.ru

Abstract. Currently used recycling technologies have limitations on the composition of recyclable waste, which makes them specialized. Thus, the preliminary sorting of municipal solid waste is a necessary step, increasing the efficiency of using municipal solid waste as a resource. To sort municipal solid waste we developed a method for detecting and classifying waste on a conveyor line using neural network image processing. Images from a camera are fed to a neural network input, which determines the position and type of detected objects. To train the neural network a database of more than 13,000 municipal solid waste images was created. Mean-Average Precision for the neural network model was 64%.

1. Introduction

Municipal solid waste (MSW) is one of the types of post-consumer waste which includes biowaste, end-of-life goods and products, used packaging and etc. The continuous increase in MSW generation is associated with a population growth as well as with an increase of waste amount generated per person, especially among urban residents. This increases the importance of building an effective process chain for the disposal of MSW. Recycling technologies that are currently employed have limitations on the composition of recyclable waste. In the following paragraphs some of the most common methods of MSW disposal in Russia will be briefly observed.

Landfilling is technologically the easiest way of waste disposal. Only non-recyclable and non-hazardous wastes landfilling can be justified, otherwise the resource potential of the waste is lost, which makes landfill an inefficient way of waste disposal. In case of efficient waste management, the proportion of MSW sent to landfills may be reduced to 10% or less. For example, landfill reduction to a maximum of 10% of municipal waste by 2030 is one of the aims of European Commission Circular Economy Package.

Composting is a technology of recycling based on biological decomposition, with the aim of obtaining mineral fertilizers. It is mainly used for green waste, such as leaves and mowed grass. The quality of the compost is directly dependent on the purity of the feedstock. The presence of a large amount of non-decomposable components, such as glass, metal, most of the plastics, makes the resulting fertilizers unsuitable for further use [1].

Incineration of waste allows using it as a fuel for energy generation. There is a number of related technologies, including incineration on grates, plasma incineration, and so on. The composition of MSW determines its specific heat of combustion and ash content. For example, food waste, paper and cardboards, textiles, leather, rubber, most of the plastics have a high calorific value (more than 10 MJ
per kg of dry weight) and low ash content. At the same time incineration of some plastics and rubbers is hazardous as they release large amount of poisonous volatiles when burnt at a temperature lower than 850 °C. In addition, recycling is more productive way of disposal than incineration for a significant part of plastics. Glass, metals, electronic scrap, construction waste, lamps and batteries have high ash content and only reduce the quality of MSW as a fuel.

As one can see, the type of input material affects the environmental impact and efficiency of common recycling methods. Thus, the preliminary sorting of MSW is a necessary step to increase the efficiency of using MSW as a resource.

In this paper, a general design of the novel automated sorting unit is described and a neural-network based method for detecting and classifying waste on a conveyor line, meant for embedding into the sorting unit, is proposed and tested.

2. General design of the sorting unit
The sorting unit performs detection and classification of waste components on the conveyor belt and transports objects from the belt to the collection bins according to the type of waste. The sorting unit design assumes that unpacked waste is supplied to the conveyor belt as a sparse feed so that cases of objects overlapping are rare. Individual objects are detected and classified by a machine vision system and their position on the belt is evaluated. In addition, height profile of the objects on the conveyor belt is measured with a height profile scanner. Coordinates, height of an object and classification results are then supplied to the collection module, which moves objects from the belt to one of the collection bins by a handling device.

Height profile scanner is the first system along the conveyor path. The system consists of a laser sheet and a camera. The laser is directed vertically downwards, and the laser sheet is directed across the conveyor belt. The camera that captures the shape of the laser sheet on the surface of the object is located above the conveyor belt (Fig. 1).

![Figure 1. Setup of a height profile scanner. The laser source is in the upper left corner, the camera is in the upper left corner.](image)

Before measuring the height profile, the camera is calibrated: a calibration target is placed in the plane of the laser sheet and the calibration program calculates all parameters necessary to rectify an image and determine the height of the laser line in millimeters. As a result, the height profile of objects on a conveyor belt along the cross-section of the laser sheet can be evaluated from each frame. From the sequence of measured height profiles the heights map is compiled.

The classification module is the core module of the system. Hardware part of the classification module includes only a camera, which continuously takes images of the conveyor belt, and a computer
with a neural-network-based classifier software. Classifier acquires images from the camera and determines the position and type of objects in the images. The neural network model and training dataset and results achieved in waste objects classification are described in the following chapters of the paper.

The collection module is the last module along the conveyor belt. Proposed design assumes that a fast delta-robot is used as a handling device that moves objects from the belt to the collection bins. The delta-robot uses a vacuum grip as an end effector. Comparing to a mechanical grip, the vacuum grip is more tolerant to the shape and form of the object, and, as most of the MSW objects are quite light, its gripping strength is enough to pick and transport them. The list of the detected objects and all the related data from the other modules, such as coordinates, height and type, are fed to the robot’s controller, which schedules robot’s movement. Obviously, another type of the robot can be used for the collection module. Considerations about robotic systems architecture for waste sorting can be found in [2].

3. Neural network model for objects classification

Currently machine learning algorithms are one of the main tools for solving classification tasks [3]. Obviously, an idea of employing this tool for waste sorting is not new: examples of testing different neural network models for this task can be found in [4 - 6].

Generally speaking, there are two key points in such works: a model, that is tested, and a database for training and testing of the network. Earlier publications reported testing of classical convolutional neural networks [4], [5]. Although some models demonstrated quite good classification results, a major drawback of such solution is that a whole image is classified at once. This means that an input image can contain only one object. It is hardly possible to provide such images from the MSW stream on a conveyer belt, or at least some additional measures should be taken to isolate objects in images.

A more recent advancement in neural networks design is so called region proposal convolutional neural network (R-CNN). This model incorporates different blocks for detection of objects in images and classification of the detected objects. In this case, multiple objects of different type can be recognized in one image. Currently there are several ‘branches’ of such networks that have different architecture, some of those are presented in [7] and [8]. There have been few examples of successful employment of region-proposal neural networks for waste sorting. One of such works is [6], in this work authors tested a R-CNN-type neural network which was trained on a dataset of 1332 scenes with multiple labeled objects in the scene. Authors reported classification precision of 78.5 %, however, the paper contains very few details about the dataset, training and testing procedures.

In the present work one of the models from the Tensorflow detection model zoo [9], a Faster R-CNN ResNet101, initially trained to determine the position and class of objects of the COCO database, was adopted for the task of waste objects detection and classification.

A major issue in adaptation of CNN - or R-CNN-type neural network for a new task is that it requires a large training set that would, ideally, fully represent all possible features inherent in classes of images under consideration. In the reported cases, training was typically performed on a relatively small datasets. The small dataset can lead to a bias in estimation of models’ performance. In this regard an own database of MSW images was created. In addition a method of making large synthetic training data set from single MSW objects images is discussed in details further.

4. Training database

Waste recovered directly from separate waste collection bins in Novosibirsk region was used for the database. The waste was carefully pre-sorted to avoid errors in objects labelling. Images of MSW objects were taken on a stationary conveyer belt to resemble a situation of actual waste sorting at the sorting facility. More than 14,000 images of individual MSW objects were taken. An archive with the database can be downloaded from the webpage https://sites.google.com/view/lfeit/trash-sorting-database. Three types of objects were used for the database: polyethylene terephthalate (PET) bottle, a high-density polyethylene (HDPE) bottle and an aluminum can. In addition, a data set for the “other”
type, which contained MSW of various types that differ from the listed above, was created. Although it is virtually impossible to cover all types of waste within the ‘other’ type, such data set, which includes most common sorts of objects, can be very useful in classification. Another kind of the images was group-of-objects images. In this case three or more objects of different types were put together and captured in an image. More than 1000 of such groups were captured for the database. All individual objects and objects in groups were labelled according to their material. Additionally, PET bottles were labeled by color. Detailed information on the numbers of images in the database is given in Table 1 and examples are shown in Fig. 2.

![Table 1. The number of images of each class in the database.](image1)

| Material       | Number of Images |
|----------------|------------------|
| HDPE           | 4978             |
| Aluminum       | 3411             |
| PET Transparent| 1749             |
| PET Dark       | 1477             |
| PET Green      | 511              |
| PET Blue       | 1041             |
| PET Teal       | 309              |
| PET Multicolor | 52               |
| Other          | 1007             |
| **Total**      | **14535**        |

![Figure 2. Examples of waste objects images from the created database.](image2)

- a) HDPE;
- b) PET green;
- c) PET dark;
- d) PET transparent;
- e) PET blue;
- f) PET teal;
- g) PET multicolor;
- h) ‘other’;
- i) aluminum can;
- j) multiple objects scene.

![Figure 3. Example of correct region detection (first row) and incorrect region detection (second row).](image3)

5. Model training and validation

A selected model was fine tuned (additionally trained) to detect 4 classes (HDPE, aluminum, PET without color distinction, and “other”). Different strategies exist for making a training dataset: training on images of individual objects, training on a labeled dataset of groups of objects, and training on synthetic groups of objects, compiled from several images of single objects. Training on the labeled dataset of real images containing groups of objects is most relevant to the application case, but, at the same time, manual labeling of a large number of scenes with multiple objects is extremely laborious and prone to errors. In this regard a decision to begin with other strategies of making a training dataset was made. Initially, a training on single objects images was tested, but this approach proved to be ineffective. Then a dataset, consisting of scenes with groups of objects, that were cropped from the
single object images and assembled into the synthetic group-of-objects image, was created. Single objects were cropped along the borders of the objects’ rectangular bounding box.

The trained model was validated (tested) on a set of 729 real multi-object scenes that were labelled manually. Examples of the classification, performed by the model are shown in Fig. 3 and the number of false positives and true positives for each class is shown in Fig. 4.

To measure an accuracy of object position estimation together with quality of classification mean Average Precision (mAP) metrics with Intersection over Union (IoU) were used [10]. The dependence of precision (p) on recall (r) for different threshold values are shown in Fig. 5. In this case, the value of AP is defined as the area under the graph of the curve p(r). The maximum AP score was achieved for the HDPE class. For the ALLUM class, the classification accuracy p is quite high (close to 100%), but half of all aluminum cans were not detected on real images, therefore the AP value for this class was only 0.48. Final mAP over all classes was 55%.

![Figure 4. True positive and false positives on the validation samples.](image)

A particular problem associated with the training data set was revealed in neural network model tests on the synthetic group-of-objects images. The model tended to detect rectangular frames, along which objects were cropped from the original single object images, as real bounding boxes. This was most likely due to the difference in lighting conditions for single object images, which resulted in sharp gradients at the boundary when placing a cropped object into the scene. Though it is difficult to assess the impact of this effect on a final result, a decision to modify the scheme of the data set preparation was made. For the next training dataset single objects were cropped from the initial images using watershed segmentation algorithm, which allowed cropping objects by their natural boundary rather than by a rectangular bounding box. In addition, data augmentation by variation of brightness, contrast and saturation was included into the training procedure. Although such training set still suffered from occasional errors, such as faulty segmentation, a model, trained on this set yielded a mAP score of 64.1 % when performing detection and classification on the set of real images.
Figure 5. Average Precision graph and mean Average Precision for each class.

Conclusion
A general design of a robotic MSW sorting unit with the neural network-based waste objects classifier is presented. The waste objects classifier based on a model of region-proposal type convolutional neural network is tested. To perform training and testing of the model a database of waste objects images has been created, and a scheme for training set preparation is developed. The developed method of the training data set preparation allows us to avoid laborious manual labeling of images that contain multiple objects. A final mAP score of 64% is reached in the tests, which is quite promising, taking into account shortcomings of the training and testing sets.

Acknowledgement
The work was supported by the Ministry of Science and Higher Education of the Russian Federation (grant Agreement No. 14.604.21.0203, unique identification number of the project RFMEFI60418X0203).

References
[1] Zabolotskikh V V, Gomonitskaya A O, Kutmina S V 2016 Izvestiya Samarskogo nauchnogo tsentra Rossiyskoy akademii nauk 18 5 (3) 437
[2] Hancu O, Rad C R, Lapusan C, Brisan C 2018 Aspects concerning the optimal development of robotic systems architecture for waste sorting tasks IOP Conf. Ser. Mater. Sci. Eng. ID 444052029
[3] Witten I H, Frank E, Hall M A, Pal C J 2017 Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques (Burlington: Morgan Kaufmann)
[4] Thung G, Yang M 2016 Classification of trash for recyclability status. CS229 Project Report URL: cs229.stanford.edu/proj2016/report/ThungYang-ClassificationOfTrashForRecyclabilityStatus-report.pdf
[5] Bircanoglu C, Atay M, Beser F, Genc O, Kızrak M 2018 RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks in proc. INISTA

[6] Zhihong C, Hebin Z, Yan W, Yanbo W, Binyan L 2018 Multi-task Detection System for Garbage Sorting base on High-order Fusion of Convolutional Feature Hierarchical Representation in proc. 37th Chinese Control Conference

[7] Ren S, He K, Girshick R and Sun J 2015 Faster r-cnn: Towards real-time object detection with region proposal networks. in proc. NIPS

[8] Huang J, Rathod V, Sun C, Zhu M, Korattikara A, Fathi A, Fischer I, Wojna Z, Song Y, Guadarrama S, Murphy K 2017 arXiv arXiv:1611.10012v3

[9] Tensorflow detection model zoo URL: github.com/tensorflow/models/blob/master/research/object_detection/g3doc/detection_model_zoo.md

[10] Everingham M, Van Gool L, Williams C K I, Winn J, Zisserman A 2010 Int. J. Comput. Vis. 88 (2) 303