PROJECT OF SORTING SYSTEM FOR PLASTIC GARBAGE IN SORTING PLANT BASED ON ARTIFICIAL INTELLIGENCE

Janusz Bobulski and Mariusz Kubanek

Department of Computer Science,
Czestochowa University of Technology, Poland

ABSTRACT

An important element of a complex recycling process that is an integral part of municipal waste management is the sorting of materials that can be re-used. Manual sorting of garbage is a tedious and expensive process, which is why scientists create and study automated sorting techniques to improve the overall efficiency of the recycling process. An important aspect here is the preliminary division of waste into various groups, from which detailed segregation of materials will take place. One of the most important contemporary environmental problems is the recycling and utilization of plastic waste. The main problem under consideration in this article is the design of an automatic waste segregation system. A deep convoluted neural network will be used to classify images.

KEYWORDS

Convolutional Neural Network, Deep Learning, image processing, waste management, environmental protection, recycling.

1. INTRODUCTION

Plastics are synthetic organic matter obtained from e.g. natural gas, petroleum or cellulose as a result of the polymerization process. They are used for the production of packaging, clothes, some building and car components, toys and home appliances. It has been estimated that the total amount of manufactured plastic products in the world increased from 1.5 million tonnes (1950) to 245 million tonnes (data for 2008) [1]. The most important types of plastics include thermoplastics and thermosetting plastics. Thermoplastics can be subjected many times to heat treatment, during which they soften and melt, and after cooling again become hard. This property means that old products can be used for recycling. Examples are products made of PE (polyethylene), PS (polystyrene), PET or PVC (polyvinyl chloride). They constitute 80% of all plastics used. On the other hand, thermosetting plastics (duroplastics) can be melted and formed, but after forming they remain hard and do not soften under the influence of heating. These materials include: epoxides, phenoplast, polyurethanes, polytetrafluoroethylene (PTFE), polyester resins. These materials constitute 20% of all plastics used. Landfilling is undoubtedly the worst way to dispose of waste, both from an economic and ecological point of view. Research on plastics recycling has been conducted for years.

In order to facilitate recycling processes, the obligation to label the type of waste was introduced all over the world, and in the case of plastic waste - the type of plastic. Thanks to this, we know, for example, whether a given packaging can be recycled and what method. Plastic waste was divided into seven groups and marked with numbers from 1 to 7 in a triangle consisting of

David C. Wyld et al. (Eds): ICAITA, CDKP, SAI, NCO, CMC, SOFT, MLT, AdNLP - 2020
pp. 27-35, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100903

arrows. They are: 1-PET - polyethylene terephthalate, 2-PE-HD - high-density polyethylene, 3 - PVC - polyvinyl chloride, 4 - PE-LD - low-density polyethylene, 5 - PP - polypropylene, 6 - PS - polystyrene, 7 - Other. Unfortunately, manual sorting of waste is expensive and laborious, and existing mechanical methods are inefficient. With this in mind, we decided to build a plastic waste sorting system using image processing methods and artificial intelligence. In this article, we present the results of using convolution neural network and deep learning techniques to recognize waste type based on the garbage image. There is a continuation of previous work that gave promising results in the classification into the four most common types of plastic in household waste. In this work, we present results for all seven groups [2]. The results allow the conclusion that the proposed system can be used in real conditions.

2. BACKGROUND

In many countries in the world waste segregation has already been introduced at the beginning of the recycling path, i.e. at home. Just people divide waste into groups such as plastic, metal, glass and organic/bio. The use of selectively automated techniques for these groups is easier than for municipal solid waste (MSW). Unfortunately, a large part of the waste is still collected in the form of the MSW, which is why the countries strive for the most effective reprocessing of waste materials. In order to do this, you should effectively sort the rubbish into individual factions and materials. Therefore, an important task is to isolate individual types of materials from the MSW. Therefore, techniques and procedures for segregating waste are used for the main groups of materials such as paper, glass, metal, wood, plastic and biomass by property system [3]. The biggest challenge, however, is the separation of various types of materials within a given group, i.e. sorting different colour of glass or different types of plastic. The problem of plastic garbage is interesting and at the same time important due to the possibility of recycling only some types of plastic (e.g. PET). To simplify the recycling process, international labelling of various types of plastics was introduced. These are:

- 1 - PET - polyethylene terephthalate,
- 2 - HDPE - high-density polyethylene,
- 3 - PVC - polyvinyl chloride,
- 4 - LDPE - low-density polyethylene,
- 5 - PP - polypropylene,
- 6 - PS - polystyrene,
- 7 - other.

The whole process of automatic sorting of materials suitable for reprocessing from MSW is complicated. There are many methods of waste sorting depend on type of material, that is: mechanical, electromagnetic, X-ray, grinding and the use of rotary equipment, manual, optic based and many others [3].

Commonly techniques often used physical features but ignored visual properties like colour, shapes, texture and size for the sorting of waste. In optical sorting, camera based sensors are used for the identification of waste fractions. In this section we present optical sorting techniques.

Sorting technique based on features like shape and colour was proposed by Huang et al. [4]. This method combines a 3D colour camera and laser beam over the conveyor belt. This technique formed triangles over the image from the camera on the base laser beam, so is called triangulation scanning. The technique achieves an accuracy of 99% for plastic fractions.

Spectral imaging is a combination of spectral reflectance measurement and image processing technologies. We may found several spectral imaging methods using NIR (near infrared), VIS (visual image spectroscopy) and HSI (hyperspectral imaging) [5].

A hyperspectral sensor produces images over a continuous range of narrow spectral bands and next system analysis the spectroscopic data. The conveyor system moves the waste fractions beneath the spectral camera acquires images. At the second stage data is pre-processing and reduction. Next to perform material classification special algorithm is applied. A set of compressed air nozzles is mounted at the end of the conveyor belt and depending upon the classifier decision, one of nozzles are triggered the waste into particular bins [6].

In the case of techniques based on spectroscopy, light is illuminated on plastic waste. Each type of plastic reflects a unique wave range, therefore NIR and laser sensors are used to read the wavelengths reflected from the material being tested. Then, based on the unique signature, it is classified by the processor unit.

Safavi et al. developed a technique that uses reflectometric spectroscopy to identify PP plastic in mixed wastes. The identification unit uses a spectrometer to analyse the reflected light from the sample and determine the type of material, and the compressed air nozzle ejects the elements to the appropriate boxes [7].

The HSI approach is used to classify high purity PP and PE plastics from mixed waste using near-infrared light NIR (1000-1700 nm) [8]. A typical spectroscopic system is equipped with a movable conveyor belt and a sensor system including a backlight and a NIR spectral camera. The image of the materials in the control zone is acquired by the NIR camera, and then it is processed by the classification algorithm.

In order to improve the efficiency of the classification algorithm, the principal components analysis (PCA) is used to reduce the data classification dimensions obtained from the spectral images [9]. Kassouf et al. [10] developed a quick way to classify plastics with the combination of MIR spectroscopy and independent component analysis (ICA). In addition, a more accurate classification is obtained by separating some plastic waste, e.g. LDPE and HDPE.

3. PROPOSED SYSTEM

After extracting plastic garbage from the MSW, a computer system based on image processing can be used to divide it into different types (Fig. 1). The method we propose uses an RGB digital camera and a computer with software for classifying plastic waste. In contrast, an air stream is used to direct the waste to a specific container. The software used in this system uses image processing techniques in the process of image pre-processing. However, convolution neural networks and deep learning [11] are used to recognize objects.

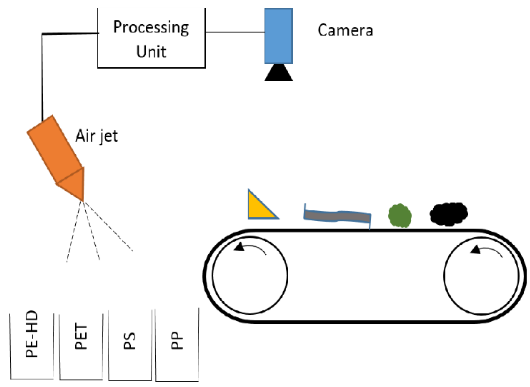

Fig. 1. Proposed system for plastic waste sorting.

3.1. Convolutional Neural Network

Convolutional neural network (CNN) is a feed-forward artificial neural network in which the organization of neurons is similar to the animal visual cortex. In order to recognize the shape of an object, the local arrangement of pixels is important. CNN starts with recognition of smaller local patterns on the image and concatenate them into more complex shapes. CNN was proved to be efficient especially in object recognition on an image. CNNs might be an effective solution to the waste sorting problem.

CNN explicitly assumes the input is an image and reflects it onto its architecture. CNN usually contains Convolutional layer, Pooling layer and Fully-connected layer. Convolutional layers and Pooling layers are stacked on each other, fully-connected layers at the top of the network outputs the class probabilities.

Convolutional layer

Convolutional layer consists of neurons connected to a small region of pixels (also called the receptive field) of previous layer. The neurons in same feature map share the same weights. Convolutional layer (CL) contains a set of learnable filters. One filter activates when a specific shape or blob of colour occurs within a local area [12]. Each CL has multiple filters F. Filter $f_i \in F$ is a set of learnable weights corresponding to the neurons in previous layer. Filter is small spatially (along width and height) and extends along the full depth c of the previous layer.

Pooling layer

Pooling layer (PL) is an effective way of non-linear down-sampling. It has as the convolutional layer receptive field and stride, however is not adding any learnable parameters. PL layer is usually put after the CL.

The receptive field r_n of neuron n in PL is 2 dimensional. It extends over a square of neurons in one feature map (or the neurons sharing the same third coordinate in case of not CL layer). Let us denote $width^{i-1}$ and h^{i-1} the spatial width and height of the previous layer. For each feature map there is P neurons in PL.

$$P = \frac{width^{i-1} \cdot h^{i-1}}{s_w \cdot s_h} \tag{1}$$

Max-pooling layer is the most frequently used pooling layer. Each neuron outputs the maximum of its receptive field. Usually the stride is the same as the size of receptive field. The receptive fields do not overlap, but touch. In most cases stride and size of receptive field are 2 x 2.

The output (x_i, y_j, f_k) of the max-pooling neuron n with position (x_n, y_n, z_n) and receptive field r_n of size $f_w \times f_h$ in PL with stride (s_w, s_h):

$$y(x_n, y_n, z_n) = max_{0 \le l < f_w} max_{0 \le m < f_h} v(x_n + l, y_j + m, f_k) \tag{2}$$

Inspired by a processes in the visual cortex of animals MAX-pooling layer amplifies the most present feature (pattern) of its receptive field and throws away rest. The intuition is that once a feature has been found, its rough location relative to other features is more important than its exact location. The pooling layer is effectively reducing the spatial size of the representation, does not add any new parameters – reducing them for latter layers, making the computation more feasible.

The idea of pooling layer was created back in the time with lack of computational power. Due to its destructiveness – throwing away 75% of input information in case of small 2 x 2 receptive field, the current trend prefers stacked convolutional layers eventually with stride and uses pooling layers very occasionally or discards them altogether [13].

CNN Structure

As a feed-forward artificial neural network, the CNN consists of neurons with learnable weights and biases. CNN's neurons still contains activation function and the whole network expresses single differentiable score function. The position of the pixel matters in comparison with MLP. It receives 3 dimensional space input (x, y, z) – the value of z-th channel of the pixel or occurrence of z-th feature of CL at position (x, y). One pixel is usually made of three channels – red, green and blue.

The convolutional layer and pooling layer are locally connected to the outputs of the previous layer, recognizing or magnifying local patterns in the image. Pooling layer is usually put after the convolutional layer. This pair of layers is repeatedly stacked upon each other following with the fully connected layers at the top.

Usual architecture can be: input layer (IL), CL, PL, CL, PL, full-connected layer (FC), FC. Recent studies suggest stacking many CLs together with fewer PLs.

The fully connected layer is connected to all outputs of last pooling layer. The outputs of last pooling layer should already represent complex structures and shapes. The fully connected layer follows usually with another one or two layers finally outputting the class scores.

Back-propagation

The single evaluation is completely consistent with the feed-forward neural network. The input data or activations are passed to next layers, dot product is computed over which activation function is applied. Down-sampling the network using pooling layer might be present. At the end two or three fully connected layers are stacked. In order to use gradient descent learning algorithm, the gradient must be computed.

The usual back-propagation algorithm is applied with two technical updates. Classical back-propagation algorithm would calculate different partial derivatives of weights belonging to the neurons in same filter, however these must stay the same. Therefore, derivatives of loss function with respect to weights of neurons belonging to the same feature map are added up together.

The update of back-propagation itself is when dealing with max-pooling layers. The back propagating error is routed only to those neurons which have not been filtered with max-pooling. It is usual to track indices of kept neurons during forward propagation to speed up the back-propagation.

Autoencoder

Autoencoder is a feed-forward neural network where expected output is equal to the input of the network – its goal is to reconstruct its own inputs. Therefore, autoencoders are belonging to the group of unsupervised learning models [14]. Usually autoencoder consists of an input layer, one or many hidden layers and output layer. Since the idea of autoencoders is very similar to Restricted-Boltzman Machine, it is common for the structure of autoencoders to follow the rule: $|l_i| = |l\ (k\text{-}i)|$. Let us denote the layer l_c, such that the number of neurons in l_c is lower than in any other layer. The l_c is an encoding layer. The feed-forward neural network consisting of layers l_i; $i <= c$, is called encoder. Expectedly, stacked layers l_j, such as $j >= c$ is called decoder. Each autoencoder consists of encoder and decoder. The encoder can be used for compression. Unlike Principal Component Analysis analysis restricted to linear mapping, the encoder represents non-linear richer underlying structures of the data [15]. The activations of the l_c layer can be further used for classification. Fully-connected layers are appended with the size of the last corresponding to the number of labels. In our system, the autoencoder is used to encode the input signals.

Deep Autoencoder consists of many layers stacked on each other allowing to discover more complicated and non-linear structures of the data. Since it may be complicated to tune deep autoencoder network, commonly the training procedure is made of two steps:

Pre-training, each layer $l_1,...,l_c$ is pre-trained. Firstly the pair l_0 as an example and l_1 as encoder is used. The goal is to find representation of l_0 in l_1 using the right optimizer. The weights l_0 to l_1 and l_1 to l_0 may be tied up representing Restricted Boltzman-Machine. When good representation of l_0 inputs is encoded in l_1 the pair l_1, l_2 is pre-trained further till pair l_{c-1}, l_c is reached.

Fine-tuning, the full network is *connected* and fine-tuned. In case of classification, the encodings of input data points can be used for classification training or the whole network $l_0 \rightarrow \ldots \rightarrow l_c \rightarrow f_{c1} \rightarrow f_{c2}$ is part of the supervised learning.

4. EXPERIMENT

4.1. Structure of the Network

A number of important factors had to be taken into account when working on the appropriate selection of the network structure. First of all, the size of the input image was an important element. Too high resolution resulted in increasing the number of calculations, which resulted in fairly frequent overload of memory available computing unit. But, too low resolution of the input data could have prevented the achievement of theexpected performance. Determined to conduct research for images with a resolution of 60 x 120 pixels. Another important element was the selection of the number and types of layers of the CNN network. Our network contained 16 layers. The first convolution layer consisted of 64 convolution filters with dimensions 9 x 9. Three layers of convolution encode information, transferred to a two-layer fully connected layer. The network diagram for 60 x 120 pixel images is shown in Table 1.

Table 1. Structure of the proposed network.

No	Name of layer	Parameters
1	Image input layer	60 x 120 x 3
2	Convolution Layer	64 filters, size 5 x 5
3	Max Pooling Layer	
4	ReLU Layer	
5	Cross Chanel Normalization Layer	
6	ReLU Layer	
7	Max Pooling Layer	
8	Convolution Layer	64 filters, size 5 x 5
9	ReLU Layer	
10	Max Pooling Layer	
11	Convolution Layer	64 filters, size 5 x 5
12	Fully connected layer	Inputs 4992, outputs 64
13	ReLU Layer	
14	Fully connected layer	Inputs 64, outputs 7
15	ReLU Layer	
16	Classification layer	7

4.2. Input Data

Preparation of input data for the learning and testing phase was important in the context of correct classification of objects in natural working conditions. In the case of deep neural networks, as many data as possible should be collected for each identified class. In our case, it was necessary to collect photos of classified waste. We adopted a simplified model where there could be only one waste within the camera lens. This approach does not reflect the natural working conditions, but for research needs it gives sufficient opportunities to generate a properly functioning network. All collected images represented objects classified into seven considered classes: PET, PE-HD,PVC, PE-LD, PP, PS and Other. These images came from the commercial Garbage database[16], and their samples maybe seen on Fig. 2. To increase the number of images in individual classes, we have modified existing images by flipping and rotating. Images from all classes were rotated by angle 18 degrees, In this way, we obtained over 200 000 images. We chose randomly for teaching of 10,000 per class and for testing of 1,000 for each class.

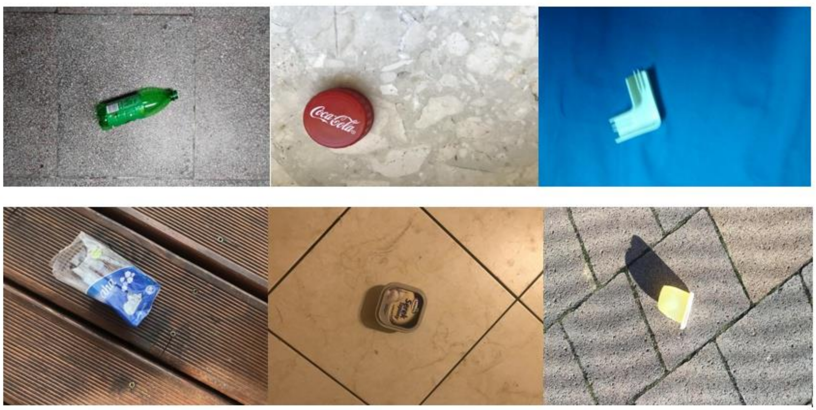

Fig. 2. Samples images of plastic waste

Table 1**.** Learning results of a 16-layer network (image resolution 60 x 120 pixels)

No.	Initial learn rate	Learned rate per period	Max epoch	Regularization	Accuracy [%]
1	0.0001	10	20	0.1	85
2	0.001	5	20	0.1	90
3	0.0001	5	20	0.1	88
4	0.0001	15	20	0.1	86
5	**0.001**	**10**	**30**	**0.1**	**91**
6	0.01	10	30	0.1	34
7	0.001	10	30	0.1	85
8	0.001	10	30	0.1	89
9	0.001	15	30	0.01	88
10	0.001	15	30	0.01	87

Table 2 presents learning stages conducted for our network using images with a resolution of 60 x 120 pixels. Analysing the obtained results, it we see our network reach good results for fifth stage, when it achieved 91%. Ten epochs were sufficient to obtain an acceptable level of accuracy. Further learning, even with a reduced learning rate, no longer significantly affects accuracy. Average accuracy of 91% is a very good result as a fairly small number of iterations. Regarding the other learning parameters, the best results were obtained for Initial learn rate 0.01, Learned rate per period 10, Max epoch 30 and Regularization 0.1. The last one parameter didn't make strong influence on accuracy. The biggest impact on the change of accuracy was the change Initial learn rate to 0.1.

We used five sets of images, were each class contained 1000 images, and we obtain the highest average result 91% for parameters presented at line 5 in Table 3. Accuracy at presented level is acceptable for the proper functioning of the system in real conditions.

5. Conclusion

The conducted research has shown that the 15-layer network proposed by us allows achieving high efficiency for images with a resolution 60 x120. Classification of segregated waste into seven main classes takes place in most cases without error. Of course, this is to a certain extent caused by the artificially increased number of individual class representatives. Further work will mainly consist of extending the database of segregated waste images with photos of waste in more realistic conditions. Hence, efforts to obtain recordings of waste on a conveyor belt from enterprises dealing with waste segregation. Our research in the future will assume the possibility of training the network while working in real conditions, which is possible to implement with our proposal. After introducing modifications to the training database, we also want to determine the accuracy for real images of waste taken from the conveyor belt during the segregation process.

6. Acknowledgements

Supported by the project financed under the program of the Polish Minister of Science and Higher Education under the name "Regional Initiative of Excellence" in the years 2019 – 2022. Project number 020/RID/2018/19 the amount of financing 12,000,000 PLN.

REFERENCES

1. GUS homepage, https://stat.gov.pl/, last accessed 2019/09/25.
2. Bobulski, J., Kubanek, M.: Waste Classification System Using Image Processing and Convolutional Neural Networks. In: Rojas, I., Joya, G. (eds.), Advances in Computational Intelligence, pp. 350—361 (2019).
3. Gundupalli, S.P., Hait, S., Thakur A., A review on automated sorting of source-separated municipal solid waste for recycling. Waste Management 60, pp. 56–74 (2017).
4. Huang, J., Pretz, T., Bian, Z.: Intelligent solid waste processing using optical sensor based sorting technology. 3rd International Congress on Image and Signal Processing (CISP), vol. 4. IEEE, pp. 1657–1661, (2010).
5. Pieber, S., Meirhofer, M., Ragossnig, A., Brooks, L., Pomberger, R., Curtis, A.: Advanced Waste-Splitting by Sensor based Sorting on the Example of the MTPlantOberlaa. Tagungsbandzur 10, DepoTech Conference, pp. 695–698, (2010).
6. Tatzer, P., Wolf, M., Panner, T.: Industrial application for inline material sorting using hyperspectral imaging in the NIR range. Real-Time Imag. 11 (2), pp. 99-107, (2005).
7. Safavi, S.M., Masoumi, H., Mirian, S.S., Tabrizchi, M.: Sorting of polypropylene resins by color in MSW using visible reflectance spectroscopy. Waste Manage. 30 (11), pp. 2216-2222, (2010).
8. Serranti, S., Gargiulo, A., Bonifazi, G.: Characterization of post-consumer polyolefin wastes by hyperspectral imaging for quality control in recycling processes. Waste Manage. 31 (11), pp. 2217-2227, (2011).
9. Serranti, S., Gargiulo, A., Bonifazi, G.: Classification of polyolefins from building and construction waste using NIR hyperspectral imaging system. Resour. Conserv. Recycl. 61, pp. 52-58, (2012).
10. Kassouf, A., Maalouly, J., Rutledge, D.N., Chebib, H., Ducruet, V.: Rapid discrimination of plastic packaging materials using MIR spectroscopy coupled with independent components analysis (ICA). Waste Manage. 34 (11), pp. 2131-2138, (2014).
11. Wang, M., Wang, Z., Li, J.: Deep Convolutional neural network applies to face recognition in small and medium databases. 4th International Conference on Systems and Informatics, ICSAI 2017 Volume 2018-January, 4 January 2018, pp. 1368-1372, (2018).
12. Convolutional Neural Networks (CNNs / ConvNets). 2016. url: http://cs231n.github. io/convolutional- networks/, last accesses on 09/01/2019.
13. Goodfellow I., Benigo Y., Courville A.: Deep Learning, The MIT Press, 800 (2016)
14. Kaiming He.: Deep Residual Learning for Image Recognition, url: https://arxiv.org/abs/1512.03385 last accesses on 12/15/2018.
15. Hinton G.E., R. R. Salakhutdinov R.R., Reducing the Dimensionality of Data with Neural Networks, url: http://science. sciencemag.org/content/313/5786/504, last accesses on 15/07/2019.
16. Garbage Image Database, url: http:// garbage.eu5.com last accesses on 01/09/2019.