Towards Cleaner Environments by Automated Garbage Detection in Images

Chapter · October 2020

DOI: 10.1007/978-3-030-62098-1_5

Aleksandar Despotovski
Ss. Cyril and Methodius University in Skopje

Filip Despotovski
Ss. Cyril and Methodius University in Skopje

Jane Lameski
Technische Universität München

Eftim Zdravevski
Ss. Cyril and Methodius University in Skopje

Towards cleaner environments by automated garbage detection in images

Aleksandar Despotovski[1], Filip Despotovski[1], Jane Lameski[2], Eftim Zdravevski[1], Andrea Kulakov[1], and Petre Lameski[1]

[1]University of Ss. Cyril and Methodius in Skopje, Faculty of Computer Science and Engineering
Ruger Boskovic 16, 1000 Skopje, North Macedonia
https://www.finki.ukim.mk
[2]Technical University of Munich, Germany, Faculty of Informatics and Mathematics
https://www.tum.de/

Abstract. The environment protection is becoming, now more than ever, a serious consideration of all government, non-government, and industrial organizations. The problem of littering and garbage is severe, particularly in developing countries. The problem of littering is that it has a compounding effect, and unless the litter is reported and cleaned right away, it tends to compound and become an even more significant problem. To raise awareness of this problem and to allow a future automated solution, we propose developing a garbage detecting system for detection and segmentation of garbage in images. For this reason, we use deep semantic segmentation approach to train a garbage segmentation model. Due to the small dataset for the task, we use transfer learning of pre-trained model that is adjusted to this specific problem. For this particular experiment, we also develop our own dataset to build segmentation models. In general, the deep semantic segmentation approaches combined with transfer learning, give promising results. They show great potential towards developing a garbage detection application that can be used by the public services and by concerned citizens to report garbage pollution problems in their communities.

Keywords: Image segmentation · Environment protection · Deep Learning · deep semantic segmentation

1 Introduction

Littering and garbage pollution is one of the oldest problems in any human settlement. The increased garbage production is connected to the increased population and the increased development of the living, commercial, and industrial areas [4]. Garbage pollution and collection problems are especially emphasized in developing countries where the garbage collection services are overwhelmed. Furthermore, the lack of environmental awareness further increases garbage pollution, especially in areas with lower development. The process of garbage collection automation is already present in different developed areas in the world, and there are quite a few examples of this automation. In [13] authors use computer vision techniques to estimate the volume of the dump based on eight different perspectives with about 85% accuracy. Authors in [16] use a web camera to estimate the garbage quantity in garbage cans for a smart garbage collection system. In [11] authors use Arduino and ultrasound sensor to estimate how full are the garbage collection units and alert the driver which bins they should visit and empty. [14, 15] use deep learning architecture for segmentation and classification of aerial scenes.

Different approaches are used to increase garbage collection and recycling efficiency. In [3], authors propose a gamification approach to motivate society to be more involved in garbage recycling. Their case study proved that gamification approaches increase recycling and involvement significantly. To increase the awareness and help the garbage collection services, the new developments in computer vision can be of significant aid. In [13] is presented the architecture behind a smartphone app that detects and coarsely segments garbage regions in geo-tagged user images. It utilizes deep architecture of fully convolutional networks for detecting garbage in images with sliding windows, achieving a mean accuracy of 87.69 on the GINI dataset.

In this paper, we use machine learning techniques to design a system that would detect and segment garbage in images. We try to resolve the problem of detection by using image segmentation and deep semantic segmentation methods to achieve as good segmentation for garbage in images as possible. A successful segmentation would allow us to not only detect the garbage in the images but also allow the system to find the exact location of the garbage and possibly determine the type of garbage.

The motivation behind this work is to build a system that would allow the users to automatically detect garbage in images and send the images to local and state authorities. The goal is to increase public awareness about misplaced garbage and garbage pollution in urban areas, especially in the developing countries where the legal authorities involvement and public awareness is on a superficial level. Such a system would ease the automation in garbage detection and increase the possibilities for applications that would allow reducing the littering problem in urban and suburban areas. Our approach can also be used to classify the types of garbage and with further development, estimate the quantity of garbage in the area based on a single image.

The main contribution of this paper is the novel dataset for garbage detection that contains labelled pixels of several garbage classes and the application of deep learning segmentation architectures to obtain initial models and results.

We organized the paper as follows: In Section 2, we describe the used dataset, the data augmentation methods and model architecture for our experiments. In Section 3 we present the obtained results and finally in Section 4 we shortly describe the conclusion of our work and the future work.

2 Methods

2.1 Dataset

One of the essential resources for building a machine learning model is the data. For garbage detection and segmentation in images, at the moment of performing the experiments, there were no publicly available dataset, to the best of our knowledge. Although there are many garbage images on the Internet, there were no adequate annotations for those images which mark the garbage segments at a pixel level. For this reason, we gathered a collection of 100 images from the Internet and our environment that contain garbage. The images are diverse; some of them capture pollution in nature, while others contain garbage on the streets. The images are in different sizes and resolutions. Before the labelling and training process, we resized all of the images to 640 × 480 pixels.

We manually segmented and labelled every image in the dataset into the garbage and non-garbage regions. The segmentation is needed so that the machine learning models can learn which parts of the images represent garbage and even to which garbage type. Some such examples where the garbage content is manually segmented are given in Figure 3 and Figure 1.

Additionally, we manually classified the segmented images to the seven classes: Non-garbage (NG), Plastic (PL), Metal (ME), Paper (PA), Glass (GL), Other (OT). Figure 1 shows one example of a labelled image from the dataset.

Fig. 1. An image from the dataset containing multiple class labeled segments and the corresponding labels.

The significantly lower number of pixels representing the garbage class compared to the non-garbage class presents another issue that increases the complexity of the problem. Namely, the dataset would be highly imbalanced, which is expected considering the nature of the problem. The percentages of each class are shown in Figure 2.

To further increase the number of samples in the training set, we performed data augmentation. We used the Image Augmentation module implemented in the Keras [7] and Tensorflow [1] libraries. This approach transforms the images by slight rotation, scaling, zooming, adding noise, skewing, random rotation,

Fig. 2. Class distribution of the train (left) and test (right)

Fig. 3. An image from the dataset, next to its segmentation. The white segments represent garbage regions.

horizontal and vertical translation, flipping and zooming to obtain new images that are somewhat different than the original. We used data augmentation technique only for the training set. The remaining images represented the test set and were used to evaluate the trained models.

2.2 Models

For tackling the problem of detecting garbage in images, we used a deep learning approach, which utilizes deep neural networks for predicting the most plausible image segmentation.

Several recent state-of-the-art approaches commonly used for semantic segmentation, could be applied to this problem [2] [5] [10] [12] [18] [19]. The success shown in several competitions, including [5], motivated us to choose the TernausNet deep learning architecture [10]. TernausNet is based on pre-trained VGG16 as an encoder linked with U-net [17] decoder layers. Similarly to the TernausNet, we combined VGG16 with deep CNN architecture SegNet [2]. The described model here is similar to [12], where it was applied for segmentation of skin lesions, but without using the pooling indices. For this experiment, we used a convolutional neural network with 13 convolutional layers combined in 5 blocks, each having a max-pooling and down-sampling layer. Additionally, we added 6 blocks of 2 convolutional layers with a dropout rate of 0.2 to regulate overfitting [19]. After each of the five blocks used to up-sample the output, another layer is added for concatenation of the transposed convolutional layer and the convolutional layer with the corresponding dimension, resulting in a so-called skip connection that combines the knowledge of the two levels and improves the model performance. This choice is in agreement with other studies where the combined potential of non-linear encoding features have shown increased accuracy in predictive tasks [6] [8] [9]. The segmentation model is shown in Figure 4.

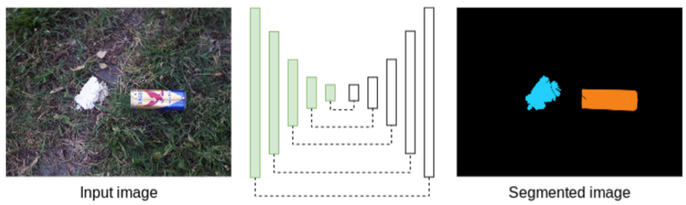

Fig. 4. Diagram of the fully convolutional neural network. The green blocks of layers are pre-trained and we only train the white layers. The dotted lines depict the skip connections between the layers

The results in the next section were achieved using 300 epochs with the Adam optimizer and 0.01 learning rate for training the model. We used categorical cross-entropy loss function, which is usually used for segmentation and classification tasks.

3 Results

Based on the experimental setup described in the previous section, we trained the models and used the test set to evaluate them. For the garbage vs non-garbage segmentation, we obtained a DICE coefficient score of 0.62. As expected, the segmentation based on pre-trained deep learning yielded acceptable results. Figure 5 shows one example of the results to illustrate the effectiveness of the approach. The picture shows that the garbage is localized correctly and the segmentation successfully covers most of the garbage pixels. The DICE coefficient of this segmentation is 0.86, which is better than the result obtained for the whole dataset.

Fig. 5. Example result of the segmentation. Input image (left), Ground truth (middle), Obtained result (right)

Segment class	Dice coefficient
Non-garbage (NG)	0.99
Plastic (PL)	0.58
Metal (ME)	0.34
Paper (PA)	0.58
Glass (GL)	0.25
Other (OT)	0.35

Table 1. Evaluation of the segmentation for each label class

For the multi-label segmentation the results are shown in Table 1. The model is quite successful in segmenting the garbage from plastic and paper material. However, the most difficult to segment is the glass type of garbage is the most difficult to segment. The reason for this could be the lack of samples, considering that the glass material garbage class is under-represented in the dataset comparing to the top-scoring classes. A sample of the segmentation results is shown in Figure 6.

Fig. 6. Example of segmentation results for glass-based garbage class on the test set. Original image (left), Ground truth segmentation (Middle), Segmentation result (right)

4 Conclusion

In this paper, we present a novel dataset for multi-label segmentation of garbage, and we tackle the problem of garbage segmentation. The primary motivation is to create a system that would allow automatic detection of garbage in images that can be used from the public organizations and citizens. For this purpose, considering the lack of available datasets, we have created our own dataset and used it to test several approaches. To increase the number of images, we used data augmentation, and we applied transfer learning to improve the model performance. The results show that our approach is suitable for garbage detection in images. However, it needs further improvements for detecting different types of garbage. Several ideas could be attempted in future works. One is to improve the dataset class balance because the current class imbalance causes an obstacle when training any machine learning model. Another idea is to add weights to the trainers and see if the models can be further improved.

Something that is missing in our experiments due to performance constraints, but could significantly improve the results is hyperparameter tuning. The learning rate, the number of training epochs, as well as the dropout rates were chosen on intuition and empirical results during the experiment. With proper hyperparameter optimization, it is reasonable to expect better results.

An additional idea is to increase the dataset size by non-standard augmentation techniques, such as, to use artificially created data using the placement of known segments over different, natural and artificial backgrounds, to increase the number of samples further. Obtaining and labelling additional images is always an option and could improve the performance.

Acknowledgments

The work presented in this paper was partially funded by the Ss. Cyril and Methodius University in Skopje, Faculty of Computer Science and Engineering. We also gratefully acknowledge the support of NVIDIA Corporation through a grant providing GPU resources for this work. We also acknowledge the support of the Microsoft AI for Earth for providing processing resources.

References

1. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X.: TensorFlow: Large-scale machine learning on heterogeneous systems (2015), http://tensorflow.org/, software available from tensorflow.org
2. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence **39**(12), 24812495 (2017)
3. Briones, A.G., Chamoso, P., Rivas, A., Rodríguez, S., Prieta, F.D.L., Prieto, J., Juan M. Corchado", editor="Uden, L., Hadzima, B., Ting, I.H.: Use of gamification techniques to encourage garbage recycling. a smart city approach. In: Knowledge Management in Organizations. pp. 674–685. Springer International Publishing, Cham (2018)
4. Brown, D.P.: Garbage: How population, landmass, and development interact with culture in the production of waste. Resources, Conservation and Recycling **98**, 41 – 54 (2015), http://www.sciencedirect.com/science/article/pii/S0921344915000440
5. Carvana: Carvana image masking challenge automatically identify the boundaries of the car in an image. https://www.kaggle.com/c/carvana-image-masking-challenge/ (last accessed 30.05.2019)
6. Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. Lecture Notes in Computer Science p. 833851 (2018)
7. Chollet, F., et al.: Keras. https://github.com/fchollet/keras (last accessed 30.05.2019) (2015)

8. Corizzo, R., Ceci, M., Japkowicz, N.: Anomaly detection and repair for accurate predictions in geo-distributed big data. Big Data Research **16**, 18 – 35 (2019)
9. Corizzo, R., Ceci, M., Zdravevski, E., Japkowicz, N.: Scalable auto-encoders for gravitational waves detection from time series data. Expert Systems with Applications **151**, 113378 (2020)
10. Iglovikov, V., Shvets, A.: Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. ArXiv e-prints (2018)
11. Kumar, N.S., Vuayalakshmi, B., Prarthana, R.J., Shankar, A.: Iot based smart garbage alert system using arduino uno. In: 2016 IEEE Region 10 Conference (TENCON). pp. 1028–1034 (Nov 2016)
12. Lameski, J., Jovanov, A., Zdravevski, E., Lameski, P.L., Gievska, S.: Skin lesion segmentation with deep learning. In: IEEE EUROCON 2019 - 18th International Conference on Smart Technologies. IEEE (2019). https://doi.org/10.1109/EUROCON.2019.8861636
13. Mittal, G., Yagnik, K.B., Garg, M., Krishnan, N.C.: Spotgarbage: Smartphone app to detect garbage using deep learning. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. pp. 940–945. UbiComp '16, ACM, New York, NY, USA (2016)
14. Petrovska, B., Atanasova-Pacemska, T., Corizzo, R., Mignone, P., Lameski, P., Zdravevski, E.: Aerial scene classification through fine-tuning with adaptive learning rates and label smoothing. Applied Sciences (2020)
15. Petrovska, B., Zdravevski, E., Lameski, P., Corizzo, R., Stajduhar, I., Lerga, J.: Deep learning for feature extraction in remote sensing: A case-study of aerial scene classification. Sensors **15**(1), 1–1 (2020)
16. Prajakta, G., Kalyani, J., Snehal, M.: Smart garbage collection system in residential area. IJRET: International Journal of Research in Engineering and Technology **4**(03), 122–124 (2015)
17. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention MICCAI 2015 p. 234241 (2015)
18. Ryan, S., Corizzo, R., Kiringa, I., Japkowicz, N.: Pattern and anomaly localization in complex and dynamic data. In: 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA). pp. 1756–1763 (2019)
19. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research **15**(1), 1929–1958 (2014)