



Deep Learning Based Garbage Detection for Autonomous Garbage Collection Vehicles

Süleyman Uzun^{1*}, Dilara Karaca²

^{1*} Sakarya University of Applied Sciences, Faculty of Technology, Department of Computer Engineering, Sakarya, Turkey, (ORCID: 0000-0001-8246-6733), suleymanuzun@subu.edu.tr

² Sakarya University of Applied Sciences, Graduate Education Institute, Department of Biomedical Engineering, Sakarya, Turkey, (ORCID: 0000-0002-4352-4141), dilakaraca@subu.edu.tr

(International Conference on Design, Research and Development- 15 – 18 Aralık 2021)

(DOI: 10.31590/ejosat.1045510)

ATIF/REFERENCE: Uzun, S., Karaca, D., (2021). Deep Learning Based Garbage Detection for Autonomous Garbage Collection Vehicles. *European Journal of Science and Technology*, (32), 1194-1198.

Abstract

With the rapid technological advances and the increasing human population, the need for more production has emerged and consumption has increased accordingly. As a result of this increased consumption, more garbage has been generated. Environmental pollution caused by these garbages emerges as a problem that people have to overcome both in Turkey and in the world. Many studies have been conducted to overcome this problem. Especially today, with the development of autonomous vehicles and artificial intelligence, the solutions using these technologies have increased. In this study, a new data set was created for autonomous garbage collection vehicles and a model was proposed in which these vehicles can be used. The data set was prepared with images of garbage with paper cups, which is one of the most polluting garbage, taken in different places, and images consisting of different garbage without paper cups. Paper cups were detected from these images with pre-trained SqueezeNet, VGG-19 and GoogLeNet convolutional neural networks. The performance rate of the SqueezeNet, GoogLeNet and Vgg-19 networks used in the study was found as 97.77%, 96.44% 94.66%, respectively.

Keywords: Transfer learning, Garbage detection, Garbage dataset.

Otonom Çöp Toplama Araçları İçin Derin Öğrenme Temelli Çöp Tespiti

Öz

Günümüz teknolojilerin gelişimi ve artan insan nüfusu ile birlikte daha fazla üretim ihtiyacı ortaya çıkmakta ve buna bağlı olarak da tüketim artmaktadır. Artan bu tüketim beraberinde daha fazla atık oluşturmaktadır. Bu atıkların sebep olduğu çevre kirliliği gerek ülkemizde gerekse dünyada insanların üstesinden gelmesi gereken bir sorun olarak ortaya çıkmaktadır. Bu sorunun üstesinden gelmek amacıyla pek çok çalışma yapılmaktadır. Özellikle günümüzde otonom araçların ve yapay zekanın gelişimi ile bu teknolojilerin kullanıldığı çözümler giderek artmaktadır. Yapılan bu çalışmada, otonom çöp toplama araçları için oluşturulan yeni veri seti kullanılarak bu araçların kullanılabileceği bir model hazırlanmıştır. Veri seti, farklı yerlerde çekilen, çevreyi en çok kirleten çöplerden biri olan kâğıt bardakların olduğu çöp görüntüleri ve kâğıt bardağın olmadığı farklı çöplerden oluşan görüntüler ile hazırlanmıştır. Bu görüntülerden ön eğitilmiş SqueezeNet, VGG-19 ve GoogLeNet evrişimli sinir ağları ile kâğıt bardaklar tespit edilmiştir. Çalışmada kullanılan ağlar ile SqueezeNet %97.77, GoogLeNet %96.44 ve Vgg-19 %94.66 başarımla elde edilmiştir.

Anahtar Kelimeler: Transfer öğrenme, çöp tespiti, çöp veri seti.

* Corresponding Author: suleymanuzun@subu.edu.tr

1. Introduction

One of the most important causes of environmental pollution is population growth. Depending on the consumption need of the increasing population, environmental pollution has also increased rapidly. Household waste is one of the leading factors in the formation of pollution. Environmental pollution has been one of the important problems for people from the past to the present. Especially with the increase of the garbages such as paper, plastic, glass, etc., many studies have been carried out in order to make the cleaning process of the environment easier and more practical. Today, with the advances in the field of artificial intelligence, autonomous vehicles have started to be used in garbage collection processes (Akanksha, Gupta, & Tripathi, 2021).

Artificial intelligence is a very broad and important field that includes machine learning and deep learning. It was first developed in the 1950s. Alan Turing, the leader of the concept of artificial intelligence, has created the basis of Artificial Intelligence by introducing the "Turing test". The questions of "Can machines act, choose like humans?" have been influential in the emergence of machine learning. Today, these and other new questions like this have contributed to the development of both machine learning and deep learning (Coşkun & Gülleroğlu, 2021). Deep learning is a sub-title of machine learning that processes input data in its layered architecture and produces more accurate results. This layered structure includes the process that the input data is multiplied by the transpose of the weight vector in each layer, added to the bias value, and then passed through a nonlinearity activation to the output neuron.

In the literature, various studies have been carried out on garbage detection. Panyavaraporn et al. [3] presented a prototype of a water surface garbage collection robot. The robot was designed using Raspberry Pi and was wirelessly controlled by an android device. The garbage collection performance of the automated process, guided by the robot's deep learning models, was measured with an accuracy of 83.25%. Bai et al. [4] also performed garbage recognition using deep neural networks to collect garbage on the grass surface. This study shows that garbage collection accuracy can be up to 95%. Ultrasonic sensors are used in robots for garbage collection, but deep neural networks are preferred because they provide much more information than these sensors. In studies where ultrasonic sensors are used, a web camera placed on the robot detects the garbage in the images. Kang et al. (Kang, Kim, & Lim, 2020) used the video image data set created using CCTV cameras in order to detect the garbage existing in the environment where objects and people are located, and classified them with K-Nearest algorithms among machine learning algorithms. The study presents a new technique for the detection of garbage in crowded environments. Bircanoglu et al. (Bircanoglu, Atay, Beser, Genc, & Kizrak, 2018) identified and classified recycling garbages in recycling facilities. In their study, they used Inception-Resnet, Inception-v4 and DenseNet121 among the pre-trained networks. They achieved 90% of test accuracy in Inception-Resnet and Inception-v4 networks and 95% with DenseNet121. In order to increase their performance, they proposed a new network called RecycleNet by changing the connection models of the skip connections in dense blocks in the DenseNet121 structure. In this way, they presented the RecycleNet network, which reduced the number of parameters of the DenseNet121 network, which has 7 million parameters, to 3 million. Cireşan et al. (Cireşan, Meier, & Schmidhuber, 2012) analyzed transfer learning algorithms with deep neural networks

in order to recognize various characters. Pre-trained network was used to learn Chinese characters. As a result of the study, they achieved much better results with less training time. Singh (Singh, 2021) conducted the classification of polyethylene bags based on deep learning. The polyethylene bags were classified with 88.90% performance. Rao et al. (Rao, Rao, & Ranjan, 2020) developed IoT-based smart garbage detection systems. They measured garbage levels with their proposed neural network model. As a result of the study, they were able to accurately measure the garbage levels with 80.33% performance rate.

In this study, 500 images of paper cups taken from different angles in streets, gardens and avenues, and 250 images of other garbage were used. This new data set is planned to be shared globally later. While preparing the data set, different locations were selected and the places where the most garbage can be found in the environment were preferred.

2. Material and Method

2.1. Convolutional Neural Networks

In Convolutional Neural Networks (CNN), the main goal is to classify each pixel into a fixed category class (Khairandish, Sharma, Jain, IRBM, & 2021, 2021). Each image is a matrix. CNNs are 3D matrices of depth, height and width in Convolutional meshes (Lee, Lee, Kim, ..., & 2019, n.d.). The CNN model consists of two different stages: training and inference. The basic structure of CNNs consists of convolution and pooling layers as shown in Figure 1 (Siavashi, Najafi, Ebadi, Fuel, & 2022, n.d.).

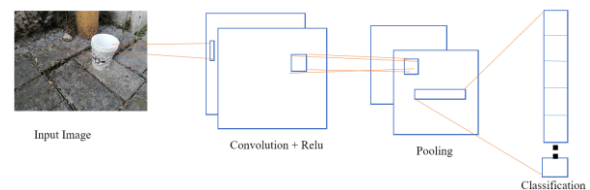


Figure 1. CNN network architecture.

The convolution layer can be thought of as a matrix and the output matrix is calculated by performing linear operations on the input matrix with the filter function, which is a different matrix (Xiao et al., n.d.). In other words, image features are determined in this layer (Mahiba, Measurement, & 2019, n.d.). The pooling layer is a non-activation layer. Since the previous layer is linear, nonlinear features need to be added in this layer. In addition, this layer reduces the weight value by combining image features with higher level features (Min, Lee, & Yoon, 2017). The purpose of the pooling layer is to obtain a smaller matrix or image without losing the properties of the existing data. In the next step, the output matrix is transferred to the neural networks. Inputs updated with weight values in deep networks are passed through activation functions and transferred to the next neuron (Jaloli, Choudhary, & Cescon, 2020). The loss function is computed in the output neuron. The aim is to reduce the loss function to the minimum level, and for this purpose, the operations are differentiated and brought back to the input level of the neural networks. This process is called backpropagation (Guo et al., 2020). The process continues until the loss function is minimized. As a result, the output value is as much as the number of objects to be classified. Assuming that these classified objects are given a value in the range of 1-0, the output that produces the result closest to 1 is

accepted as the attribute to be predicted by the network (Polsinelli, Cinque, & Placidi, 2020).

2.2. Transfer Learning

Transfer learning is a reused design for a trained network to perform new tasks. There is no need to restart the training for each new scenario encountered. In this way, both resources and time are saved (Liu et al., n.d.). Transfer learning can be divided into three categories (Fu, Xue, Gao, Computing, & 2021, n.d.). In the conventional machine learning method, the domains and task of the sources and targets are the same. In inductive transfer learning, the domains of the sources and targets are the same but their tasks are different. In transductive transfer learning, while the source and target domains vary, their tasks are the same (Chiba & Sasaoka, 2021). When there is more similarity between the source task and data referenced in the trained network model and the target data and task to be trained, the better results are obtained. In addition, if there is enough data when training a model, it is better to train the model using a trained network model rather than training from scratch (Elmas, 2021).

In this study, GoogLeNet ImageNet, which is one of the transfer learning techniques, was trained using images in the dataset. GoogLeNet includes convolution and pooling layers (Lloyd, Marshall, Moore, & Rosin, 2016). It is an advantageous network compared to other convolutional neural networks with fewer parameters and fewer layers. The neural network connected layer is completed with softmax and classification layers (Deepak, Medicine, & 2020, n.d.). VGG-19 is another transfer learning technique. The prefix of 19 in its name denotes the number of layers. In other words, it is a 19-layer CNN model. It consists of convolution and pooling layers, which are also ordered in these layers (Lloyd et al., 2016). The classification and location finding capability of the VGG19 neural network has been proven (Zhou et al., n.d.). In the SqueezeNet architecture, the input images are generalized by convolution and pooling is applied (Ucar, hypotheses, & 2020, n.d.). It is an 18-layer convolutional neural network. It aims to achieve high accuracy and use few network parameters. After the convolution layer, all image features are combined and the result is averaged in the pooling layer (Ucar et al., n.d.).

2.3. Data Set

The data set used in the article consists of 750 images in jpg format in total. The images were taken on the side of the road where it is possible for autonomous vehicles to go on the road route. 500 images contain paper cups. 250 images consist of food and beverage packaging, mask, bag garbage in the environment. With pre-trained networks, which of these images belong to paper cup litter and which belong to other litter categories are separated? Data with a resolution of 3456 x 4608 pixels were used in each sample image.

Table 1. Confusion matrix

Ağ	True Positive (TP)	True Negative (TN)	False Positive (FP)	False Negative (FN)
SqueezeNet	147	73	3	2
VGG-19	150	63	0	12
GoogLeNet	143	74	7	1

3. Results and Discussion

3.1. Experimental Results

With the widespread use of autonomous vehicles, the importance of data used in this field has also increased. Especially the success of the studies in this field depends on the new environmental image data created about the area where autonomous vehicles are used. In this context, a data set was created for autonomous garbage vehicles to detect and collect garbage.



Figure 2. Sample images taken from the dataset with paper cups.

In this data set, paper cups that pollute the environment the most were selected as the first class and 500 images were created. Sample images taken from this data set are shown in Figure 2. As the second class, 250 images of different garbage without paper cups were used. Sample images taken from this data set are shown in Figure 3. These images were used with SqueezeNet, VGG-19 and GoogLeNet pre-trained networks.



Figure 3. Sample images taken from the dataset without paper cups.

The confusion matrix of the pre-trained networks used in the study is shown in Table 1. According to this table, it is seen that the Vgg-19 network correctly predicted all of the images with paper cups, and it incorrectly predicted 12 of the images as paper cups in the images without paper cups. The SqueezeNet network, which gave the most successful results, correctly predicted 147 of the data with paper cups and misclassified 3 of them as images without paper cups.

It also predicted correctly 73 of the images in which there was no paper cup, and it predicted incorrectly 2 of the images as paper cups. Although Vgg-19 successfully predicted the dataset with paper cups, it was observed that it failed to predict the dataset without paper cups compared to other networks. Classification performances of the pre-trained networks used in the study are

shown in Table 2. It is seen that the network with the highest classification performance is SqueezeNet with 97.77% success rate, followed by GoogLeNet with 96.44% success rate and Vgg-19 with 94.66% success rate. The Precision and Specificity results were seen to be 100% since VGG-19 correctly predicted the entire dataset with paper cups.

Table 2. Performance results

Ağ	Accuracy	Precision	Sensitivity	Specificity
SqueezeNet	0.977778	0.980000	0.986577	0.960526
VGG-19	0.946667	1.000000	0.925926	1.000000
GoogLeNet	0.964444	0.953333	0.993056	0.913580

4. Conclusion

Millions of garbage is produced every day in the world. Autonomous vehicles will be of great benefit to human beings in the separation and collection of this garbage. Within the scope of the study, 500 images of paper cups and 250 images of other garbage were classified by deep learning algorithms. Literature studies show that pre-trained networks are used in various decomposition processes. A new data set was proposed in the study. SqueezeNet, VGG-19 and GoogleNet pre-trained networks were used to classify the paper cup in the data set. SqueezeNet was the most successful network with a classification success of 97.77%. These results show that successful results will be obtained when the trained network model is used in autonomous vehicles.

References

- Akanksha, Gupta, M., & Tripathi, M. M. (2021). Smart robot for collection and segregation of garbage. *Proceedings of International Conference on Innovative Practices in Technology and Management, ICIPTM 2021*, 169–173. <https://doi.org/10.1109/ICIPTM52218.2021.9388369>
- Bircanoglu, C., Atay, M., Beser, F., Genc, O., & Kizrak, M. A. (2018). RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks. *2018 IEEE (SMC) International Conference on Innovations in Intelligent Systems and Applications, INISTA 2018*. <https://doi.org/10.1109/INISTA.2018.8466276>
- Chiba, S., & Sasaoka, H. (2021). Basic Study for Transfer Learning for Autonomous Driving in Car Race of Model Car. *6th International Conference on Business and Industrial Research, ICBIR 2021 - Proceedings*, 138–141. <https://doi.org/10.1109/ICBIR52339.2021.9465856>
- Cireşan, D. C., Meier, U., & Schmidhuber, J. (2012). Transfer learning for Latin and Chinese characters with deep neural networks. In *Proceedings of the International Joint Conference on Neural Networks* (pp. 1–6). IEEE. <https://doi.org/10.1109/IJCNN.2012.6252544>
- Coşkun, F., & Gülleroğlu, H. D. (2021). Geçmişten Günümüze Yapay Zekânın Gelişimi ve Eğitim Alanında Kullanılması. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 1–20. <https://doi.org/10.30964/auebfd.916220>
- Deepak, S., Medicine, P. A.-C. in B. and, & 2020, undefined. (n.d.). Retrieval of brain MRI with tumor using contrastive loss based similarity on GoogLeNet encodings. *Elsevier*. Retrieved from
- <https://www.sciencedirect.com/science/article/pii/S0010482520303243>
- Elmas, B. (2021). Identifying species of trees through bark images by convolutional neural networks with transfer learning method. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 36(3), 1253–1269. <https://doi.org/10.17341/gazimmfd.689038>
- Fu, W., Xue, B., Gao, X., Computing, M. Z.-A. S., & 2021, undefined. (n.d.). Transductive transfer learning based Genetic Programming for balanced and unbalanced document classification using different types of features. *Elsevier*. Retrieved from <https://www.sciencedirect.com/science/article/pii/S1568494621000958>
- Guo, Y., Chen, J., Du, Q., Hengel, A. Van Den, Shi, Q., networks, M. T.-N., & 2020, undefined. (2020). Multi-way backpropagation for training compact deep neural networks. *Elsevier*, 126, 250–261. <https://doi.org/10.1016/j.neunet.2020.03.001>
- Jaloli, M., Choudhary, D., & Cescon, M. (2020). Neurological Status Classification Using Convolutional Neural Network. *IFAC-PapersOnLine*, 53(5), 409–414. <https://doi.org/10.1016/j.ifacol.2021.04.193>
- Kang, M. S., Kim, P. K., & Lim, K. T. (2020). A Simple and fast method to detect garbage dumping using pedestrian attribute. *2020 IEEE International Conference on Consumer Electronics - Asia, ICCE-Asia 2020*. <https://doi.org/10.1109/ICCE-ASIA49877.2020.9276940>
- Khairandish, M., Sharma, M., Jain, V., IRBM, J. C.-, & 2021, undefined. (2021). A Hybrid CNN-SVM Threshold Segmentation Approach for Tumor Detection and Classification of MRI Brain Images. *Elsevier*. <https://doi.org/10.1016/j.irbm.2021.06.003>
- Lee, M., Lee, J., Kim, J., ... B. K.-2019 I. S., & 2019, undefined. (n.d.). The Sparsity and Activation Analysis of Compressed CNN Networks in a HW CNN Accelerator Model. *Ieexplore.Ieee.Org*. Retrieved from <https://ieeexplore.ieee.org/abstract/document/9027643/>
- Liu, J., Zhang, Q., Li, X., Li, G., Liu, Z., Xie, Y., ... 2021, undefined. (n.d.). Transfer learning-based strategies for fault diagnosis in building energy systems. *Elsevier*. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0378778821005405>
- Lloyd, K., Marshall, D., Moore, S. C., & Rosin, P. L. (2016). Detecting Violent and Abnormal Crowd activity using Temporal Analysis of Grey Level Co-occurrence Matrix

- (GLCM) Based Texture Measures.
<https://doi.org/10.1007/s00138-017-0830-x>
- Mahiba, C., Measurement, A. J.-, & 2019, undefined. (n.d.). Severity analysis of diabetic retinopathy in retinal images using hybrid structure descriptor and modified CNNs. *Elsevier*. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0263224118311771>
- Min, S., Lee, B., & Yoon, S. (2017). Deep learning in bioinformatics. *Briefings in Bioinformatics*, 18(5), 851–869. <https://doi.org/10.1093/BIB/BBW068>
- Polsinelli, M., Cinque, L., & Placidi, G. (2020). A light CNN for detecting COVID-19 from CT scans of the chest. *Pattern Recognition Letters*, 140, 95–100. <https://doi.org/10.1016/J.PATREC.2020.10.001>
- Rao, P. P., Rao, S. P., & Ranjan, R. (2020). Deep Learning Based Smart Garbage Monitoring System. *MPCIT 2020 - Proceedings: IEEE 3rd International Conference on "Multimedia Processing, Communication and Information Technology,"* 77–81. <https://doi.org/10.1109/MPCIT51588.2020.9350390>
- Siavashi, J., Najafi, A., Ebadi, M., Fuel, M. S.-, & 2022, undefined. (n.d.). A CNN-based approach for upscaling multiphase flow in digital sandstones. *Elsevier*. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0016236121019232>
- Singh, D. (2021). Polyth-Net: Classification of Polythene Bags for Garbage Segregation Using Deep Learning. *2021 International Conference on Sustainable Energy and Future Electric Transportation, SeFet 2021*. <https://doi.org/10.1109/SEFET48154.2021.9375766>
- Ucar, F., hypotheses, D. K.-M., & 2020, undefined. (n.d.). COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images. *Elsevier*. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0306987720307702>
- Xiao, X., Xiao, W., Zhang, D., Zhang, B., Hu, G., Security, Q. L.-C. &, & 2021, undefined. (n.d.). Phishing websites detection via CNN and Multi-Head Self-Attention on imbalanced datasets. *Elsevier*. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0167404821001966>
- Zhou, J., Ren, K., Wan, M., Cheng, B., Gu, G., Optik, Q. C.-, & 2021, undefined. (n.d.). An Infrared and Visible Image Fusion Method Based on VGG-19 Network. *Elsevier*. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0030402621016363>