Research Paper
Journal of Smart Systems Research (JOINSSR) 4(2),160-176,2023
Received: 13-November-2023 Accepted: 08-December-2023
Doi: 10.58769/joinssr.1390459

SAKARYA UNIVERSITY OF APPLIED SCIENCES

A Yolov3-Based Garbage Detection Systems

Dilara KARACA[1*], Süleyman UZUN[2], Sezgin KAÇAR[3]

[1] Department of Biomedical Engineering, Graduate Education Institute, Sakarya University of Applied Sciences, dilarakaraca@subu.edu.tr
[2] Department of Computer Engineering, Faculty of Technology, Sakarya University of Applied Sciences, suleymanuzun@subu.edu.tr
[3] Department of Electrical and Electronics Engineering, Faculty of Technology, Sakarya University of Applied Sciences, skacar@subu.edu.tr

ABSTRACT

Today, the increase in the number of people and advances in industry and technology cause an increase in the number of wastes generated with the acceleration of production. It is important for the future of our country and the world that these wastes are more easily identified and recycled. In the process of recycling wastes, the classification of wastes as well as their collection require costly energy and manpower. Wastes are basically separated into paper, plastic, glass, and metal. Various studies have been carried out to complete these processes in a shorter and easier way with technologies such as artificial intelligence, deep learning, and image processing. In this study, waste detection was performed using YoloV3, an artificial intelligence network model frequently used in object detection, using a specially created dataset and global datasets. Also, a dataset of paper, plastic, and food and beverage wastes that are common in the environment was created. In this dataset, paper cups, plastic water bottles, and fast food wastes were detected from different locations in nature and photographed. These images were labeled, trained, and tested with YoloV3 deep learning algorithms. In addition, in order to compare the performance of the new dataset, studies were conducted on a global dataset used in the literature. As a result of the studies, it was observed that it was successful in classifying the newly created dataset and the global dataset.

Keywords: YoloV3, Garbage detection, Garbage dataset.

1. Introduction

Since the beginning of nature, various wastes have developed while living things on earth continue to carry out their essential functions. The development of robots, the creation of intelligent machines, and advancements in engineering all contribute to making life easier for people as time goes on and

* Corresponding Author's email: dilarakaraca@subu.edu.tr

technology advances. As in all facets of life, these developments have begun to be applied to the identification, collection, and recycling of waste. Waste is currently gathered and dumped in landfills. In these areas, a variety of reclamation projects are being completed. Examples of reclamation work done in landfills include recycling, incineration, and composting, which is the decomposition of organic wastes by microorganisms when they are reacted with in an oxygenated environment [1]. Some of these areas have waste areas covered and have been planted. Methane gas is now created after waste is gathered and handled in various ways. This gas is used to create electricity. The majority of the electricity requirements in the areas where these facilities are located are satisfied in this way. These storage spaces also provide a solution to the issue of waste collection. Additionally, avoid environmental harm from waste and the issue of unsightliness. People classify waste in these facilities. Workers separate different types of waste, including glass, paper, plastic, batteries, etc. The system suggested in this paper offers a solution for the problem of decomposition [2]. Wastes of all kinds, including toxic wastes, medical wastes, industrial wastes, and others, are created. In our nation, waste workers are responsible for collecting trash. Workers in the waste industry can be categorized as either permanent employees or temporary employees. The technique used for waste collection directly relates to the city's level of development [3]. On Earth, millions of tons of waste are produced every day. According to studies, paper and plastic make up the majority of the waste that is gathered. According to the research done in our nation in 2014, paper and cardboard make up 40.7% of the domestic waste collected from municipalities, and plastics make up 30% [4]. There are nations in the world that use the technique of sorting and getting rid of waste during the collection phase. While this process is currently carried out by humans, it is anticipated that, within the parameters of the study and with technological advancements, robots will be able to perform it. According to technological advancements, mechanization will increase in many areas, including the process of collecting waste [5]. Through artificial intelligence, machines are being taught to perform their tasks more humanely. Both supervised and unsupervised methods are used to accomplish this. The first approach involves teaching by naming various phenomena. The machine predicts a phenomenon based on what it has previously observed. The second approach is to instruct without defining the phenomenon. By grouping related features, learning occurs. Semi-supervised learning is possible by combining the two techniques. In this learning process, the learning process is realized, and 50% of the data is labeled. K-Nearest Neighbor (KNN) and SVM (Support Vector Machine) are two artificial neural network techniques and applications that have been developed. Typically, these techniques fall under the category of machine learning algorithms [6]. Machine learning algorithms make a prediction on the created model [7]. The neural networks in the human brain serve as the foundation for machine learning. A subset of machine learning known as "deep learning" was developed as artificial neural networks were further developed [8]. Convolutional Neural Networks (CNN), which stand for deep learning, have been used in many computer vision applications, including face recognition, text detection, target detection, etc. The process of detection using techniques derived from mathematical models based on human experience is known as machine vision [9]. Deep learning utilizes a layered architecture to process input data and produces more accurate results.

The studies conducted by designing autonomous vehicles and robots for waste sorting and waste sorting using deep learning models in the literature were analyzed. Alawi et al. [10], classified waste using CNN in their study. In their study, they used Alexnet, DenseNet121, and SqueezeNet algorithms from pre-trained networks. In the study, they used the waste classification data dataset made available for open access on the Kaggle website and used a total of 22564 images in two classes: recyclable and organic wastes. They divided 10% of the images as test data, 20% as validation data, and 70% as training data. According to the performance results obtained at the end of the study, the highest classification performance was achieved with the DenseNet121 pre-trained network, with an accuracy of 94.10%. With the other networks used in the study, SqueezeNet and Alexnet achieved 91% and 92% accuracy, respectively. Ramsurrun et al. [11], used a deep learning model to classify waste images into five classes: plastic, metal, paper, cardboard, and glass. It uses a trained machine-learning algorithm to perform object recognition. In their study, they used the Trashnet dataset, consisting of 2527 images divided into 6 classes. 70% of the dataset was used for training and 30% for testing. The proposed SVM, Sigmoid, and Softmax classifiers are tested with 12 different transfer learning techniques, and the highest result is obtained in the VGG-19+Softmax algorithm with an accuracy of approximately 87.9%. Assis et al. [12], in their study, developed a system that recognizes waste and collects it through an autonomous waste collection vehicle. The robot is connected to a Raspberry Pi device, ultrasonic sensors, and an Arduino UNO. The robot is programmed to stop when an object is detected at a distance of 20 cm or less from the ultrasonic sensor, and when the robot stops, the object detection module activates the robot arm. The object image is detected by a web camera. They used the YoloV3 real-time deep learning algorithm to distinguish the detected object image from other objects. In the study, they used a self-selected dataset divided into 5 classes: bottle, can, food package, paper ball, and plastic ball. The images were manually labeled with LabelImg. As a result of the study, they achieved 93% accuracy.

Koganti et al. [13], in their study, divide waste into biodegradable and non-biodegradable. In the study, images in the outdoor environment are taken with a Raspberry Pi (a SSD-Mobilenet trained model is pre-loaded) and the camera connected to it. Using the MobileNet deep learning algorithm in Rasperry Pi, the detected objects are classified into two different classes: biodegradable and non-biodegradable. The model gives results only for non-biodegradable waste. The test accuracy of the previously trained MobileNet network model was 99%. Zhihong et al. [14], realized a robotic sensing system for waste sorting in their study. The system consists of three main structures: a recognition module for waste detection, a transportation module for waste transportation, and a capture module for waste collection. The two main networks that make up the Fast-RCNN architecture were implemented with Region Proposal Network (RPN) and VGG-16 models, and these models were used for object recognition (bottle) in the study.

Different object detection studies with deep learning algorithms were also conducted, and model results were compared. In addition, different autonomous vehicles were designed, and different data sets were used in these studies. Khanum et al. [15], trained a ResNET convolutional neural network to drive an autonomous vehicle by lane control on a Udacity simulation. The study provided a network model trained to prevent traffic accidents that may occur in situations such as driver inattention by lane control, resulting in high achievements. In the study, the Udacity simulator was used by considering driving situations and target situations such as steering angle throttle, brake and speed. The images stored in the folder are saved as a csv file. The data is decomposed as 80% training and 20% testing. The model result showed that 0.81% of the time the vehicle correctly planned its movement to avoid leaving the lane. Turgut et al. [16], conducted a study for search and rescue activities under the rubble caused by natural disasters. It is very dangerous to work in risky areas as a result of natural disasters. Semantic classification of the ramps and obstacles in the tracks facilitated the robots' operations in these areas. The data obtained with 2b laser gives information about the height at which the ramps are located. In this study, semantic classification of planes such as ramps, floors and walls were performed with PointNET, Dynamic Graph Convolution Network (DGCNN), PointNet++ and PointCNN deep learning models. ESOGU RAMPS dataset was used. 581 data were used for training while 100 were used for testing. High performance was obtained in all CNN models used. The highest performance was obtained in the DGCNN model with 99.9%. Yao et al. [17], proposed a machine vision-based vehicle detection system for smart cities in their study. In the study, it was determined that the traditional YoloV3 algorithm is not a preferred method for fast-moving and small-sized vehicles. For this reason, the model created by adding one more convolution layer to the original YoloV3 layer was used. The study uses 1642 photographs of traffic vehicles taken from the internet. Labeling was done in Pascal VOC format.

The model achieved an accuracy of 91.01% (mAP). Rajesware et al. [18], used transfer learning methods to detect Alzheimer's disease, which is very difficult to identify and diagnose. In this context, they used techniques such as VGG-19, VGG-16, ResNet50, and Xception. The dataset used for the study is ADNI brain MRI images. 80% of the data was used for training and 20% for testing. The highest accuracy was obtained in the VGG-19 model with 98%.

In the literature, there are also studies that use data sets generated by waste but use algorithms other than deep learning models. In the studies, systems proposed for waste sorting, waste collection, and controlling waste levels in waste bins were mentioned.

Ravanan et al. [19], proposed an IoT-based smart waste monitoring system. The system monitors waste bins with ultrasonic sensors and provides information about the amount of waste through a web page. Their proposed system consists of a PIC microcontroller, GSM module and wifi module. When the measured waste level in the waste bins is equal to the width of the waste bins, the data sent to the server alerts the authority via the internet. The system proposed by Amitha et al. [20] determines the level of waste in the waste bin with the help of sensors. It checks whether the waste bin is full or empty. It classifies the waste as dry or wet and sends a message to the user when it detects that the waste bin is full. Authorities have access to the information that the system is full or empty at any time. Thus, the need for unnecessary use of waste collection vehicles is eliminated. Bai et al. [21] designed a deep learning-based robot that automatically collects waste from lawns. GPS measurement module, navigation module, sensing module, map module, and driving module are used in the robot. In the map module, the environment to be cleaned is defined for the robot. The sensing module is used to understand whether the detected object is an object or ground. ResNet-34 is used in the architecture of the network. If the detection module does not work correctly, it will perceive a non-waste object as waste. For this reason, a data set consisting of 5 waste classes and 1 non-waste class was used. It consists of 40,000 training images and 7,000 test images. The robot was tested in a playground. Experimental results show that waste recognition accuracy can reach up to 95%. Yuan et al. [22] designed a robot for cleaning swimming pools. It is important to clean swimming pools that have been in continuous use for a short time. The robot can be controlled by a wired remote control with a wheeled movement feature. It also consists of hydraulic push and movement mechanisms. The control box provides control of the robot, while the power supply provides power for the robot. The monitor displays the working environment in time with the camera of the robot body. The water passing through the filter is sent out by the pump of the robot body.

In the second part of the study, the dataset created specifically for the study and the global dataset used for comparison purposes are analyzed in detail. This section also focuses on the data labeling processes and the complexity matrix. Furthermore, this section provides a detailed review of the Yolo v3 network structure. The third section presents the results of the experimental studies and describes the results in detail. The fourth and final section summarizes the overall conclusions of the study.

2. Material and Method

2.1. Custom Dataset

In this study, two distinct datasets are used to compare deep network model outputs and draw conclusions using performance metrics. The first data set was specifically created for the study, whereas the second data set was created using data that was widely available. Paper cups, plastic bottles, and waste food and beverage packaging make up the majority of the waste images in the dataset created for the study. Three classes make up the data set that has been prepared for the study. Paper cups, plastic bottles, and other trash are separated from it. The data set for other wastes includes bags, cardboard, food and beverage packaging, etc. 1225 images in total were used. These photos are unique ones that were taken in the study's setting. It consists of 250 waste images, 500 plastic bottle images, and 500 paper cup images. The most frequent items found in the environment, such as food and beverage packaging and fast food packaging, are included in the data set for the other waste class. The images have a resolution of 1200x1599 and are in jpg format. The waste found in places like playgrounds, street sides, and school gardens was photographed to create the images. The dataset created especially for this study will be made public and shared on a global scale.

Figure 1: Examples from the custom dataset

2.2. Global Dataset

The publicly accessible Trashnet dataset was also used to compare the outcomes of the networks used in the study. [23]. There are 2527 images total in this dataset. The 512x384 dataset was captured in natural light and under artificial lighting, and it was printed on a white poster board. Images from six categories of waste-glass, paper, cardboard, plastic, and metal—are included in the dataset. 1000 images from the Trashnet dataset were used to create a more precise comparison with the dataset produced for the purposes of the study. Only images from the global dataset that belonged to the classes of paper, glass, and metal were chosen.

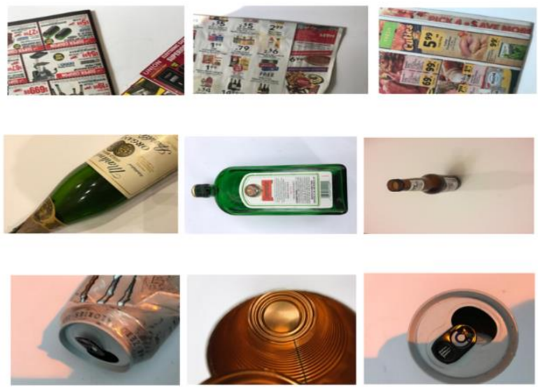

Figure 2: Examples of the global dataset's images

2.3. Labeling The Dataset

Paper cups, plastic bottles, and other waste were the three categories into which the data set that was created for the study was labeled. Glass, paper, and metal were assigned to images from the Trashnet dataset. MakeSense.AI (MSA) was used to manually label the data [24]. 70% of the labeled data was used for training, 20% for testing, and 10% for validation. After the labeling process, the program generates the dataset for the study as a.txt file. The values of the image are stored in this file as matrices. The data consists of the object's ID, the x and y axes' centers, and its width and height.

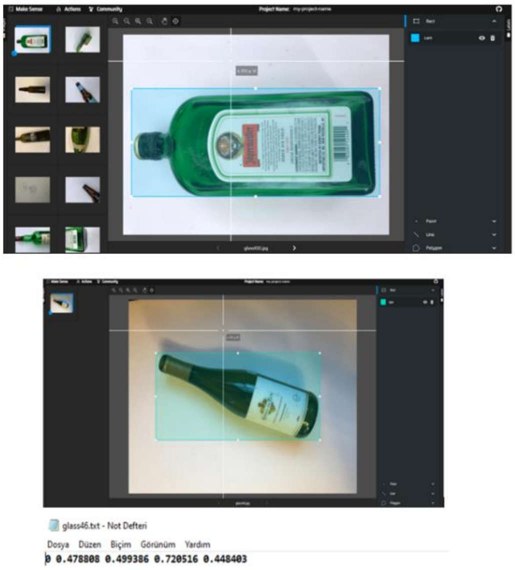

Figure 3: Labeling images with MSA

2.4. Confusion Matrix

In machine learning and deep learning algorithms, where results are obtained through mathematical calculations from a set of outputs, the confusion matrix is a performance metric. The estimated and actual values from the algorithm are combined in four different ways in the matrix. A dataset of 400 images, 200 of which included paper cup images and 200 of which did not, was created. The confusion matrix depicted in Figure 7 was used to calculate the deep learning network's performance after training and testing on this dataset.

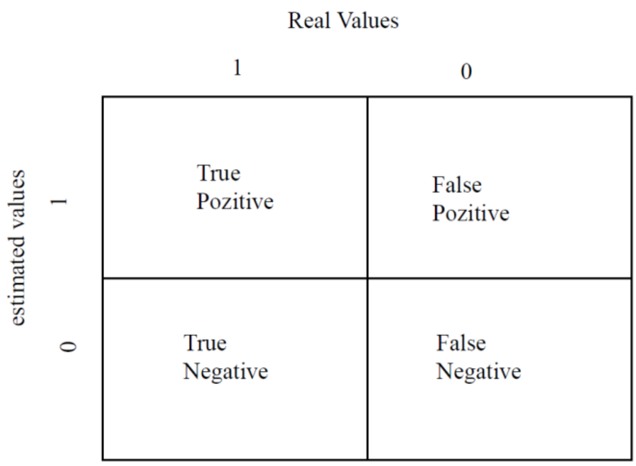

Figure 4: Confusion matrix

The matrix represents the following four states:

True Positive (TP): The prediction of the data that includes a picture of a paper cup is referred to as "the one with a paper cup" in this situation.

True Negative (TN): "No paper cup" is another name for the prediction of the data that does not include a picture of a paper cup.

False Positive (FP): "With a picture of a paper cup" is the prediction made for data that does not contain a paper cup.

False Negative (FN): Data with an image of a paper cup is predicted to be "an image without a paper cup."

These numbers are used to calculate various performance metrics in the confusion matrix [25].

$$\begin{aligned} Sum &= TP + TN + FP + FN \\ Real\ Positive &= TP + FN \\ Real\ Negative &= TN + FP \end{aligned} \tag{1}$$

Accuracy: Refers to how accurate the model is all in all. It is calculated by multiplying all of the matrix's data by the accurate forecasts. In Equation 2, it is specified.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{2}$$

Recall: The percentage of correct values that are classified as correct is known as sensitivity, and it is expressed as a recall metric. It appears in equation three.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

Precision: a way to gauge how well and how poorly all outcomes are predicted. In Equation 4, it is specified.

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

F1- Score: The harmonic mean of the sensitivity and sensitivity of the score is F1. In Equation 5, it is specified.

$$F1 - Score = \frac{2*Precision*Recall}{Precision+Recall} \tag{5}$$

2.5. Artificial Intelligence

With the introduction of the "Turing test" in the 1950s, Alan Turing first laid the groundwork for artificial intelligence, which now encompasses the fields of machine learning and deep learning [26].

The idea of artificial intelligence is based on the artificial working structure of neuron cells in the human brain and is derived from the idea of intelligence. Science has greatly benefited from understanding how the human brain functions in order to train machines [27]. According to a recent report, thinking robots could be created by imbuing machines with emotions as a result of the human brain's similarities to machines [28]. With its historical development, artificial intelligence has become specialized in machine learning and deep learning. By processing large amounts of data collected on a variety of topics in machine learning algorithms, intelligent results can be produced [29]. Deep learning performs training and results with its multi-layered architecture, which is acknowledged to produce more accurate results than machine learning [30].

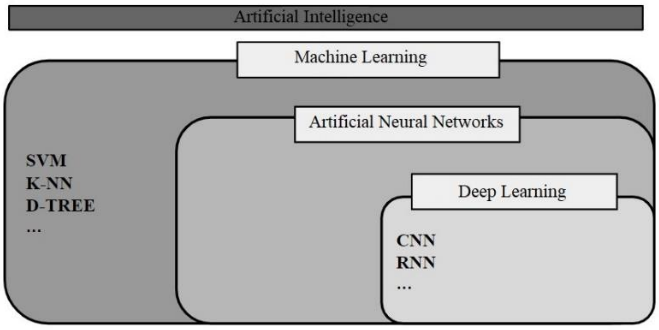

Figure 5: Artificial Intelligence and subclasses

2.6. Yolo version 3

Different deep learning algorithms have been developed up to this point. Yolo is one of the most successful and quick algorithms. This is because the algorithm only takes into account the image's bounding boxes as a whole [31]. A real-time object detection algorithm called YoloV3 was created by Redmon et al. in 2018 [32]. Both the performance and the object detection rate are quite high. YoloV3 was developed using the Yolo and YoloV2 network models [33]. Furthermore, Yolo V3 [34] used Darknet-53 as a backbone network to replace Darknet-19 [35], [36]. Yolo networks perform very well in fast object detection and especially in real-time object detection [37]. The YOLOv3 network was chosen because it uses Darknet-53, which is a better backbone than YoloV2, and uses a single-stage CNN network and a feature extraction network [38], [39].

The COCO dataset is utilized by the initial YoloV3 architecture. TensorFlow and the DARKNET libraries' networks were used to train this dataset. The 200,000 images in the COCO dataset, which includes 80 different objects, were prepared for detection, tracking, and classification. An open source dataset is COCO [40]. The object is divided into regions by the algorithm. Bounding boxes, which are drawn around the object in a region, are created. In order to determine the likelihood of discovering the object, calculations are performed on the marked boxes. For each object, the marked boxes determine a confidence score. The percentage of time that the algorithm is confident that the data is present in the grid is represented by the confidence score [41]. An image is provided to the algorithm as input data. The image is set to be 32 pixels wide and has multiples. The Darknet-53 CNN model is utilized by the YoloV3 algorithm as a feature extractor. Consequently, an additional 53 layers are added to the object detection algorithm [42]. The Yolo layer makes use of the features that were extracted from the convolution layer (darknet-53). By using pre-trained weights for object detection, the network in the

Darknet 53 layer is taught how to extract features. There are a total of 106 convolutional layers in the YoloV3 architecture. As seen in Figure 5, YoloV3 employs feature pyramid networks (FPN) to find objects at three different scales. The input image is divided into 52×52 boxes at scale 3, 26×26 boxes at scale 2, and 13×13 boxes at scale 1. This results in the detection of the smallest objects at scale 3 and the medium-sized objects at scale 2. The algorithm executes the output process by executing specific steps on each of these three scales. The overlap between the estimated bounding box value (Intersection over Union, IoU) and the actual area value of the object is used to calculate the confidence score. The same object is enclosed several times in the marked boxes if more than one output vector detects it. Application of the Non-Maximal Suppression (NMS) algorithm The calculated threshold value is used to delete checked boxes that fall below it.

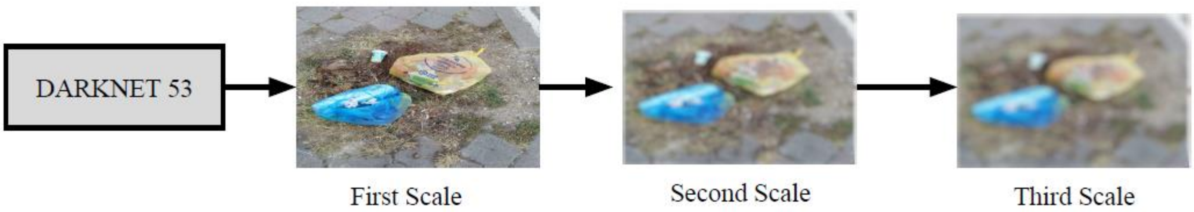

Figure 6: Output of the YoloV3 network

3. Experimental Results

3.1. YoloV3 Training Phase

To train the YoloV3 network, files called "obj.name," "obj.data," "generatetrain.py," and “generatevalidation.py” are created. The names of the classes used in training are listed in the “Obj.names” file in the labeling order. The number of classes, the training and validation file paths, the path to the file containing the class names (obj.names), and the path to the folder where the weights produced as a result of the training will be saved are all contained in the “object.data” file. The training data set access codes for DarkNet are written in a file called "Generate_train.py." The code for DarkNet to access the validation file is contained in the file "Generate_validation.py." The preparation of the file with a ".cfg" extension is the final step in the algorithm's preparation. By reading this file, the algorithm creates the network. The batch parameter is the parameter that specifies the number of images processed at the same time. It is set to 1 to increase the training speed. The "batch" parameter divides the data set into portions for training. We train using these chunks. The batch parameter specifies the part size and is assumed to be 1 during the test phase. The max batch value is used to prevent overfitting by controlling the number of batches during the training phase. The value of Max_batches is determined. It conveys the parameter size's maximum value. "class_number * 2000" is the formula used to calculate this value. In single-class algorithms, it is interpreted as 4000. Calculating "Steps = 4800, 5400" involves subtracting 80% and 90%, respectively, of the value of max_bathces. " (Number of classes + 5) * 3" is the formula for the filter parameter. The final three convolution layers at the bottom have their filter values changed. The DarkNet weights to be used as a feature extractor are first downloaded to begin training in the Colab environment, which enables us to run Python codes directly through the browser used in studies such as machine learning and data analysis. This comes after the file preparations are complete. The Open Source Computer Vision Library (OPENCV) is configured for the Graphics

Processing Unit (GPU). By connecting to the Drive environment, images and labeling files that have been prepared for training and validation are moved to the data folder in DarkNet. "obj.names, obj.data," generate train and validation files, and the ".cfg" file created at the start of the training are transferred to DarkNet.

Obj.names

```
Paper cup
Pet bottle
Other garbage
```

(a)

Obj.data

```
Classes = 3
Train = data/train.txt
Valid = data/validation.txt
Names = data/obj.names
Backup =
/content/gdrive/MyDrive/yolov3/backup/
```

(b)

```
[net]
# Testing
# batch=1
# subdivisions=1
# Training
…..
max_patches = 6000
#(number of class * 2000)
policy=steps
steps=4800,5400 #
max_batches * 80 / 100 ve 90
/ 100
…..
[convolutional]
size=1
stride=1
pad=1
filters=24 #(number of class +
5 ) * 3
activation=linear
```

(c)

Figure 7: Example of project files. (a) Obj.names file. (b) Obj.data file. (c) cfg file.

3.2. Training Results

The Mean Average Precision (mAP) values obtained from the original dataset created for this study using the YoloV3 algorithm using 3 different classes, including paper cups, pet bottles, and other wastes, are provided in Table 1.

Table 1: *Performance results of the classes as a result of training from custom dataset.*

	mAP	TP	FP
Paper Cup	100.00%	201	3
Pet bottle	98.34%	203	2
Other trash	76.91%	96	38

The most precise prediction value for each class in the specially created data set was 203 for the pet bottle class, according to Table 1. When the results from images with and without objects are compared, it can be seen that the other waste class has the highest error value. 38 is this value. The paper cup class received the best performance value of 100% out of the three classes' AP scores. The F1 Score, Recall, and Precision values of the network after training all classes with the Yolov3 algorithm are shown in Table 2.

Table 2: *Overall performance result obtained as a result of training from custom dataset.*

	Recall	Precision	F1-Score
Custom Data Set	0,94	0,92	0,93

In the study, 500 data points were correctly predicted. 43 data points were found to be present even though they were not objects. In 31 image data points, although there was a labeled object, it was found to be absent. In the testing process of the YoloV3 network model trained with the dataset prepared for the study, 175 test data sets are used, different from the data used for training and validation. The results obtained from these data are given in Table 3.

Table 3: *Performance results of the classes as a result of test from custom dataset.*

	F1-Score	Precision	Recall
Paper Cup	0,9902	0.9712	1,0000
Pet bottle	0,9804	0.9712	0,9902
Other trash	0,3902	0,2682	0,7164

The highest accomplishments belong to the paper cup class, per Table 3's results. The pet bottle class had an accuracy value of 0.98, the paper cup class had an accuracy value of 0.99, and the other waste class had an accuracy value of 0.39. The high values obtained in the test results are a result of the patterns on the paper cups being similar to the data in the paper cup dataset. Table 4 provides the test performance results for the entire YoloV3 network.

Table 4: *Overall performance result obtained as a result of test from custom dataset.*

	F1-Score	Precision	Recall
Custom Data Set	88.111	73.682	90.220

In the study, the network's mAP value of 86.45 was calculated from test results using various images from 175 training and custom datasets. There are tables of the results obtained using the deep learning network models and the custom dataset. Below are some images that were validated using the YoloV3 algorithm using a custom dataset. Figure 8 displays the results of testing the YoloV3 network model that had been trained on a single image.

Figure 8: Images from test results produced with a custom dataset.

According to these values, the highest value belongs to the images belonging to the paper cup class and then to the pet bottle class. The high performance in the paper cup class is due to the similar external designs of the paper cup objects in the images. The high number of similar images resulted in high performance when the network was tested with new data that was similar to the images in the trained dataset. The lowest value belongs to the food packaging belonging to the other waste category. This is due to the fact that the other waste dataset consists of multiple different types of data with low similarity. The fact that this data set category does not belong to a single standard in terms of similarity means that training with very different data and testing with very different data again reduces the network performance by one level. According to the training results, it is seen that higher achievements are obtained when the real area overlaps with the predicted area. The higher achievements obtained from the paper cup class are due to the fact that the objects in the training and test images are similar; the difference is due to the object shape, color, and pattern, and these options contain less different formatting compared to other datasets. The data for the YoloV3 architecture training and testing process with the global data set are given in the tables below. Performance values of 700 data points used in training are given in Table 5.

Table 5: *Performance results of the classes as a result of training from global dataset.*

	mAP	TP	FP
Glass	%95. 11	42	9
Metal	%95.83	35	3
Paper	%96.13	33	0

When the training data set was used, the best performance was obtained from the paper data set when the correct prediction for each class was proportional to the number of data sets. None of the images without the paper dataset were identified as having paper. The highest AP value was obtained from the paper dataset.

Table 6: *Overall performance result obtained as a result of training from global dataset.*

	Recall	Precision	F1-Score
Global Data Set	0.90	0.90	0.90

Table 6 shows the training network performance metrics of the YoloV3 algorithm. Precision, recall, and F1-score values were obtained as 0.90. In the results of the study, 122 correct predictions were made. For 12 images with the labeled object, the prediction was made as no object.

Table 7: *Performance results of the classes as a result of test from global dataset.*

	F1-Score	Precision	Recall
Glass	0.8636	0,7600	1,0000
Metal	0.8510	0,7407	1,0000
Paper	0,9050	0,9048	0,9048

The highest success rate of the test result of the global dataset is obtained from the glass class. The mAP value of the glass class was obtained as 1.0000. The results of the test dataset, consisting of images different from the images in the training and validation datasets, are given in the tables below.

Table 8: *Overall performance result obtained as a result of test from global dataset.*

	F1-Score	Precision	Recall
Global Data Set	87.722	80.183	96.825

Figure 9: Test result images obtained using global dataset.

.

As a result of the test study with the global dataset, the individual results of the prediction of the sample images of the metal, paper, and glass classes that are not included in the dataset are given in Figure 9. The sample image belonging to the glass class was predicted correctly at a rate of 0.50. The sample image belonging to the metal class was predicted correctly with a prediction rate of 0.79, and the paper class was predicted correctly with a prediction rate of 0.95. The highest prediction success was obtained in the paper class.

Table 9: *A comparison table of the results obtained in the study with the literature*.

Authors	Year	Network Model	Data set	mAP(%)
Carolis et al. [43]	2020	YoloV3	Trashnet	59.57
Lin et al. [44]	2022	Yolo-Green	Trashnet	78.04
Mao et al. [45]	2022	YoloV3	Trashnet	81.36
Zhang et al. [46]	2022	Yolo-Waste	Trashnet	93.12
Proposed Method	2023	YoloV3	Trashnet	96.67
Proposed Method	2023	YoloV3	Custom	86.45

The comparison of the obtained results with the literature is shown in Table 9. When the table is analyzed, it is seen that Yolo-based deep learning networks have achieved quite high performance compared to the studies on the Trashnet dataset.

4. Conclusion

In this study, a system that classifies the wastes that are found in settlements has been developed. In the developed system, an algorithm that includes transfer learning techniques from deep learning network models and the CNN algorithm is proposed. The networks used in the study were trained and tested using two types of data sets. The first data set used is a data set consisting of glass, plastic, and paper waste prepared for the study. The other dataset used is the globally published Trashnet dataset. It consists of six classes: glass, paper, cardboard, plastic, metal, and waste. Considering the performance results, the results obtained with the dataset prepared for the study are generally higher than the performance metrics obtained from the Trashnet dataset. The reason is that the integrity of the images in the dataset prepared for the study is higher. In this sense, the dataset prepared for the study has a privileged aspect in the category of images that increases its performance.

5. Declarations

5.1. Study Limitations

The findings of the study show that nowadays it is of great importance to automatically detect, collect, and recycle waste. With the development of artificial intelligence, this is now possible. However, a data set is needed to train the AI. Considering that the waste data set is quite small today, it is thought that the waste data set created in this study will contribute significantly to the field.

5.2. Acknowledgements

There are no acknowledgements available.

5.3. Funding source

There is no funding source available.

5.4. Competing Interests

There is no conflict of interest in this study.

5.5. Authors' Contributions

Define the contribution of each researcher named in the paper to the paper.

Corresponding Author Dilara KARACA: Developing ideas or hypotheses for the research and/or article, planning the materials and methods to reach the results, taking responsibility for the experiments, organizing and reporting the data, taking responsibility for the explanation and presentation of the results, taking responsibility for the literature review during the research.

2. Author's Süleyman UZUN: Developing ideas or hypotheses for the research and/or article, Contribution to the article. Organizing and reporting the data, taking responsibility for the explanation and presentation of the results, taking responsibility for the literature review during the research, taking responsibility for the creation of the entire manuscript or the main part, reworking not only in terms of spelling and grammar but also intellectual content or other contributions.

3. Author's Sezgin KAÇAR: Contribution to the article. Organizing and reporting the data, taking responsibility for the explanation and presentation of the results, taking responsibility for the literature review during the research, taking responsibility for the creation of the entire manuscript or the main part, reworking not only in terms of spelling and grammar but also intellectual content or other contributions.

6. Human and Animal Related Study

The work does not involve the use of human/animal subjects.

6.1. Ethical Approval

No ethical approval is required for this project.

6.2. Informed Consent

All authors consent to the publication of the study.

References

[1] A. Sağlık, Y. Selim Domaç, Ş. N. Reyhan, F. Avcı, F. Kartal, and D. Şenkuş, "Akademia Doğa ve İnsan Bilimleri Dergisi Academia Journal of Nature and Human Sciences Katı Atık Depolama Alanlarının Islahı ve Analizi Çanakkale Onsekiz Mart Üniversitesi Örneği," vol. 7, no. 1, pp. 105–125.

[2] R. Erdoğan and G. Uzun, "Katı Atık Depolama Alanlarının Bitkisel Islahına Bir Örnek: Adana-Sofulu Çöp Depolama Alanı," *Akdeniz Üniversitesi Ziraat Fakültesi Derg.*, vol. 20, no. 1, pp. 71–82, 2007.

[3] N. Özgen, "Kent ve çöp," *TBB Mesleki Sağlık ve Güvenlik Derg.*, vol. 7, no. 8, pp. 10–12, 2006.

[4] T. Ç. M. Odası, "Dünya çevre günü Türkiye raporu," TMMOB Çevre Mühendisleri Odası. [Online]. Available: http://www.cmo.org.tr/resimler/ekler/0d4a5b926c005a6_ek.pdf, Erişim Tarihi: 12.10.2021

[5] P. P. Rao, S. P. Rao, and R. Ranjan, "Deep Learning Based Smart Garbage Monitoring System," *MPCIT 2020 - Proc. IEEE 3rd Int. Conf. "Multimedia Process. Commun. Inf. Technol.*, pp. 77–81, Dec. 2020, doi: 10.1109/MPCIT51588.2020.9350390.

[6] A. Datumaya Wahyudi Sumari, R. Andrie Asmara, D. Rossiawan Hendra Putra, and I. Noer Syamsiana, "Prediction Using Knowledge Growing System: A Cognitive Artificial Intelligence Approach," *Proc. - IEIT 2021 1st Int. Conf. Electr. Inf. Technol.*, pp. 15–20, Sep. 2021, doi: 10.1109/IEIT53149.2021.9587367.

[7] E. Saygin *et al.*, "Karaciğer Yetmezliği Teşhisinde Özellik Seçimi Kullanarak Makine Öğrenmesi Yöntemlerinin Başarılarının Ölçülmesi," *Fırat Üniversitesi Müh. Bil. Derg. Araştırma Makal.*, vol. 33, no. 2, pp. 367–377, 2021, doi: 10.35234/fumbd.832264.

[8] F. Doğan and İ. Türkoğlu, "Derin Öğrenme Modelleri ve Uygulama Alanlarına İlişkin Bir Derleme," *DÜMF Mühendislik Derg.*, vol. 10, no. 2, pp. 409–445, 2019, doi: 10.24012/dumf.411130.

[9] G. Liu, "Surface Defect Detection Methods Based on Deep Learning: A Brief Review," *Proc. - 2020 2nd Int. Conf. Inf. Technol. Comput. Appl. ITCA 2020*, pp. 200–203, Dec. 2020, doi: 10.1109/ITCA52113.2020.00049.

[10] A. E. Ba Alawi, A. Y. A. Saeed, F. Almashhor, R. Al-Shathely, and A. N. Hassan, "Solid Waste Classification Using Deep Learning Techniques," *2021 Int. Congr. Adv. Technol. Eng. ICOTEN 2021*, Jul. 2021, doi: 10.1109/ICOTEN52080.2021.9493430.

[11] N. Ramsurrun, G. Suddul, … S. A.-… zooming innovation in, and undefined 2021, "Recyclable waste classification using computer vision and deep learning," *ieeexplore.ieee.orgN Ramsurrun, G Suddul, S Armoogum, R Foogooa2021 zooming Innov. Consum. Technol. Conf. (ZINC), 2021•ieeexplore.ieee.org*, pp. 11–15, May 2021, doi: 10.1109/ZINC52049.2021.9499291.

[12] A. Assis, A. R. Biju, N. A. Alisha, A. Dhanadas, and N. Kurian, "Garbage Collecting Robot Using YOLOv3 Deep Learning Model," *10th Int. Conf. Adv. Comput. Commun. ICACC 2021*, 2021, doi: 10.1109/ICACC-202152719.2021.9708298.

[13] S. K. Koganti, G. Purnima, P. Bhavana, Y. V. Raghava, and R. Resmi, "Deep Learning based Automated Waste Segregation System based on degradability," *Proc. 2nd Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2021*,

pp. 1953–1956, Aug. 2021, doi: 10.1109/ICESC51422.2021.9532837.

[14] C. Zhihong, Z. Hebin, W. Yanbo, L. Binyan, and L. Yu, "A vision-based robotic grasping system using deep learning for garbage sorting," in *Chinese Control Conference, CCC*, 2017, pp. 11223–11226. doi: 10.23919/ChiCC.2017.8029147.

[15] A. Khanum, C. Y. Lee, and C. S. Yang, "End-to-end deep learning model for steering angle control of autonomous vehicles," *Proc. - 2020 Int. Symp. Comput. Consum. Control. IS3C 2020*, pp. 189–192, Nov. 2020, doi: 10.1109/IS3C50286.2020.00056.

[16] K. Turgut and B. Kaleci, "Comparison of Deep Learning Techniques for Semantic Classification of Ramps in Search and Rescue Arenas," *Proc. - 2020 Innov. Intell. Syst. Appl. Conf. ASYU 2020*, Oct. 2020, doi: 10.1109/ASYU50717.2020.9259851.

[17] X. Yao *et al.*, "Traffic vehicle detection algorithm based on YOLOv3," *Proc. - 2021 Int. Conf. Intell. Transp. Big Data Smart City, ICITBS 2021*, pp. 47–50, Mar. 2021, doi: 10.1109/ICITBS53129.2021.00020.

[18] S. S. Rajeswari and M. Nair, "A Transfer Learning Approach for Predicting Alzheimer's Disease," *2021 Int. Conf. Nascent Technol. Eng. ICNET 2021 - Proc.*, Jan. 2021, doi: 10.1109/ICNTE51185.2021.9487746.

[19] V. Ravanan, R. Subasri, M. G. Vimal Kumar, K. T. Dhivya, P. S. Kumar, and K. Roobini, "Next Generation Smart Garbage Level Indication and Monitoring System using IoT," *Proc. - 1st Int. Conf. Smart Technol. Commun. Robot. STCR 2021*, Oct. 2021, doi: 10.1109/STCR51658.2021.9588961.

[20] S. Amitha *et al.*, "Segregated waste collector with robotic vacuum cleaner using internet of things," *Proc. - 2020 IEEE Int. Symp. Sustain. Energy, Signal Process. Cyber Secur. iSSSC 2020*, Dec. 2020, doi: 10.1109/ISSSC50941.2020.9358839.

[21] J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu, "Deep Learning Based Robot for Automatically Picking Up Garbage on the Grass," *IEEE Trans. Consum. Electron.*, vol. 64, no. 3, pp. 382–389, Aug. 2018, doi: 10.1109/TCE.2018.2859629.

[22] F. C. Yuan, H. L. Sun, S. J. Hu, and L. Z. Wang, "Design of cleaning robot for swimming pools," *2011 Int. Conf. Manag. Sci. Ind. Eng. MSIE 2011*, pp. 1175–1178, 2011, doi: 10.1109/MSIE.2011.5707629.

[23] G. Thung, "Trashnet Dataset." [Online]. Available: https://github.com/garythung/trashnet/blob/master/data/dataset-resized.zip, Erişim Tarihi: 10.08.2021

[24] M. Hewitt, "Make Sense," Victorian Studies. [Online]. Available: https://www.makesense.ai/, Erişim Tarihi: 10.08.2021

[25] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, "The impact of class imbalance in classification performance metrics based on the binary confusion matrix," *Pattern Recognit.*, vol. 91, pp. 216–231, Jul. 2019, doi: 10.1016/J.PATCOG.2019.02.023.

[26] J. Mccarthy, "What Is Artificia Intelligence?," in *What Is Artificia Intelligence?*, philpapers.org, 2004.

[27] J. H. Fetzer, "What is Artificial Intelligence?," pp. 3–27, 1990, doi: 10.1007/978-94-009-1900-6_1.

[28] K. Arslan, "Eğitimde Yapay Zekâ ve Uygulamaları," *Batı Anadolu Eğitim Bilim. Derg.*, vol. 11, no. 1, pp. 71–80, 2020.

[29] J. G. Carbonell, R. S. Michalski, and T. M. Mitchell, "An Overview of Machine Learning," *Mach. Learn.*, pp. 3–23, 1983, doi: 10.1007/978-3-662-12405-5_1.

[30] E. S. Brunette, R. C. Flemmer, and C. L. Flemmer, "A review of artificial intelligence," *ICARA 2009 - Proc. 4th Int. Conf. Auton. Robot. Agents*, pp. 385–392, 2009, doi: 10.1109/ICARA.2000.4804025.

[31] T. Dergisi, A. Altay, S. Yilmaz, K. Tarihi, A. Kelimeler, and Y. Algoritmas, "YOLO Algoritması Kullanılarak T Hücrelerinin Sınıflandırılması Classification Of T Cells Using YOLO Algorithm Abstract," vol. 3, no. 2, pp. 66–81, 2023.

[32] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *arxiv.org*, Apr. 2018, Accessed: Apr. 17, 2023. [Online]. Available: https://arxiv.org/abs/1804.02767

[33] A. Aktaş, B. Doğan, and Ö. Demir, "Tactile paving surface detection with deep learning methods," *J. Fac. Eng. Archit. Gazi Univ.*, vol. 35, no. 3, pp. 1685–1700, 2020, doi: 10.17341/gazimmfd.652101.

[34] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in *IEEE conference on computer vision 2017, Undefined*, 2017, pp. 7263–7271.

[35] S. K. Tai, C. Dewi, R. C. Chen, Y. T. Liu, X. Jiang, and H. Yu, "Deep Learning for Traffic Sign Recognition Based

on Spatial Pyramid Pooling with Scale Analysis," *Appl. Sci. 2020, Vol. 10, Page 6997*, vol. 10, no. 19, p. 6997, Oct. 2020, doi: 10.3390/APP10196997.

[36] C. Dewi, R. C. Chen, and S. K. Tai, "Evaluation of Robust Spatial Pyramid Pooling Based on Convolutional Neural Network for Traffic Sign Recognition System," *Electron. 2020, Vol. 9, Page 889*, vol. 9, no. 6, p. 889, May 2020, doi: 10.3390/ELECTRONICS9060889.

[37] C. Dewi, R.-C. C. Chen, X. Jiang, and H. Yu, "Deep convolutional neural network for enhancing traffic sign recognition developed on Yolo V4," *Multimed. Tools Appl.*, vol. 81, no. 26, pp. 37821–37845, Nov. 2022, doi: 10.1007/s11042-022-12962-5.

[38] H. Hussein, Y. Zhu, R. F. Hassan, S. Teng, Z. Liu, and X. Li, "Improved YOLOv3-Based Bridge Surface Defect Detection by Combining High- and Low-Resolution Feature Images," *Build. 2022, Vol. 12, Page 1225*, vol. 12, no. 8, p. 1225, Aug. 2022, doi: 10.3390/BUILDINGS12081225.

[39] S. Teng, Z. Liu, G. Chen, and L. Cheng, "Concrete Crack Detection Based on Well-Known Feature Extractor Model and the YOLO_v2 Network," *Appl. Sci. 2021, Vol. 11, Page 813*, vol. 11, no. 2, p. 813, Jan. 2021, doi: 10.3390/APP11020813.

[40] E. Ulutaş, H. Cengİz, and M. C. Yazicioğlu, "Derin öğrenme temelli nesne tespiti algoritmaları kullanılarak kişiye özgü reklam sunulması," vol. 24, no. 1, pp. 10–28, 2022, doi: 10.25092/baunfbed.878224.

[41] T. Shi, Y. Niu, M. Liu, Y. Yang, C. Wang, and Y. Huang, "Underwater dense targets detection and classification based on YOLOv3," *IEEE Int. Conf. Robot. Biomimetics, ROBIO 2019*, pp. 2595–2600, Dec. 2019, doi: 10.1109/ROBIO49542.2019.8961615.

[42] A. Karacı, "Detection and classification of shoulder implants from X-ray images: YOLO and pre-trained convolution neural network based approach," *J. Fac. Eng. Archit. Gazi Univ.*, vol. 37, pp. 283–294, 2022, doi: 10.17341/gazimmfd.888202.

[43] B. De Carolis, F. Ladogana, and N. Macchiarulo, "YOLO TrashNet: Garbage Detection in Video Streams," in *2020 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, IEEE, May 2020, pp. 1–7. doi: 10.1109/EAIS48028.2020.9122693.

[44] W. Lin, "YOLO-Green: A Real-Time Classification and Object Detection Model Optimized for Waste Management," in *2021 IEEE International Conference on Big Data (Big Data)*, IEEE, Dec. 2021, pp. 51–57. doi: 10.1109/BigData52589.2021.9671821.

[45] W.-L. Mao, W.-C. Chen, H. I. K. Fathurrahman, and Y.-H. Lin, "Deep learning networks for real-time regional domestic waste detection," *J. Clean. Prod.*, vol. 344, p. 131096, Apr. 2022, doi: 10.1016/j.jclepro.2022.131096.

[46] Q. Zhang *et al.*, "A multi-label waste detection model based on transfer learning," *Resour. Conserv. Recycl.*, vol. 181, p. 106235, Jun. 2022, doi: 10.1016/J.RESCONREC.2022.106235.

© 2020 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).