



International Journal of Marketing Management

ISSN 2454 - 5007



www.ijmm.net

Email ID: editor@ijmm.net , ijmm.editor9@gmail.com

The Use of Convolutional Neural Networks (CNN) for the Identification of Recycling and Garbage

¹ S. Akhila, ² G. Varshni,

¹Assistant Professor, Megha Institute of Engineering & Technology for Women, Ghatkesar.

² MCA Student, Megha Institute of Engineering & Technology for Women, Ghatkesar.

Abstract—

Using deep learning and image processing algorithms to better identify waste in public settings, this study aims to enhance municipal trash collection. Better waste management techniques and the development of smart cities are both aided by this study. Two CNNs were trained using the AlexNet architecture; one was trained to detect garbage in images, and the other to distinguish between landfill rubbish and recyclables. To demonstrate the idea, the two-stage CNN system was first trained and evaluated on the industry-standard TrashNet indoor picture dataset, where it obtained impressive results. The authors then trained and evaluated the system using outside photographs captured in the actual setting of use. The first convolutional neural network (CNN) was able to obtain a preliminary accuracy of 93.6% in identifying garbage and non-trash objects on a picture database of various types of rubbish using the outdoor image dataset. Accuracy levels ranged from 89.7 percent to 93.4 percent, with a total of 92.2 percent, achieved by training a second CNN to differentiate between landfill waste and recyclables. The eventual aim is to include this garbage recognition system that is based on image processing into a smart garbage can robot that can identify and gather garbage in its immediate vicinity by taking pictures in real-time. Cullowhee, North Carolina's Yanjun Yan School of Engineering and Technology Dear Sir/Madam, unwelcoming to strangers, and particularly harmful to children and the elderly. In industrialized nations, inhabitants face a major hygienic, aesthetic, and environmental protection issue with garbage and refuse that goes uncollected along roads and other locations. Worldwide, "unsafe water supply, sanitation and hygiene" are responsible for an estimated 842,000 fatalities annually, according to the World Health Organization [1]. Under five-year-olds make up 361,000 of them, with the majority residing in nations with low per capita income. Trash collection costs a pretty penny in both wealthy and developing nations,

but automatic methods will cut that cost in half while also boosting public health. For instance, according to CBSNewYork [2], the annual cost of garbage collection in New York City is \$300 million.

Keywords—CNN, AlexNet, Image Classification, Deep Learning, Object detection.

I. INTRODUCTION

When a city is clean and well-maintained, its residents enjoy it the most. It is becoming increasingly difficult to keep cities clean in this period of rapidly expanding populations because more and more people are relocating to urban areas, which results in an increase in garbage production. Its difficulty is readily apparent when we consider the South Asian nations. Developing nations, which are home to the vast majority of the world's population, lack the resources to implement and sustain an effective waste management system, in contrast to first-world nations that have long since perfected this art. For that reason, waste management has been an important concern on a global scale. The majority of underdeveloped nations often have garbage cans that are overflowing. People in these nations also have a habit of throwing their garbage outside the can instead of inside. Germs thrive in the soil and debris around the garbage can. This is horribly unsanitary and uncomfortable. People are unwelcoming to strangers, and it's particularly harmful for children and the elderly to go by a garbage can in such a state. In industrialized nations, inhabitants face a major hygienic, aesthetic, and environmental protection issue with garbage and refuse that goes uncollected along roads and other locations. Worldwide, "unsafe water supply, sanitation and hygiene" are responsible for an estimated 842,000 fatalities annually, according to the World Health Organization [1]. Under five-year-olds make up 361,000 of them, with the majority residing in nations with low per capita income. Trash collection costs a pretty penny in both wealthy and

developing nations, but automatic methods will cut that cost in half while also boosting public health. For instance, according to CBSNewYork [2], the annual cost of garbage collection in New York City is \$300 million. Unprecedented advancements in computer vision have been substantially facilitated by recent developments in deep learning research. Among the most effective deep-learning algorithms, convolutional neural networks (CNNs) find widespread use in picture categorization, segmentation, and detection [3-6]. Thus, convolutional neural networks (CNNs) are suggested for garbage detection and identification in this work. To sort garbage that people throw out in public places in cities, Chu et al. [7] suggested a multilayer hybrid deep-learning system (MHS). Trash cans and other garbage can contents can be automatically sorted by the system. In order to extract important picture characteristics, they used the AlexNet CNN [3], while optical sensors were employed to identify additional numerical feature information. To categorize the rubbish item, this system used data from many sources using multilayer perceptrons (MLP). Although the suggested MHS managed an average accuracy rate of above 90%, it is limited to the classification of 22 fixed items of public area waste. In their method, other kinds of litter on the road or in a park would not be included. A garbage collection robot that can correctly and independently locate rubbish on the lawn was demonstrated by Bai et al. [8]. For garbage detection and robot navigation, they used a ResNet [9] deep neural network. The robot can autonomously clean parks and schools of waste thanks to its garbage detection and navigational capabilities. Their accuracy in identifying garbage exceeded 95%. However, the robot is limited to detecting waste on grass. This meant the robot couldn't pick up litter in parking lots or on roads.

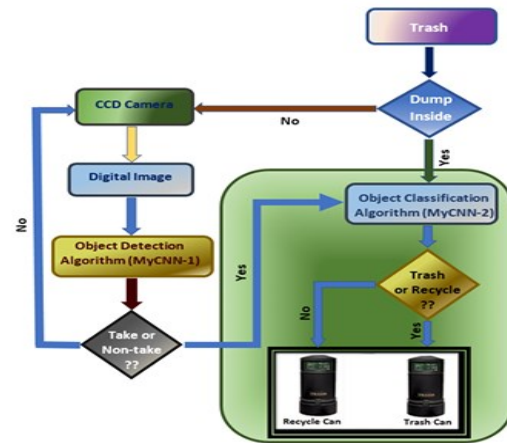


Figure 1: Flow Chart of the Decision-Making Process

II. APPROACH (THEORY)

An approach used in Deep Learning is a Convolutional Neural Network, abbreviated as CNN. It starts with an input picture and uses it to identify features and objects in the image by assigning them learnable weights and biases. In order to identify and categorize objects in digital photos, other researchers have used convolutional neural networks (CNN) [3-6]. Following training, convolutional neural networks (CNNs) may filter and classify pictures into multiple classes based on the input-output pairings, replacing rudimentary approaches like filtering by a hand-engineered procedure. A convolutional neural network (CNN) is designed in a way that mimics the way neurons in the human brain are connected inside. It consists of normalizing layers, fully connected layers, pooling layers, and convolutional layers [3-6]. By convolving the input picture with moving kernels that have a certain window size and stride size, each kernel in the convolutional layers creates a feature map. In order to prevent gradient vanishing, the output is processed using the rectifier linear unit (ReLU). Subsequently, pooling is used to decrease noise levels and feature dimensions. Following a series of convolutional layers and pooling, the features are smoothed out before being input into the fully connected layers. These layers are composed of sets of artificial neurons, or nodes, arranged in columns, and the output of each layer's activation neuron is connected to the input of every layer below it. The following equations represent the completely linked function in mathematical terms, according to the laws of forward pass and backward pass propagation [9]. With $j=1, 2, 3, \dots, n1$ and $n1=n$

being the dimensionality of x , the components of the input vector x are the outputs of layer 1, which may be written as $(1) = xb$. Neuron i in layer l 's computation is provided by

$$z_i(l) = \sum_{j=1}^{n_{l-1}} W_{ij}(l) a_j(l-1) + b_i(l) \quad (1)$$

The net input to neuron i in layer l , which is produced utilizing all outputs from layer $l-1$, is represented as (l) , where $i=1,2,3,\dots,n_l$ and $l=2,\dots,L$. The bias value linked to the i th neuron in the l th layer is denoted as (l) . If $i=1,2,3,\dots,n_l$, then the activation value of neuron i in layer l is $ai(l) = h(zi(l))$. Here, h is an activation function. For any integer i from 1 to n_O , the value of the network output node i is $(L) = h((L))$. In a fully linked feedforward network, these are the only operations needed to convert input to output. Any neuron's output and net input in any layer (apart from the first layer's $j(1)$) is directly related to each other, and the same holds true for any hidden layer's node j .

$$\delta_j(l) = \frac{\partial E}{\partial z_j(l)} \quad (2)$$

where the j -th neuron's mistake is denoted by E . $j(l+1)$ as we will be working backward in the network. δ between $j(L-2)$, and ultimately reach layer 2, we need the relationship $j(L)$ and $\delta_j(l)$ and δ in order to begin with find is δ a possible way to describe the equation $j(L-1)$,

$$\delta_j(l) = h'(z_j(l)) \sum_i W_{ij}(l+1) \delta_i(l+1) \quad (3)$$

Through some algebra it can be shown that the rate of change of error with respect to network weights is:

$$\frac{\partial E}{\partial w_{ij}(l)} = a_j(l-1) \delta_i(l) \quad (4)$$

Similarly, the rate of change of error with respect to biases is:

$$\frac{\partial E}{\partial b_i(l)} = \delta_i(l) \quad (5)$$

These rates of change are used in the backpropagation model to update the weights and biases:

$$w_{ij}(l) = w_{ij}(l) - \alpha a_j(l-1) \delta_i(l) \quad (6)$$

$$b_i(l) = b_i(l) - \alpha \delta_i(l) \quad (7)$$

Tuning dimensional parameters and local architectural structure may improve CNN accuracy [8]. In recent years, a wide variety of convolutional neural network (CNN) designs have appeared [3]. The in-field processing capabilities and cheap computing cost of AlexNet [3] make it an ideal choice for our study. It was during the 2012 ImageNet Challenge (ILSVRC) where AlexNet [3] was first shown. The top-5 mistake rate for picture categorization was down from 26% to 15.3% thanks to it. Its very capable architecture has made it well-known. The MATLAB® version of AlexNet, which has 25 layers (five convolutional and three fully connected) as described in Table 1, was employed in this work. The necessary picture characteristics are extracted using a number of Convolutional Kernels. In a convolutional layer, several identically sized kernels are used. The 1000-way softmax function, which corresponds to 1000 class labels, is fed the output of the final Fully Connected Layer. The fourth and eighth levels are linked to the Cross-Channel Normalization layer. Layers for max-pooling are stacked after the sixteenth and cross-channel normalization layers. Each convolutional layer is followed by the ReLU nonlinearity. Every one of the 4096 neurons in a fully connected layer is linked to every single neuron in the layer below it [3]. As we saw in Section III, we set the number of neurons in the final Fully Connected Layer (layer 23) to the number of categories (02).

Table 1: MATLAB AlexNet Layer Configuration

Layer	Type
1	Data (227x227x3 Size Images)
2	96 kernels of size 11x11x3 Convolutions
3	ReLU
4	Cross Channel Normalization
5	3x3 Max Pooling
6	256 kernels of size 5x5x48 Convolutions
7	ReLU
8	Cross Channel Normalization
9	3x3 Max Pooling
10	384 kernels of size 3x3x256 Convolutions

11	ReLU
12	384 kernels of size 3x3x192 Convolutions
13	ReLU
14	256 kernels of size 3x3x192 Convolutions
15	ReLU
16	3x3 Max Pooling
17	4096 Fully Connected Layer
18	ReLU
19	50% Dropout
20	4096 Fully Connected Layer
21	ReLU
22	50% Dropout
23	1000 Fully Connected Layer
24	Softmax
25	Classification Output

III. METHOD AND RESULTS

We created a series of steps to teach a CNN to sort things into three categories: garbage, recyclable, and other. For this smart robot garbage can, we trained it using AlexNet CNN architecture and public space take or not capture photographs so it can determine whether to seize an item or not. In particular, "take" indicates that the object in question is garbage and should be taken, but "non-take" indicates that the thing in question is not trash and should not be taken. Here is a rundown of the four tests that we taught AlexNet to do: 1. AlexNet was trained using photos from TrashNet [11] and the CNN that was produced was evaluated using a subset of TrashNet images classified into five groups: metal, plastic, glass, paper, and cardboard. 2. Put the identical CNN through its paces with an indoor camera, focussing on garbage cans in real time. 3. Trained AlexNet to distinguish between "take" and "non-take" based on outside photographs. I used a subset of outside photos to test the CNN that was generated. 4. Used outside photos to train AlexNet to distinguish between recyclables and rubbish for the landfill. I used a subset of outside photos to test the CNN that was generated. To ensure the correctness of the AlexNet CNNs, testing 1 and 2 served as preliminary testing. For our method development in MATLAB, we used a publically accessible CNN known as Deep Learning Toolbox Model for AlexNet Network [12]. We adjusted the amount of neurons in the final Fully Connected Layer of the AlexNet design to meet our needs. In addition, to train the CNN, we extracted over 2000 photographs from a publicly accessible

database called TrashNet [11] that included rubbish photos captured inside and classified them into five groups: metal, plastic, glass, paper, and cardboard. Initial training results on TrashNet indoor photos accurately demonstrated CNN's suitability for this task. The finished garbage robot design might undergo tests 3 and 4, which are practical testing. We shot all of the outside "take" and "non take" photos in and around human habitation zones on a university campus. All of the images used to teach the CNN are actual rubbish from our neighborhood. To complete the second job, we used the pictures of the "take" goods to train an additional AlexNet CNN that could further distinguish between recyclable and garbage products for the landfill. What follows is a full description of the steps to take in order to complete each of these exams. First Experiment: Five Groups in a Managed Indoor Environment Using MATLAB's Convolutional Neural Network (CNN) implementation and deep learning techniques, AlexNet was trained to distinguish between five picture categories: metal, plastic, glass, paper, and cardboard. Applying our trained AlexNet model to a subset of the TrashNet photos, we evaluated its accuracy. Table 2 displays the results. All five categories have detection accuracy levels over 80%. All of the photos included for this test were shot in a carefully controlled indoor setting with uniform lighting. That provides some context for the high accuracy of the data shown above. Figure 2 displays several instances of the test photographs.

Table 2: Results of CNN classification using TrashNet indoor images

Category	Total count of images	Count of correctly detected images	Accuracy (%)
Metal	41	39	91.68
Plastic	48	38	81.25
Paper	59	53	89.83
Cardboard	40	37	92.5
Glass	50	46	92
Overall	238	213	89.50

ISSN 2454-5007,

www.ijmm.in

Vol. 13 Issue. 2, 2025

Category: Glass**Category: Metal**

(a) Example of training images used as metal



(e) Example of training images used as glass

Category: Plastic

(b) Example of training images used as plastic

Category: Paper

(c) Example of training images used as paper

Category: Cardboard

(d) Example of training images used as cardboard

Figure 2. Examples of training images used to detect five categories of recyclable materials.

Test 2 – Five Categories using a Real-Time Camera

Going on to the next stop, we arrived at an inside workplace with a camera trained on a white display board backdrop, photographing thirteen distinct items. To get a total of 260 photos, we rotated each item to get it from 10 different perspectives. Figure 3 shows a series of ten photos that were used to evaluate a single item. In Test 1, we taught AlexNet to classify images into one of five categories: metal, plastic, glass, paper, and cardboard. We then put each picture through its paces. Table 3 displays a synopsis of the test findings. The trained AlexNet achieved an accuracy of 90% or higher when identifying 7 of the 13 things, and an accuracy of 70% to 80% when identifying 3 of the 13 objects. Three items caused the CNN some trouble: orange plastic box, clear glass, and brown paper. A mix-up between cardboard and plastic led to the brown paper being labelled as 70% cardboard and 30% plastic. The fact that the CNN cannot see the difference between paper and cardboard because of their similar colors is one probable explanation. The CNN showed that there was 60% glass and 40% plastic when it came to clear glass. One possible explanation is that plastic and glass both reflect light in the same way. A transparent plastic cup also displayed the same issue. An example of the real-time picture with the identified item labeled in the figure title is provided by the MATLAB software that implemented the real-time detection. Images of the thirteen items used to train the convolutional neural network (CNN) are shown in figure 4. Here are a few examples of the outcomes. As an example, the plastic bag in Figure 4 object 1 is appropriately classified as plastic. Figure 4 displays a number of items. A few have been accurately named, but others have not.

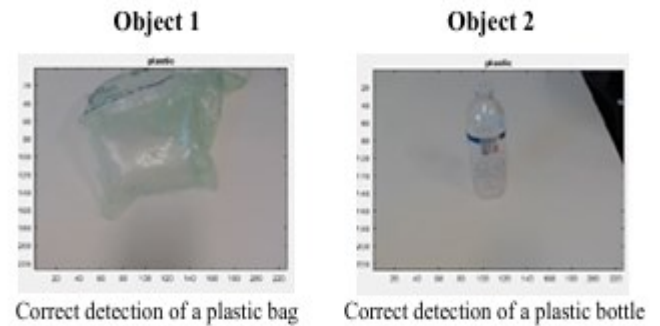


Figure 3: One full test procedure for an object

Table 3: Results of AlexNet CNN classification using indoor camera, tra TrashNet images

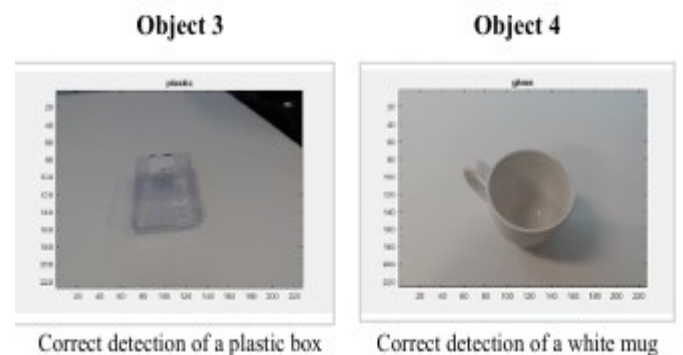
Object number	Object	Actual Category	Detected Category
1	Green Plastic bag	Plastic	100% Plastic
2	Plastic bottle	Plastic	100% Plastic
3	Plastic box	Plastic	100% Plastic
4	White glass mug	Glass	100% Glass
5	Red Plastic cup	Plastic	80% Plastic, 20% n
6	Brown paper	Paper	70% cardboard, 30%
7	White paper with writing	Paper	70% Paper, 30% card

8	Clear glass	glass	60% glass, 40% pl
9	green soda can	metal	90% metal, 10% pl
10	clear plastic cup	plastic	70% plastic, 30% g
11	3D printed green object	plastic	100% plastic
12	Box	Cardboard	100% Cardboar
13	Orange plastic box	Plastic	20% plastic, 80% pa cardboard



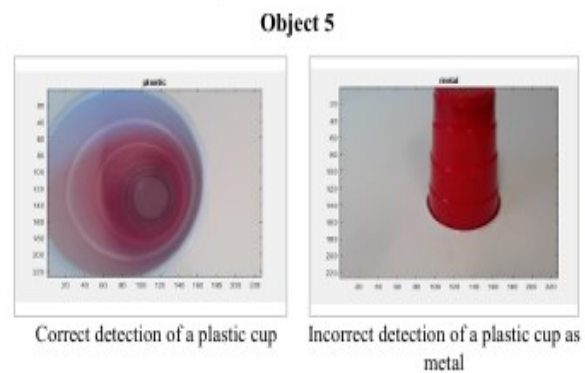
Correct detection of a plastic bag

Correct detection of a plastic bottle



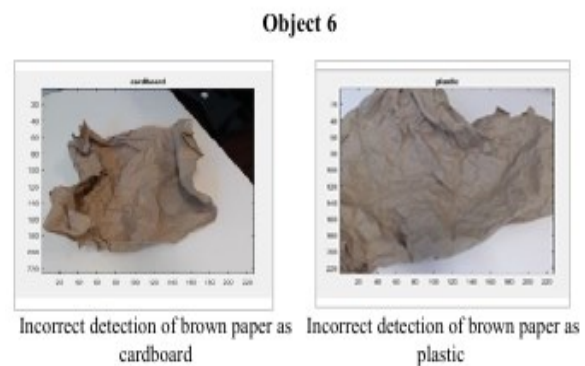
Correct detection of a plastic box

Correct detection of a white mug



Correct detection of a plastic cup

Incorrect detection of a plastic cup as metal



Incorrect detection of brown paper as cardboard

Incorrect detection of brown paper as plastic

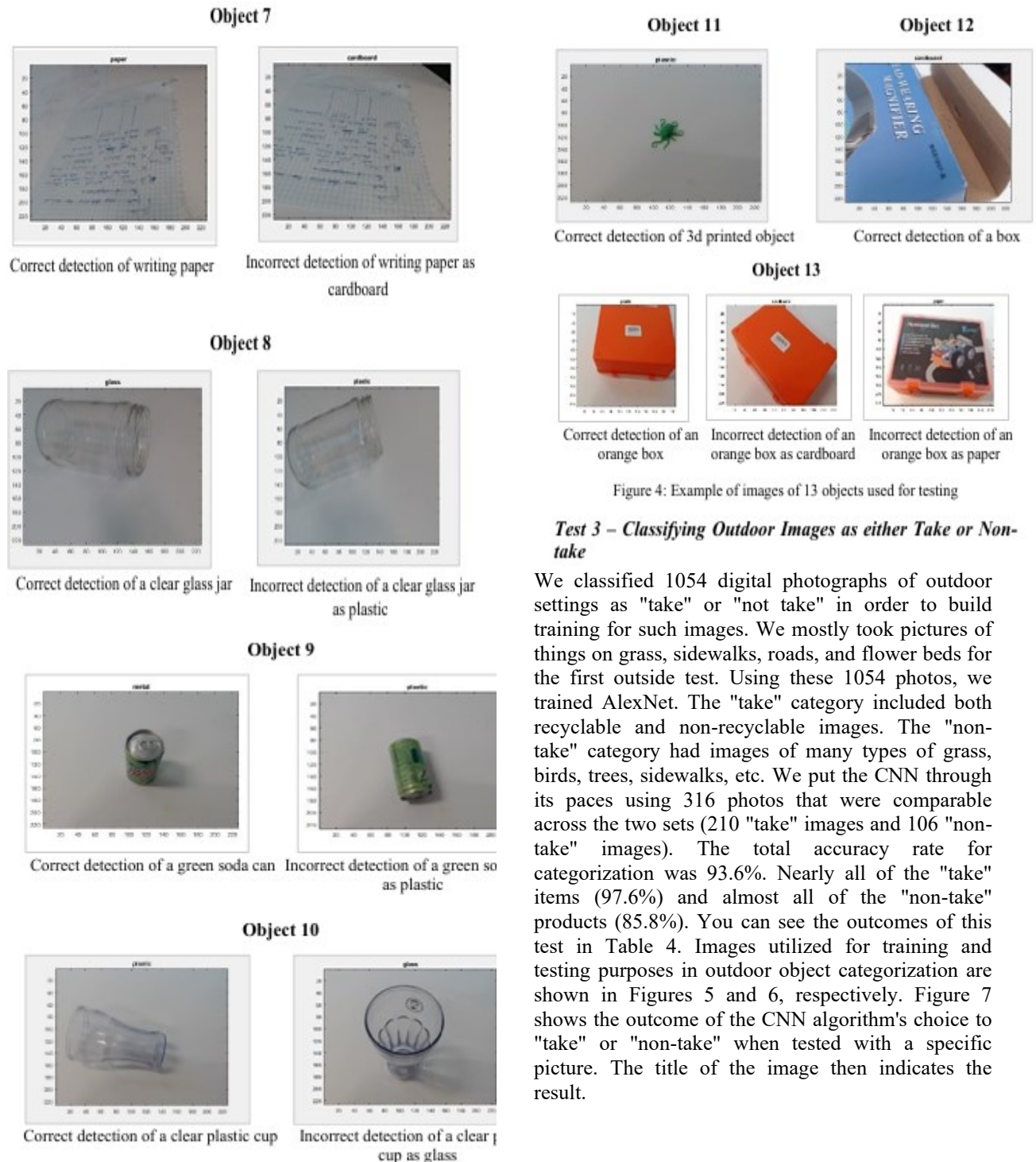


Figure 4: Example of images of 13 objects used for testing

Test 3 – Classifying Outdoor Images as either Take or Non-take

We classified 1054 digital photographs of outdoor settings as "take" or "not take" in order to build training for such images. We mostly took pictures of things on grass, sidewalks, roads, and flower beds for the first outside test. Using these 1054 photos, we trained AlexNet. The "take" category included both recyclable and non-recyclable images. The "non-take" category had images of many types of grass, birds, trees, sidewalks, etc. We put the CNN through its paces using 316 photos that were comparable across the two sets (210 "take" images and 106 "non-take" images). The total accuracy rate for categorization was 93.6%. Nearly all of the "take" items (97.6%) and almost all of the "non-take" products (85.8%). You can see the outcomes of this test in Table 4. Images utilized for training and testing purposes in outdoor object categorization are shown in Figures 5 and 6, respectively. Figure 7 shows the outcome of the CNN algorithm's choice to "take" or "non-take" when tested with a specific picture. The title of the image then indicates the result.

Table 4: Results of CNN classification using outside images with 2 categories

Category	Total count of images	Count of correctly detected images	Accuracy (%)
"take"	210	205	97.6
"non-take"	106	91	85.9
Overall	316	296	93.6



Figure 5: Example of training images for outdoor object classification



Figure 6: Example of test images for outdoor object classification

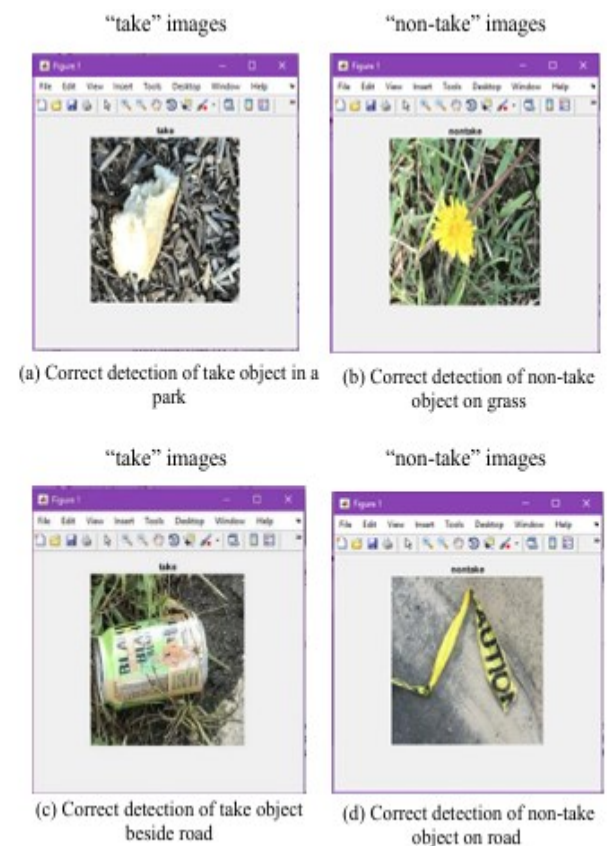


Figure 7: Sample output of test images for outdoor object detection

The fourth test is to determine if outside images are recyclable or trash. Afterwards, we separated the "take" picture database into the "trash" and "recyclable" sections. Using 700 outdoor photos from the "take" category, we trained an additional AlexNet CNN. Afterwards, we put the CNN through its paces on 175 comparable pictures sourced from the "trash"

and "recycle" categories. As previously demonstrated in Table 5, the findings were correct. The total accuracy rate for categorization was 92%. Correct identification of "Recycle" goods was 89.7 percent, while "Trash" items were 93.5 percent accurate. Several examples of output graphics with the choice "trash" or "recycle" specified in the title are shown in Figure 8.

Table 5: Results of CNN classification using outside images with 2 categories

Category	Total count of images	Count of correctly detected images	Accuracy (%)
"recycle"	68	61	89.7
"trash"	107	100	93.5
Overall	175	161	92

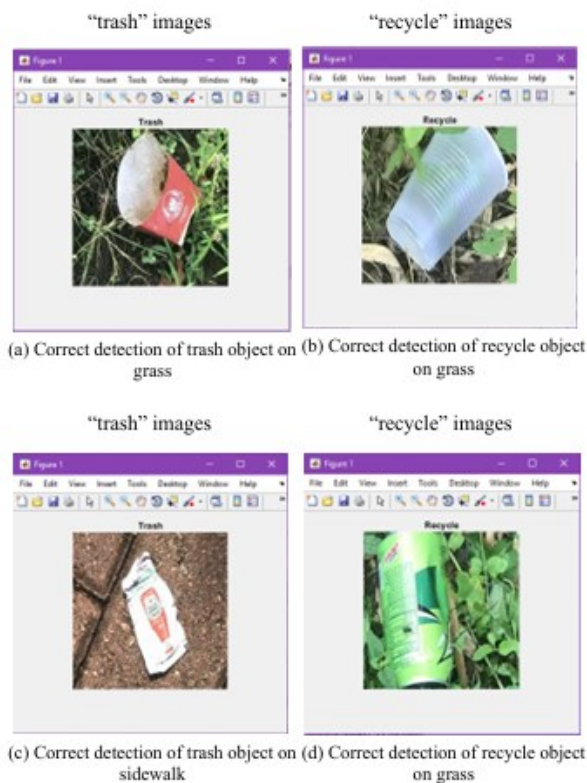


Figure 8: Sample output of test images for outdoor object classification

IV. CONCLUSION

In this work, we aimed to design an autonomous garbage collection system, therefore we built a CNN-based algorithm to recognize trash and non-trash, and to further distinguish between landfill and recyclable

materials in the waste category. The results showed a good trend, with a detection accuracy between 89.7 and 93.5 percent. Compared to the current methods utilized by road sweeper trucks or vacuum cleaners, smart trash cans that include this categorization based on image processing are more suited to cleaning public places of rubbish. The results of the experiments shown that the suggested algorithm is capable of properly distinguishing between trash and recycled materials. This algorithm may be a useful resource for creating a garbage can robot that can mow a large grass at a school or park. In the future, we want to implement our two-stage trained CNN into an algorithm that can communicate with a microcontroller and a camera to guide a garbage can robot around a public area, detect objects on the ground, and collect recyclables and trash accordingly.

REFERENCES

- [1] World Health Organization Website, https://www.who.int/water_sanitation_health/diseases-risks/diseases/en/
- [2] CBSNewYork, <https://newyork.cbslocal.com/2018/01/29/nyc-garbage-pickup-charges/>
- [3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proceedings of Neural Information Processing System Conference, Lake Tahoe, CA, USA, December 2012.
- [4] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. Lecun, "OverFeat: integrated recognition, localization and detection using convolutional networks," in Proceedings of International Conference on Learning Representations, Banff, Canada, April 2014.
- [5] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional neural networks," in Proceedings of 13th European Conference on Computer Vision (ECCV), Zurich, Switzerland, September 2014.
- [6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of Conference on Computer Vision and Pattern Recognition, Las Vegas Valley, NV, USA, June 2016.
- [7] Chu, Y., Huang, C., Xie, X., Tan, B., Kamal, S., & Xiong, X. (2018). "Multilayer Hybrid Deep-Learning Method for Waste Classification and Recycling," Computational Intelligence and Neuroscience, 2018.

- [8] J. Bai, S. Lian, Z. Liu, K. Wang and D. Liu, "Deep Learning Based Robot for Automatically Picking Up Garbage on the Grass," in IEEE Transactions on Consumer Electronics, vol. 64, no. 3, pp. 382-389, Aug. 2018.
- [9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, 2016, pp. 770–778.
- [10] Gonzales, R.C. and Woods, R.E., "Digital Image Processing," 4th edition, Pearson, 2018.
- [11] Trashnet, Dataset of images of trash, <https://github.com/garythung/trashnet>.
- [12] AlexNet Tool Box, Mathworks, <https://www.mathworks.com/matlabcentral/fileexchange/59133-deep-learning-toolbox-model-for-alexnet-network>.