



ISSN: 2454-9940



**INTERNATIONAL JOURNAL OF APPLIED
SCIENCE ENGINEERING AND MANAGEMENT**

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

Convolutional Neural Networks for the Identification of Waste and Recycled Materials

¹Mrs. G. Swarnalata, ²Nellimarla Manichandana, ³Nalla Vineela, ⁴Sivala Vasavi,

¹Associate Professor, Dept. of CSE, Rajamahendri Institute of Engineering & Technology,
Bhoopalapatnam, Near Pidimgoyyi, Rajahmundry, E.G. Dist. A.P. 533107.

^{2,3,4} Students, Dept. of CSE, Rajamahendri Institute of Engineering & Technology,
Bhoopalapatnam, Near Pidimgoyyi, Rajahmundry, E.G. Dist. A.P. 533107.

Abstract—

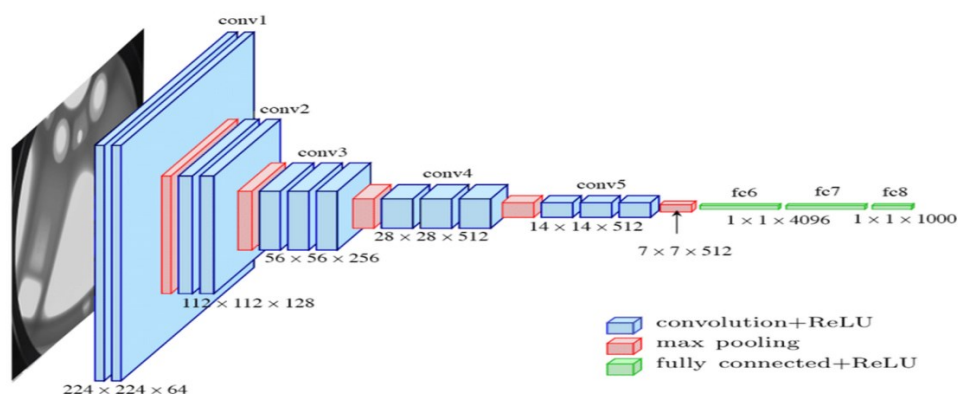
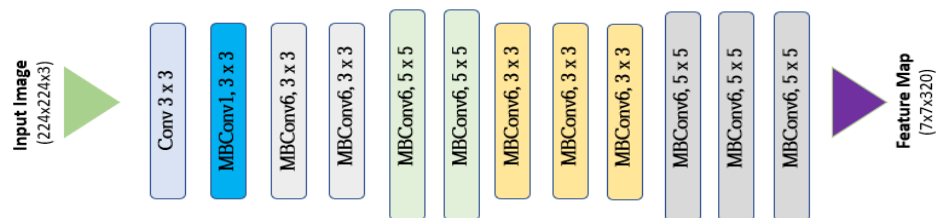
The aim of this research is to improve municipal trash collection using image processing algorithms and deep learning technologies for detecting trash in public spaces. This research will help to improve trash management systems and help to create a smart city. Two Convolutional Neural Networks (CNN) were developed to search for trash objects in an image and separate recyclable items from the landfill trash objects, respectively. The two-stage CNN system was first trained and tested on the benchmark TrashNet indoor image dataset and achieved great performance to prove the concept. Then the system was trained and tested on outdoor images taken by the authors in the intended usage environment. Using the outdoor image dataset, the first CNN achieved a preliminary 93.6% accuracy to identify trash and nontrash items on an image database of assorted trash items. A second CNN was then trained to distinguish trash that will go to a landfill from the recyclable items with an accuracy ranging from 89.7% to 93.4% and overall 92%. A future goal is to integrate this image processing-based trash identification system in a smart trash can robot with a camera to take real-time photos that can detect and collect the trash all around it.

INTRODUCTION

People prefer living in a city the most when it is clean and well-maintained. However, in this period of rapidly expanding populations, more and more people are settling into urban areas, leading to an increase in waste production and making it more challenging than ever to keep cities clean. Its difficulty is readily apparent when we consider the South Asian nations. The bulk of the world's population lives in developing nations, which are ill-equipped to handle waste management, in contrast to first-world nations that have a robust system in place thanks to their ample financial resources. Which is why waste management has been an important topic all across the globe. In the majority of poor nations, garbage cans are often seen overflowing. People in these nations also have a habit of throwing their garbage outside the can instead of inside. Germs thrive in the filth that surrounds garbage cans. This is highly uncomfortable and unsanitary. Seeing a garbage can in such a state is clearly not pleasant for anybody, but it is particularly hazardous for children and the elderly, and it sends the wrong message to potential visitors. 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. An estimated 842,000 people die each year as a result of "unsafe water supply, sanitation and hygiene," 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. The expense of garbage collection is substantial in both rich and developing nations; however, automatic waste collection systems will not only improve public health but also significantly lower this cost. For instance, according to CBSNewYork [2], the annual cost of garbage collection in New York City is \$300 million. Unparalleled advancements in computer vision have been considerably facilitated by recent advances 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]. So, to accomplish garbage detection and identification, CNN is suggested in this article. A multilayer hybrid deep-learning system (MHS) was suggested by Chu et al. [7] to sort garbage that people throw out in public spaces in cities. Trash cans and other garbage can contents can be automatically sorted by the system. They used optical sensors to identify additional numerical feature information and a CNN [3] to extract important picture characteristics. Through the consolidation of data acquired from several sources, this system classified the waste item using multilayer perceptrons (MLP). Although the system can only identify 22 fixed pieces of public area garbage, the suggested MHS nonetheless managed to obtain a mean

accuracy greater than 90%. They would not include other types of litter in their system, whether it be on the road or in a park. An autonomous garbage collection robot with good grass-detection capabilities was introduced by Bai et al. [8]. To recognize garbage and direct the robot's movements, they used a navigation technique using a deep neural network called ResNet [9]. The robot can autonomously clean parks and schools of waste thanks to its garbage detection and navigational capabilities. Their accuracy in rubbish identification surpassed 95%. Grass is the only surface that the robot can detect garbage on. Consequently, the robot was unable to distinguish between different types of litter in public spaces. Both of the aforementioned studies used CNN architectures to great effect, yielding very accurate results. Our proposed system is based on these papers and can detect rubbish in any public place, including streets, parking lots, parks, leisure areas, community centers, etc. Our long-term objective is to construct a garbage-collecting robot that can autonomously explore a public area in search of dropped items. Object recognition from photographs is the end objective of this work. You can see the decision-making process shown in Figure 1. When passers-by toss items into the garbage robot, it sorts them into two inner bins according to whether they are recyclable or rubbish. Aside from a garbage can, the robot will also have a camera that it can use to take pictures of objects and determine whether or not to take them. It will collect everything it determines to be recyclable or garbage. The garbage robot would take an item it has picked up and bring it inside for a closer look. Once the robot has a better look at the item, it can decide if it's garbage or recyclable. Metal, plastic, glass, and fiber are the four categories into which the robot will sort recyclables. Fiber encompasses all things made of paper or cardboard. In this study, we taught a CNN [3] to distinguish between "take" and "non-take" photos, and then to sort them into "trash" and "recyclable" categories. Results from testing these CNNs with actual outdoor photographs taken in public places were surprisingly accurate.

EfficientNet Architecture



VGG 16 - ARCHITECTURE

II.APPROACH (THEORY)

An approach used in Deep Learning is a Convolutional Neural Network, abbreviated as ConvNet or CNN. It begins with an input picture, then uses learnable biases and weights to distinguish between distinct characteristics and objects within the image. Other researchers have used convolutional neural networks (CNN) to classify or identify objects in digital photos [3-6]. After training, convolutional neural networks (CNNs) may filter/classify pictures into multiple classes based on the input-output pairings, replacing crude approaches like filtering with a hand-engineered procedure. Convolutional neural networks (CNNs) are designed to mimic the way neurons in the human brain communicate with one another. Layers for normalizing, pooling, complete connectedness, and convolution are all part of it [3-6]. Every kernel in the convolutional layers creates a feature map by passing the input picture through a convolving process using a moving kernel that has a certain window size and stride size. After applying the rectifier linear unit (ReLU) to the output, which prevents the gradient from disappearing, we use the pooling algorithm to lower the feature dimensions and noise. Following the use of pooling and numerous convolutional layers, 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 is connected to the input of the 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]. The outputs of layer 1 may be represented as $(1)=xb$, where $j=1, 2, 3, \dots, n1$, and $n1=n$ is the dimensionality of x . These components make up the input vector x . The output of neuron i 's calculation in layer 1 is provided by

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

This is created by adding together all the outputs from layer $l-1$ and represents the net input to neuron i in layer l , where $i=1, 2, 3, \dots, n_l$ and $l=2, \dots, L$. A neuron in the i th layer is linked to a bias value denoted as (l) . With $i=1, 2, 3, \dots, n_l$, the activation value of neuron i in layer l may be expressed as $a_i(l)=h(z_i(l))$. For all integers i from 1 to n_l , the value of the network output node i is $a(L)=h(z_i(L))$. In a fully linked feedforward network, these are the only operations needed to convert input to output. $j(l)$ for any node j in any hidden layer. δ Any neuron in any layer (apart from the first layer) has the same relationship between its net input and output, which may be represented as $j(l)$ is as δ One way to represent

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

E represents the j th neuron's mistake. $j(l+1)$, $\delta_j(l)$ and $\delta_j(l-2)$, and ultimately reach layer 2 in the network, we need the relationship between $\delta_j(l-1)$, $\delta_j(l)$, locate δ To begin with One way to represent the equation is as follows:

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

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 number of CNN designs with varying degrees of complexity have surfaced [3]. It had a considerable impact, lowering the top-5 mistake rate in picture categorization from 26% to 15.3%. Its very capable architecture has made it well-known.

III. METHOD AND RESULTS

To train a Convolutional Neural Network (CNN) to distinguish between recyclable, garbage, and other types of things, we came up with a set of approaches. In order to teach a smart robot garbage can to determine whether to seize an item or not, we used a CNN architecture and public space take or not capture photographs. 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. This is how we trained our CNN to complete a battery of four tests: 1. We used TrashNet photos for training [11]. Then, we evaluated the CNN using a subset of TrashNet images categorized as follows: metal, plastic, glass, paper, and cardboard. 2. Put the identical CNN through its paces with an indoor camera, focusing on garbage cans in real time. 3. Used outside photos to train the CNN to distinguish between "take" and "non-take" scenarios. 4. Used outside photos to train the CNN to distinguish between recyclables and rubbish for the landfill.

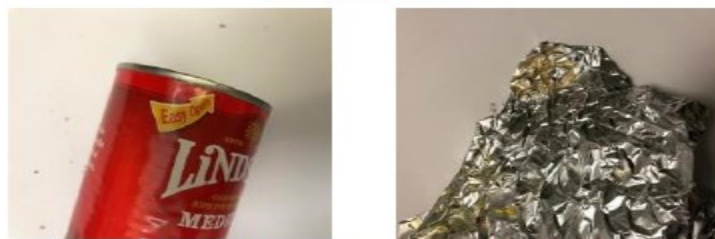
The correctness of the CNNs was confirmed in the early tests 1 and 2. To build an algorithm in MATLAB, we used a publicly accessible CNN model available through the Deep Learning Toolbox [12]. We adjusted the amount of neurons in the final Fully Connected Layer of the network design to meet our needs. 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 the CNN's suitability for this task. The finished garbage robot design might undergo tests 3 and 4, which are practical testing. All of the outside "take" and "non-take" photos were shot by us in the immediate vicinity of residential areas on a university campus. The CNN is trained using real-life waste scenarios in our neighborhood. Second, we used the "take" item photos to train an additional CNN to further sort the objects into recyclables and rubbish for the landfill.

Below, you will find a thorough description of the protocol for each of these tests. First Test—Five Groups in a Managed Indoor Environment For the main test, we trained the CNN in the five picture categories (metal, plastic, glass, paper, and cardboard) using deep learning techniques for MATLAB's Convolutional Neural Networks (CNN). Using a selection of the TrashNet photos, we evaluated the efficacy of our trained CNN model. The findings may be seen in Table 2. Detection accuracy was more than 80% across all five groups. It should be mentioned that the test photographs were captured in a controlled indoor setting with uniform illumination. That provides some context for the high accuracy of the data shown above. Figure 2 displays a number of screenshots that were used throughout the test.

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

Category: Metal



(a) Example of training images used as metal

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

Category: Glass



(e) Example of training images used as glass

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

Test 2 – Five Categories using a Real-Time Camera

Next, we went to an inside workplace with a camera that was focusing on a white display board backdrop, taking pictures of thirteen different things. A total of 260 photos were obtained by rotating each item to capture it from 10 distinct points of view. Figure 3 shows a collection of ten photos used to evaluate a single item. We ran each picture through our trained CNN model (from Test 1), which sorted them into one of five groups: metal, plastic, glass, paper, and cardboard. Table 3 displays a synopsis of the test findings. Seven of the thirteen things were accurately identified

by the trained CNN with a 90% or greater accuracy, while three of the thirteen objects were successfully identified with an accuracy of 70% to 80%. Three items—brown paper, transparent glass, and an orange plastic container—caused the CNN some trouble. The brown paper was mistaken for plastic and cardboard, leading to a 70% cardboard and 30% plastic classification. One probable explanation is that the CNN cannot tell the difference between paper and cardboard because of the similarity in color perception. The CNN showed that there was 60% glass and 40% plastic when it came to clear glass. Light reflected off of plastic and glass might be the culprit. A transparent plastic cup was also experiencing the same issue. The real-time detection is implemented in the MATLAB software. The figure title lists the identified item in the real-time picture. Figure 4 displays the recorded pictures of the thirteen items used to evaluate the convolutional neural network (CNN). These graphs illustrate a few representative 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, trained on 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% metal
6	Brown paper	Paper	70% cardboard, 30% plastic
7	White paper with writing	Paper	70% Paper, 30% cardboard

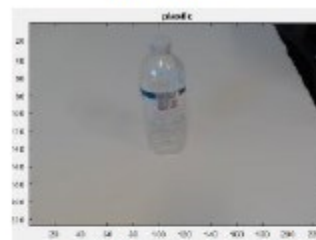
8	Clear glass	glass	60% glass, 40% plastic
9	green soda can	metal	90% metal, 10% plastic
10	clear plastic cup	plastic	70% plastic, 30% glass
11	3D printed green object	plastic	100% plastic
12	Box	Cardboard	100% Cardboard
13	Orange plastic box	Plastic	20% plastic, 80% paper or cardboard

Object 1



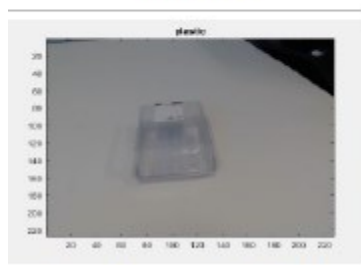
Correct detection of a plastic bag

Object 2



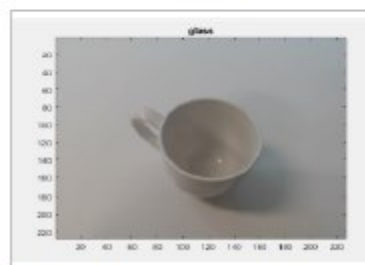
Correct detection of a plastic bottle

Object 3



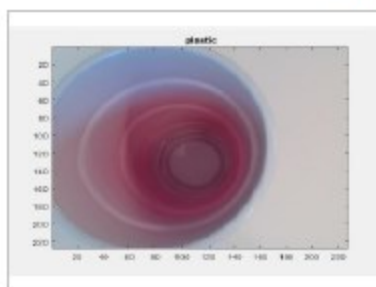
Correct detection of a plastic box

Object 4

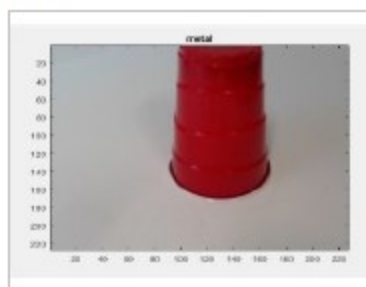


Correct detection of a white mug

Object 5

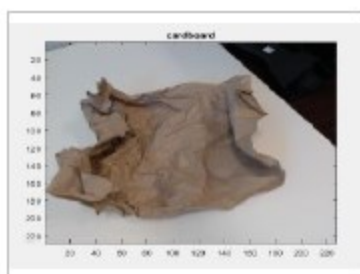


Correct detection of a plastic cup

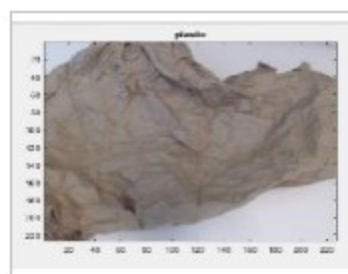


Incorrect detection of a plastic cup as metal

Object 6

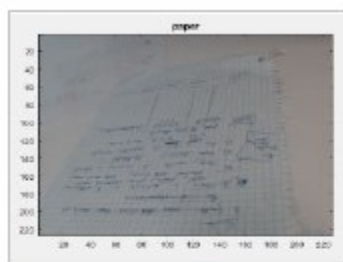


Incorrect detection of brown paper as cardboard

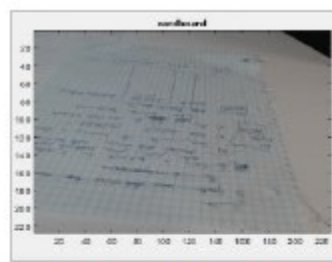


Incorrect detection of brown paper as plastic

Object 7

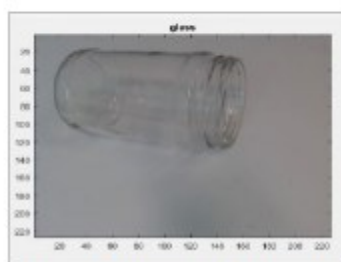


Correct detection of writing paper

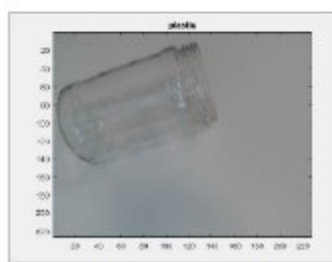


Incorrect detection of writing paper as cardboard

Object 8

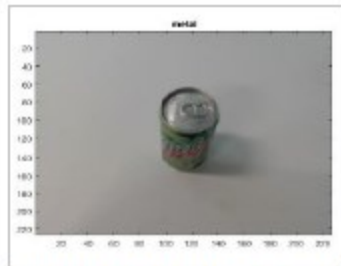


Correct detection of a clear glass jar

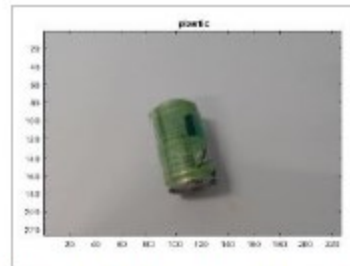


Incorrect detection of a clear glass jar as plastic

Object 9

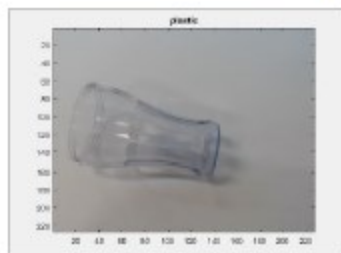


Correct detection of a green soda can

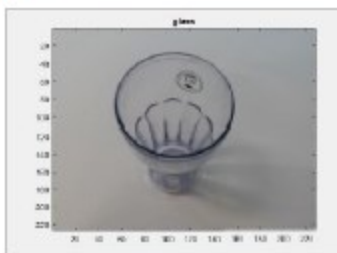


Incorrect detection of a green soda can
as plastic

Object 10

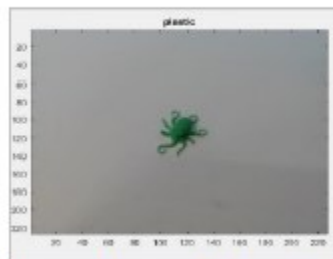


Correct detection of a clear plastic cup



Incorrect detection of a clear plastic
cup as glass

Object 11



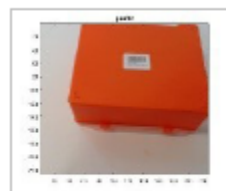
Correct detection of 3d printed object

Object 12

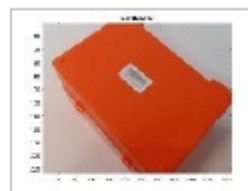


Correct detection of a box

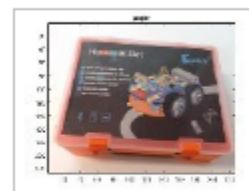
Object 13



Correct detection of an
orange box



Incorrect detection of an
orange box as cardboard



Incorrect detection of an
orange box as paper

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

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

In order to create outdoor image training, we sorted 1,054 digital photographs of outside situations into "take" and "not take" folders. In the first outdoor trial, we mostly photographed objects in natural settings, such as grass, sidewalks, roads, and flower beds. After that, we used these 1,054 photos to train the CNN. Images tagged with "take" featured both recyclable and non-recyclable materials. Pictures of grass, birds, trees, sidewalks, etc., were included in the "non-take" category of images. Two hundred and ten "take" photographs and one hundred and six "non-take" images were used to train the CNN. A total of 93.6% of the time was correctly classified. Accurate identification was achieved for 97.6% of the "take" goods and 85.8% of the "non-take" items. The test's outcomes are shown in Table 4. Examples of training and test photos used for outdoor item categorization may be seen in Figures 5 and 6, respectively. As seen in Figure 7, the CNN algorithm decides whether to "take" or "non-take" a picture and then shows the image with a caption that indicates the choice.

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

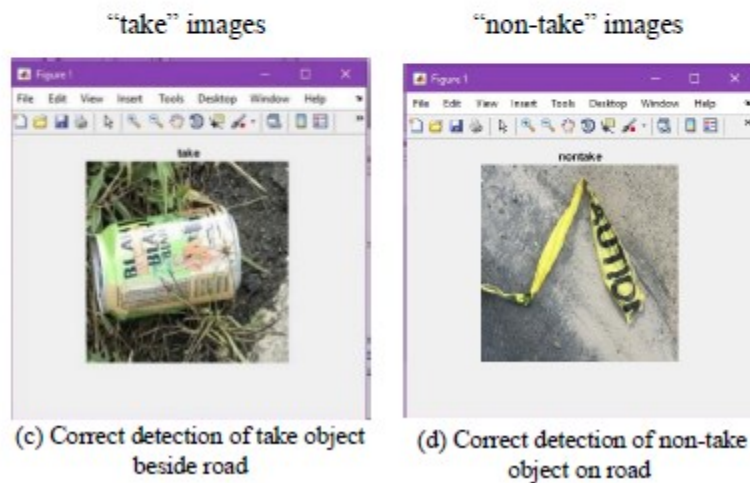
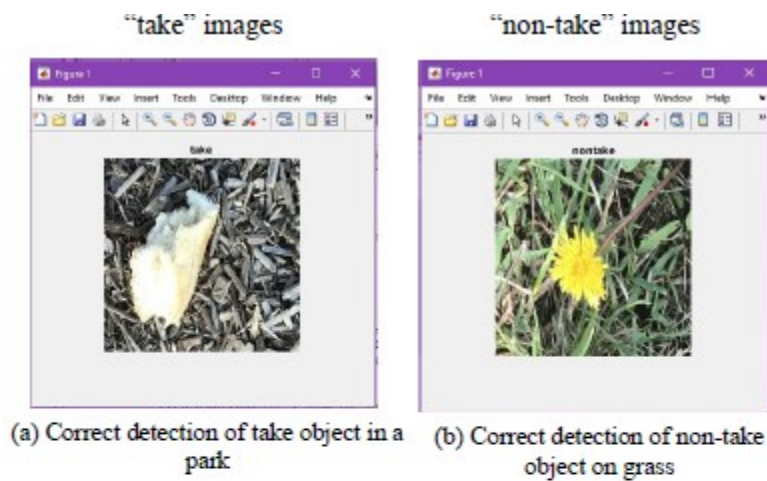


Figure 7: Sample output of test images for outdoor object detection

Test 4 – Classifying Outdoor Images as either Trash or Recyclable

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 CNN. Following this, we put the CNN through its paces on 175 comparable photos spanning the two sets (107 "trash" images and 68 "recycle" images). As previously

demonstrated in Table 5, the findings were correct. Accuracy in categorization was 92% overall. Correct identification of "Recycle" goods was 89.7 percent, while "Trash" items were 93.5 percent accurate. Images with the option to "recycle" or "trash" labeled in the title are shown in Figure 8 as examples.

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 research, we aimed to design an autonomous garbage collection system by creating an algorithm based on convolutional neural networks (CNNs) that could recognize trash and non-trash, and differentiate between landfill and recyclable objects within the waste category. The detection accuracies ranged from 89.7 percent to 93.5 percent, and

the results were positive. Smart trash cans equipped with image processing-based categorization will outperform road sweeper vehicles and vacuum cleaners when it comes to cleaning public places of rubbish. Experiments validated the accuracy of the suggested algorithm in distinguishing between trash and repurposed 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. The next step is to implement our two-stage trained CNN into an algorithm that can communicate with a microcontroller and a camera. This will allow the robot to navigate public spaces, detect objects on the ground, and collect and sort waste according to recyclables or landfills.

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] Agrawal, K. K. ., P. . Sharma, G. . Kaur, S. . Keswani, R. . Rambabu, S. K. . Behra, K. . Tolani, and N. S. . Bhati. "Deep Learning-Enabled Image Segmentation for Precise Retinopathy Diagnosis". *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 12s, Jan. 2024, pp. 567-74, <https://ijisae.org/index.php/IJISAE/article/view/4541>.
- [11] Samota, H. ., Sharma, S. ., Khan, H. ., Malathy, M. ., Singh, G. ., Surjeet, S. and Rambabu, R. . (2024) "A Novel Approach to Predicting Personality Behaviour from Social Media Data Using Deep Learning", *International Journal of Intelligent Systems and Applications in Engineering*, 12(15s), pp. 539-547. Available at: <https://ijisae.org/index.php/IJISAE/article/view/4788>
- [12] Trashnet, Dataset of images of trash, <https://github.com/garythung/trashnet>.
- [13] AlexNet Tool Box, Mathworks, <https://www.mathworks.com/matlabcentral/fileexchange/59133-deep-learning-toolbox-model-for-alexnet-network>.