Smart Waste Management System Using Machine Learning

Shivam Shivam Sharma, Utkarsh Dwivedi and Veenu Singh

May 31, 2023

Smart Waste Management System using Machine Learning

Shivam Sharma
Department of Computer Science and Engineering
Amity University, Uttar Pradesh
Noida, India
Shivam2001sharma07@gmail.com

Utkarsh Dwivedi
Department of Computer Science and Engineering
Amity University,Uttar Pradesh
Noida, India
rhythmakd07@gmail.com

Ms. Veenu Singh
Department of Computer Science and Engineering
Amity University, Uttar Pradesh
Noida, India
vsingh22@amity.edu

***Abstract*— Large waste segregation has always been a major problem. It is done to classify the objects before they are processed as garbage. Waste segregation is typically done manually by labourers or workers. Many issues arise because of workers becoming ill or contracting infections. This has an impact on the worker's life, performance, and total time taken. Manual waste segregation takes a significant amount of time and effort. To solve this problem, we can leverage deep learning algorithms and automate the process by finding an efficient waste segregation algorithm, which can be implemented easily in real time systems with appropriate accuracy. Industries have been transformed by artificial intelligence. However, while most people are familiar with many of us are excited about self-driving cars and face recognition software but still ignorant of the significant influence AI had on the waste segregation and recycling industries. Considera garbage disposal site. Do you see rows of machines in your mind's eye? That's a lot closer to the truth than you might believe. The procedures we use to transport, collect, process, and sort various forms of garbage are constantly improving thanks to artificial intelligence and robotics. According to USEPA (Environmental Protection Agency) 264.4 million tonnes of municipal garbage were generated in the year of 2017 only, or 4.51 pounds an average per person, every twenty-four hours. To put it another way, waste segregation facilities could use each and every assistance available.**

***Keywords*-**Deep Learning, Waste classification, YOLO, CNN

I. INTRODUCTION

Increasing cases of labours getting illness while manually segregating the waste has always been a major issue in the world. Not only labours but companies in the recycle business are also getting affected with this problem. Companies are facing issues while segregating the waste manually and lowering the profit margins. Also, we need these recycling businesses to grow at a certain level that the number of recycled products is equal to the waste generated. Waste segregation has always a labour-intensive operation. Machine learning, Artificial intelligence, computer vision, robots, and someother cutting-edge technologies have enabled towns to remove much of the need for human labour, therefore lowering costs and increasing efficiency. Such hi-tech technology would revolutionise way we manage various sorts of garbage. Start Rocket,for example, intends to remove space trash with space foam. Numerous outstanding contributors left an indelible mark on trash management through machine learning andIoT. We will be determining the efficiency and comparing the deep learning models to findout the best algorithm for segregating waste. The models trained in this project will have the sole purpose of identifying non - biodegradable waste from biodegradable waste. Currently, this is implemented in the form of different coloured dustbins for recyclable material. But still, people may mix all the garbage. This means that companies have to hire manual workers to sort out the garbage after collection. But doing this manually is a cumbersome process and still possess some margin of error – human error, composite materials etc. Such actions are more expensive for the companywhen compared to mass dumping or incineration. Therefore, our project will beimplemented on the level where all the collected garbage is brought for disposal. The models developed in this project will be used solely to classify waste into 6 categories.The six categories are:

1. Paper
2. Glass
3. Cardboard
4. Metal
5. Trash
6. Plastic

Doing this manually is a time-consuming process with some margin of error – human error, composite materials, and so on. When compared to mass dumping or incineration, such actions are more expensive for the company. Therefore, our project will be implemented on the level where all the collected garbage is brought for disposal. Using a camera to capture the images of garbage, our models will classify the objects into 6 classes. The model does not work for any other type of class which does not belong to the 6 classes discussed above for example electronic waste, etc.

1.1. Motivation

Tumors, a neurological disorder that poses a serious risk to life, occur when aberrant nerve cells grow uncontrollably in the human nervous system. All ages have shown an increase in the prevalence of brain tumors during the past 20 years. It is now predicted to be the third most common cancer, mostly affecting adults and teens. Over 136000 people worldwide are investigated annually for brain tumors, and 87000 of them died in 2017 due to brain tumors [8]. Despite the inconsistent efforts of medical experts to address brain tumor concerns, the World Health Organization (WHO) predicts that 251,329 individuals will die from malignant brain illness each year. Hence, a proper diagnosis of a brain tumor is crucial to the patient's survival and the delivery of high-quality medical care. Recently, numerous researchers developed various multidisciplinary models to comprehend the condition and find better surgical techniques. These models included medical scientific knowledge, mathematical understanding, and computer science. To analyze brain image data (CT), many imaging modalities are employed, including, single photon emission computer tomography (SPECT), magnetic resonance imaging (MRI), biopsy scan, and computer tomography. The most popular imaging methods used for brain scans are MRI and CT, which may both detect the presence of brain tumors and pinpoint their location for the purposes of desired expert treatment. The location, size, shape, category, and grade of the brain tumor determine the best course of treatment. Options for medical care also depend on whether the tumor is growing and impacting the body's other organs or a region of the central nervous system (CNS).

II. LITERATURE REVIEW

In a following study, the scientists developed a waste classification system using image processing and Convolutional Neural Networks. Throughout their investigation, they had focused their efforts on detecting polyethylene. In addition, the authors performed various tests to detect polyethylene, terephthalate, high-density polyethylene, polystyrene and polypropylene. For solid waste management, the authorsof a study [1] used a Capsule Neural Network (Capsule-Net). The Capsule-Net could distinguish between plastic and non-plastic products. The researchers looked at two publicly available data sets and found that 96.3 percent and 95.7 percent of theinformation was correct. The integrated system was designed and validated on a varietyof hardware components.

In a study published in 2019 [2], the author suggested a novelclassification model for classifying garbage categories using deep learning methods. Additionally, the technique was utilised for rubbish recycling. The article [3] proposes a method for automatically identifying garbage using a deep learning framework. Additionally, the authors claimed that the model was used for recycling trash Sorting. The researchers developed a method for categorising garbage using a pre-trained CNN model, also known as ResNet-50 or Support Vector Machine, in their study [4]. (SVM). When tested against a publicly available dataset, the model obtained anaccuracy of 87 percent.

The authors of the paper [5] examined a novel method for identifying and classifying electronic garbage, or e-waste. The model classified using a CNN and identified various forms of e-waste using an RCNN model. The accuracy of detection and classification was determined to be between 90% and 97 percent by the authors.

The authors of the paper [6] examined a novel method for identifying and classifying electronic garbage, or e-waste. The model classified using a CNN and identified various forms of e-waste using an RCNN model. The accuracy of detection and classification was determined to be between 90% and 97 percent by the authors. [13] In a paper study, they introduced an intelligent and price effective garbage bin for waste management. The writers added a variety of gadgets, including an Arduino small, an ultrasonic sensor, and a GSM module, to enable IoT functioning of the system. The system employed the GSM module to send an SMS to the User's mobile phone when the trash level exceeded a particular minimum level. This major problem motivated usto think and come up with a technical solution where least human intervention is required and businesses can also increase their growth.

III. METHODOLOGY

The main theory used in CNNs is that of convolution. Convolution is the process of identifying certain "features" of an image, like the eyes of an animal, or the loops in a number; and recording their position with respect to the actual image so as to classify it. For a n x n image, a relatively smaller "filter" is used to identify these said features.The filters themselves are smaller matrices representative of the values in each pixel ofthe image. Instead of using the whole image, like in ANN, here, we try to divide the picture into smaller patterns that are broken down with the help of filters. The actualprocess involves creating a feature map; this is done by mapping the whole image withthe filter and multiplying the obtained smaller matrix with it. If the value in the featuremap comes out to be "1", then that portion of the image matches with the feature we want to find. The feature maps themselves then contain the position of these features with respect to each other.

After convolution, the next layer is the activation layer. For any neural network model, the activation layer applies its activation function to the output from the previous layer. In this case, the ReLU activation.

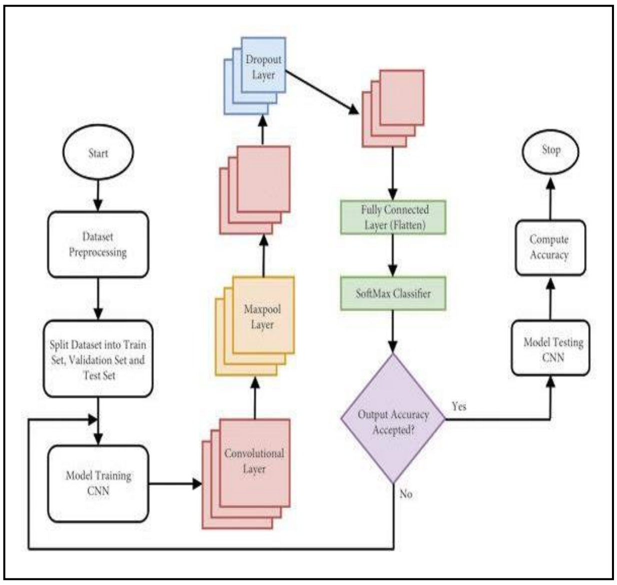

Fig 1. Max Pooling

function is used on the previous output given by the convolutional layer. It replaces all negative values of the feature map with “0” so as to make the model less linear.

The final layer, i.e., the Pool Layer takes the output from the Activation Layer, and thenmakes it smaller. This helps in increasing performance and decreasing the load of the system. The process is called pooling, and it can be implemented in two different ways, namely – Max Pooling and Average Pooling. The former is morepopular, as we just take the max value from a given portion of the matrix for the new output. Average Pooling, on the other hand, uses the Average of all the present values.With the help of pooling, we get a reduced matrix, as well as the model performs betterin terms of position invariance and overfitting of the model.

The final layer, i.e., the Pool Layer takes the output from the Activation Layer, and thenmakes it smaller. This helps in increasing performance and decreasing the load of the system. The process is called pooling, and it can be implemented in two different ways, namely – Max Pooling and Average Pooling. The former is morepopular, as we just take the max value from a given portion of the matrix for the new output. Average Pooling, on the other hand, uses the Average of all the present values.With the help of pooling, we get a reduced matrix, as well as the model performs betterin terms of position invariance and overfitting of the model.

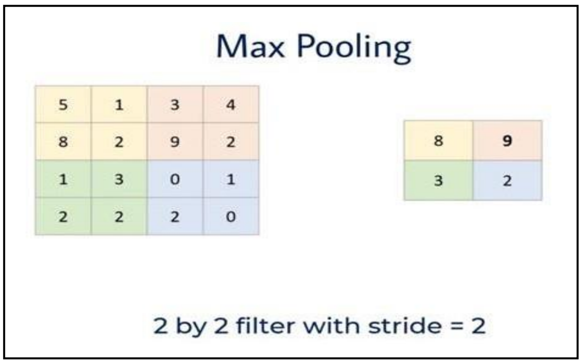

Fig 2. Max Pooling

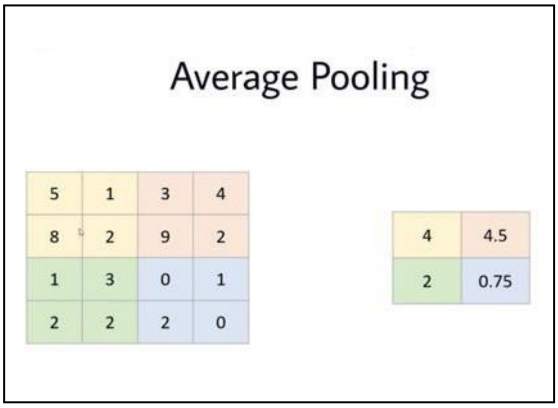

Fig 3. Average Pooling

YOLO v3 uses both binary cross-entropy loss and logistic classifiers which are independent for the prediction of classes while training with the help of such changes it is possible while training to use complex datasets like OID (Open Images Dataset) which have many labels overlapped with each other for the images used in the dataset. The classes in YOLOv3 are more specific because of the multilabel approach used by it whereas the previous version which is YOLOv2 used softmax which is a function that converts number vectors into probability vector. The softmax detects each bounding box as it belongs to a single class which many times might not be the case this case arises in case of datasets like Open Images Dataset.

The main advantage of using Yolov3 is that it is faster than many other networks and it is still able to maintain accuracy. Model is looked at once while testing the image and hence the image’s globalcontext informs about the prediction at first it separates the image into grid after that the number of bounding boxes is predicted the bounding boxes have their own confidence score depicting how accurate the prediction should be and at a time it detectsonly a single object. YOLO performs classification and regression of bounding box at the same time. The main difference between YOLOv3 and the version before it is that is more precise, faster and more specific in terms of determining classes and there is a huge difference in terms of accuracy between YOLOv3 and its previous versions and darknet-19 was used by YOLOv2 for feature extraction whereas YOLOv3 used darknet-53 for the same purpose which was also created by Ali Farhadi and Joseph Redmon as the name suggests the darknet –53 has 53 convolutional layers whereas darknet-19 has only 19 hence it becomes more efficient and powerful than darknet-19.

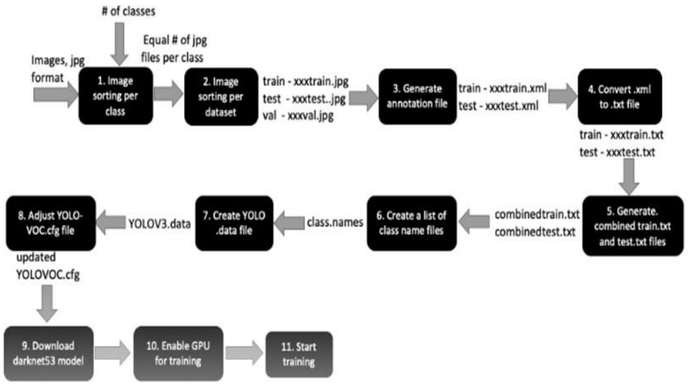

Fig 4. Timeline

The main disadvantage for which YOLO is known for is its inaccuracy for detecting small objects YOLOv2 can't even be compared to other algorithms while detecting small objects having an AP of 5.0 which was very less when compared to Retina Net which has an AP OF 21.8 and SSD513 which has an AP of 10.2 with the invention of YOLOv3 this problem was resolved to some extent having an AP of 13.3 which is muchmore when we compare it to YOLOv2.

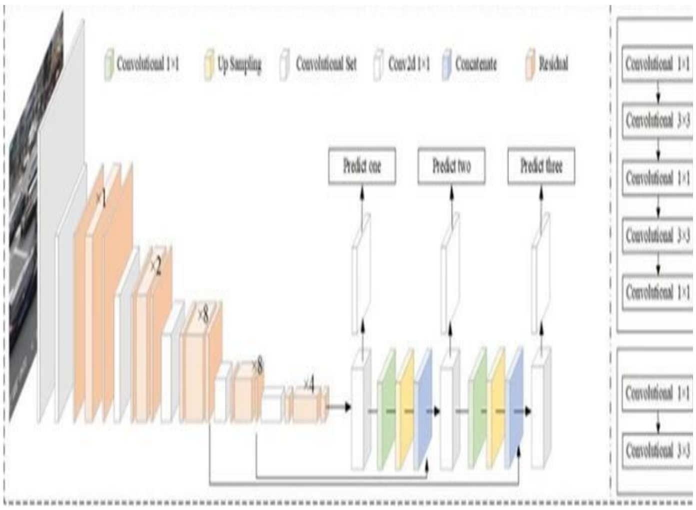

Fig 5. YOLO framework

YOLOv3 uses both binary cross-entropy loss and logistic classifiers which are independent for the prediction of classes while training with the help of such changes itis possible while training to use complex datasets like OID (Open Images Dataset) which have many labels overlapped with each other for the images used in the dataset.The classes in YOLOv3 are more specific because of the multilabel approach used by it whereas the previous version which is YOLOv2 used softmax which is a function that converts number vectors into probability vector. The softmax detects each bounding box as it belongs to a single class which many times might not be the case this case arises in case of datasets like Open Images Dataset.

The steps followed to build the model are as follows. Annotated the dataset which contained images for glass, plastic, paper,cardboard, metal, trash for Yolov3 using LabelImg Tool.

Training of the model on google collab for which we used darknet and OpenCVversion 3.2.0.

Imported time library in the detect.py file so that when a video is passed as inputa specific time interval of five seconds is there while detecting the images.We used the Cv2 .VideoCapture function to capture the video.

We used the cap.read function so that the images in the video are read by theProgram. For optional channel swapping and mean absraction scaling we usedcv2.dnn.blobFromImage function.

The process of detecting objects from garbage dataset has been divided into:

1. Data Collection and Cleaning
2. Image Augmentation
3. Image Annotation
4. Model Training
5. Model Evaluation

3.1 Data Collection and Cleaning

Annotated the dataset which contained images for glass, plastic, paper,cardboard, metal, trash.In this phase, we collected a dataset containing six categories: glass,cardboard, paper, metal, trash, and plastic. We addedmore images by manually clicking images, resizing them, and then adding them to the dataset. The cleaning process involved removing blur images from the dataset,resizing it to (720, 960,3) format from (3000,4000,3), and labeling them with correct classes for bounding box and mask region training.We used the matplotlib library to read and save the images. cv2 library to change the resolution of the images. Imported constants.py, which had the information about the dimensions.We used os library to iterate over the images and np to rotate the images by 90 degrees.

3.2 Image Augmentation

In this phase, we increase the data size by adding images manually. We have also used certain augmentation techniques such as random horizontal flip, scaling, and random resized crop. This is done to avoid the problem of overfitting because of small dataset as pre-trained models need a bigger dataset. We Converted the images to grayscale. Flipped images along the horizontal axis Performed random image rotation and skewand Transformed images to numpy array using numpy library.

3.3 Image Annotation

This phase is required only for YOLO v3 deep learning algorithms. This phase includes annotating the images for the bounding box coordinates using the LabelMe Tool. It is an open-source software used to label images. It saves annotation information in JSON format and has different JSON file for each image. The JSON filecontains the coordinates of the bounding box and uses this information to train the convolutional neural network for mask prediction bounding box prediction. This is a manual process and is done to train the model for bounding box prediction on the test images. The task of manually annotating the images for the bounding boxes was time consuming.

3.4 Model Training

I. CNN (Convulaional Neural Network)

This phase includes the model training of CNN model on our garbage dataset. The data set was divided into three sections, which are as follows:

1. Training Dataset
2. Testing Dataset
3. Validation Dataset

Split the dataset to train, test, and validation in 50/25/25 ratio. Used the trainedresnet34 CNN model on our test dataset. A residual neural network is a (CNN) convolutional neural network with multiple layers of convolutional neurons used in artificial intelligence. The CNN resnet34, which includes a total of 34 layers and has been pretrained on the ImageNet database, is one example. As a result, because it has previously acquired some visual features and can transfer that information, a pretrained CNN will perform considerably better on fresh photo classification tasks. Deep neural networks are expected to outperform shallow neural networks when it comes to training data because they are suitable for describingmore levels of complexity. Deep neural networks, on the other hand, tend to perform measurably worse than shallow neural networks in practise. Resnets were developed inorder to go around this bug by utilizing a hack known as shortcut connections. The weights and bias of some nodes in a layer can be adjusted; however, if a node has ideal values (i.e., its residual is zero). Shortcut connections employ the identity function to convey information to the next tier in the hierarchy when changes are required. Shortening the neural network, when possible, allows resnets to have deep architectures and perform more like shallow neural networks than they would otherwise. The number 34 in resnet34 merely represents the network's layer count.

Identifying a learning rate

Finding a learning rate for gradient descent is critical for ensuring that the neural network converges quickly while still achieving the best error.

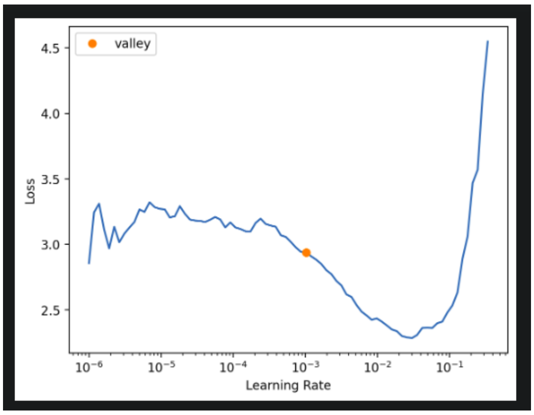

Fig. 6. Learning Rate vs Loss Plot

Training

Ran the model for a total of 20 epochs. We get closer and closer to the optimum levelwith each epoch because the learning rate decreases with every epoch.
First and foremost, we can examine which images were classified incorrectly the mostfrequently.

epoch	train_loss	valid_loss	error_rate	time
0	1.513418	0.651505	0.238095	07:12
1	1.016411	0.607873	0.182540	07:11
2	0.978930	0.788354	0.246032	07:13
3	1.060076	1.012475	0.295238	07:14
4	1.160038	1.044769	0.269841	07:12
5	1.139708	0.811752	0.263492	07:11
6	0.950835	0.765725	0.239683	07:12
7	0.982388	0.635572	0.204762	07:11
8	0.695169	0.538402	0.177778	07:12
9	0.711271	0.595780	0.185714	07:11
10	0.495293	0.556972	0.161905	07:12
11	0.501241	0.406515	0.126984	07:11
12	0.415007	0.367834	0.111111	07:12
13	0.354920	0.390749	0.112698	07:11
14	0.240962	0.333982	0.107937	07:12
15	0.199863	0.353655	0.100000	07:11
16	0.157238	0.334674	0.087302	07:12
17	0.220873	0.338140	0.085714	07:11
18	0.124813	0.329583	0.082540	07:12
19	0.130007	0.325397	0.082540	07:11

Fig. 7. Trained the model for 20 epochs

Visualizing the majority of incorrect images

The images on the recycler did not perform well on were actually degraded. The photosappear to have received an excessive amount of exposure, so this isn't model's fault.

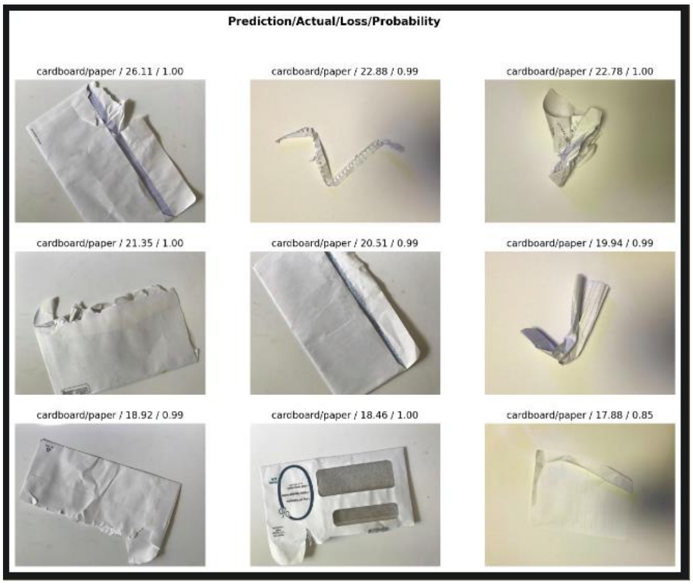

Fig. 8. Prediction / Actual Loss / Probability

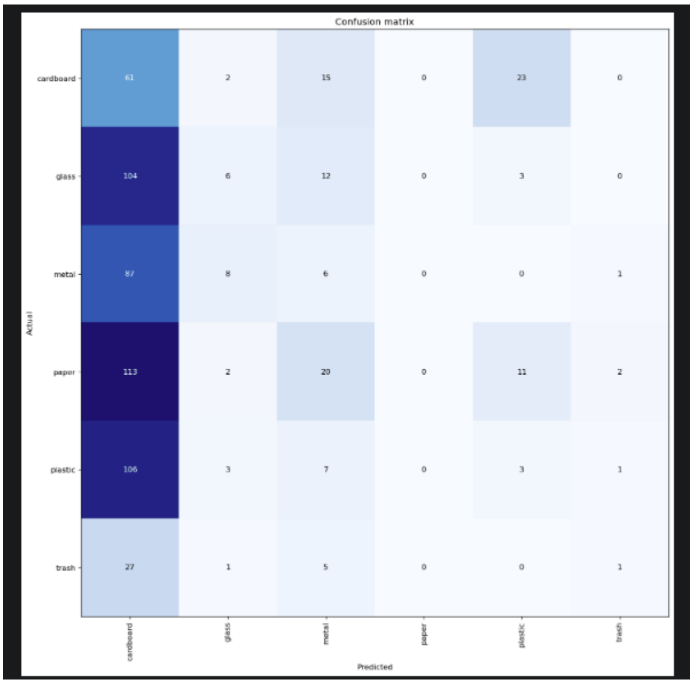

Fig. 9 Confusion Matrix for CNN

Found out the most incorrect images using the top_losses() function from the fastai module. Displayed them using the plot_top_losses() function to show which type of images were wrong.

Visualized the performance using a Confusion matrix. This model frequently confused cardboard for paper, plastic for metal and plastic for paper, among other things.

The following is a list of the images that have caused the most confusion.

```
[('paper', 'cardboard', 113),
 ('plastic', 'cardboard', 106),
 ('glass', 'cardboard', 104),
 ('metal', 'cardboard', 87),
 ('trash', 'cardboard', 27),
 ('cardboard', 'plastic', 23),
 ('paper', 'metal', 20),
 ('cardboard', 'metal', 15),
 ('glass', 'metal', 12),
 ('paper', 'plastic', 11),
 ('metal', 'glass', 8),
 ('plastic', 'metal', 7),
 ('trash', 'metal', 5),
 ('glass', 'plastic', 3),
 ('plastic', 'glass', 3),
 ('cardboard', 'glass', 2),
 ('paper', 'glass', 2),
 ('paper', 'trash', 2)]
```

Fig. 3.23 Confused Images

II. YOLO v3 (You Only Look Once)

YOLOv3(You Only Look Once) is a recent algorithm that is introduced in deep learning to remove the slow speed and increase in complexity during optimization caused by various other Algorithms such as R-CNN, fast R-CNN and faster R-CNN. Using YOLO object detection is transformed into a single regression problem.

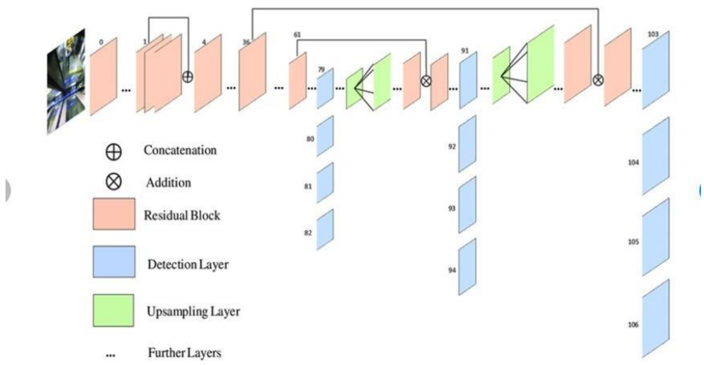

Fig 10. Architecture of YOLOv3

The YOLOv3 algorithm takes an image as input and extracts a single image which getsdivided into various grids which can be 3X3,4X4,19X19etc and the prediction of the bounding boxes is done with the help of anchor boxes which utilize the dimension clusters.4 coordinates(tx,ty,tw,th) are used for prediction purpose for each bounding box.

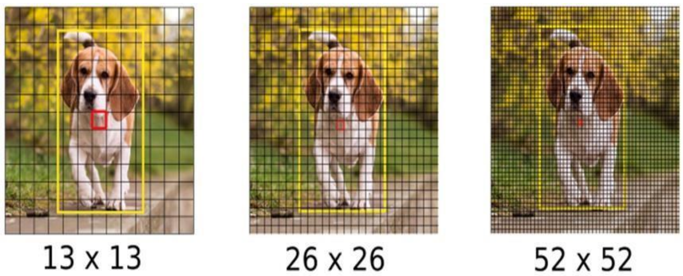

Fig. 11. Grid

After that each object in the image gets surrounded by a bounding box which has some width(bw) and height(bh) and classes(C) which can be a person, dog, car etc and the centre of bounding box(bx,by).Generally this algorithm works well when several objects are there in a single image it creates bounding box across each object, problem accours when multiple bounding boxes are used but to overcome that problem this algorithm uses Intersection Over Union(IOU) which gives an output box surrounding the object perfectly. Grid boxes are responsible for predicting the score of confidence of each and every bounding box.

IV.RESULTS

I. CNN

Tested the trained resnet34 CNN model on our test dataset. Inference is done to check if the trained model is working as it should. Also, it gives the opportunity to improve the accuracy of the model by tweaking some hyper parameters.

```
torch.Size([635, 6])
tensor([[9.6581e-01, 2.4697e-02, 8.0555e-05, 9.2086e-03, 7.0032e-06, 1.9963e-04],
        [9.8796e-01, 1.7847e-04, 5.4057e-06, 5.1378e-05, 9.3261e-03, 2.4821e-03],
        [9.9999e-01, 1.4395e-09, 3.0080e-10, 3.8624e-06, 1.3739e-06, 2.7315e-06],
        ...,
        [5.1053e-06, 7.6157e-07, 1.4170e-07, 2.9622e-01, 3.1686e-04, 7.0346e-01],
        [3.9680e-04, 3.8644e-05, 7.9046e-07, 6.6174e-01, 2.8764e-05, 3.3780e-01],
        [2.5746e-02, 8.4096e-05, 5.9605e-04, 6.5018e-02, 1.2686e-03, 9.0729e-01]])
```

Fig. 12. Grid

For each image, these are the anticipated probability. There are 365 rows in this tensor,one for each image, and six columns, one for each material type.
Now we'll convert the probabilities in the tensor above into a vector of predicted classnames.

```
['cardboard',
 'cardboard',
 'cardboard',
 'cardboard',
 'cardboard',
 'glass',
 'cardboard',
 'cardboard',
 'cardboard',
 'trash',
```

Fig. 13. Grid

Now comparing it with the first image and check whether it is actually a cardboard ornot.

```
learn.data.test_ds[0][0]
```

```
print(yhat[520])
plastic
```

Fig. 14. Predicting water bottle

The above figures shows the predicted outputs of different wastematerials like cardboard, plastic, paper, metal, glass and trash by using CNN resnet34 model.

Now comparing it with the first five labels of the dataset

So, first five predicted labels matched with the labels of the dataset.

```
['cardboard', 'cardboard', 'cardboard', 'cardboard', 'cardboard']
['cardboard', 'cardboard', 'cardboard', 'cardboard', 'cardboard']
```

Fig. 15 Predicted Table

II. YOLO

Here we created a detect.py file to test the trained model in which we passed the images which are not trained but were from the same dataset in this we included the cv2.video capture function which takes a video as input which should contain images of differentkinds of waste materials like plastic, metal, glass, paper, cardboard and other trash. We also included the cap.read function so that it could read the images in the video. We used the cv2.dnn.blobFromImage function for optional channel swapping and mean abstraction.

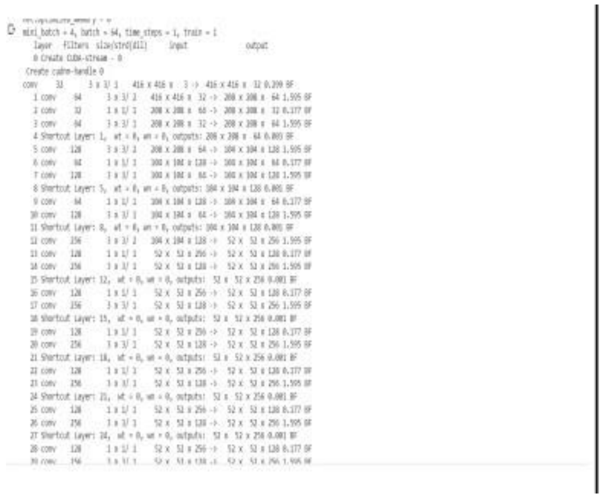

Fig 16. Model Training

Fig. 17. Predicted Plastic Image

Fig. 18. Predicted Metal Image

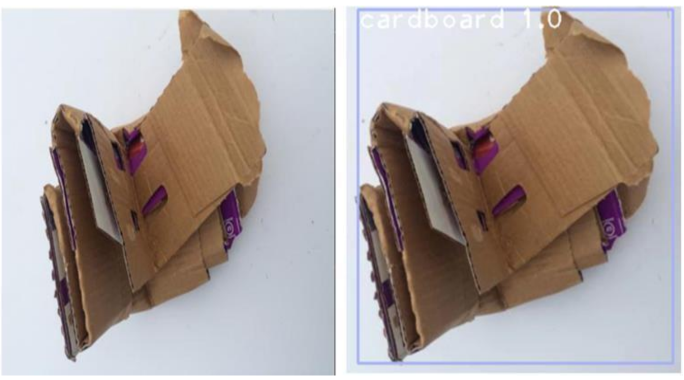

Fig. 19. Predicted CardBoard

The above figures shows the predicted outputs of different waste materials like cardboard, paper, metal, glass and plastic by using YOLOv3 object detection technique. In this phase, we test the trained model on our test dataset. Inference is done to check if the trained model is working as it should. Also, it gives the opportunity to improve the accuracy of the model by tweaking some hyper parameters. In this phase, we test the trained model on our test dataset.

Based on the accuracy and recall values provided, YOLO v3 seems to be slightly more accurate and have a higher recall compared to CNN for the specific smart waste management system scenario you mentioned.

YOLO v3 achieves an accuracy of 85.29%, whereas CNN achieves an accuracy of 82.75%. This indicates that YOLO v3 has a slightly higher ability to correctly classify and localize waste items compared to CNN.

Additionally, YOLO v3 has a recall of 79.1%, while CNN has a recall of 77%. Recall represents the ability of the model to detect relevant objects, in this case, waste items. YOLO v3 demonstrates a slightly better performance in detecting waste items, ensuring fewer false negatives compared to CNN.

However, it's important to note that the difference in accuracy and recall between YOLO v3 and CNN is relatively small. The performance of these algorithms can vary depending on the specific implementation, training setup, and dataset used. It is recommended to consider other factors such as computational efficiency, speed, and specific requirements of the smart waste management system when choosing between YOLO v3 and CNN.

V. CONCLUSION

We worked on this project to see which algorithm out of the two algorithms that were CNN and YOLOv3 ,also doing the comparative analysis discussed above, we came to a conclusion that YOLO v3 outperforms the CNN algorithms by a significant margin because YOLO can detect multiple objects in a single pass, even in complex scenes with overlapping or closely positioned waste items. This capability is advantageous in waste management systems, where multiple waste items may need to be identified and sorted simultaneously. YOLO v3 is an object detection algorithm that utilizes a CNN as its backbone. It divides the input image into a grid and predicts bounding boxes and class probabilities for objects within each grid cell. YOLO v3 is known for its real-time object detection capability and overall speed. The YOLO model is getting an accuracy of 85.29%.

After working on this project, we can also conclude that this model will only work on images of plastic, metal, glass, paper, cardboard and other trash this modelwill not be suitable for any other kind of waste material like electronic waste etc.

Our algorithms uses deep learning to identify the type of waste there is in the garbage pile and predict the with a descent accuracy in types of garbage it is trained. Using IOT,we can make it for industrial use. Our model can predict 6 categories of Which will reduce the human labour and hazardous locations increasing human safety. If used in hybrid these can increase the efficiency of the work and further research can cause moreautomation reducing cost. YOLOv3 can be used for industrial purposes as it has anaccuracy which can have practical relevance.

VI. REFERENCES

[1] G. Sai Susanth, L. M. Jenila Livingston, and L. G. X. Agnel Livingston, "Garbage waste segregation using deep learning techniques," IOP Conf. Ser. Mater. Sci. Eng., vol. 1012, no. 1, p. 012040, 2021.

[2] "Mask R-CNN: A beginner's Guide," viso.ai, 19-Mar-2021. [Online]. Available: https://viso.ai/deep-learning/mask-r-cnn/. [Accessed: 23-Mar-2022].

[3] R. Girshick, "Fast R-CNN," arXiv [cs.CV], 2015.

[4] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama, et al. Speed/accuracy trade-offs for modern convolutional object detectors. In CVPR, 2017

[5] R.S.Sandhya Devi, Vijaykumar VR , M.Muthumeena, "Waste Segregation using Deep Learning Algorithm," Ijitee.org, Dec-2018. [Online]. Available: https://www.ijitee.org/wp-content/uploads/papers/v8i2s/BS2676128218.pdf. [Accessed: 10-Apr-2022].

[6] F. M. Abien, "Deep Learning using Rectified Linear Units (ReLU)," Arxiv.org. [Online]. Available: https://arxiv.org/pdf/1803.08375.pdf. [Accessed: 28-Apr-2022].

[7] Thung, Gary and Ming Xiang Yang. 2016. "Classification of Trash for Recyclability Status.", Corpus Id-27517432.

[8] Sreelakshmi, K. & Soman, Akarsh & Ravi, Vinayakumar & Kp, Soman. (2019). Capsule Neural Networks and Visualization for Segregation of Plastic and Non-Plastic Wastes. 631-636. 10.1109/ICACCS.2019.8728405.

[9] Huiyu, Liu & O., Owolabi & Kim, Sung-Ho. (2019). Automatic Classifications and Recognition for Recycled Garbage by Utilizing Deep Learning Technology. 1-4. 10.1145/3377170.3377190.

[10] O. Adedeji and Z. Wang, "Intelligent waste classification system using deep learning convolutional neural network," Procedia Manuf., vol. 35, pp. 607–612, 2019.

[11] D. D. Sri, P. Gagana, H. Anarghya, S. M. L. Pannaga, and S. V. Sheela, "Towards a Greener Future: Sustainable Development and Green IT," Int. J. Eng. Res. Technol. (Ahmedabad), vol. 10, no. 6, 2021.

[12] Misra, Nisha & Kumar, Sandeep & Jain, Arpit. (2021). A Review on E-waste: Fostering the Need for Green Electronics. 1032-1036. 10.1109/ICCCIS51004.2021.9397191. 78