Disposable Beverage Cup Classification for Waste Recycling: A Mask R-CNN based Instance Segmentation

Jerome E. Lacieras
College of Engineering
Batangas State University
Batangas City, Philippines
jerome.lacieras@g.batstate-u.edu.ph

***Abstract*— One of the most significant environmental issues the world faces today is waste pollution. Disposable beverage cups are said to be one of the top 10 contributors in this problem. As the usage of disposable beverage cups become more convenient in all sectors of society. The problem on how it will be dump arises. This study proposed an object detection model for the classification of disposable beverage cup whether it is paper or plastic cup, that can be utilized by projects that functions as smart collection system and the likes. The developed classification model for paper cup and plastic cup uses Mask R-CNN architecture and ResNet-101 as the backbone. This deep learning algorithm obtain an impressive result as it produced a model with 98.75% of mAP which suggest that Mask R-CNN is a good option to use when detecting such objects.**

Keywords—beverage cups, cups, detection, mask r-cnn, recycling, waste.

I. Introduction

One of the most significant environmental issues in the world today is Waste Pollution. It contaminates water source, harming plants and animals, and threaten human health directly [1]. The massive production of disposable goods over the past few years has significantly increased the amount of waste generated. By 2050, the amount of garbage produced annually is predicted to exceed 3 billion tons, according to World Bank statistics. They also state that only 13.5% of global waste is recycled, while about 33% of garbage is thrown away openly without any preliminary classification[2].

Disposable cups are one type of container which is designed for single-use to serve beverages such as water, cold drinks, hot drinks and alcoholic beverages [3]. The first part of the 20th century has seen the emergence of disposable cups, which have now penetrated in all sectors of the society. For consumers on the go, many eateries and kiosks serve beverages in disposable cups (e.g., by commuters, shopping public, or beach visitors). Disposable cups are also frequently used in locations where the lack of cleaning facilities and large numbers of customer throughput make the use of reusable cups nearly impractical. This applies to huge public events like festivals and events. Medium organization such as schools and universities, where consumption peaks during breaks also applies to this sake. Additionally, the usage of disposable cups in office-type settings is rising, frequently in combination with vending machines [4]. With this usage of Single-use or disposable plastic cups it is undeniably that it can be one of the top 10 that contributes to global waste as per [5]. The use of beverage cup for the coffee alone according to [6] is definitely high. Per year, US's coffee cup usage tallies about 50 billion per year while UK and Australia marks with 2.5 billion and 1 billion cups per year, respectively [6].

In order to reduce the volume of waste about disposable cups and prevent further pollution of the environment, recycling is one of the remedies that can help with this issue [7]. However, recycling disposable cups is challenging but there are studies that can provide solutions in recycling or reusing this type of waste. The study [8] suggested resolutions on how waste disposable paper cups (WDPC) through laboratory study and commercial approach. The laboratory study includes using WDPCs as a carbon source to form graphene sheets. WPDCs is also used in Vermicomposting technology to make it as a nutrient-rich manure. Lastly, by applying pyrolysis to WPDCs [8]. The result of pyrolysis of paper cup waste constitute to the study of [9] as the obtained hydrocarbon liquid and it is examined to determine its feasibility as a commercial fuel. Result indicates that the hydrocarbon liquid could serve as a feedstock for further upgrading or as a diesel fuel substitute when combined with lighter chemicals [9].

Now that the ways on to how this waste can be reduced and be recycled in a safe, efficiently, and cost-effective manner is provided. Smart collection and segregation of garbage should take over first. Since the volume of disposable beverage cup usage is high, it can be the type of waste to focus to. This study aims will help to classify whether the beverage cup is paper or plastic that will help the existing automatic waste sorting machine such as [10] in supporting wise segregation and use the disposable beverage cups waste in more valuable manner such as [8][9]. Only plastic cup and paper cup are can be detected using this model. No comparison of algorithms occurs in this study since there are no other existing beverage cup detection model. It only focuses on using Mask R-CNN as the architecture and ResNet-101 as the backbone which is discussed in section II along with the ways on how the system is created. The

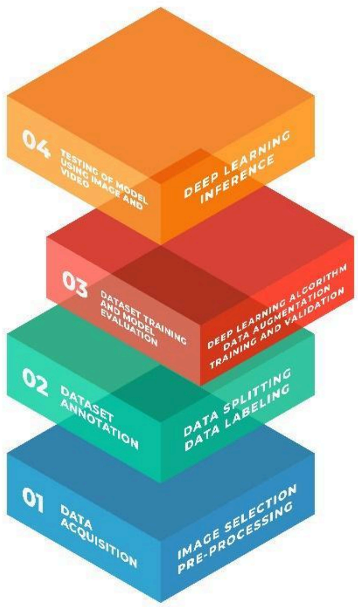

Fig. 1. Diagram of the system work flow.

experimental results and discussions are covered in section 3, while the conclusions are covered in section 4.

II. Methodology

The work flow of the system is shown in Fig. 1. The development or creation of the system passed through different phases or stages, including dataset preparation, data annotation, model training and evaluation, and finally, model inferencing.

A. Data Acquisition

The dataset utilized in this study was obtained from Kaggle [11]. It contains synthetic images of paper cups and plastic cups as show shown in Fig. 2. There are 15,000 pictures of cups in the collection that it is not yet categorized whether it is a paper cup or plastic cup. Since the dataset contains synthetic images, there are just limited number of cup samples that putted into different background, angle, lighting, and distance.

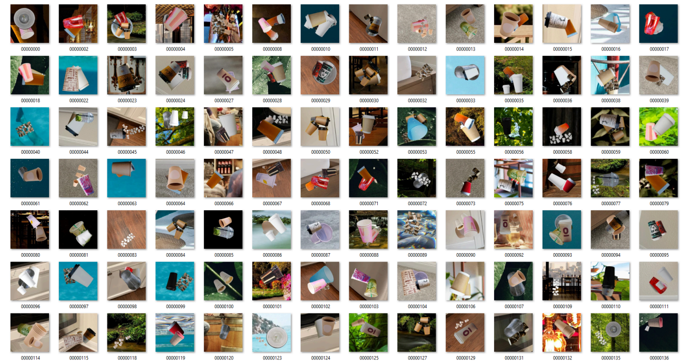

Fig. 2. Sample image from the cups dataset.

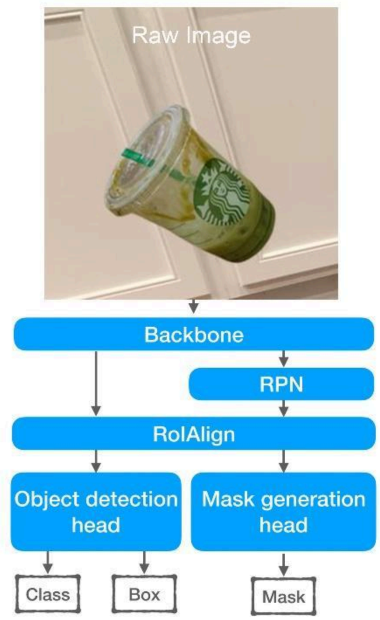

Fig. 3. Mask R-CNN architecture.

B. Data Annotation

In order to proceed in training of the data, it must be labeled first. Mask R-CNN algorithm required to have labeled training sets which will be created using the LabelMe software. It is an open source annotation tool based on http://labelme.csail.mit.edu. It was written in Python to support manual image polygonal annotation for object detection, classification, and segmentation [12].

C. Mask R-CNN

The deep learning algorithm used in this project is Mask-RCNN. it is a variant of a Deep Neural Network that detects objects in an image and generates a high-quality segmentation mask for each instance which makes it as the state-of-the-art in terms of image segmentation [13]. This algorithm was used to further classify paper cups from plastic cups and could be divided into five main structures which can be shown in fig 3. The five structures of Mask R-CNN are explained as follows [14]:

1. ResNet-101 as Backbone Framework.

This study used Resnet-101 as the backbone framework and extracted the image features of a paper cup and plastic cup using feature pyramid network (FPN). ResNet-101 is widely used convolutional neural network that is capable of extracting features from images. In addition, FPN was utilized to fuse the feature images from the bottom to the top to maximize the effectiveness of the extraction process. FPN can boost productivity and speed while also creating higher-quality feature map pyramids.

2. Regional Proposal Network (RPN)

The images are divided into two categories while considering also the background: paper cup and plastic cup. The RPN will scan the feature map and proposing regions that may have objects in them (RoI). The RPN prediction will preserve the box with the highest foreground score and discard the other boxes if the predicted boxes in a region overlap too much. This process will now make the paper cup, plastic cup and the background be distinguished from one another.

3. Region of Interest (ROI)

The RPN provides the region of interest using ROI align. The feature map of the region of interest is cut by pooling and sent to two branches using bilinear interpolation. One branch network consists of a region of interest classifier and boundary regression, while the other branch network is a mask generating network built from full convolutional networks.

4. Object detection branch

Recognizing paper cup and plastic cup needs the employment of ROI classifier and border regression; both are composed of a full connection layer. To identify paper cup and plastic cup more correctly, the border regression alters the region of interest's aspect ratio and center point position.

5. Mask generation branch

The fully connected network makes up the mask-generating network. The network can segment the images of both cups and build a mask that fits the size and shape of the paper cup and plastic cup. The mask pictures are combined with the recognition results to generate an image containing the paper cup and plastic cup category and segmentation mask.

D. Model Evaluation

The trained models were compared using the mean Average Precision (mAP) to make sure that the best-trained model was picked for model inference detection. It will create files that will be used in the testing procedure. As the mAP, which is the average of the AP, rises, the model's detection accuracy also improves.

Average Precision (AP): To eliminate the effect of curve wiggles, interpolate the accuracy during successive recall stages until the AP (1) is determined. The area under the interpolated curve is defined as AP, and it can be calculated as follows. The interpolated precision P_{interp} (2) specifies a specific degree of recall r as the highest level of accuracy discovered at any step of recall $r'\geq r$.

Mean Average Precision (mAP): The average precision (*AP*) across all tests is referred to as mean average precision (*mAP*) (3), where 0 is the number of queries in the set and AP i represents the average precision (AP) for a given query, 0.

$$AP = \sum_{i-1}^{n-1}\left(r_{i+1} - r_i\right)p_{interp}\left(r_{i+1}\right) \quad (1)$$

$$p_{interp}\left(r_{i+1}\right) = p\left(r'\right) \quad (2)$$

$$mAP = \frac{\sum_{i=1}^{o} APi}{o} \quad (3)$$

E. Model Inference and Testing

For the model inferencing, the proponents used the h5 file of the best-trained model and validate it in two ways; via image detection and video detection. The accuracy is based on the output given by the instance segmentation code that executed using Google Collaboratory. The image and video that will be used in the detection is obtained outside the dataset to prevent biases.

For the video detection, the total accuracy is computed by (4):

$$Accuracy = \frac{Total\ of\ the\ result\ per\ frame}{Total\ number\ of\ frames \times 100} \times 100 \quad (4)$$

III. Results and Discussion

This section discussed the obtain results from dataset preparation, data annotation, model training and evaluation, and model inferencing.

A. Data Preparation

Categorizing the images in the dataset were done in the data preparation. A total of 800 images were categorize in two class which show in Fig. 4: 400 for paper cups (a) and 400 for plastic cups (b). After that the new collection were split for training in validation which ideally have an splitting ratio of 80% (640 in total, 320 images from each class)for the training and 20% (160 in total, 80 images from each class) for the validation.

B. Data Labeling

Fig. 5 shows the how the images from datasets are annotated and labeled using LabelMe. It was done by constructing a bounding box using polygon on the disposable cups and label it whether it is a paper cup or plastic cup. A JSON file corresponding to the original image was used to store the labeling information.

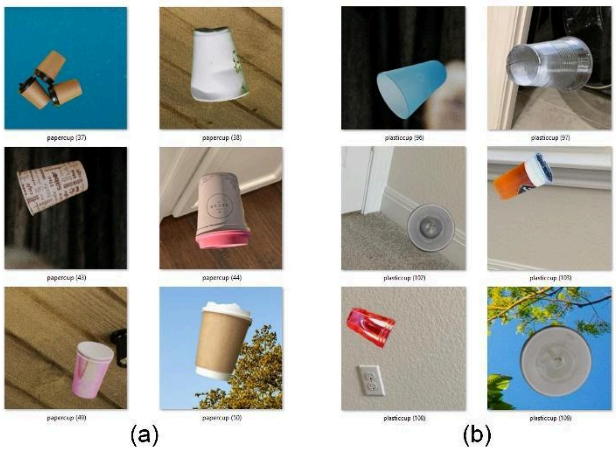

Fig. 4. Sample datasets for paper cup (a) and plastic cup (b).

Fig. 5. Data annotations of paper cup (left) and plastic cup (right).

C. Training Results

The proponent trained the data using Mask R-CNN algorithm by setting the number of epoch to 50. Result obtained is showed on Fig. 6, where the different parameters is represented in different color. The parameter loss and val_loss of the epochs started 0.616 and 0.253 respectively, and ends at 0.052 and 0.0528 in epoch 50 respectively. The lowest value of val_loss obtained is from epoch 38 which got 0.051. Although the set number of epoch is 50, the proponents only saved the model that creates improvement in terms of val_loss which will be evaluated afterwards.

D. Model Evaluation

Figure 7 shows the result of evaluation models. It can be noticed that there a just only 17 models saved from the 50 epochs that has been trained, it is because these models show improvement in terms of val_loss. The mAP, which is presented as a percentage, represents the accuracy of the dataset validation.

Fig. 6. Training results of Mask R-CNN architecture.

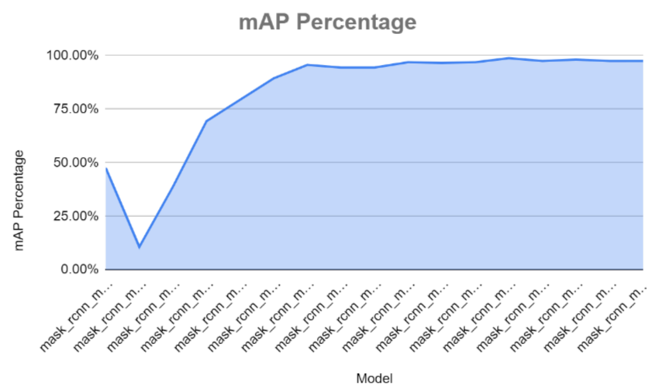

Fig. 7. Model evaluation result.

Since the mAP is equal to 1.00 (100%), the validity number is near to its maximum possible value. The result started as the first model obtained 47.5% but it quickly drop at 10.6% at the second model. After this the mAP started to progress again from model 003 to model 038. The best-model generated for this study is model 030 which got an mAP of 98.75% (0.9875).

E. Model Testing and Inference

To test the chosen model's performance, which is the Model 030, it was applied in an image and video. Model 030 was selected since it got a mAP of 98.75% and had a loss and val_loss of 0.068 and 0.058 respectively. Inferencing of the model is shown as follows.

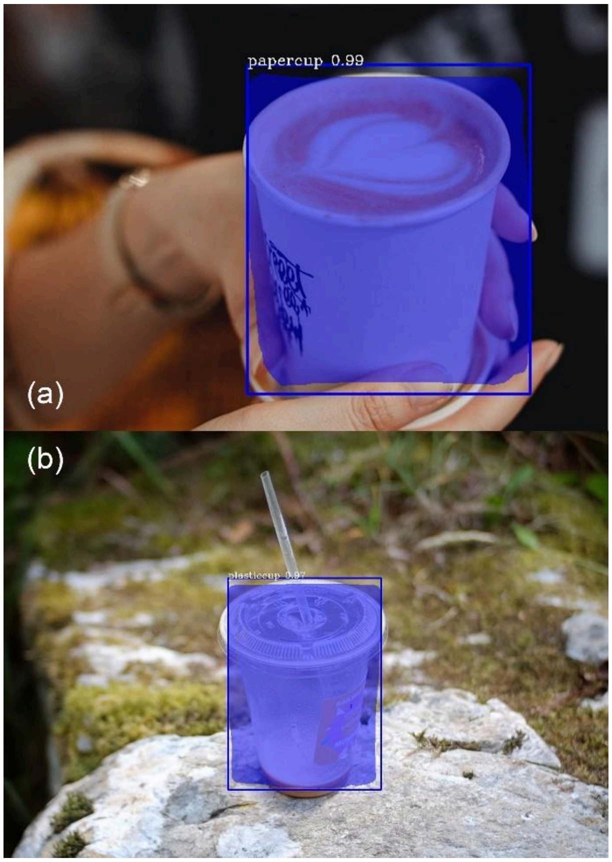

Fig. 8. Image inference result for paper cup (a) and plastic cup (b).

Fig. 9. First frame of video inference result for paper cup (a) and plastic cup (b).

1. Image Inference

The proponent tests the model through image detection by providing 1 image in each class. The image used for the testing is obtained at pexels.com and unsplash.com so that the result of being bias is can be prevented. Figure 8 shows the output (a)(b) of image testing using the best-model which got an accuracy of 99% for paper cup (a) and 97% for plastic cup (b).

2. Video Inference

Figure 9 shows the result of a single frame from the video tested using the model. The proponent provides 1 video per class which is obtained at pexels.com and fed it in the model testing. The video for paper cup got an accuracy of 98.36% based on calculations performed on 306 frames while the video for plastic cup got an accuracy of 99.99% on 240 frames. In fig. 9, a single frame from the video of paper cup and plastic cup got an accuracy 99% and 100% respectively. The chosen frame is both first frame of the videos.

IV. Conclusions

The usage disposable tableware such as beverage cups has been essential today. More and more disposable beverage cup are used per year but discarded after which often contribute to waste that sooner become waste pollution. Disposable cups wastes' do have other use once recycled. In this study, the proponents focused on creating an object detection model for disposable cups as it is one of the top 10 contributors to waste pollution. Automatic detection of disposable cups can help existing smart sorting trash can to collect garbage and used it and more meaningful way. Therefore a paper cup and plastic cup classification model was developed. The model used Mask R-CNN architecture, together with ResNet-101 as the backbone, to create a model which can have a good accuracy in classifying those disposable beverage cups. The best model produced has a mAP of 98.75% and also obtained impressive output during the testing stage. This suggest that Mask R-CNN is a good option to use when detecting such objects.

V. Recommendations

For future work, the proponent recommends to use authentic images for the datasets. Also, be mindful of the data that will be used since the class specially about plastic cups do have semi-transparent objects that sometimes the machine tends to learn what's inside of the cup and eventually results to false detection. Future researchers may implement the said model to their future work regarding on smart bin collection and the likes. Lastly, explore other different deep learning algorithm and compare it.

Acknowledgment

The completion of this paper could not be possible without the presence and assistance of the people throughout the writing. Likewise, the author would like to express his sincerest gratitude to the professor, department, institution, or university which he is connected.

References

[1] H. Li, H. Guo, N. Huang, and J. Ye, "Health risks of exposure to waste pollution: Evidence from Beijing," *China Econ. Rev.*, vol. 63, no. May, p. 101540, 2020, doi: 10.1016/j.chieco.2020.101540.

[2] S. Majchrowska *et al.*, "Deep learning-based waste detection in natural and urban environments," *Waste Manag.*, vol. 138, no. June 2021, pp. 274–284, 2022, doi: 10.1016/j.wasman.2021.12.001.

[3] "Disposable Cup Definition | Law Insider." https://www.lawinsider.com/dictionary/disposable-cup (accessed Dec. 11, 2022).

[4] J. Potting and E. van der Harst, "Facility arrangements and the environmental performance of disposable and reusable cups," *Int. J. Life Cycle Assess.*, vol. 20, no. 8, pp. 1143–1154, 2015, doi:

10.1007/s11367-015-0914-7.

[5] V. Šuškevičė and J. Kruopienė, “Improvement of packaging circularity through the application of reusable beverage cup reuse models at outdoor festivals and events,” *Sustain.*, vol. 13, no. 1, pp. 1–18, 2021, doi: 10.3390/su13010247.

[6] “The Problem with Disposable Cups - Plastic EDU.” https://plastic.education/the-problem-with-disposable-cups/ (accessed Dec. 10, 2022).

[7] M. A. Bilek, H. J. Salem, R. Korehei, and J. A. Olson, “Recycling Paper-Plastic laminate coffee cups using a Single-Disk Refiner: Energy requirements and recovered fiber quality,” *Waste Manag.*, vol. 136, no. September, pp. 104–112, 2021, doi: 10.1016/j.wasman.2021.10.008.

[8] Y. Ma, “Problems and resolutions in dealing with waste disposable paper cups,” *Sci. Prog.*, vol. 101, no. 1, pp. 1–7, 2018, doi: 10.3184/003685017X15129981721365.

[9] B. Biswal, S. Kumar, and R. K. Singh, “Production of Hydrocarbon Liquid by Thermal Pyrolysis of Paper Cup Waste,” *J. Waste Manag.*, vol. 2013, pp. 1–7, 2013, doi: 10.1155/2013/731858.

[10] S. Puig and N. Foukia, “CleverTrash: an IoT system for waste sorting with deep learning,” pp. 1–8, 2022, doi: 10.1109/ithings-greencom-cpscom-smartdata-cybermatics55523.2022.00016.

[11] “Plastic and Paper Cups Synthetic Image Dataset | Kaggle.” https://www.kaggle.com/datasets/vencerlanz09/plastic-and-paper-cups-synthetic-image-dataset (accessed Dec. 11, 2022).

[12] “Labelme: the Basics and a Quick Image Segmentation Tutorial.” https://datagen.tech/guides/image-annotation/labelme/ (accessed Dec. 11, 2022).

[13] “Everything about Mask R-CNN: A Beginner’s Guide - viso.ai.” https://viso.ai/deep-learning/mask-r-cnn/ (accessed Dec. 11, 2022).

[14] “Understanding Mask R-CNN Basic Architecture - ResNet backbone - RPN - RoIAlign - object detection branch - mask generation branch.” https://www.shuffleai.blog/blog/Understanding_Mask_R-CNN_Basic_Architecture.html (accessed Dec. 11, 2022).