IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-11, Nov 2018 E-ISSN: 2347-2693

A Prototype for Mobile Application of Garbage Quantification using Mask R-CNN

B Venugopal[1*] , L S Chakravarthy[2] , L Praveen[3] , Ashish Kumar Dwivedi[4]

[1,2,4] Department of CSE, Gayatri Vidya Parishad College of Engineering(A), Visakhapatnam, AP, India
[3]Director of Peritus Technologies, It Sez, Madhurawada, Visakhapatnam, AP, India

**Corresponding Author: venugopalboddu88@gmail.com Tel: +91-8897917920*

Available online at: www.ijcseonline.org

Accepted: 20/Nov/2018, Published: 30/Nov/2018

***Abstract*—** In cities especially some areas have garbage regions, So people are affected by several health issues. The main problem is authorities will not clean on time due to lack of information. Sometimes, authorities have also highly impossible to track these areas. Garbage quantification is an important step in improving the cleanliness of the cities. This paper presents one mobile application for garbage images with GPS locations to send authorities directly. When the user clicks the garbage image using through this app, then it will send that image to the server for automatic garbage detection with quantification by using the deep learning in computer vision techniques. Convolutional Neural Network (CNN) algorithms will be used to garbage detection with quantification in an image for accurate results.

Keywords— Garbage Quantification, Garbage Detection, Deep Learning, Computer Vision, Convolutional neural networks.

I. INTRODUCTION

The scrap in urban can be interpreted as the useless or wastage of items are dumping improperly in cities[1]. In recent days, most of the researchers are found that urban littering is the main concern in our country. As public parks and street roads are making as garbage dumps, causing threats to the people health. The primary thing to properly maintain urban cleanliness is to implement a continuous improvement management system[2]. In the past, there were several mobile apps that allow people to report authorities by uploading images of garbage with manual verification. So, there is a possibility to upload illegal images in those apps. That's why this proposed application is very useful to people for uploading every image to a server for automatic garbage detection using deep learning. But, before that should train the machine to sample data set with garbage and non-garbage images.

II. REALTED WORK

In the past, there were several mobile applications for uploading garbage images manually without the help of the machines. But they were failed to produce accurate results. As there is a scope to upload illegal images such as people and selfies, this paper proposes an effective solution for detecting garbage in urban areas with more accuracy.

One such type of most used effective algorithm in data science is Convolutional neural networks. An effective application can be designed with the adequate user interface that shows How much amount of garbage and detection of garbage in an image.

The following SIFT algorithm was used by many researchers for detecting garbage that has relatively accurate.

Scale Invariant Feature Transform was invented by David G. Lowe in 1999. The SIFT is a technique for portraying and detecting regional features in a given image[6],[7],[19].
The following steps are the classification and detecting objects of images in SIFT.

Scale-space Extrema Detection used to detect reliable key point locations in the given range of dimensions.

Key point localization used to design capable model to find scale values and region and select key points with the parameter called stability.

Orientation assignment used to compute the best position for key point region.

Key-point Descriptor will Use regional image ascents at selected position and transposition to describe each keypoint region.

Indexing and Matching used to create a hash table with labels of sample images.

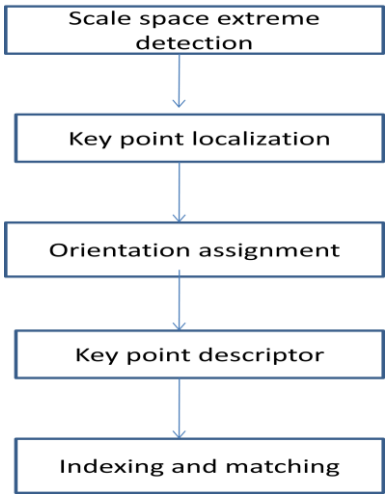


Figure 1: SIFT Architecture

The drawback is that it is mathematically over complex and needs heavy assessment and is not effective for low powered devices. So that, to improve processing power, we will use CNN technique for classifying an image.

III. METHODOLOGY

A. Convolutional Neural Network(CNN)

Convolutional neural networks are one of the main algorithms in the area of deep learning and very effective in areas such as image recognition and classification. It contains four layers compared to artificial neural networks as it has only three layers like an input layer, hidden layer, and output layer. CNN use relatively little pre-processing compared to other image classification algorithms[8],[9].

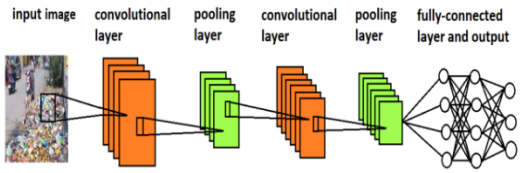


Figure 2: An Exampe of a convolutional neural network.

Traditional methods like SIFT and HOG models could not achieve accurate results of detection in standard datasets such as PASCAL VOC. These models conceal a very low-level characteristic of the objects and therefore are not able to distinguish well among the different labels.

Deep learning (Convolutional neural networks) based methods have become the state of the art in garbage detection in an image[10],[11].

Classification is named as a supervised learning problem. Since it identifies a different category of objects like garbage in a given image and accustomed for training the model to classify the image according to labels of garbage images.

Object Detection is one of the classical problems of computer vision and is often described as a difficult task. It is the process of spotting the instances of objects like faces, bicycles, and buildings in an image. But, this paper describes that detecting object like garbage particles in an image.

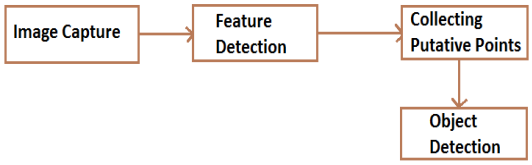


Figure 3: Block Diagram for Object Detection

Segmentation is the process of separating something into parts or segments. In the case of garbage image, it is the process of dividing of an image into multiple constituents that share similar attributes like color, to simplify the representation and making it more useful for the evaluating the performance.

Figure 4: Garbage Detection on Segmentation

B. Mask R-CNN

Mask R-CNN is a detailed segmentation model that allows users to determine the pixel-wise region for our class that has been classified using classification[14]. It is very distinct from classical garbage detection models like Faster R-CNN as it is where, in addition to finding the class and location of bounding box[15],[16]. It will work in two phases. a) scanning the image with object b) classification step. Mask R-CNN is a fusion of a Faster R-CNN and Fully Convolutional Network (FCN). Faster R-CNN does garbage detection and FCN does pixel-wise boundary in an image. To do this Mask R-CNN uses the Fully Convolutional Networks(FCN)[17].

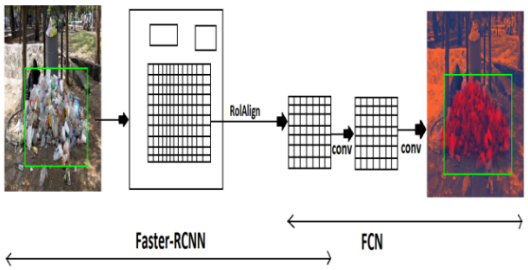


Figure 5: Mask R-CNN → Faster R-CNN + FCN

Fully Convolutional Network (FCN) is a sequence of convolutional network layers and it is effectively suitable for doing semantic segmentation. It is built only from locally connected layers, such as convolution, pooling and upsampling. This network uses various blocks of convolution and max pool layers to first decompress an image to 1/32th of its original size. It then makes a class prediction at this level of coarseness. Finally, it uses upsampling and deconvolution layers to resize the image to its original dimensions.

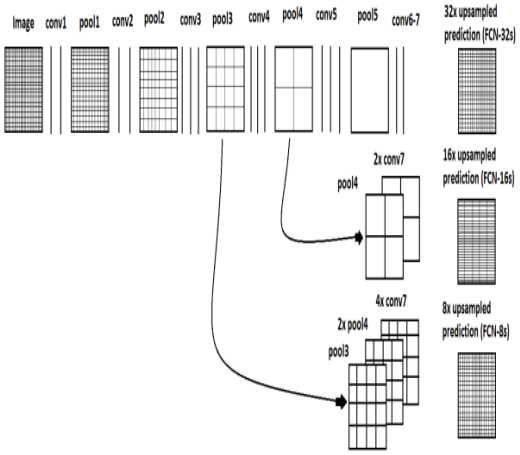


Figure 6: FCN Architecture

There are 3 categories of FCN architecture, which mainly differ in the spatial precision of their output. Above figure shows the FCN-32, FCN-16, FCN-8 variants[18].

FCN-32 will produce the segmentation map from conv7 directly by using a transposed convolution layer with the sliding size of 32.

FCN-16 performs Sums of 2x upsampled prediction from conv7 with pool4 and then produces the segmentation map, by using a transposed convolution layer with stride 16 on top of that.

FCN-8 also will Sum up the 2x upsampled conv7 with pool4, upsamples them with stride 2 transposed convolution and sums them with pool3, and applies a transposed convolution layer with stride 8 on the resulting feature maps to obtain the segmentation map.

IV. DATASETS

A. COCO Dataset

Common Objects in Context(COCO) is a dataset that aims to enable future research for object detection, instance segmentation, image captioning, and person keypoints localization.

The annotations include pixel-level segmentation of object belonging to 80 categories, keypoint annotations for person instances, stuff segmentations for 91 categories, and five image captions per image.

B. PASCAL VOC

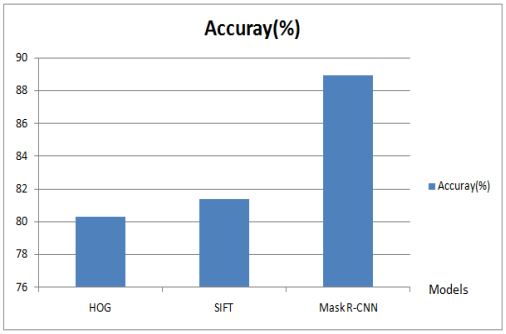
The PASCAL VOC datasets were provided as part of the PASCAL Visual Object Classes challenge from 2005 to 2012. The goal of the datasets is to recognize objects from a number of visual object classes in realistic scenes. It is fundamentally supervised learning images are provided.

C. GARB Dataset

This paper uses GARB dataset which is gathered a number of garbage and non-garbage images with size 200x256. The effective model Mask R-CNN trained on this dataset for classifying and detecting garbage in images.

Figure 7: The sample of garbage images

V. EXPERIMENTAL RESULTS



Graph 1. Comparison of different models to detect garbage

Model	Accuracy (%)
HOG	80.32
SIFT	81.4
Mask R-CNN	88.9

Table 1: Comparison of different models to detect garbage

The above table 1. describes of the different models for garbage detection in images with accuracy. The traditional models HOG and SIFT are could not give better results and failed to give accurate results in garbage detection. The Mask R-CNN model has been successfully generated accuracy with 88.9% for classifying and detecting garbage in images.

Figure 8: Left - input image showing garbage, Right - coarsely segmented output image

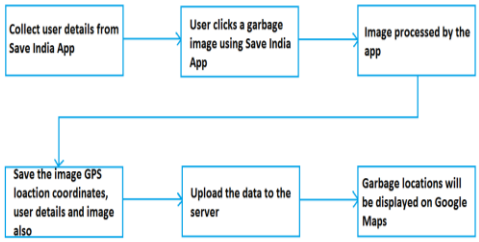


Figure 9: Save India app workflow.

Save India mobile application is designed to upload garbage images to the server for automatic garbage detection in an image. Figure 9 describes the workflow of the application and the following steps describe in detail, the process of our mobile application.

Step 1: To upload garbage images into our mobile application, before users must and should enter their details like name and phone number.

Step 2: After collecting details, users can allow taking a garbage image through our mobile application.

Step 3: This step takes a garbage image for processing of the application.

Step 4: In this step, the application does save the Geo-tagged (GPS) location coordinates, user details, and garbage image locally.

Step 5: Here, the application does upload the entire saved data of garbage image to the server. All classification and detection part is done with Mask R-CNN model in the server. The required output will be displayed in the web application.

Step 6: The garbage image location markers will be plotted on Google Maps. But these locations can be accessed by the admin only in the web application.

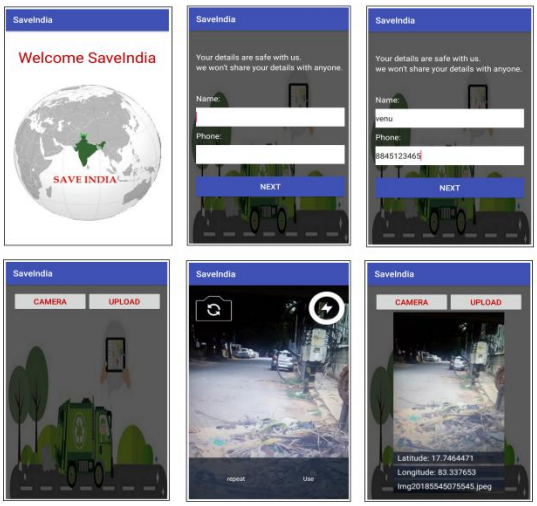


Figure 10. Save India app screenshots.

VI. CONCLUSION AND FUTURE WORK

Nowadays, this mobile application is very useful to people for maintaining cleanliness of the cities. The user clicked garbage images are processed by Mask R-CNN model which is a submodel of Convolutional neural networks. This model has been successfully generated accurate results of detecting garbage and it's quantification in an image for comparing other traditional models. In our future work, we can add more garbage and non-garbage images to the dataset to improve more accurate and better results for garbage detection with quantification in an image.

ACKNOWLEDGMENT

The authors are very grateful to donors who helped to give the garbage and non-garbage images dataset[1] for this research.

REFERENCES

[1] Mittal, Gaurav, et al. "SpotGarbage: smartphone app to detect garbage using deep learning." Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM, 2016.

[2] Rad, Mohammad Saeed, et al. "A Computer Vision System to Localize and Classify Wastes on the Streets." International Conference on Computer Vision Systems. Springer, Cham, 2017.

[3] Wang, Liming, et al. "Object detection combining recognition and segmentation." Asian conference on computer vision. Springer, Berlin, Heidelberg, 2007.
[4] Erhan, Dumitru, et al. "Scalable object detection using deep neural networks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014.
[5] Lowe, David G. "Distinctive image features from scale-invariant keypoints." International journal of computer vision60.2 (2004): 91-110.
[6] Brown, Matthew, and David G. Lowe. "Invariant features from interest point groups." BMVC. Vol. 4. 2002.
[7] Abdel-Hakim, Alaa E., and Aly A. Farag. "CSIFT: A SIFT descriptor with color invariant characteristics." Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. Vol. 2. IEEE, 2006.
[8] Simard, Patrice Y., Dave Steinkraus, and John C. Platt. "Best practices for convolutional neural networks applied to visual document analysis." null. IEEE, 2003.
[9] Ji, Shuiwang, et al. "3D convolutional neural networks for human action recognition." IEEE transactions on pattern analysis and machine intelligence 35.1 (2013): 221-231.
[10] Kalchbrenner, Nal, Edward Grefenstette, and Phil Blunsom. "A convolutional neural network for modelling sentences." arXiv preprint arXiv:1404.2188 (2014).
[11] Rastegari, Mohammad, et al. "Xnor-net: Imagenet classification using binary convolutional neural networks." European Conference on Computer Vision. Springer, Cham, 2016.
[12] Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." arXivpreprint arXiv:1704.04861 (2017).
[13] Shi, Wenzhe, et al. "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
[14] He, Kaiming, et al. "Mask r-cnn." Computer Vision (ICCV), 2017 IEEE International Conference on. IEEE, 2017.
[15] Ren, Shaoqing, et al. "Faster R-CNN: towards real-time object detection with region proposal networks." IEEE Transactions on Pattern Analysis & Machine Intelligence 6 (2017): 1137-1149.
[16] Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems. 2015.
[17] "Building a Mask R-CNN Model for Detetcting Car Damage " at https://www.analyticsvidhya.com/blog/2018/07/building-mask-r-cnn-model-detecting-damage-cars-python/
[18] "Fully Convolutional Networks (FCN) for 2D segmentation" at http://deeplearning.net/tutorial/fcn_2D_segm.html
[19] "Introduction to SIFT (Scale-Invariant Feature Transform)" at https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_sift_intro/py_sift_intro.html

Authors Profile

B Venugopal completed his B.Tech in Computer Science and Engineering from Qis college of engineering and technology, Ongole, Prakasam(d.t), AP, India(2015). He is currently pursuing M.Tech in Computer Science and Engineering in Gayatri Vidya Parishad college of engineering(A), Visakhapatnam, Andhra Pradesh. His main area of research interest is Artificial Intelligence.

L S Chakravarthy had received his Doctoral degree from Dept. of Computer Science & Systems Engineering, Andhra University, Visakhapatnam, AP, India(2015). Masters in Computer Science Engineering from Andhra University, AP, India (2009). Masters in Applied Mathematics from Andhra University, AP, India (1997). AMIETE in Computer Science & Engineering from Institute of Engineers, New Delhi (2016). He is currently working as an Associate Professor in the Department of Computer Science & Engineering, Gaytri Vidya Parishid (A), Visakhpatnam, AP, India. His main research interests are Game Theory, Data Engineering, Privacy, Artificial Intelligence, Machine Learning, Cognitive & Soft Computing Techniques, Computer Vision.

L Praveen completed his M.Tech in Computer Science and Engineering from Andhra University, Visakhapatnam, AP, India. He is director of Peritus Technologies. It is located in It Sez, Madhurawada, Visakhapatnam, AP, India.

Ashish Kumar Dwivedi completed his M.Tech in Computer Science and Engineering, National Institute of Technology, Rourkela, India(2014). Ph.D. in Computer Science and Engineering, National Institute of Technology, Rourkela, India, (He submitted thesis on 2 nd August, 2017). He is currently working as an Assistant Professor in the Department of Computer Science & Engineering, Gayatri Vidya Parishad college of engineering(A), Visakhapatnam, AP, India.

