Grenze International Journal of Engineering and Technology, July Issue

Geo Mapping of Waste: A Bottom-Up Approach using Android and SVR

T.R. Krishnaprasad[1], Ravipati Jhansi[2], Sai Keerthana[3] and K Namitha[4]
[1-4] Amrita School of Engineering/Dept of CSE, Amritapuri, India
Email: krishnaprasadtr@am.amrita.edu, rjhansillakshmi@am.students.amrita.edu,
csaikeerthana@am.students.amrita.edu, namithak@am.amrita.edu

***Abstract*—With the increase in industrial output, a large amount of garbage results as a byproduct. The question of how to collect garbage, particularly in areas where proper waste bins are not present requires serious consideration. This paper explores the issue of the mapping of waste that one encounters during daily walks onto a publicly available mapping platform. Our proposed model also uses a machine learning model to predict the volume of waste so that the municipal authorities can decide on the efficient placement of garbage at a very local level. The results of the simulated experiments are presented along with.**

***Index Terms*— Android application, mapping, classification, machine learning, waste**

I. INTRODUCTION

Increase in demand for goods around the world has also led to an increase in waste getting littered around. In India alone, more than ten thousand tons of garbage is produced every day [1]. Rapid urbanization along with unavailability of proper dumping grounds where the waste can be recycled are also one of the factors of pollution. Waste recycling and disposal of such a volume as mentioned earlier, will be effective only if the people of the respective locations also contribute and not merely leave the duty of waste collection to the municipal authorities [2]. This paper proposes a model of community participation in efficient waste identification. As more and more technological applications like Android are affordable, they are increasingly utilized to reduce a lot of problems facing the country today [3]. Smartphones based on Android OS made up close to 86 percent of total sales in the world [4]. Hence, we chose to develop our application on the Android OS platform to enable wide reach. Along with mobile handheld devices gaining popularity, applications that look to automate stages of their entire process if not the whole processes are also being rapidly incorporated into the real-world devices. One of the features of automation is the ability called as machine learning where the computers learn from inputs, resultant outputs and the expected outputs [5]. based on training the data set. Various machine learning models are available in the literature and one model may not work for every purpose and every data set. Each machine learning model must be evaluated for its effectiveness on a particular data set and the scenario under consideration. The proposed model presented in this paper also uses machine learning to classify waste and to predict the concentration of waste at a given point in the future.
The rest of the paper is structured as follows: section II looks into the related works, section III provides the details of the proposed model, section IV mentions the experiments and results while section V provides the future work.

Grenze ID: 01.GIJET.6.2.9
© Grenze Scientific Society, 2020

II. RELATED WORK

TABLE I. RELATED WORK SUMMARY

Name	Purpose	Method	Disadvantage
Mission Swachhta[6]	Bridge gap between end users and municipal authorities to attain a clean environment	Mobile Cloud based smart phone application, Firebase database	Does not determine the type of waste
Cloud Based Architecture for Solid Waste Garbage Monitoring and Processing [7]	Collect information of presence of air pollutants in garbage waste contributing air pollution	Deployed sensors at garbage. Process the collected data using AWS kinesis	No information about mapping of waste
Serious Game on Recognizing waste category[2]	Trains how to identify and sort waste into different categories using a virtual environment	Microsoft Kinesis SDK is used to track the movement of user's wrist	Does not show the count of waste items
Perazuhan[8]	To sell recyclable solid waste to junk collecting shops	Mobile application with officials as intermediaries	It is focused on selling waste in return for money
Mobile application design of geolocation to collect solid waste[9]	To geo locate any approaching waste gathering vehicles	Android app and Firebase as database	Is not meant for tracking waste in discrete and isolated locations

The work mentioned in [2] trains people to identify and sort waste by introducing a game. The trainee will find a red ball in the middle of the screen. A particular waste drops down from the top of the screen and the dustbins with proper labels are at the four corners of the screen. The trainee has to catch the waste with the help of ball and put the waste in dustbin where it belongs to. A mobile application where user(s) can register complaint regarding garbage in this app along with pictures or videos if needed, and the action by the authorities will be taken based on the gravity of the complaint is given in [6]. It updates the user regarding the status of complaint via SMS. In [7] a cloud-based solution to monitor and process solid waste garbage odor in crowded cities is provided. The proposed architecture helps in collecting the data from the garbage locations through sensors and process the data collected using AWS Kinesis. Some solid wastes can be recyclable and by collecting and selling solid wastes to junk shops people can earn money, which can also reduce the solid waste in the surroundings. So, they have proposed a mobile application in which users can register and login to their application and can sell recyclable solid wastes like bottles, newspapers, plastic to junk shops [8]. In [9] the application focuses on locating the waste collecting vehicles from the users.

III. PROPOSED MODEL

An overview of proposed model is given in Fig 1. This model has been categorised into four parts, namely, the user application part, the image classification part, comparison of algorithms and prediction of volume of waste at a given location.

A. *User Application*

The user application is developed using Android Studio and serves three primary purposes achieved via three UI's.

- The first purpose is to send it to the server.
- The second purpose is to display a map with markers over the area where the uncollected waste is located using the 'view garbage' button.
- Third purpose is to predict the waste in the locality by clicking on the marker of a given place

The XML activity class of Android is used to create the UI of the app with each widget assigned a unique ID. The coding of the user application is done in the Main activity class. Each button is associated with a onClickListener interface where the code required for the particular action is written. Fig 4 represents the scenario where the user, after capturing a waste (electronic waste) is about to send the picture to the server. Upon clicking the 'Submit' button (Fig 4), photo of the waste and the GPS location of the user will be sent to the server, to be stored into the database.

B. *Image Claasification Algorithm*

Once the image of waste is sent by the user through the mobile application, both the image and the GPS co-ordinates are received by the server. We use a Bag of Features (BoF) model to classify the image (waste) into a particular category. In the given visual context, a bag of words model seeks to create a 'vocabulary' of

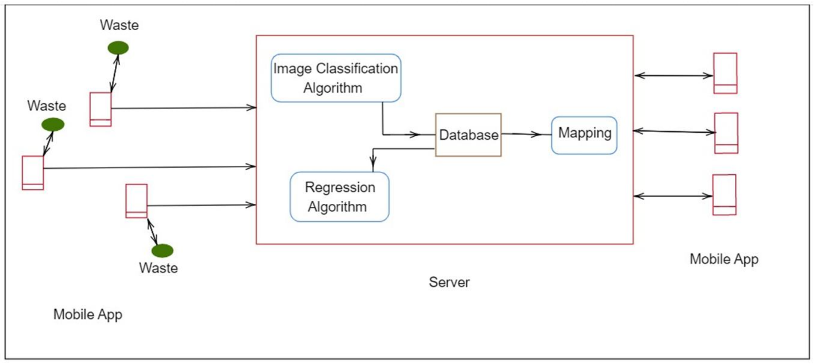

Figure 1. Overview of Proposed Model

features of an image classification. Initially the features are extracted using the pixels around a particular point of interest. Speed Up Robust Features (SURF) technique is employed to extract features from all images in each category. Then, each of these points become the central location in different neighborhoods. Taking all such constructed features into 'vocabulary' will lead to huge overload and hence the bag of words model uses K-means clustering algorithm to reduce the number of features. After reducing the number of features, the model represents the image as a histogram in which each word's frequency of appearance is calculated within a given image. This encoded image, in the form of a feature vector, from each category is given as input into a multiclass linear Support Vector Machine (SVM) classifier to classify the images.

C. Comparision of Algorithms

One of the objectives of our model is to predict the volume of waste at a municipal level but also at the level of a local residential group of households. For this we gather data from people through the user application and run the better fitting regression algorithm amongst the Linear Regression algorithm and the Support Vector Regression (SVR) algorithm.
In Linear Regression, the objective is trying to predict the value of a dependent variable given one or more independent variables. As the term denotes, linear regression can only be applied in scenarios where there is a directly or inversely proportional relationship between the dependent variable (y) and the independent variable(s) (x). Quite often y is a factor of multiple variables (x_1, x_2...) and is denoted by the equation:

$$y = m * x_1 + m * x_2 + c \qquad (1)$$

where x_1 and x_2 are the factors upon which influences the variable y, m is the slope and c is the error margin. There can be more than two factors depending upon the scenario involved. Now, linear regression tries to draw a line that best fits the delineation of the points plotted in the graph.

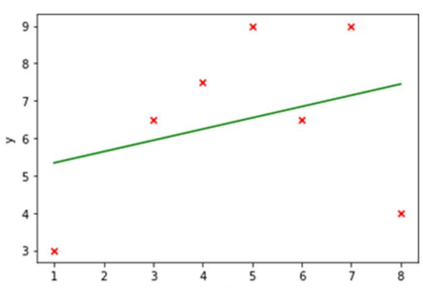

Figure 2. Sample Linear Regression Plot

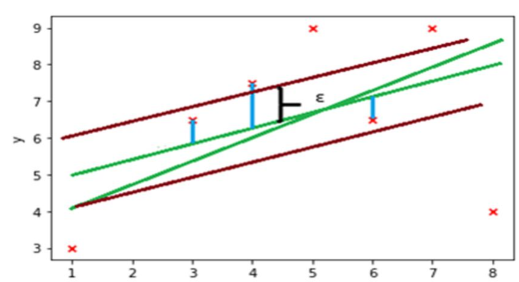

Figure 3. Sample SVR Plot

The best fit line is that line which reduces the error between the points and the line as much as possible. A sample representation of a linear regressive line is given in Fig 2. We also compare the performance of the SVR model against the Linear Regression model. SVR is based on the Support Vector Machine model. Although SVM is primarily used for classification, it is also used for regression purposes. SVM draws margins to classify groups and it tries to maximize the margin from the nearest points such that the groups are clearly delineated. Normally in a regression view, the margin of error is tried to be reduced, in SVR the idea is to minimize the coefficients and by limiting the error within a specified margin ε as shown in the sample plot given below in Fig 3. As shown in Fig 3, ε represents the error margin on either side of the best fitting line. We can see that one of the points is lying outside the ε margin and thus is excluded.

IV. Simulation and Results

Our simulation of the application involves four stages:

- Mapping of waste
- Dataset analysis
- Categorizing the waste
- Estimating the quantity of waste
- Displaying the estimated waste quantity in the map

A. *Mapping of Waste*

For this work, we have used waste which are delineated into three categories: plastic, e-waste and cardboard. The user uploads an image to the server as shown in Fig 4. The image of waste sent by the user is received by the server along with the GPS co-ordinates of the user. The location from which each image is uploaded is geo-mapped by the server. Any user who wishes to view the geo-mapped locations of waste needs only to click the 'View location info' button and the map as shown in Fig 5 will be displayed onto the screen. When a user clicks on the 'View location info' button all the images which are mapped already, are displayed to the user.

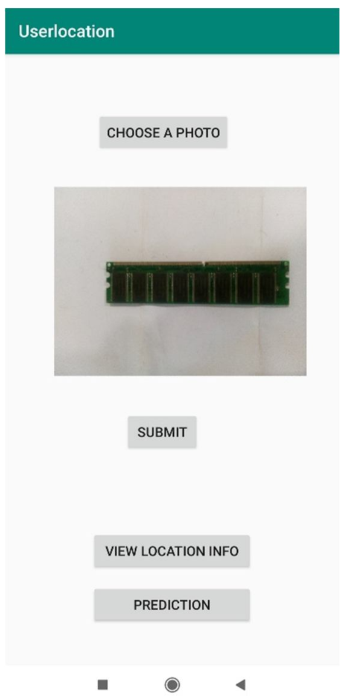

Figure 4. Select image

Figure 5. View map

TABLE II. Average Weight in Grams

Waste Category	Average weight
Plastic	0.003
e-waste	0.152
Cardboard	0.013

TABLE III. Sample Data from the City Level Data Set

Population	MSW
1033756	124538
1001694	216504
268243	29200

B. Data Set Analysis

We calculated the weight of waste material and have tabulated the average weight of a material in the Table II. Upon similar exercise for other waste materials it is possible to come up with the net weight of waste materials accumulated at a precise location.

The average weight of plastic cover, e-waste and cardboard is given in Table II. We use the data from Worldbank[10] which provides a city wise waste collected information. In our case, the waste produced within a much smaller locality can be measured as mentioned in Table II, whereas the Worldbank[10] data set covers a city. We thus, use the data set from Worldbank due to it being a prior measured real-life data source which provides for efficient evaluation. We have only included those localities which are within India and which do not have any empty or zero values in food, glass, metal, paper and plastic waste categories. A sample data from the data set featuring only the population and waste columns is given in Table III.

MSW stands for Municipal Solid Waste which is represented in tons. The idea is that just like the city level measuring of waste, we can do the same in municipal level, or panchayat level or even in residential colonies so that each individual becomes aware of the waste generated by her/ him. To determine which machine learning algorithm to apply, we observe the pair plot as given in Fig 4 below:

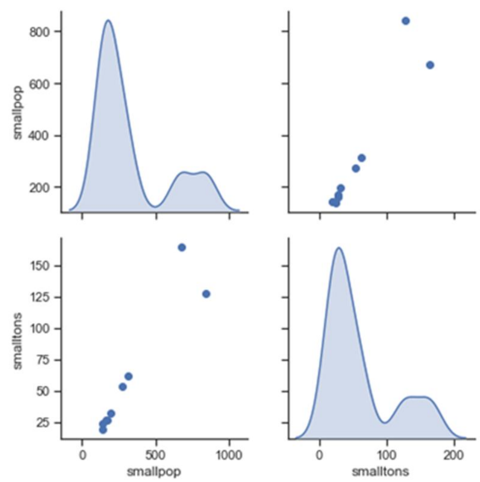

Figure 4. Pair Plot

As we can see from Fig.4 the waste produced is more or less directly proportional to the population of the locality. Hence, we ran the dataset on both the linear regression and the SVR algorithms. The software environment utilized for this project involves Anaconda[11], Python[12], Scikit-learn[13], matplotlib[14] and seaborn[15].

C. Categorising the Waste

The BoF encodes each image into a feature vector and represents it as a histogram of visual word occurrences. These encoded training images from each category are used as input to a classifier. We evaluate the classification results of images in three categories: cardboard waste, e-waste and plastic waste using confusion matrix and metrics such as precision, recall and F1-score.

(a) Plastic (b) Plastic (c) Cardboard (d) E-waste (e) E-waste (f) E-waste

Figure 7. Sample training images in different categories of waste.

(a) Plastic (b) Plastic (c) Cardboard (d) E-waste

Figure 8. Sample testing images in different categories

Sample images used for training and testing are given in Fig 7 and Fig 8, respectively. The confusion matrix obtained for the three classes is shown in Table IV. The average accuracy achieved for the classification is 78%. The precision, recall and F1 scores computed for the test data is presented in Table V. We can see that the test images in each class achieve high scores for precision, recall and F1. The output category is then fed into the database for later use in the machine learning algorithm.

TABLE IV. CONFUSION MATRIX FOR THE TEST DATA

	Known		
Predicted	Cardboard	E-Waste	Plastic
Cardboard	1	0	0
E-Waste	0	0.33	0
Plastic	0	0.67	1

TABLE V. PRECISION, RECALL, F1 SCORES FOR ALL CLASSES

Class	Precision	Recall	F1 Score
Cardboard	1	1	1
E-Waste	1	0.33	0.4962
Plastic	0.5988	1	0.7491

D. Estimating the quantity of waste

Using the data set, we try to predict the quantity of waste which will be accumulated in a given locality using Linear Regression and SVR algorithms. We divide the training and testing data in 80:20 ratio for both the machine learning algorithms. The linear regression plot obtained is given below in Fig 9 and the result of

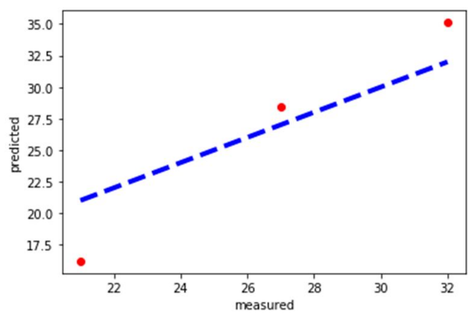

Figure 9. Result of Linear Regression plot

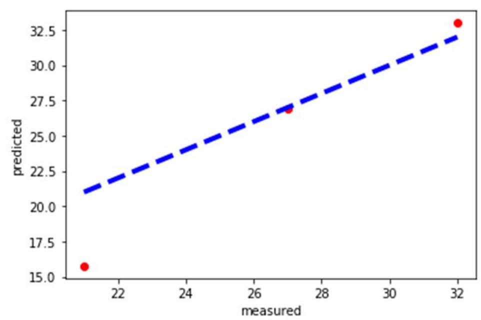

Figure 10. Result of SVR plot

TABLE VI. PERFORMANCE COMPARISON OF THE TWO MODELS

Model	MAE	RMSE
Linear Regression	3.1	3.4
SVR	2.1	3.0

SVR is given in Fig 10. By comparing both the Fig 9 and Fig 10 we can observe that SVR regression delineates the data points better than the Linear Regression plot. This is even further reinforced when we look at the performance metrics of the two algorithms. Metrics that are very common for evaluating the accuracy are the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). A comparison of MAE and RMSE of the two machine learning models are given in Table VI. From Table VI, we can infer that SVR has much less error relative to linear regression. The predicted output of both Linear regression and SVR models are shown in Fig 11 given below.

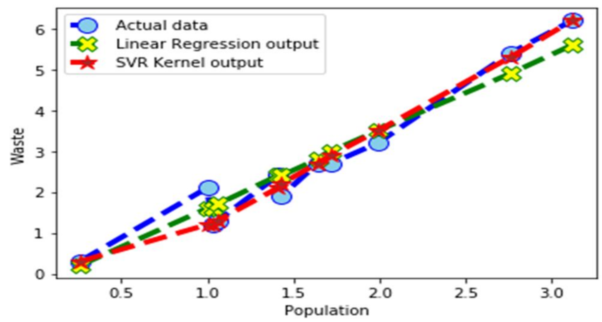

Figure 11. Actual values and Predicted values

From Fig 11 also we can observe that SVR is better in predicting values that are much closer to the actual values relative to the predicted output of Linear regression, both when the population count is low and even when the population count is high. Hence, we include the SVR algorithm in this model.

E. Displaying the estimated waste quantity in the map

The predicted waste by SVR is then sent to the mobile app as showing in Fig. 12, upon the user clicking the marker of the location.

Figure 12. Predicted value of the volume of waste

F. Conclusion and Future work

Waste littered in the area is uploaded onto the server by the user using the Android application. The uploaded waste is then geo mapped for the benefit of other individuals. The SVR algorithm is chosen over the linear regression algorithm after due analysis of data set for predicting the amount of waste at a particular locality. Once the user gets to know the predicted volume of waste generated at a given locality, then the individual will be more responsible towards to waste dumping and will contribute more towards waste recycling and reuse. As part of our future work, we would like to include more complex and varied waste categories, like drainage blocking a canal, classification of individual waste from a collection of heterogeneous waste etc.

We would also like to provide a real time web-based application that would communicate to the authorities of any serious obstacles at a particular locality along with increased interactivity on maps.

REFERENCES

[1] M. V. Ramesh *et al.*, "Water quality monitoring and waste management using IoT," 2017, doi: 10.1109/GHTC.2017.8239311.

[2] B. M. Menon, R. Unnikrishnan, A. Muir, and R. R. Bhavani, "Serious game on recognizing categories of waste, to support a zero waste recycling program," 2017, doi: 10.1109/SeGAH.2017.7939292.

[3] A. Jayakumar, M. Raghunath, M. S. Sakthipriya, S. Akhila, A. Sadanandan, and P. Nedungadi, "Enhancing speech recognition in developing language learning systems for low cost Androids," 2016, doi: 10.1109/ICCTICT.2016.7514556.

[4] Gartner.com, "Gartner Says Worldwide Sales of Smartphones Recorded First Ever Decline During the Fourth Quarter of 2017," *Gartner*. pp. 1–6, 2018, [Online]. Available: https://www.gartner.com/newsroom/id/3859963.

[5] K. V. S. Reddy, K. S. Nihith, M. S. Chowdary, and T. R. K. Prasad, "An Efficient Word Embedded Click-Bait Classification of YouTube Titles Using SVM," 2020, doi: 10.1007/978-981-15-4301-2_15.

[6] A. Sharma, "Mission swachhta: Mobile application based on mobile cloud computing," 2020, doi: 10.1109/Confluence47617.2020.9057926.

[7] G. Rashmi, M. Ameenulla, and S. Sathish Kumar, "Cloud Based Architecture for Solid Waste Garbage Monitoring and Processing," 2019, doi: 10.1109/CSITSS47250.2019.9031040.

[8] P. M. Ballaran, C. B. A. Corpuz, L. A. R. Paras, B. S. Fabito, and E. R. Rivera, "Perazuhan: A Mobile Application for Solid Waste Micro-Management Framework," 2019, doi: 10.1109/SCORED.2019.8896301.

[9] A. Delgado and J. Sosa, "Mobile application design of geolocation to collect solid waste: A case study in Lima, Peru," 2019, doi: 10.1109/INTERCON.2019.8853594.

[10] The World Bank, "What a Waste Global Database | Data Catalog." 2018, [Online]. Available: https://datacatalog.worldbank.org/dataset/what-waste-global-database.

[11] Anaconda, "vers. 2-2.4.0, Anaconda Software Distribution. Computer software," *Anaconda Software Distribution. Computer software*, 2016. .

[12] F. L. Van Rossum, Guido and Drake, "Python 3 Reference Manual." 2009.

[13] P. Pedregosa, F; Varoquaux, G; Gramfort, A; Michel, V; Thirion, B; and Grisel, O. and Blondel, . and Prettenhofer *et al.*, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, 2011, doi: 10.1007/s13398-014-0173-7.2.

[14] MATLAB, *version 7.10.0 (R2010a)*. 2010.

[15] M. W. And *et al.*, *mwaskom/seaborn: v0.8.1 (September 2017)*. 2017