ISSN No:-2456-2165

Machine Learning Approach for Solid Waste Categorization in Ethiopia

Sebahadin Nasir, Zewd Ayalew, Md Nasre Alam, Demeke Getaneh, Anteneh Tiruneh
Woldia University, Department of Computer Science, Woldia, Ethiopia

Abstract:- Solid waste categorization is challenging process but it is an important for recycle and disposal wastes in a proper way. Currently in Ethiopia solid waste are categorized manually as recyclable, combustible and compostable. We propose the machine learning approaching order to categorize the solid waste as recyclable, combustible and compostable through the machine. The scope of the study is to detect and categorize the solid waste. The solid waste collected from household, street sweeping, hotels, industries and other industries from Addis Ababa, Ethiopia. For experimental purpose we collected total 2445 images. Among these we found 650images are recyclable,865 images are combustible and 930 images are compostable category. The overall accuracy of our designed system is 89% and the model achieved 89%, 82% and 96%inrecyclable, combustible and compostable category respectively. The designed approach accuracy result is compared with manually identified categories and the average percentage error is 10.82% and the designed approach performs closer to the ground truth categorized manually.

Keywords:- *Solid Waste, Digital Image Processing, Image Segmentation, Machine Learning, Artificial Neural Network, Object Classification.*

I. INTRODUCTION

Waste is a resource if it is managed and used in its proper place otherwise it disturb the society and it causes different diseaselike flue, cholera, typhoid, etc. especially in undeveloped countries. In Ethiopia, open dumping; burning is a common practice, consequential in endless odor, flooding, and outbreak of diseases [1]. In Addis Ababa Ethiopia, Repiie or Koshe is the biggest cite and growing until it became a mountain as shown in Figure 1.1[2], in the city waste collectors are collecting wastes from household, street sweeping, Hotels, Industries and other institutions. Then they are dumping in local temporarily sites to identify recyclable and other before they are taking it to main dumping cite Koshe.

Currently Ethiopia is building the first waste-to-energy project in Africa [3]. However there is no computer assisted system that identify waste as recyclable, combustible and compostable. So we proposea Machine Learning Approach design for Solid Waste Categorization in Ethiopia using digital image processing methods. We collected total 2445 images that contains 650 images for recyclable, 865 images for combustible and 930 images for compostable category. As shown in Table 1, the recyclable category contains the waste that can be recycled like metal, plastic and paper, the combustible category contains the waste that must be burn and converted in to energy like plastic, leather, medicine and textiles. The other waste category is compostable containing damaged vegetables, foods and garden waste that can be used in agriculture to fertile the soil after some process.

Fig 1:- Koshe Waste Dumping Site

The captured images are preprocessed using median filter because itis capable for reducing *salt-and-pepper* noise formthe image[4]. Threshold segmentation algorithm is selected as segmentation algorithm that divide the image pixels with respect to their intensity level and are used over images having lighter objects than background[5], so that it is selected since the region of interests are in collected images to be detected are light than the background region. Among presently available different classification method such as support vector machine

(SVM), Bayesian classifier, K-Nearest Neighbor (KNN), Artificial neural network (ANN) and so on. For this study, ANN is selected because of it is very fast in detection and classification even it is slow in training [6].

II. RELATED WORKS

In the beginning of our study, in the literature review some related works done by different authors that detect and label solid waste according to their categories using image processing techniques. In robot based system, the hardware and the software working in collaboration to categorize the solid waste. Solid waste categorization can be done in two different ways: first way is the incoming solid waste in the garbage collector can be detected and categorized according to its class and in the second way solid waste are collected from household and etc. after they mixed together in the garbage and in this way waste are affecting each other in color and particles.

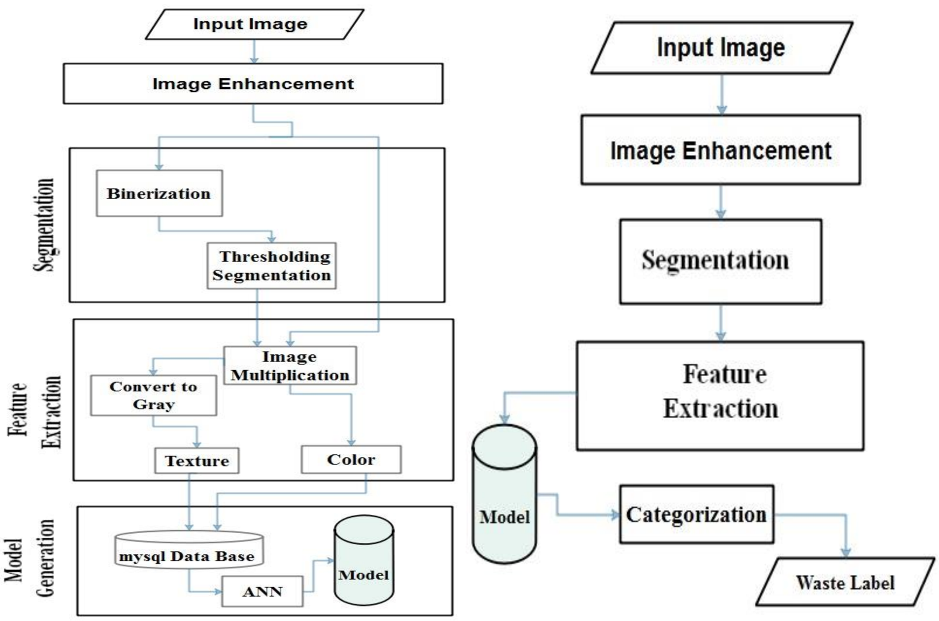
Yijian Liu and *et al.*[7] did a study on smart waste sorting system based on image processing and it contain both hardware and software part. In their work the used they used SURF-BOW feature extraction algorithm and multiclass SVM for classification and they achieved the overall classification accuracy of 83.38 %. However their system was designed to sort the solid waste from the beginning and to put it in different place in the same garbage collector, it is not after the waste is combined together. Once the solid waste is mixed together their visual appearance is changed that can affect the other by color and particles.

The intelligent waste separator study was conducted by Andres T and *et al.* [8] to separate the incoming waste and places it directly in different containers by using a multimedia embedded processor, image processing, and machine learning methods. However their system was designed to classify only plastic bottles, plastic cutlery and aluminium cane. It will be better if they consider all types of waste categories like recyclable, combustible, compostable and etc. An intelligent garbage classifier done by Alvaro S *et al.*, their system analyzesimages from camera and the robot arm and conveyor belt for visual classification [9]. The most relevant character they consider was shape and they used watershed for splitting an overlapping waste and K-NN for classification. Though they didn’t state the accuracy of the classifier. Using only shape is not good enough to detect objects because the same class of waste can have different size and shape. Since they are using shapes feature for classification, watershed has a problem of over segmentation.The general garbage detection system in the street was designed for intelligent urban management by Ying W. and Xu Z [10]. The objective of the study was to implement a deep learning strategy for automatic garbage detection in the street by training a Faster R-CNN open source framework with region proposal network and ResNet network algorithm. But the system is still not able to separate the solid waste formmixture of waste.Thesystem cannot classify the waste in different class, it can only detect the waste from other objects in the street.

We revieweda number of papers those are related to solid waste categorization, but all the above works has not addressed how to categorize solid waste from the mixture of waste as recyclable, combustible and compostable. In our study we designed a machine learning system that categorizes solid waste as recyclable, combustible and compostable using digital image processing techniques.

III. THE DESIGNED APPROACH FOR SOLID WASTE CATEGORIZATION

The designed system architecture for Machine Learning Approach for Solid Waste Categorization depicted in Figure 2 (a),(b).



(a) (b)

Fig 2:- The system Architecture: (a) Model Generation (b) Waste Categorization

4.1 Image Collection

An image collection is the first steps in the process of any image processing applications design. In this study we collect image data from Ethiopia, Addis Ababa City West Management Authority Agency Koshe dump site.

4.2 Image Enhancement

In designing any image processing applications, all images must pass through the image preprocessing and for waste categorization; we used median filtering operation which perform the image matrix value using the default 3 x 3 neighborhood pixels as shown in the equation (1).

$$I\_filterd(x,y) = \sum_{k=1}^{n} I\_original(x,y) \quad (1)$$

Where$n$ is the total number of pixels in 3 x 3 matrix region, x and y are integer numbers starting from 1, $I\_filterd$ is the enhanced image of an image and$I_original$.

4.3 Image Segmentation

Image segmentation is an important and challenging task of an image processing systems. We used global threshold segmentation algorithm [4].In global threshold segmentation methods, as shown in equation (2), pixels labeled 1 corresponding to objects, whereas pixels labeled 0 corresponding to the background and T has is fixed value [4].

$$I(x,y) = \begin{cases} 1 \ if \ f(x,y) \geq T \\ 0 \ if \ f(x,y) < T \end{cases} \quad (2)$$

Where I pixels arein threshold image and f pixels are in an input image to the segmentation algorithm, x and y are integers. The threshold point$T$ can be selected by visual inspection [4] as shown the image histogram in Figure 3 [4], on which the threshold value shown with clear distinction.

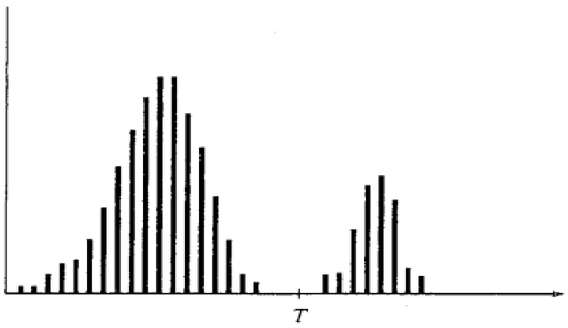


Fig 3:- The Image Histogram

Threshold value can also be selected by trial and error until the region of interest is retrieved [4]. Thus we choose 0.002 threshold value T by repeatedly tryingand observing the result using the algorithm depicted in Algorithm 4.1 [4]. In the algorithm, the initial thresholdvalue T can be computed by calculating the mid and max intensity value of an image.

1. Compute an initial T_0= mid_point_image_intensity_(min , max)
2. Segmentthe image using T and group pixels to G1 if f(x,y) ≥ T and group pixels to G2 otherwise
3. Compute average intensity(μ_1 and μ_2) for pixels in region G1 and G2.
4. Compute T_new value $T\_new = \frac{1}{2}(\mu_1 + \mu_2)$
5. Return to step 2 until T_new is less than an initial T_0

Algorithm 4.1 Threshold Value selection

4.4 Feature Extraction

For solid wastes detection and categorization we used two groups of features color and texture. There is total 6 color and 6 texture features are extracted using the algorithm depicted in Algorithm 4.2. These extracted features are stored in mysql database.

Feature_extraction (enhamnced_image I_filterd, I_bw)
{
For Each segmented region in an image
//Color Feature
I_RGB = Array multiply (I_filterd, I_bw)
R= extract red layer(I_RGB);
G=extract Green layer(I_RGB);
B=extract blue layer(I_RGB);
Compute R_avg=mean(R);
G_avg=mean(G);
B_avg=mean(B);
[H,S,V]= Convert RGB to HSV
Calculate H_mean=mean(H);
H_mean=mean(H);
S_mean=mean(S);
V_mean=mean(V);
I_gray=RGB_2_gary(I_RGB)
//Texture features
Compute statistical measure using statexture method
T1 = Average gray level (I_gray)
T2 = Average contrast (I_gray)
T3 = Measure of smoothness (I_gray)
T4 = Third moment (I_gray)
T5 = Measure of uniformity(I_gray)
T6 = Entropy(I_gray)
End
}

Algorithm 4.2 Feature Extraction

The Algorithm 4.2 extracts 12 features; it accepts filtered image from image enhancement step and binary image from threshold segmentation. Then, it performs element-by-element binary array multiplication operation to get the color version of the segmented region. Two set of features extracted in Algorithm 4.2 are described below: Color feature and Texture Features extraction.

4.4.1 Color feature Extraction

From the segmented RGB image region, the algorithm extracts R, G and B layers from the RGB color image in the given region. After extracting each layer it finds the average value of R, G and B of the image region using the equation (3), (4) and (5) respectively.

$$R = \frac{1}{N}\sum_{k=1}^{N} r(x,y) \quad (3)$$

$$G = \frac{1}{N}\sum_{k=1}^{N} g(x,y) \quad (4)$$

$$B = \frac{1}{N}\sum_{k=1}^{N} b(x,y) \quad (5)$$

Where x, y, k, and N are positive integers. In addition to RGB, the HSV color model is extracted from RGB color model and each H, S and V are extracted from HSV color model for the given region and then, mean values of each H, S and V are calculatedby using Equation 6, 7 and 8 respectively.

$$H = \frac{1}{N}\sum_{k=1}^{N} h(x,y) \quad (6)$$

$$S = \frac{1}{N}\sum_{k=1}^{N} s(x,y) \quad (7)$$

$$V = \frac{1}{N}\sum_{k=1}^{N} v(x,y) \quad (8)$$

Where x, y, k, and N are positive integers.

4.4.2 Texture Features Extraction

First the RGB image is converted in to gray and six statistical measure of texture are used to extract the texture of the region such as average gray level, average contrast, measure of smoothness, third moment, measure of uniformity and entropy.

4.4.3 Model Generation

The artificial neural network is used to training the system using the extracted feature images of solid waste. The model is generated be loading the features from mysql database. The artificial neural network design is depicted in Figure 4.

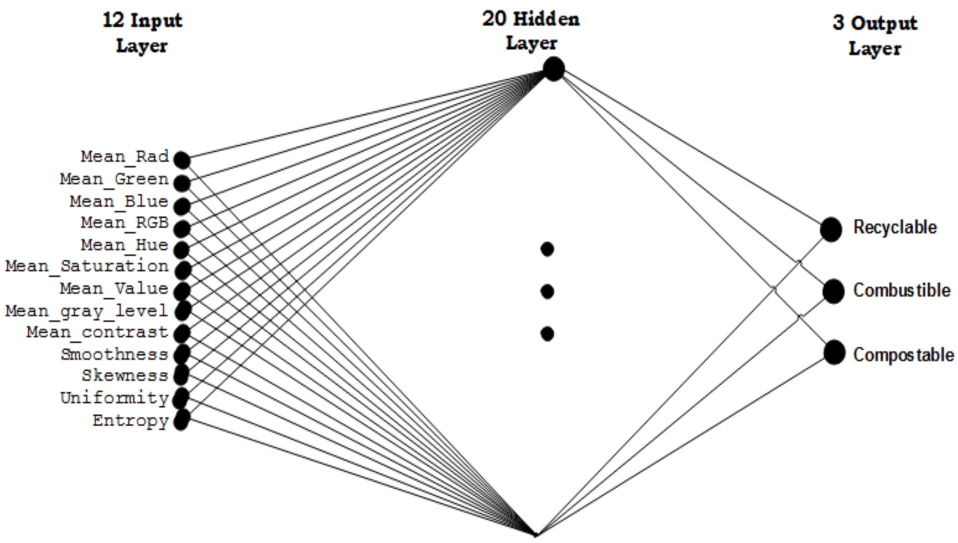


Fig 4:- The Artificial Neural Network Design

4.4.4 Categorization

In all image processing application after training and model is generated. In order to perform the classification based on the model generated, first the input image preprocessed followed by segmentation and feature extraction. Then the classification isperformed by the knowledge retrieved from the model as shown in the Figure 2 (b). In Figure 2 (b), the same process is repeated as shown in Figure 2 (a), until model generation. Dynamic model simulation is to test the response of the model generated to an input images. In MathLab *sim* command takes the trained model and the input parameters to test and give the result [ ]. The overall accuracy of solid wastecategorization is 89%. Furthermore in each specific category, the model achievedan accuracy of 89%, 82% and 96% for recyclable, combustible, and compostable respectively. We compared the accuracy of the system with manually categorized waste from Addis Ababa City West Management Authority Agency, Koshe dump siteas shown in Table 1.

IV. EXPERIMENTAL RESULT AND DISCUSSION

We used total of 2445 solid waste images captured using Redmi Note 6 Pro camera in four categories depicted in Table 1. In the Table, 650 images for recyclable, 865 images for combustible and 930 images for compostable categories is captured. Among this images 70% are used for model generation and 30% used for testing the model.The user interface is designed using MATLABR2012a as shown in Figure 5, and we used mysql from wamp server version 2.5.

ISSN No:-2456-2165

Solid Waste General Category	Solid waste Sub-category	No. of Images/sub-category	No. of Images/General category	Image File Extension	Image Size
Recyclable	Metal	200	650	Jpg	960 x 720
	Plastic	150			
	Paper	300			
Combustible	Plastic	400	865	Jpg	960 x 720
	Leather	90			
	Medicine	235			
	Textiles	140			
Compostable	Vegetables	300	930	Jpg	960 x 720
	Foods waste	450			
	Garden waste	180			
Total Number of Images in all category			2445		

Table 1:- Image Data Summary

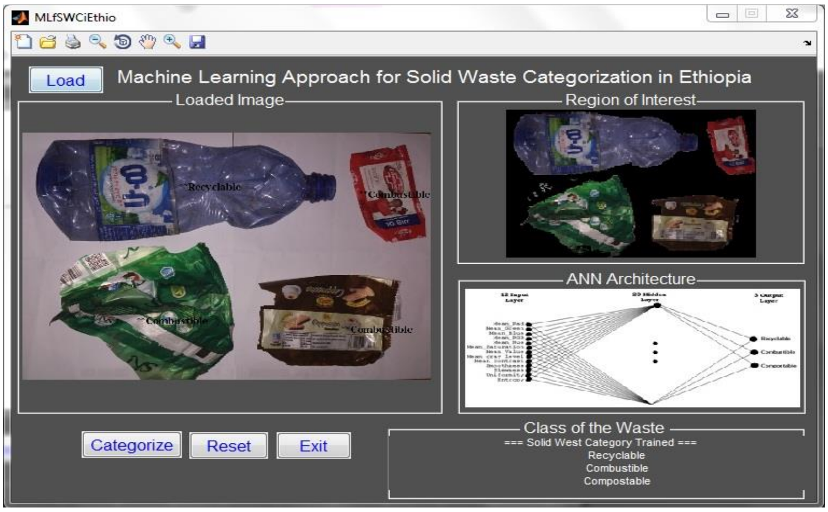


Fig 5:- The design user interface for solid waste Categorization

5.1 Comparing with Manual Categorization

The accuracy of thedesign system is compared with the ground truth categorized manually as shown in Table 2 below.Percentage error in class recyclable, combustible and compostable is10.77%, 17.76% and 3.94% respectively. Moreover the average percentage error in all waste category classification is 10.82%. Our system is very close to the ground truth and this shows that it performs solid waste categorization in Ethiopia very close to the manual waste categorization.

S. No	Solid Waste Class	Total No of Images	No of Correctly Categorized waste images	No of incorrectly-categorized Categorized waste images	Error in %
1	Recyclable	195	174	21	10.77
2	Combustible	259	213	46	17.76
3	Compostable	279	268	11	3.94
Average Percentage Error					10.82

Table 2:- Percentage error comparison between the accuracy of the system with the ground truth

V. CONCLUSION AND FUTURE WORKS

In this study, we designed a Machine Learning Approach for Solid Waste Categorization in Ethiopia. For the experiment we used total of 2445 images and 650, 865, 930 in recyclable, combustible and compostable category respectively. The experimental result shows that we achieved the accuracy of 89% for recyclable category, 82% for combustible, 96% compostable category and average of 89% accuracy for all class. We compared the accuracy of our designed system with the ground truth categorized manually and average percentage error in all waste category classification is 10.82%. The result shows that we achieve a good result in solid waste categorization.In the future we will do further to categorize mass/crowded solid waste, to integrate the designed system with hardware robot and to achieve better accuracy result.

REFERENCES

[1]. "Waste-To-Energy Has No Place In Africa", Global Alliance for incinerator Alternatives(GAIA), 2018.
[2]. Ethiopia's rubbish Policies-African, available at : https://africanarguments.org/2017/04/11/ ethiopias-rubbish-policies, last accessed on January 02, 2019.
[3]. The Reppie waste-to-energy facility, available at https://africannews.com/african_economy/ 2019, last accessed on February 23, 2019.
[4]. Rafael C. Gonzalez, Richard Eugene Woods, and Steven L. Eddins. Digital Image Processing using MATLAB. Pearson Education India, 2004.
[5]. Rafael C.Gonzalez and Rechard E.woods, Digital Image Processing, 3rd edition, Pearson International Edition prepared by Pearson Education, *Prentice Hall.*
[6]. Altman, N. S. (1992). "An introduction to kernel and nearest-neighbor nonparametric regression". The American Statistician. 46 (3): 175–185. doi:10.1080/00031305.1992.10475879.
[7]. Computer and information Science, " Novel Smart Waste Sorting System based on Image Processing Algorithms: SURF-BoW and Multi-class SVM ", *cis.ccsenet.org*, vol. 11, No. 3;ISSN 1913-8989, Canada, 2018.
[8]. Andres Torres-García, Oscar Rodea-Aragón, Omar Longoria-Gandara, Francisco Sánchez-García, Luis Enrique González-Jiménez, "Intelligent Waste Separator" *Computer Syetemas*, Vol. 19, No. 3, pp.487-500, 2015.
[9]. Alvaro Salmador, Javier Pérez Cid, Ignacio Rodríguez Novelle, "Intelligent Garbage Classifier, "International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 1, No. 1, ISSN 1989-1660.
[10]. Ying Wang and Xu Zhang,"Autonomous garbage detection for intelligent urban management "MATEC Web of Conferences 232, 01056 (2018) EITCE 2018.
[11]. MatLab Product help file, R2012a (7.14.0.739), Licenced with no: 161052