International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET

Smart Garbage Classifier

Prof. M. Premalatha, Syeda Zainab, Arpita Sahu, Anushka Bohara, Vijaya Lakshmi
School of Computing Science and Engineering Vellore Institute of Technology, Chennai, Tamil Nadu, India

Article Info
Volume 8, Issue 3
Page Number: 555-560
Publication Issue :
May-June-2021
Article History
Accepted : 20 June 2021
Published: 30 June 2021

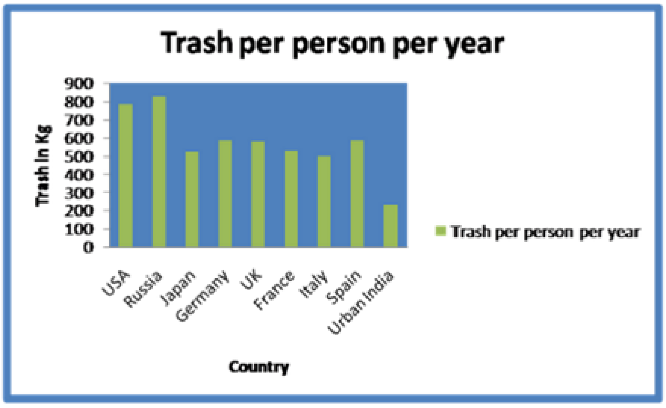
ABSTRACT

Smart garbage classifier classifies the given solid garbage into paper, metal, trash, glass and cardboard. It analyses the image using a camera and preprocesses the images using image processing algorithms followed by machine learning algorithm to do the classification. So that it will be easier for further recycling.

Keywords : Garbage classification, computer vision, shape descriptor ,object classification, k nearest neighbour.

I. INTRODUCTION

Land pollution is becoming a major problem that we face nowadays in our day to day life. In spite of numerous efforts to combat the issues associated with this like landfills, the problem still prevails. An average Indian generates nearly 200 kilos of garbage in a year. The world's population has increased drastically and so has the garbage produced by this population, which has become a major issue. Many measures have been taken by the Indian Government to stop this like the "Swatch Bharat Abhiyan". Most of the present waste disposal systems are manual and hence, require human intervention in some form or the other.



Sources: US EPA website, World Bank report on solid waste management

II. LITERATURE SURVEY

According to studies, improvement in outlook of people towards a better environment leads to a better quality of life and smart waste management strategies play a major role in this [1]. Along with efficient use of available resources, appropriate waste disposal is also as much important for sustainable development of any country. Statistically, out of the 7.4 million tonnes of e-waste is generated every year, approximately 0.7 million tonne of e-waste ends up in landfills [2]. This is all due to improper garbage segregation, which is usually not considered as a serious issue. Solid Waste management in urban cities is one of the most neglected aspects in our country. This problem is further magnified by population growth and industrialization [3].

Feature selection methods are generally used to reduce the dimensionality of the data along with its feature space. K-nearest neighbor classifier is usually employed in medical reports to distinguish between several parameters. [4]

III. PROPOSED WORK

Garbage disposal is a major problem all over the world. The first step to reduce the impact of this problem is to separate the garbage carefully as per the difference in their recycling processes. This makes the recycling process becomes easier. The garbage segregation starts from home but segregation of garbage with 100% accuracy is not always possible manually. Often times, this leads to problems in the recycling process. The proposed system helps in separating different types of solid wastes without any human intervention.

The computing speed and accuracy of a computer can never be surpassed by that of a human. This factor plays a major role in the smart garbage classifier. All the collected garbage is sent to recycling plant where the garbage separation takes place in the first step of recycling. The garbage separation then takes place in various steps. The first steps aims at computer vision followed by machine learning.

A. Computer Vision

In order to classify the image captured by the camera into paper, cardboard, metal, glass or trash, first we need to separate the object from their background. All the relevant characteristics are considered and it is found out that shape descriptors give the best possible output. We find out the shape of the object and characterize it by comparing its shape with the given dataset.

i. Segmentation

• gray scaling

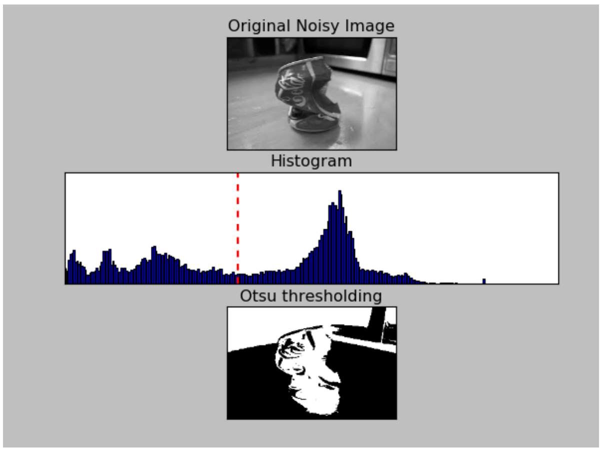
In grayscale images, the thresholding algorithms are fairly easy to conceptualize. This converts the image intensity values ranging from 0 to 255.

• gaussian blur

It is used to remove noisy signals. It is a type of image blurring feature that uses gaussian function.

• Otsu algorithm

Otsu algorithm is used to convert an image into its binary form. First the histogram of the current image is formed. It usually forms a bi- modal histogram with two peak points. One out of the two peak points is assigned 0 as the value and the other as 1 which results in the formation of a binary image. The object is white and the image black or vice-versa.



Segmentation process

After forming the binary image, we go for contouring. Contouring is the first step in feature extraction in our project.

ii. Contouring:

It is a method for extracting the boundary of an object in an image. It is different from edge detection as it focuses on the closed shape rather than every high contrast detection. Before finding the contour in any image it should be ensured that the image is in binary form. The object in the binary image should be white and background should be black. After this step, findContours and drawContours functions are used in order to get object. findContours helps in retrieving all the contours as vector points followed by drawContours which draws the contour outline in the image. The first -1 in the drawContours function indicates that we have retrieved all the possible contours. The second -1 specifies the thickness of the contour. After finding all the contours, we have eliminated the small contours so that focus will be on the large object present on the foreground.

One limitation of this method is that if the object touches the boundary of the image or if the whole object is not inside the image then it is very difficult to find the contours.

Contoured Image

iii. Moments

An image moment is a certain particular weighted average (moment) of the image pixels intensities or a function of such moments, usually chosen to have some attractive property or implementation. Hu-moments are rotation invariants .As shown in the works of Hu etal invariants with respect to transaction ,scale and rotation can be constructed .The first one, Ii , is analogous to the moment of inertia around the image's centroid , where the pixel's intensities are analogous to physical density.

The second one is skew invariant, which enables it to distinguish mirror images of otherwise identical images.

J.Flusser showed that the traditional set of Hu moment invariant is not independent nor complete and later specialized the theory for N-rotationally symmetric shapes case.

• Translation invariants

The central moments μ_{ij} of any order are, by construction, invariant with respect to translation.

• Scale invariants

Invariants η_{ij} with respect to both translation and scale can be constructed from central moments by dividing through a properly scaled zero-th central moment:

$$\eta_{ij} = \frac{\mu_{ij}}{\mu_{00}^{\left(1+\frac{i+j}{2}\right)}}$$

where i + j ≥ 2. Note that translational invariance directly follows by only using central moments.

• Rotation invariants

As shown in the work of Hu invariants with respect to translation, scale, and rotation can be constructed:

$$I_1 = \eta_{20} + \eta_{02}$$
$$I_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$
$$I_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$
$$I_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$
$$I_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$
$$I_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$
$$I_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].$$

These are well-known as Hu moment invariants.

B. Learning and Classification

The processed image of the garbage is taken and then it has to be classified as four types namely- plastic, trash, glass, paper. To classify the image of the garbage we use a Machine Learning Algorithm called KNN (k- Nearest Neighbour). This is a supervised learning algorithm that is applied on the final hu moments of the image, which are obtained after applying several image processing techniques. In KNN classification, the distance of the image to be classified with images stored is calculated. We calculate Euclidean distance of the hu moments of the test image and other training images, which is given as.

$$d(p,q) = d(q,p)$$
$$= \sqrt{(q_1 - p_1)^2 + (q_3 - p_3)^2 + \cdots + (q_n - p_n)^2}$$
$$= \sqrt{\sum_{i=1}^{n} (q_i - p_i\ )^2}$$

Euclidean distance is suitable for input variables of similar type. Though we are using Euclidean distance, We can use other types of distance function like Manhattan Distance, or Minkowski distance. Manhattan distance is used when input variables are not of similar type. These three distance functions can be used only when the variables are continuous. In the case of discrete variable Hamming distance is used.

It also brings up the issue of standardization of the numerical variables between 0 and 1 when there is a mixture of numerical and categorical variables in the dataset.

After this we select the value of k to get the best possible fit for the dataset. Only k values having least distance among the distances calculated, are considered for deriving further references for classification. For example, if the distances calculated are:

1.8 for trash
8.3 for plastic
9 for glass, and

19 for paper (in increasing order) If the selected value of k is 3, then the first three values are kept for classification and rest of the values are eliminated. Usually, the most optimal value of k for most of the dataset is between 3-10. Based on the value of k, the garbage is classified into the class whose number of occurrences are found to be maximum. Also, a small value of k provides more accuracy as it has low bias and high variance. On the other hand, if we take higher value of k then it averages the result of prediction and therefore is more resilient to the outliers. Thus, a higher value of ‘k’ means increase in bias but lower variance.

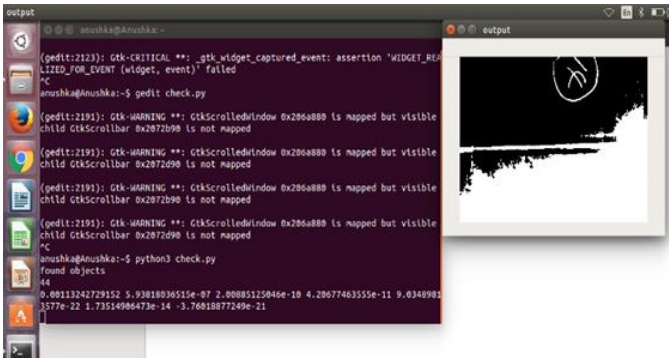
We have preferred KNN classification over other classification techniques of Machine Learning classification algorithms because this is very easy to implement and simple to train.

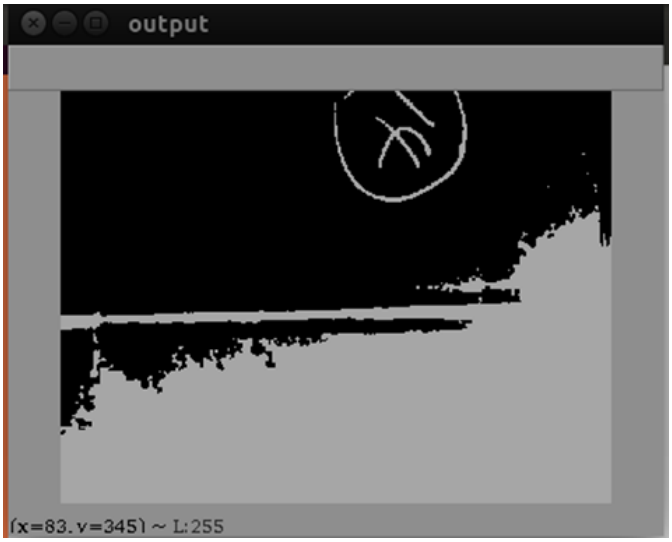
The advantage of KNN is that it keeps all the training data and no retraining is required if the new training pattern is added to the existing training set. It also keeps appending the result of test dataset to the training dataset so that it constantly gets better at classification. Don't need any prior knowledge about the structure of data in the training set. Though there are other disadvantages of KNN classification. In case we have a large dataset, it takes more space. Also, distance has to be calculated between test data and training dataset, hence it will take a lot of time for classification.

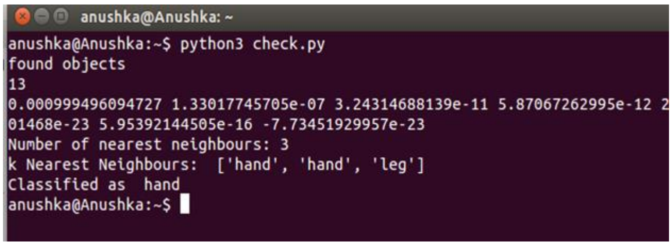
IV. ALGORITHM

1. Get and resize training images. 2: Convert image to gray scale.
2. Convert gray scale image to binary using thresholding.
3. Get and store the Hu moments for training data.
4. Repeat steps 1-3 for test data also.
5. Calculate Euclidean distance between test image's hu moments and training data's hu moments
6. Select a value for k.
7. Retrieve k least-distance values.
8. Classify the test image as the class with maximum occurrences.

V. IMPLEMENTATION







VI. CONCLUSION

Given the current global scenario and the extent of land pollution, this is a problem that needs to be dealt with. Garbage segregation helps in proper and efficient recycling of the garbage and is very important for protection of the environment if done with high accuracy. The accuracy of Smart Garbage Classifier can be increased by increasing the number of images in the training dataset. By improving and adding to these algorithms over time, we can make the classification process better and more accurate.

VII. REFERENCES

[1]. Adam Garcia, Tom Jacobson, "Smart waste disposal in Guam", Journal of Waste Disposal and management (2017)

[2]. Wang, F., Kuehr, R., Huisman, J. "The Global E-waste Monitor" Institute of Advanced Study of Sustainability (2014)

[3]. Bhavik Gupta, Shakti Kumar Arora "A study on management of municipal solid waste in Delhi" Journal of Environment and Waste Management 3 (2016)

[4]. Chao Li, Shuheng Zhang, et al, "Using the K-Nearest Neighbor Algorithm for the Classification of Lymph Node Metastasis in Gastric Cancer" Computational and Mathematical Methods in Medicine, Volume 2012 (2012), Article ID 876545

[5]. US EPA website: World Bank Report on Solid Waste Management

Cite this article as :

Prof. M. Premalatha, Syeda Zainab, Arpita Sahu, Anushka Bohara, Vijaya Lakshmi, " Smart Garbage Classifier, International Journal of Scientific Research in Science, Engineering and Technology(IJSRSET), Print ISSN : 2395-1990, Online ISSN : 2394-4099, Volume 8, Issue 3, pp.555-560, May-June-2021. Journal URL : https://ijsrset.com/IJSRSET2183133