International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 5, Issue 10, ISSN No. 2455-2143, Pages 208-213
Published Online February 2021 in IJEAST (http://www.ijeast.com)

HYBRID APPROACH OF GARBAGE CLASSIFICATION USING COMPUTER VISION AND DEEP LEARNING

Anish Tatke, Madhura Patil, Anuj Khot, Parul JadhavDr Vishwanath Karad's
MIT World Peace University
Pune, Maharashatra, India

Abstract - As waste segregation becomes an important issue in our lives, with the use of technology like deep neural networks and computer vision, we can make the process efficient and robust by image segmentation and classification. These systems on the rise need accurate and efficient segmentation and recognition mechanisms and this demand coincides with the increase of computational capabilities of modern computer architectures and more effective algorithms for image recognition. This paper does a comparative analysis of various different approaches and methods like Simple CNN, ResNet50, VGG16, etc in brief. The comparative analysis and study explains the performance of every approach, this paper concludes that ResNet50 gives excellent performance. VGG16 network also provides good performance which meets the needs of daily use.

Keywords - Garbage Classification, CNN, Waste Segregation, Data Augmentation

I. INTRODUCTION

The world generates 2.01 billion tonnes of municipal solid waste annually, with at least 33 percent of that—extremely conservatively—not managed in an environmentally safe manner.[1]

At this stage, due to the large population of India, our country adds about 277 million tons of municipal garbage each year[2], and with the improvement of living standards, the rate of garbage generation is still rising. The reason is that India's garbage generation is rapid, and people's awareness of garbage classification is weak.

The resource classification and recycling treatment mechanism of urban domestic garbage is an effective solution to solve the environmental damage caused by urban domestic garbage in India.

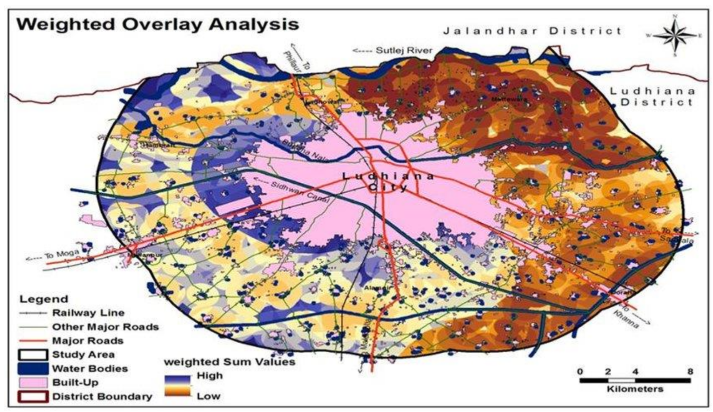


Fig 1. Weighted Sum Average of Ludhiana district Solid Waste

Due to the large increase in computer operating speed, the efficiency of computer processing of images has greatly improved. Deep learning models with CNN (Convolutional Neural Network) as the core have begun to play a pivotal role in the field of image recognition and classification.

Many new ideas like Simple CNN [3], VGG16 [4], ResNet50 [5], HOG+SVM [6] are proposed to gain accuracy in image classification and object detection. Among various deep models, convolutional neural networks(CNNs) have led to a series of breakthroughs for image classification.

The rest of the paper is organised as follows sec II explains the garbage image dataset. Sec III describes various image processing techniques. Details of studied models are described in Sec. IV. Sec. V provides comprehensive evaluation followed by related works and discussion in sec. VI. And lastly sec. VII elucidates the conclusion.

International Journal of Engineering Applied Sciences and Technology, 2021
Vol. 5, Issue 10, ISSN No. 2455-2143, Pages 208-213
Published Online February 2021 in IJEAST (http://www.ijeast.com)

II. DATASET

A lot of image data involving garbage is available and present on the internet and government based datasets. This data is present in variety with different types according to different class labels/target labels. We have explored 2-3 different datasets given below:

- Garbage Classification Dataset from Kaggle is a famous dataset pertaining to garbage segregation and collection [7]. It has almost 2467 images of garbage which has target labels given as cardboard, glass, metal, paper, plastic and trash. This amount of data is still not enough to be fed to any state-of-the-art CNN algorithm. The images in the dataset are of the certain object related to its target label which are kept in a clean background.
- Second Data set is the Garbage In Images (GINI) Dataset [8]. The dataset contains images with and without garbage in it. The maximum of the dataset is annotated and has the label and bounding box parameters in a Comma Separated file. The target labels are given as 1 (Garbage Present) and 2 (Garbage Not Present).

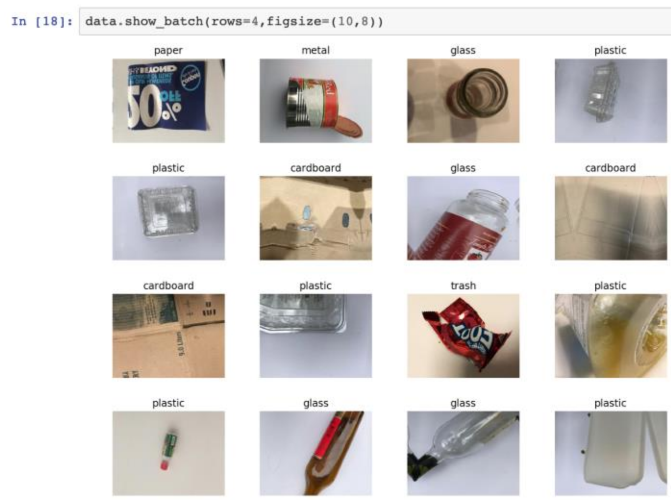
Sr No	Category	Number of Images present	Annotated
1	cardboard	393	yes
2	glass	491	yes
3	metal	400	yes
4	paper	584	yes
5	plastic	472	yes
6	trash	127	yes
	Total	**2467**	

Table1. Dataset Overview

III. IMAGE PROCESSING TECHNIQUES

The images need to be pre-processed using computer vision techniques so as to make them easy to understand for the computer so it can make a prediction on it. Basically the image needs to be brought in the format necessary for the model it needs to be fed in. This whole process is quite tedious when dealing with multiple images of different types and categories. The main aim for image pre-processing is the improvement of image data that suppresses undesired distortions or enhances some image features relevant for further processing and analysis tasks.

Fig 2. Images from the dataset



Before doing image pre-processing it is important to understand the images that we are dealing with. What steps should be taken, what image pre-processing techniques should be used depending on "how strong is your image data ?" is and "what operations are you going to perform on it ?". One important necessary step before feeding the data to a model is to standardize the images. Image standardization refers to resizing of images in a dataset to a unified dimension which is suitable for the particular model it is fed to.

Second important step is Data Augmentation[9] which is the process of scaling, rotating and performing other affine transformations (Denoising using Gaussian Blur, Segmentation, Morphological Operations) on the existing images from existing dataset. This is mainly done to enlarge the size of the dataset and expose the model to a wide variety of variations of the images. One particular paper published in 2020 called as A Study of Garbage Classification with Convolutional Neural Networks, Indo - Taiwan Conference on Computing, Analytics and Networks [3] have performed Data Augmentation of the same dataset we have used in this paper and have generated about 10108 images from the dataset after data augmentation and split the data into 9095 training set images and 1013 test set images.

Literature Review

IV. METHODOLOGY

A. Simple CNN

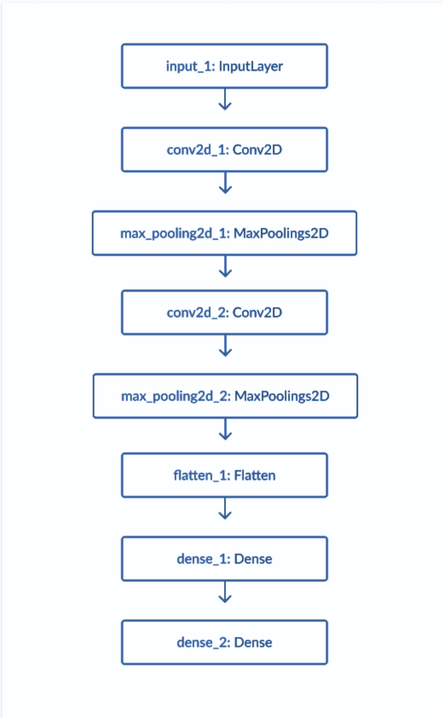


Fig 3. Architecture of Simple CNN

This is a simple architecture which is provided for base comparison with other state-of-the-art models [3]. It uses the 2D Convolutional layers to capture the features of the images. Just the basic 3x3 filters can be used. Max pooling layers are important to reduce the dimensions of the input and the number of parameters to be learned. These Max pooling layers are added between the 2D convolutional layers. Then a layer is used to flatten the feature matrix to a column matrix. Then two fully connected layers are added. The cost function used in between the convolutional layers is RELU function to avoid the problem of vanishing gradients. The second type of cost function is used in the last fully connected layer which is soft max function which fits the cross-entropy loss function.

B. VGG16 with Batch Normalization

Wang Hao proposes a model with VGG16 and Batch Normalization[4]. This network is based on Alex Net network. It can more accurately express the characteristics of the dataset when classifying images. This network consists of 13 convolutional layers, 5 pool layers and 3 fully connected layers. The filters used in the convolutional layers are 3x3 with a step of 1 while that of the pooling filters is 2x2 with step of 2.

Basically every Sigmoid and TanH cost function can be replaced with RELU functions except for in the last fully connected layer. This is done to avoid slow model convergence and improve efficiency. Batch Normalization can also be added at the end of fully connected layers to avoid the problem of gradient disappearance. The Batch Normalization layer can also avoid overfitting during the training process, thereby increasing the accuracy of recognition.

C. ResNet50:

ResNet is a classic neural network well known for its establishment since the kick start of the deep learning era [5]. It has been the backbone for various computer vision tasks. It is different from its predecessors like Alex Net because it solved the notorious problem of vanishing gradients - as the gradient is back-propagated to earlier layers, repeated multiplication may make the gradient very small. Thus as a network gets deeper, the performance starts depleting.

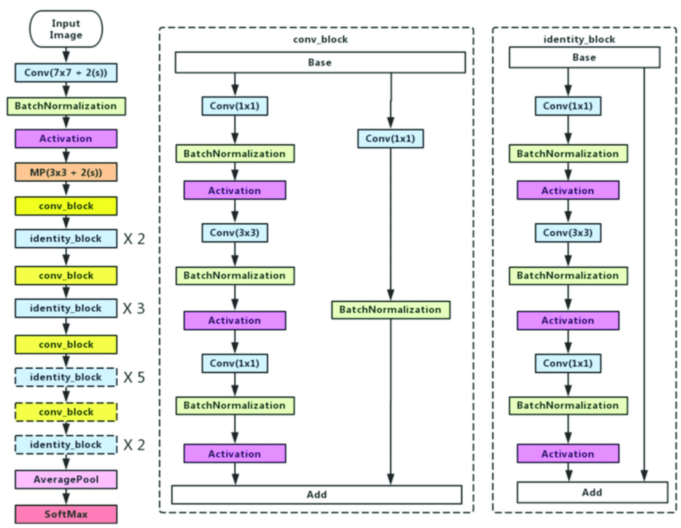


Fig 4. Architecture of ResNet50

In ResNet50, researchers use the bottleneck architecture in the residual block [3]. In each block, there are two convolutional layers with a 3x3 convolutional layer surrounded by a 1x1 filter on both sides. Thus the dimensions of the identity part and the residual part are maintained. The convolutional layers and pooling layers get through the rough features of the images. The model uses 16 residual blocks after the convolutional block. Further an average pooling layer is used to down-sample the feature matrix, a flatten layer and a dropout layer followed by a fully connected layer. The dropout is considered as regularization which helps in adding extra noise and average the overfitting errors. The cost function used is the RELU function followed by a Sigmoid SoftMax in the last layer.

D. HOG + Support Vector Machine

The images in the dataset have objects placed on a clean background. Thus using a Histogram of Oriented Gradients (HOG) suits well for the images which then can be fed to a classifier based on the Support Vector Machine (SVM) [6]. HOG is a type of a feature descriptor which does the work of finding essentially the useful information from the image that will help the classifier to predict better. In the HOG, the distributions (histograms) of directions of gradients (oriented gradients) are used as a feature descriptor. Gradients (x and y derivatives) of an image are useful because the magnitude of gradients is large around the edges and corners (regions of abrupt intensity change) thus giving shapes. The HOG is invariant to geometric and photometric transformations. The image is divided into smaller regions and the HOG features are called on each region. The oriented gradients of each cell are counted in 9 histogram channels. Then block normalization using L2 Norm is used to give feature vectors of cell histograms which are concatenated to a feature vector of an image. These extracted features are fed to a Support Vector Machine which finds a set of hyperplanes between the data points in a higher dimensional space with an aim to make the observations linearly separable.

V. EVALUATION

It is hard to evaluate which model performs better without actually performing and developing the model and running it. That being said, after studying and researching multiple papers on this topic, we evaluate certain models used and can derive a short experimental analysis on them. The models used go from the most basic model like Simple CNN to the most complex models used like ResNet50, etc. Below given table gives the overview of the training and testing results of the models studied in this paper. The two accuracies given for some models are calculated using Adam Optimizer and Adadelta Optimizer.

SrNo	Model	Reference	Training set Accuracy	Test set Accuracy
1	Simple CNN		81.98% 89.52%	82.19% 93.56%
2	VGG16		-------	75.6%
3	ResNet50		96.91% 99.27	51.67% 95.93%
4	HOG+SVM		-------	47.25%

Table1. Comparative Analysis of Studied Models

Comparative Analysis: Garbage classification is a Kaggle Challenge and performed by many researchers and developers using various datasets. The above table gives us the idea of the result we have achieved from the proposed models. We are training almost 9095 images and testing on 1013 test set images for an average of 40 epochs on all the models. We also use two different optimizers (Adam and Adadelta) on every algorithm which gives mixed results depending on the algorithm. Thus the results above are experimental and the accuracies can be improved by using various other datasets, different image pre-processing techniques and different models.
Still by looking at the above given data, ResNet50 gives us the highest accuracy on training and test sets. Data Augmentation remains a very important factor in these results and various approaches to it can improve the accuracies.

VI. RELATED WORK

Many different approaches are taken to this problem by various papers. Every approach is unique and uses its own different methods to achieve good accuracy. For example a particular paper uses HOG along with CNN [3]. The training accuracy reaches 89% and validation accuracy is over 93% with Adadelta. The dataset is from Kaggle and its publisher has constructed an SVM classifier based on Scale-Invariant Feature Transformation (SIFT) which achieves a test accuracy of 63% [10]. Another project, RecycleNet, uses DenseNet with an alteration of skip connections to achieve test accuracy of 81% after 200 epochs [11]. Kaggle's best model on this topic is "MobileNet using Transfer Learning" [12]. It uses sigmoid activation function and binary-cross entropy loss function along with transfer learning. Oluwasanya Awe et al. [13] in their paper, speak about a faster Region-based Convolutional Neural Network (Faster R-CNN) for object classification, reaching a mAP of 68.3%. Maher Arebey et al. [14] proposes to use the grey level co-occurrence matrix (GLCM) method in garbage detection which uses advanced communication mechanisms like radio frequency identification (RFID), Geographical information system (GIS), etc to strengthen the waste segregation process. The features obtained from GLCM are given as inputs to Multi-Layer Perceptron (MLP) and K-Nearest Neighbours classifier. Other papers include various objection detection frameworks which are useful here for garbage detection using deep learning methods [15].

VII. CONCLUSION

The results of this study show that the problem of garbage detection can be efficiently tackled by deep learning algorithms and computer vision techniques. Using the high power TensorFlow as the background for model training gives the models an added advantage as robust and accurate results and minimal error. The paper studies many different approaches to the perfect model like Simple CNN, VGG16, ResNet50 and HOG+SVM which give good test set results for the data. From a comparative analysis of these models we conclude that ResNet50 gives us the best performance and can be easily implemented for an application oriented approach to this topic. For future scope, methods of data augmentation and image processing still need to be refined. The dataset can be explored more and developed to get homogeneous variety and multiple other models can be proposed for this topic.

VIII. REFERENCES

[1] S. Meng and W. -T. Chu (2020), "A Study of Garbage Classification with Convolutional Neural Networks," 2020 Indo – Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN), Rajpura, Punjab, India, 2020, pp. 152-157, doi: 10.1109/Indo-TaiwanICAN48429.2020.9181311

[2] H. Wang (2020), "Garbage Recognition and Classification System Based on Convolutional Neural Network VGG16," 2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), Shenzhen, China, 2020, pp. 252-255, doi: 10.1109/AEMCSE50948.2020.00061.

[3] He, Kaiming & Zhang, Xiangyu & Ren, Shaoqing & Sun, Jian. (2016). Deep Residual Learning for Image Recognition. 10.1109/CVPR.2016.90, pp. 770-778.

[4] M. R. Guedira, A. E. Qadi, M. R. Lrit and M. E. Hassouni (2017), "A novel method for image categorization based on histogram oriented gradient and support vector machine," 2017 International Conference on Electrical and Information Technologies (ICEIT), Rabat, 2017, doi: 10.1109/EITech.2017.8255229, pp. 1-5.

[5] C.CHANG (2018), "Garbage Classification." Kaggle, 2018, doi: 10.34740/KAGGLE/DS/81794.

[6] G Mittal, K B Yagnik, M Garg, and N C Krishnan (2016), Spot Garbage: Smartphone App to Detect Garbage Using Deep Learning, ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 940-945

[7] A. Mikołajczyk and M. Grochowski (2018), "Data augmentation for improving deep learning in image classification problem," 2018 International Interdisciplinary PhD Workshop (IIPhDW), Swinoujście, 2018, pp. 117-122, doi: 10.1109/IIPHDW.2018.8388338.

[8] M. Yang and G. Thung (2016), "Classification of Trash for Recyclability Status," Stanford University, CS229, 2016. [Online].

[9] C. Bircanoğlu, M. Atay, F. Beşer, Ö. Genç and M. A. Kızrak (2018), "RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks," 2018 Innovations in Intelligent Systems and Applications (INISTA), Thessaloniki, 2018, pp. 1-7, doi: 10.1109/INISTA.2018.8466276.

[10] S. L. Rabano, M. K. Cabatuan, E. Sybingco, E. P. Dadios and E. J. Calilung (2018), "Common Garbage Classification Using MobileNet," 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology,Communication and Control, Environment and Management (HNICEM), Baguio City, Philippines, 2018, pp. 1-4, doi: 10.1109/HNICEM.2018.8666300.

[11] O. Awe, R. Mengistu, and V. Sreedhar (2017), "Smart TrashNet: Waste localization and classification," 2017.

[12] Maher Arebey, M.A. Hannan, R.A. Begum, Hassan Basri (2012), Solid waste bin level detection using gray level co-occurrence matrix feature extraction approach, Journal of Environmental Management, Volume 104, 2012, Pages 9-18, ISSN 0301-4797

[13] S. Begumpure, P.M.Jadhav (2019), “Comprehensive Review Of Generic Object Detection Frameworks Using Deep Learning Approach,” IEEE International Conference on Contemporary Engineering and Technology, April, 26th-27th , 2019, Chennai.