POLITICIAN IMAGE CLASSIFICATION USING MACHINE LEARNING

Bharat Goyal
(MT22024)
bharat22024@iiitd.ac.in

Saloni Agrawal
(MT22062)
saloni22062@iiitd.ac.in

Nidhi Verma
(MT22044)
nidhi22044@iiitd.ac.in

Abstract

Image recognition is a crucial component of image processing for machine learning without using any human input at any stage.A difficult and significant issue in content-based image retrieval is the classification of images using low-level visual features. Effective elements can be created for an image database based on these groupings.In this project we investigate the process of image classification utilising imaging backend.A large number of political figures are captured and the resulting images are split into training and test datasets for our learning model.The results are obtained using simple machine learning models which can be utilised to solve image classification problems.

KeyWords:-Image Classification, Haarcascade, Wavelets, OpenCV, Support Vector Machine, Logistic Regression, Random Forest, K-means Clustering, Machine Learning

1 Introduction

Classification aids us in taking decisions in daily life. When an object is assigned to a certain group or class based on the properties that apply to that object, classification is required. The majority of industrial issues are classification-related. To increase classification accuracy, scientists have developed sophisticated classification procedures. Numerous photographs are created every day, which makes it necessary to categorise them for quicker and simpler access. The information processing carried out during classification aids in the division of images into different groups.This type of problem solving is quite challenging due to the accuracy and distribution of the data properties and the capabilities of the model.Accurate identification of the features present in an image is the major objective of image classification [6].

Feature Extraction is a greater standard approach wherein one attempts to increase a change a input space onto the low-dimensional subspace that preserves maximum of the relevant information. In order to enhance performance, such as predicted accuracy, visualisation, and comprehensibility of learnt knowledge, feature extraction and selection algorithms are utilised alone or in combination.Features can typically be divided into three categories: relevant, irrelevant, and redundant. A subset of the available feature data are chosen during the feature selection process for the learning algorithm. The subset that contributes the least dimensions to learning accuracy is the best one[7].

OpenCV is a library utilised in computer vision. Herein our project we have utilised this powerful library for feature extraction and object detection. Haar cascade uses the cascading window, and it tries to compute features in every window and classify whether it could be an object. We are using OpenCV library which manages a repository containing all popular Haar cascades that can be used for: Human face detection, Eye detection, Smile Detection.The major objective of using Haar cascade in our project is that it can detect objects in images irrespective of the scale and location of those objects in image.

Our main goal on this project is classifying different politician images into various classes so that once faces are recognized and stored, our model is predicting the correct names for each politician. The project is being carried out to the concerns of classifying different politician images so that we can utilize it for many purposes such as National Security, Global Conferences, One to One interaction over issues, global debates, global news, etc.

2 Literature Review

The idea of image classification is not new, but it is also not its threshold. Working on the image is not that trivial. The whole process requires more work than usual textual or numerical data preprocessing. Photographs need to be converted not only into processable data but also into something that can extract features. We can create valuable images using several different methods, including unique methods like Haar cascade. With the help of this method, we filtered those images with smiley eyes and a clear vision of the face[1]. It has helped us narrow down the data and remove the noises from it, distorting the training of a high-accuracy model. Feature extraction includes several different methods. These methods vary from a very naive approach to a highly advanced technique. Some of the course uses flatting and Wavelets methods[2]. Using these techniques helps the model to get better image features, which allows the model to train themselves on those images quickly and efficiently. While dealing with feature extraction techniques, reducing features from n-dimensional to lower space is not that useful[3]. Finally, switching towards the end stage of using all original parts and using methods of Haarcascade and Wavelets, we have achieved significantly better results among all experiments and observations during the process[11]. We have tried using one of the unorthodox and unsupervised learning techniques, which is very useful and generally performs well on images. That is the K-means technique[5]. But our case has not shown that kind of stone-turning results on our data set. To make operation fast and clean, we have tried to keep things simple, using simple Machine Learning models, including Support Vector Machine, Logistic Regression, and Random Forest[4]. These models have shown excellent results with our cleaned and preprocessed data of images. Working with a variety of Techniques and algorithm have shown different aspects in terms of mathematical concepts. And it has been demonstrated that even simple techniques like Logistic regression can work very fast and help in training models with very high accuracy.

3 Methodology

3.1 Data Collection

Data Collection is the basic step in the machine learning pipeline for training the ML model.Data can be obtained in several ways such as web scrapping, Fatkun and many more. Web scraping collects and parse the raw data from the web. In our project, we are using Fatkun batch download to gather the large data with just one click. It automatically downloads the image to build dataset.Unlike web scraping, it do not allow privacy invasion while downloading the images. Our dataset includes 8 different politicians each having 90 real time images which constitutes total datasize of 200 MB such that format of all the images are JPEG, JPG and PNG. Each image in our dataset is initially 3-Dimensional such as dimension of the first image is (1390, 1140, 3) whereas the size of this image is 4753800. Maximum and minimum value of the first photo pixels are 255 and 0 respectively. Our dataset contains real time images such that integrating it with simple machine learning models can help us in classifying the pictures captured at real life.

Figure 1: Dataset

3.2 Data Pre-processing

Data pre-processing is a critical stage in machine learning that improves the quality of the data to encourage the extraction of valuable insights from the data. Preparing (cleaning and arranging) raw data in order to make it suitable for creating and training Machine Learning models. We are performing different steps in data pre-preprocessing. After applying the haar cascade to detect the eyes, smile, and face, we are cropping the images and Converting it into grey scale for valuable insights.

Initially, we are using the cascade classifier which works on positive as well as on the negative images for detecting the descriptors with the help of Haar cascade. We are classifying the dataset based on the characteristics such as frontal face, eyes , smile and cropping them. we are using

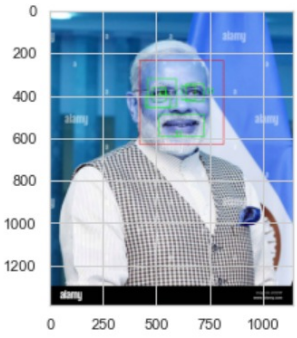

Figure 2: Face, eye and smile detection

the rectangle function of CV2 library for drawing the rectangle on these features. The dimensions for face, eyes and smile are [421,239,394,394], [197,120,65,65], [92,253,210,105] respectively. Instead of working directly on colour images, we are converting the images into grey scale for extracting characteristics because doing so simplifies the process and uses less computational power.

3.3 Data Augmentation

Data augmentation strategies target overfitting at the training dataset, which is the source of the issue. This is done with the assumption that augmentations will extract more data from the original dataset. By data warping or oversampling, these augmentations artificially increase the size of the training dataset [8].

For data augmentation, making simple alterations on visual data is popular. In this technique, instead of adding more training data, slightly modified copies of already existing data are added to the dataset. By flipping and rotating images of politicians, it is possible to present a slightly different data sample each time the model processes it. As a result, the model doesn't overfit because each data set looks unique to it. To increase the dataset, we rotate and flip the data and as a outcome our dataset will increase about three times i.e 600 MB, once all these procedures are completed.

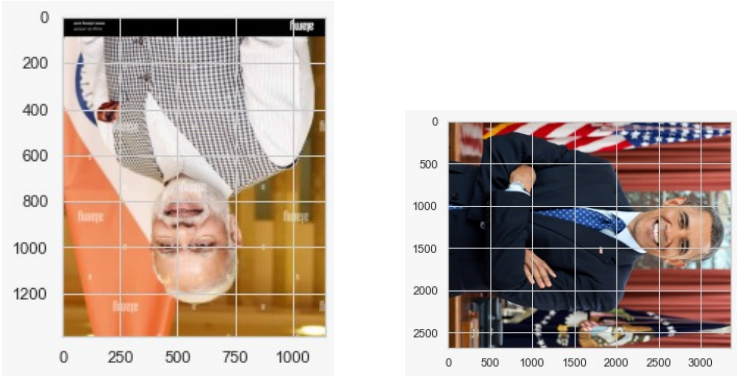

Figure 3: This figure illustrates flipped and rotated image of the politician after applying data augmentation

3.4 Feature Extraction

Features extraction modifies the original characteristics to produce more important features.Here, we are using two techniques in feature extraction:

Reshaping:- In this technique, we are reducing the dimensionality of the images in our dataset from 3-D to 1-D i.e flattening the images.

Wavelets:- Wavelets based transform are mathematical tools which are used to extract information from images [9].

In process of feature extraction our goal is to find out the most relevant feature from image which ultimately helps us in gaining useful insights. First one i.e flattening is very naive approach of converting our data from 3-D to 1-D. So we use the wavelet transformation for extracting information from politician images which relatively gives good results.

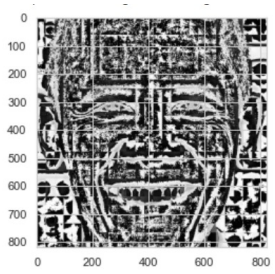

Figure 4: Wavelets

3.5 Model Building

In this project, we are using four models along with the technique named as grid search CV for tuning hyperparameters results in achieving the best accuracy. These Four models are:

Support Vector Machine:-Support Vector Machine is the classical machine learning model which is concern with data categorization issues [4]. Its primary objective is to establish the best line or decision boundary or hyperplane that can divide n-dimensional space into several classes and allows to quickly classify fresh data points in the future. We did hypertunning on hyperparameter C and kernel. The best hyperparameters for our model in process 1 is C=6 and kernel ='rbf', for process 2 we get C=16 and kernel = 'rbf' and in process 3 it is C=6, kernel = 'rbf'.

Logistic Regression:-Logistic Regression is a supervised learning technique which describes the relationship between the independent variable and dependent variable. It is used to predict categorical outcomes. We did hypertunning on hyperparameter C. The best hyperparameters for our model in process 1 is C=6 ,for process 2 we get C=21 and for process 3 C is 11.

Random Forest:-Random Forest acts as a classifier that produces numerous categorization trees. Each tree is trained using a sample of the training data. At each node, the algorithm simply looks across a random portion of the variables to find a split. Each tree receives an input to classify. The Final classification is then decided by a majority vote [10]. We did hypertunning on hyperparameter max-depth. The best hyperparameters for our model in process 1 max-depth is 16, for process 2 we get max-depth as 21 and for process 3 max-depth is 16.

K-means Clustering:-It is an unsupervised learning model which classifies the data point into various clusters or groups. It is iterative approach which helps in solving the clustering issues.

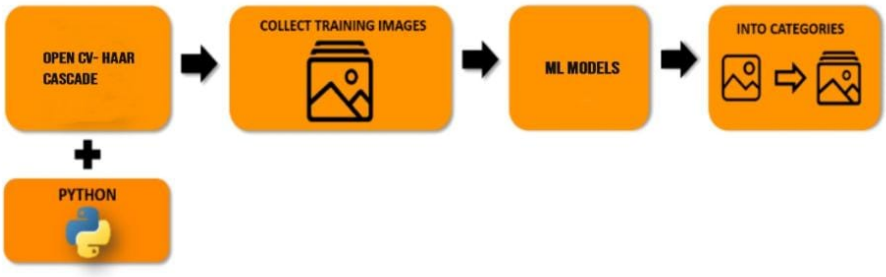

Figure 5: This Figure illustrate model building process in image classification[12]

4 Experimental Analysis and Results

After training the models on the best hyperparameters on cleaned data, we obtained significant results. We have used Accuracy, recall, precision, and F1-Score as metrics to compare the models and their working. Out of these metrics, accuracy was the primary tool of comparison. We have seen that in Process 1, all the models have performed moderately. But in process 2, our performance declined, and in process 3, suddenly, it grew a lot due to the better feature extraction technique. In Process 1, Support Vector Machine performed best because it could find the maximum margin without any outliers. On the other hand, K-means has the lowest accuracy because clusters are highly overlapping with each other. In Process 1, Support Vector Machine has an accuracy of 66.42 percent. Now, if we talk about process 2, then again, Support Vector Machine has performed best among all, and this is due to the same reason. In process 2, the accuracy of Support Vector Machine is 50.71 percent. If we talk about process 3, logistic regression has performed best. It has an accuracy of approximately 85 percent. And the reason is quite simple after feature extraction: when models are trained on it, it can find a line that can accurately separate all the classes. On the other hand, Support Vector Machine was disturbed by a lot of Outliers. But K-means was still performing very severely, and the reason is that clusters were too mixed. If we observe the classwise accuracy along all the processes, they mainly depend on the number of data points in those classes.

Process	Model	Accuracy(percent)
1	SVM	66.42
1	LR	59.28
1	RF	64.28
1	K-Means	29.21
2	SVM	50.71
2	LR	48.57
2	RF	42.85
2	K-Means	25.98
3	SVM	78.72
3	LR	85.10
3	RF	70.21
3	K-Means	18.27

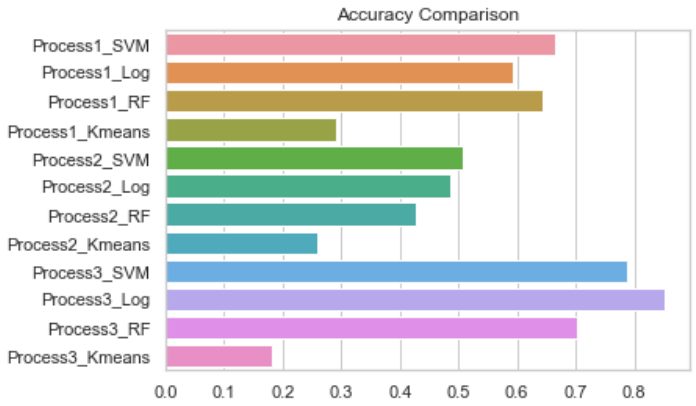

Table 1 and Figure 6: This Table and figure illustrates comparision of model on the basis of accuracies

Class	Precision	recall	f1-score
Barack Obama(0)	0.73	0.89	0.80
Donald Trump(1)	0.67	0.67	0.67
Joe Biden(2)	0.80	0.80	0.80
Justin Trudeau(3)	1.0	0.80	0.89
Narendra Modi(4)	0.50	1.0	0.67
Rishi Sunak(5)	1.0	0.91	0.95
Valdimir Putin(6)	0.83	0.83	0.83
Xi Jinping(7)	1.0	0.86	0.92

Table 2: This table illustrates classwise accuracy for the best model. In our case it is Logistic Regression

5 Scope

Moving with time, textual data is just used for processing operations and tasks. But recording and generating complete information is now shifted from Textual form to visual and audio formats, including videos, pictures, posters, audio monologue recordings, songs, etc. The most prominent example of that can be seen as the growth of Google Images and Youtube. Now processing Images makes things easy to record and automates the process. And our project includes preprocessing political leaders' images by which we extract information. This information can be used for several purposes and has scope in several places. It mainly comprises National Security, Global Conferences, One to One interaction over issues, global debates, global news, etc. All these can be achieved by extending and making things more powerful and complex.These uses are efficient for politicians and ordinary people who invest themselves in these politicians. We have used straightforward machine learning techniques for training the models instead of which, in the future, we can use powerful and complex Deep learning techniques and methods. We can use complex and efficient Computer Vision techniques even for feature extraction. This will not only improve the accuracy of models, but it can also make our work more relentless and flawless.

6 Conclusion

This project is about image classification using Simple machine-learning models. The objective is to classify the given image to its correct names, achieved in our project. All results that have been obtained showed quite impressive outcomes. We have used a series of techniques in the whole process which includes Data collection(Fatkun), Data Augmentation(OpenCV2 library), Feature Extraction(Wavelets), Model Building(Support Vector Machine, Logistic Regression, Random Forest, K-means clustering), comparing results(Accuracy, F1-score, Recall, Precision). In the data collection, we have specifically taken JPEG, JPG, and PNG images for better performance. We have mainly used the OpenCV library to detect Smile, Eyes, Face in particular images. Feature extraction with Wavelets was very fast and efficient in terms of both feature extraction and performance. Implementation of

the machine learning models for our image classification is giving good results as it is able to simulate, train and classify the images so that our best model is very high accuracy towards eight (8) different types of politicians. While comparing all the models and their performance, we have majorly used accuracy as the parameter. The model which performed best among all the processes and models is Logistic Regression, with an accuracy score of 85.10 percent. Lastly, Python has been used as the programming language throughout this project as it is easier to integrate with all the used libraries.

7 References

[1] Rafael Padilla and Marly Costa, *In the proceedings of Evaluation of Haar Cascade Classifiers for Face Detection, April 2012, Conference: ICDIP: International Conference on Digital Image Processing.*

[2] Euisun Choi, Jong Seok Lee and Joonhyun Yoon, *In the proceedings of "Feature Extraction for Bank Note Classification Using Wavelet Transform," 18th International Conference on Pattern Recognition (ICPR'06), 2006, pp. 934-937, DOI: 10.1109/ICPR.2006.553.*

[3] Yu Tao, Ernest C.M. Lam, Yuan Y. Tang, *In the proceedings of Feature extraction using wavelet and fractal, Pattern Recognition Letters, Volume 22, Issues 3–4,2001, Pages 271-287, ISSN 0167-8655.*

[4] Shan Suthaharan *In the proceedings of Machine Learning Models and Algorithms for Big Data Classification pp 207–235Cite as Support Vector Machine.*

[5] Ryan Cohn and Elizabeth Holm,*In the proceedings of Unsupervised Machine Learning Via Transfer Learning and k-Means Clustering to Classify Materials Image Data.*

[6] Siddhartha Sankar Nath, Girish Mishra, Jajnyaseni Kar, Sayan Chakraborty, Nilanjan Dey Publisher:IEEE, *In the proceedings of A survey of image classification methods and techniques.*

[7] Samina Khalid, Tehmina Khalil, Shamila Nasreen Publisher: IEEE, *In the proceedings of A survey of feature selection and feature extraction techniques in machine learning.*

[8] Connor Shorten and Taghi M. Khoshgoftaar,*In the proceedings of A survey on Image Data Augmentation for Deep Learning.*

[9] Rakesh Kumar and B.S.Saini, *In the proceedings of "Improved Image Denoising Techniques Using Neighboring Wavelet Coefficients of Optimal Wavelet with Adaptive Thresholding," International Journal of Computer Theory and Engineering, Vol.4, No.3, June 2012.*

[10] P.O. Gislason, J.A. Benediktsson, J.R. Sveinsson Publisher: IEEE, *In the proceedings of Random Forest classification of multisource remote sensing and geographic data.*

[11] Adam Harvey,*In the proceedings of OpenCV Face Detection: Visualized on Vimeo.*

[12] Mohd Azlan Abu, Nurul Hazirah Indra, Abdul Halim Abd Rahman,*In the proceedings of A study on Image Classification based on Deep Learning and Tensorflow.*