Calorie Counter

Omprakash Yadav[#1], Shreya Nair[#2], Tejaswini Pednekar[#3,] Priya Tiwary[#4]

[#]*Computer Engineering, Mumbai University*

[1]omprakash.y@xavierengg.com

[2]spn598@gmail.com

[3]tejaswinipednekar5@gmail.com

[4]priyatiwary512@gmail.com

***Abstract*— Calorie counter is a system that aims at reducing the calorie consumption and generating caloric intake report using face recognition and machine learning. The system can calculate and keep a track of the number of calories consumed by a person due to intake of beverages. The beverages may include tea, coffee, milk etc. The users will be recommended about the further number of calories which can be consumed based on the current number of calories consumed, using machine learning. On reaching the peak consumption level, if user tries to have a cup of beverage the system will send an alert message. The dataset will include images of people and the properties of different beverages. The system identifies a registered user through face recognition technique and applies machine learning when a new face is detected and creates a new id for that person and stores the calorie intake details. The system will train itself to further recognize the newly created id. Face recognition process use Cascade Classification and LBPH (Local Binary Pattern Histogram) face recognizer method based on OpenCV library and Python. In the end of the day, the user will receive an email which will provide the chart of calories consumed on intake of all beverages. The system attempts to make people conscious about their calorie intake by reducing the beverage consumption, thereby saving money of the organization spent on the beverages and making people healthy.**

***Keywords*— Face Detection, Face Recognition, Calorie intake, Python, OpenCV, Deep Learning.**

I. INTRODUCTION

It is widely known that those who are overweight often eat many more calories than those who are at their ideal weight. Calories do have a direct bearing on health and the rate that you burn calories also determines your healthy weight levels. It is difficult to set bottom line calories levels for health since everyone has a different body composition and their metabolism and activity levels are very different. However, there are some baselines that suggest 1200 calories per day for women and 1800 calories per day for men.

Eating a diet high in calories is yummy, but it can be detrimental if you are not highly active. Some high calorie diets are prescribed for active athletes or people who need to gain weight. However, if you have an average metabolism and activity level, high calorie diets will have a disastrous affect – even if you are eating nutritious foods.

Eating more calories than your body burns causes the body to either excrete the energy or store it in fat cells for the future. When you have adequate supplies of insulin your body chooses to store excess calories in fat cells. The result is weight gain and higher body fat percentages. High calorie intake will cause stress on your body. High caloric foods are high in fats and sugars and extraordinary intake of these types of foods increase your risk factors for type 2 diabetes, heart disease and cancers.

The existing vending machine systems dispense coffee or tea multiple number of times, irrespective of past intake of a person. The amount of consumption of beverages such as coffee and tea has increased in the past few years due to long hours of working and change in lifestyles. However, if you have an average metabolism and activity level, high calorie diets will have a disastrous affect – even if you are eating nutritious foods. These problems can be reduced or controlled by our system as it alerts the user on reaching the peak level consumption in a day. This in turn can be profitable to an organization that implements this system as the money spent on these resources can be reduced due to controlled consumption.

Calorie counter is a system that aims at reducing the calorie consumption and generating caloric intake report using face recognition and machine learning. The system can calculate and keep a track of the number of calories consumed by a person due to intake of beverages like tea, coffee etc.

II. LITERATURE SURVEY

Face Recognition

Face recognition is an image processing method to locate the human face which need camera to capture the image of human face. The image processing will search the important feature of a human face on the image, thus another object will be ignored. The image processing locates human face used by various algorithm and method like Haar-cascade classifier, Histogram of Oriented Gradients (HOG), Pre-processing approaches, Eigen faces, Local Binary Pattern Histogram (LBPH) and others.

Our setup of the system is simple: the camera will be pointing to a complex background. A person walks in the scene, several of their pictures are captured and compared to a previously recorded face database for identification.

A. *Face Detection using Haar-Cascade Classifier*

Haar-cascade is a method, invented by Viola and Jones (Viola & Jones, 2001) [1], can detect a human face in an Image by four key concepts:

1) *Haar features*:

Haar feature is used to find out the existence of human face on captured image. It detects the bright side and the dark side of the captured image. The existence of Haar features is determined by subtracting the average of the dark-region pixels by the average of the pixel light-region. This calculation forms a rectangle as shown in Fig. 1 around the detected face shown at Fig. 2. If the difference is above the threshold, then the Haar features are said to be "exist".

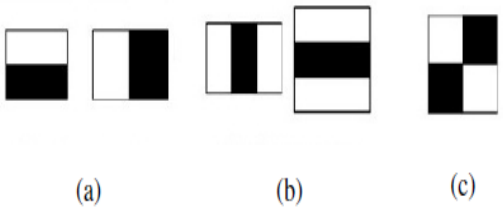

Figure 1: Haar rectangular feature [10]: (a)Edge, (b)Line, (c)Four rectangle.

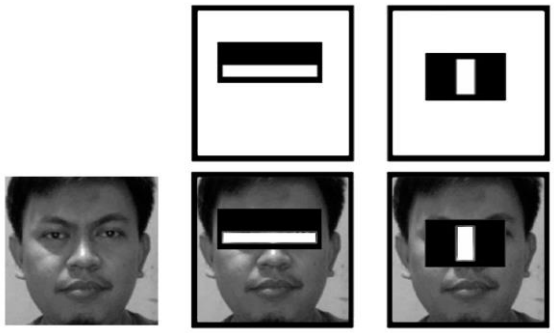

Figure 2: Haar feature result [10].

2) *Integral Image*:

The integral image is used to speed up the feature detection in a way to increase the pixel values from of original image. The integrated value already represents the sum of all pixels above the threshold and is to the left of the image. By starting the addition procedure, from the top left of the image (original pixel) become the bottom right of the image (addition pixel result) show at Fig. 3.

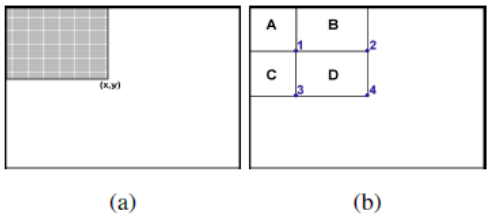

Figure 3: (a) Integration process, (b) Rectangle divided into multiple segments.

3) *AdaBoost Machine-Learning Algorithm*:

Viola and Jones used AdaBoost algorithm because it boosted the classification performance of a simple learning, in other words it gives fast and easy computation. The algorithm's learning feature gives the number of data samples, thus can be grouped into the classifier. This classifier composed of small features of the face is used as pattern detection.

4) *Cascade Classifier*:

The filter chain which was shown in Fig. 4 is made by a series of AdaBoost classifier. The sequence of filters in the cascade is based on the weight of results from the AdaBoost. The cascade eliminates a candidate if it didn't pass the first stage. Thus, a Cascade Classifier that combines many features works efficiently.

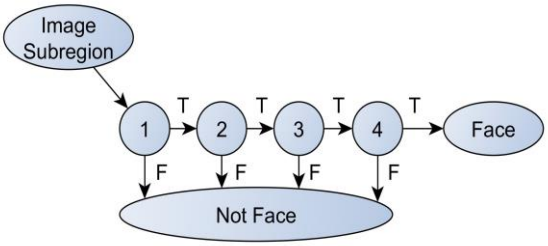

Figure 4: Cascade classifier illustration [10]

B. *Face Recognition using Local Binary Pattern Histogram (LBPH):*

This algorithm also requires grayscale pictures for processing the training. The aim of LBPH (Wagner, 2011) [6] is to work by blocks of 3x3 pixels. The pixel in the centre is compared to its neighbours. The neighbours which have smaller pixel values than the pixel value in the middle, will be indicated by a 0 value in the thresholded square that stores the results, otherwise, a 1 will be added. After completion of the comparisons, each result will be multiplied by a weight. Each pixel has a weight to the power of two from 2x to 2y and the pixel in the centre of a 3x3 square has 8 neighbours. These eight pixels represent one byte. The weights are affected in a circular order. For example, if the pixel top left has a weight of 128, it will keep this weight for all the comparisons in the picture. Then, the sum of the weights is calculated and becomes the value of the pixel in the middle of the square. When this process has been completed for each part of the picture, the picture is divided into a certain number of regions. Then, a histogram is extracted from each region and all the histograms are concatenated.

For recognizing a face, the same process is performed. The final histogram is then compared to each final histogram in the training data.

C. *Deep Learning:*

Deep learning techniques are a part of machine learning methods. These are based on learning multiple levels of representation and abstraction which help us to make sense of data such as images, sound, and text. It replaces handcrafted feature extraction with efficient algorithms for unsupervised or semi-supervised feature learning and hierarchical feature extraction. In previous few years, deep learning's performance has been exceptional in natural languages, speech recognition and computer vision. Deep learning nowadays is generally focused on multi layered neural networks. The deep neural architectures are feed forward neural networks, recurrent neural networks and convolutional neural networks. Feedforward networks as the name suggests, sends unstructured information from one end called input to the other end called output. These networks approximate some function f by defining a mapping of $y=f^*(x; \theta)$ and then learning the value of the parameters θ that best approximates f, hence are universal function approximators. Recurrent neural networks model uses self-replicated components. These are specialized for processing sequential data. They preserve some amount of memory and save long-term dependencies. RNNs are powerful computational machines. A convolutional neural network has trainable filters and local neighborhood pooling operations which are applied alternatingly on the input images. It results in a hierarchy of increasingly complex features. The pooling operations help to enlarge the input patterns. CNNs take advantage of the repetitive local input patterns across time and space, so they are translation-invariant – the capability found in visual cortex of a human. Local input patterns are small data slices, of distinct size, e.g., a group of pixels in an image.

In last few years, deep learning methods, especially CNN has succeeded in the sector face recognition in unconstrained environment. The main benefit of CNNs is that all the processing layers, even the pixel level input have configurable parameters that can be learned from data. This averts the necessity for hand crafted feature design and replaces it with supervised data driven learning of features. CNN learning based features are more reliable to complex intra-personal variations.

III. PROPOSED SYSTEM

The system program is written in Python 3.6 (with OpenCV Library, Tensorflow and SQLite3). Fig. 5 show the general process of face recognition and calculation of calories of Calorie Counter.

From the figure it can be explained that the webcam acts as the vision sensor, since it is a real-time system. There will be two types of users to this system: Employee and Guest. The working of the system for employees is based on face detection and recognition. The face detection process by using Cascade Classifier Method algorithm. The process divided into 2 sections: The first process makes a training data consisting of features extraction process and classifying human face.

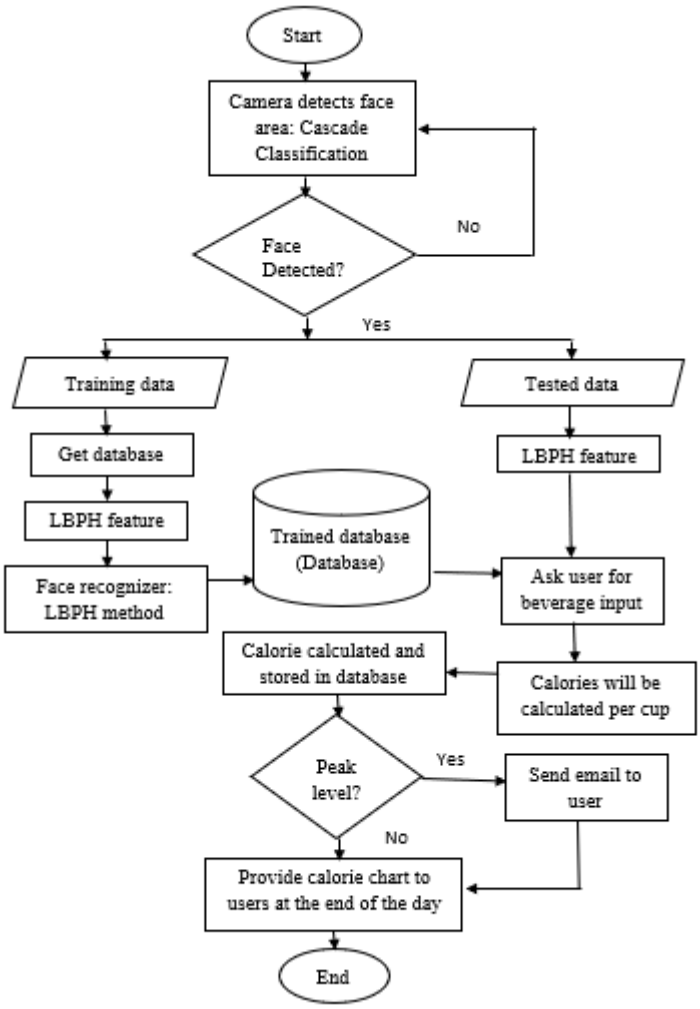

Figure 5: Flowchart of Calorie Counter

Classification using LBPH:

In this research, to capture a human face image we need a face detection algorithm which uses Cascade Classification method based on OpenCV library and Python 3.6. Fig. 6 is a database of some people images to make a database.

Figure 6: Dataset example

For the above dataset, the classification uses Local Binary Pattern Histogram (LBPH) Feature Extraction and LBPH Face Recognizer based on OpenCV library and Python 3.6. The database after the classification is now called the trained data. After that, the classified data is trained by the respondents for clarification. This method works well for a dataset consisting of about 10-12 different employees. The trained data consist of face image capture of employees, and 30-50 images per employee. The second process is a testing of trained data by trained respondents which need LBPH features extraction. Then new image which captured in this section is matched with trained Data. If the classification to identifying the person has done, the system will mention the person's name, which will be taken from the Employee database, by matching with the id returned by the system.

Classification using Deep Learning:

Deep learning is used as it provides better accuracy, speed of training and normalization for very large datasets. Training and testing data will be provided. Using deep learning method, the dataset will be in the form as shown in fig 7. The classes formed will be based on the folders present in the training data.

```
training-data
|-------------- s1
|               |-- 1.jpg
|               |-- ...
|               |-- 12.jpg
|-------------- s2
|               |-- 1.jpg
|               |-- ...
|               |-- 12.jpg
```

Figure 7: Format of storing training data

After classification and training, a model is created (.pb file). The detected faces will be predicted based on this model. When face matches, the employee name returned is the name of the folder which contains their images, i.e. the employee's name must be the name of their image folder for training as well as test data.

For a guest user, their name and email id will be asked by the system. The data will be stored in the Guest database, based on pattern matching of email id, i.e. if a record with the email id already exists, the calories values will be updated for that record and no new record will be created.

On recognizing the name of the user, the different options of beverages will be provided on the screen. The user can select any one option. On selecting the option, the user's database record will be checked to see whether they have already consumed

the maximum level of calories. If not, then they will get a message on the screen informing that their beverage is ready, the database value will be updated, and the vending machine will provide the beverage to the system. If maximum value reached, they will get a message on the screen informing that they have reached peak level, the database value won't be updated, and the program won't run further, thereby not providing beverage through the vending machine.

The system will be programmed to send an email to all the employees and the guests at the end of the day. The name, email id and calories will be stored into a text file from both the databases. By using file reading operations, the system will send email to all which includes the number of calories they have consumed that day based on the type of beverage they chose. After sending the email, the calorie attribute in the Employee database for all the records will be set to zero whereas the Guest database will be completely refreshed i.e. all the records will be deleted.

IV. CONCLUSION

This system aims at recognizing a person through face recognition and applies machine learning when a new face is detected and creates a new id for that person and stores the calorie intake details. On reaching the peak consumption level, if user tries to have a cup of beverages, the system will send an alert message and aware him/her of the health risks. In the end of the day the user will receive an email which will provide the chart of total calories consumed through beverages. The system can also be trained in future to advice people of all types of chronic effects on them in future by increasing caffeine intake, thus detecting its deadliest effects at an early stage. As a precautionary measure, they can be advised to take certain steps in advance to avoid any critical illness later. The system can be upgraded in future to calculate the optimal amount of calorie intake of a person based on their BMI, age and other health-related details.

REFERENCES

[1] Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. Computer Vision and Pattern Recognition (CVPR), 1, I--511--I-- 518. https://doi.org/10.1109/CVPR.2001.990517

[2] Geitgey, A. (2016). Machine Learning is Fun! Part 4: Modern Face Recognition with Deep Learning. Retrieved 11 May 2017, from https://medium.com/@ageitgey/machine- learning-is-fun-part-4- modern-face-recognition-with-deep-learning-c3cffc121d78

[3] Kazemi, V., & Sullivan, J. (n.d.). One Millisecond Face Alignment with an Ensemble of Regression Trees. https://doi.org/10.13140/2.1.1212.2243

[4] Turk, M., & Pentland, A. (1991). Eigenfaces for Face Detection / Recognition. Journal of Cognitive Neuroscience, 3(1), 1–11. https://doi.org/10.1162/jocn.1991.3.1.71 Department of Computer Science, Modern College of Arts, Science and Commerce, Ganesh khind Pune 16, Savitribai Phule Pune University, Maharashtra, India

[5] Jaiswal, S., Gwalior, M. I. T. S., & Gwalior, M. I. T. S. (2011). Comparison between face recognition algorithm-eigenfaces, fisherfaces and elastic bunch graph matching, 2(7), 187–193.

[6] Wagner, P. (2011). Local Binary Patterns. Retrieved 11 May 2017, from http://bytefish.de/blog/local_binary_patterns/

[7] Amos, B., Ludwiczuk, B., & Satyanarayanan, M. (2016). OpenFace: A general-purpose face recognition library with mobile applications, (June). Retrieved from http://cmusatyalab.github.io/openface/

[8] Y. Sun, X. Wang, and X. Tang. "Deeply learned face representations are sparse, selective, and robust," Proc. IEEE Conf. Comput. Vis. Pattern Recog. pp. 2892-2900 Jun. 2015.

[9] G. Hu, Y. Yang, D. Yi, J. Kittler, W. J. Christmas, S. Z. Li, and T. M. Hospedales, "When face recognition meets with deep learning: an evaluation of convolutional neural networks for face recognition," CoRR, abs/1504.02351, 2015.

[10] W.S. Mada Sanjaya, Dyah Anggraeni, Kiki Zakaria, Atip Juwardiand Madinatul Munawwaroh. "The Design of Face Recognition and Tracking for Human-Robot Interaction", from International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), pp. 315-320, 2017.