VISVESVARAYA TECHNOLOGICAL UNIVERSITY

"Jnana Sangama", Belgavi-590 018, Karnataka, India

PROJECT WORK PHASE 1 - REPORT

On

"REAL-TIME ATTENTION SPAN TRACKING IN ONLINE EDUCATION"

Submitted in Partial Fulfillment of the requirements for the award of the degree of

BACHELOR OF ENGINEERING

IN

INFORMATION SCIENCE AND ENGINEERING

Submitted By

ADARSH K R	**1SJ19IS002**
RAKSHITHA H G	**1SJ19IS128**
G YOGA SAI GANESH	**1SJ19IS036**
GOVARDHAN J A	**1SJ19IS038**

Under the guidance of

Ambika L G
Assistant Professor
Dept. of ISE,SJCIT

S J C INSTITUTE OF TECHNOLOGY
DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING
CHIKKABALLAPUR-562 101

2022-23

||Jai Sri Gurudev||

Sri Adichunchanagiri Shikshana Trust®

S J C INSTITUTE OF TECHNOLOGY

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

CHICKBALLAPUR – 562101

CERTIFICATE

This is to certify that the project work phase-1 work entitled **"**REAL-TIME ATTENTION SPAN TRACKING IN ONLINE EDUCATION**"** is a bonified work carried out by **ADARSH K R (1SJ19IS002) , RAKSHITHA H G (1SJ19IS128), G YOGA SAI GANESH (1SJ19IS036)** and **GOVARDHAN J A (1SJ19IS038)** in partial fulfillment for the award of **Bachelor of Engineering in Information Science and Engineering in Seventh Semester of the Visvesvaraya Technological University, Belagavi** during the year **2022-23.** It is certified that all corrections/suggestions indicated for internal assessment have been incorporated in the report. The project work phase-I report has been approved as it satisfies the academic requirements with respect to seventh semester project work phase-I work prescribed for the above said degree.

.....	.....	.....
Signature of the Guide	**Signature of the Coordinator**	**Signature of the HOD**
AMBIKA L G	**ARAVINDA THEJAS CHANDRA**	**SATHEESH CHANDRA REDDY**
Assistant Professor	**Associate Professor**	**Professor & HOD**
Dept. of ISE, SJCIT	**Dept. of ISE, SJCIT**	**Dept. of ISE, SJCIT**

ABSTRACT

Over the last decade, e-learning has revolutionized how students learn by providing them access to quality education whenever and wherever they want. However, students often get distracted because of various reasons, which affect the learning capacity to a great extent. Many researchers have been trying to improve the quality of online education, but we need a holistic approach to address this issue. This project intends to provide a mechanism that uses the camera feed and microphone input to monitor the real-time attention level of students during online classes. We explore various image processing techniques and machine learning algorithms throughout this study. We propose a system that uses five distinct non-verbal features to calculate the attention score of the student during computer based tasks and generate real-time feedback for both students and the organization. We can use the generated feedback as a heuristic value to analyze the overall performance of students as well as the teaching standards of the lecturers.

ACKNOWLEDGEMENT

With Great pride we would like to convey our gratitude and appreciation to our Institution **"SJC Institute of Technology"** for giving us the required platform for the fulfillment of the Phase-1 Project report on "**Real-time attention span tracking in online education**" as per the V.T.U requirements, for the seventh semester.

We express our sincere thanks to **Dr. G T Raju**, Principal of **SJCIT** for providing us with excellent infrastructure and facilities to complete the Phase-1 Project

We express whole hearted gratitude to **Prof. Satheesh Chandra Reddy,** who is the respected **HOD ofInformation Science and Engineering Department**. We wish to acknowledge his help in making ourtask easy by providing us with his valuable help and encouragement.

We express our sincere thanks to Project coordinator **Prof. Aravinda Thejas Chandra, Associate Professor, Information Science and Engineering, SJCIT** providing us with excellence infrastructure to complete Final year project phase 1.

It is our pleasure to thank our guide **Prof. Ambika L G, Assistant Professor, Department of Information Science and Engineering, SJCIT** for his guidance, encouragement and valuable suggestion from the beginning of the Phase-1 Project work till the completion without which this project work would not have been accomplished.

And last but not the least, we would be very pleased to express our heart full thanks to the teaching and non-teaching staff of the Department of Information Science and Engineering, SJCIT for their motivation and support.

We also thank all those who extended their support and co-operation for successful completion of Phase-1 Project

ADARSH K R (1SJ19IS002)
RAKSHITHA H G (1SJ19IS128)
G YOGA SAI GANESH (1SJ19IS036)
GOVARDHAN J A (1SJ19IS038)

CONTENTS

ABSTRACT **i**

ACKNOWLEDGEMENT **ii**

TABLE OF CONTENTS **iii**

LIST OF FIGURES **v**

LIST OF TABLES **v**

TABLE OF CONTENTS

CHAPTER NO	TITLE	PAGE NO
CHAPTER 1	INTRODUCTION	1
CHAPTER 2	PROBLEM IDENTIFICATION AND DEFINITION	2
CHAPTER 3	LITERATURE SURVEY	3-4
CHAPTER 4	SYSTEM REQUIREMENT SPECIFICATION	5
4.1	HARDWARE REQUIREMENT	5
4.2	SOFTWARE REQUIREMENT	5
CHAPTER 5	OBJECTIVES	6
CHAPTER 6	METHODOLGY	7-13
6.1	SOFTWARE ARCHITECTURE	7
6.2	HAAR CASCADE	9
6.3	LBPH ALGORITHM	10
6.4	CNN ALGORITHM	13
CHAPTER 7	PLAN OF EXECUTION	14-15
	CONCLUSION	16
	BIBLIOGRAPHY	17

LIST OF FIGURES

FIGURE NO.	NAME OF THE FIGURE	PAGE NO
FIGURE 6.1	SYSTEM ARCHITECTURE	7
FIGURE 6.2	FLOW DIAGRAMS	8
FIGURE 6.3	HAAR FEATURE EXTRACTION	9
FIGURE 6.4	FEATURE MAPPING	9
FIGURE 6.5	LBPH ARCHITECTURE	10
FIGURE 6.6	OPEN POSE ARCHITECTURE	13

LIST OF TABLE

TABLE NO.	NAME OF THE TABLE	PAGE NO.
TABLE 7.1	GANTT CHART	15

CHAPTER - 1

INTRODUCTION

Over the last decade, e-learning has revolutionized how students learn by providing them access to quality education whenever and wherever they need. However, students often get distracted due to various reasons, which affect the learning capacity to an excellent extent. Many researchers are tryingto enhance the quality of online education, but we'd like a holistic approach to deal with this issue. This paper intends to provide a mechanism that uses the camera feed and microphone input to watchthe real-time attention level of students during online classes. We explore various image processing techniques and machine learning algorithms throughout this study. We propose a system that uses five distinct non-verbal features to calculate the eye score of the student during computer based tasksand generate real time feedback for both students and the organization. we will use the generated feedback as a heuristic value to research the overall performance of students as well because the teaching standards of the lecturers.

The demand and need for online education are increasing rapidly. Almost all the schools and collegesthroughout the world have shifted to the online mode of lectures and exams due to the recent corona virus outbreak, and this trend will most likely continue in the upcoming years. The increasing demandfor online education opens the gate to automation in the field. One major issue in the online mode oflectures is that students tend to lose their concentration after a certain period and there is no automatedmechanism to monitor their activities during the classes. Some students tend to just start a lecture online and move away from the place, or might even use a proxy to write online tests for them. This situation also takes place in online course platforms such as EdX and Coursera where the student triesto skip lectures just for the sake of completion and certification. The loss in concentration not only affects the student’s knowledge level but also hurts the society by producing low-skilled laborers.

CHAPTER-2

PROBLEM IDENTIFICATION AND DEFINITION

2.1 Problem Identification

Some students tend to just start a lecture online and move away from the place, or might even use a proxy to write online tests for them. This situation also takes place in online course platforms such asEdX and Coursera where the student tries to skip lectures just for the sake of completion and certification. The loss in concentration not only affects the student's knowledge level but also hurts the society by producing low-skilled laborers.

2.2 Problem Definition

The demand and need for online education are increasing rapidly. Almost all the schools and collegesthroughout the world have shifted to the online mode of lectures and exams due to the recent corona virus outbreak, and this trend will most likely continue in the upcoming years. The increasing demandfor online education opens the gate to automation in the field. One major issue in the online mode oflectures is that students tend to lose their concentration after a certain period and there is no automatedmechanism to monitor their activities during the classes. Some students tend to just start a lecture online and move away from the place, or might even use a proxy to write online tests for them.

CHAPTER-3

LITERATURE SURVEY

A literature survey or a literature review in a project report shows the various analyses and research made in the field of interest and the results already published, taking into account the various parameters of the project and the extent of the project. Literature survey is mainly carried out in orderto analyze the background of the current project which helps to find out flaws in the existing system& guides on which unsolved problems we can work out. So, the following topics not only illustrate the background of the project but also uncover the problems and flaws which motivated to propose solutions and work on this project.

Literature survey describes about the existing work on the given project. It deals with the problem associated with the existing system and also gives user a clear knowledge on how to deal with the existing problems and how to provide solution to the existing problems

Objectives of Literature Survey

- Learning the definitions of the concepts.
- Access to latest approaches, methods and theories.
- Discovering research topics based on the existing research
- Concentrate on your own field of expertise– Even if another field uses the same words, theyusually mean completely.
- It improves the quality of the literature survey to exclude sidetracks– Remember to explicatewhat is excluded.

Before building our application, the following system is taken into consideration:

CHAPTER-4

SYSTEM REQUIREMENTS

System Requirement Specification (SRS) is a central report, which frames the establishment of the product advancement process. It records the necessities of a framework as well as has a depiction of its significant highlight. An SRS is essentially an association's seeing (in composing) of a client or potential customer's frame work necessities and conditions at a specific point in time (generally) before any genuine configuration or improvement work. It's a two-way protection approach that guarantees that both the customer and the association comprehend alternate's necessities from that viewpoint at a given point in time.

Software Requirements:

- **Operating System :** Windows 10,11 / Linux
- **Programming Language** : Python 3.2
- **Libraries:** Open CV, Scikit Image, NumPy
- **IDE:** Visual Studio Code
- **Browser:** Google Chrome

Hardware Requirements:

- **Processor** Intel i5 or above
- **Processor speed:** 2.44 GHz or above
- **RAM:** 8 GB or above
- **Storage space:** 120GB or above
- **Camera:** USB camera
- **Internet:** 5 Mbps

CHAPTER-5

OBJECTIVES

- The purpose of attention span detection during online classes is to gather data and analyze thestate of the student, to evaluate his performance based on concentration level, instead of just academic scores.
- To present a method that uses a camera feed and a mouthpiece contribution to monitor student's continuing attention levels during online classes.
- To improve the quality of online education.
- To tackle the issues involved in online education using five parameters. We used the facerecognition model to verify the student attending the online class.
- we have implemented and used lightweight models to reduce the processing time.
- We visualize the scores in the form of a live graph and generate automated reports.

CHAPTER-6

METHODOLOGY

System architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description and representation of a system, organized in a way that support reasoning about the structures and behaviors of the system.

1.1 System Architecture:

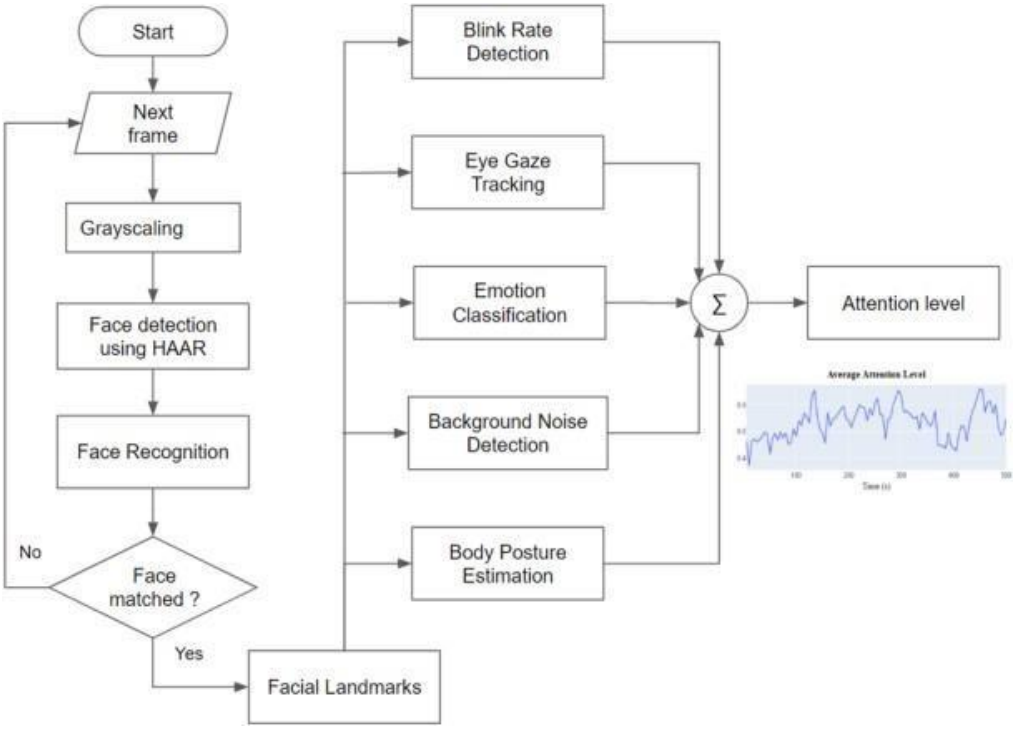


Figure 6.1 System architecture

This system makes use of five parameters to calculate the attention-span level of the student attendingthe online class. Facial recognition is used to validate the student's attendance. The attention span score is calculated using blink rate, facial expression, eye gaze, background noise, and body posture and is updated continuously for a window length of 5 seconds. Instead of sequential execution, all themodels required to calculate the attention span are executed in parallel once the online lecture starts.

SL NO	TITLE	METHDOLOGY USED	ADVANTAGES	DISADVANTAGES
1	A Survey on State-of-the-Art-Drowsiness Detection Techniques.	SVM, Hidden Markov Model, CNN.	It provides details of behavioral, vehicular and physiological parameters-based drowsiness detection.	It consumes more time to train the model. Accuracy is less is than 80%.
2	An Experimental Study on the Influence of Environmental Noise on Students Attention.	Visual tracking test.	In this paper most participants had the highest accuracy under the influence of low environmental.	Accuracy is less. It is suitable to monitor the noise in online classes.
3	Real-time Smart Attendance System using Face Recognition Techniques.	Eigenface values, Principle Component Analysis (PCA) and Convolutional Neural Network(CNN).	It give 85% accuracy in smart attendance system.	It is suitable to make automatic attendance. It is not suitable to track attention span of students.
4	Deep Sparse Representation Classifier for facial recognition and detection system.	Two-layer Convolutional Neural Network(CNN).	SRC provides better classification result even if a simple feature extraction method is used.	It is not suitable to track attention span of students.
5	An Emotion Recognition Model Based on facial Recognition in Virtual Learning Environment.	Haar cascade with CNN.	It can help teaches to change teaching strategies in virtual learning environment.	It is not suitable to track attention span of students.

6.1 Methodology:

- We are planning to implement this project using the following steps.
- To detect Facial Landmark we are using Haar cascade algorithm.
- To Face Recognition we are using LBPH algorithm
- To detect body pose we are using CNN or R-CNN

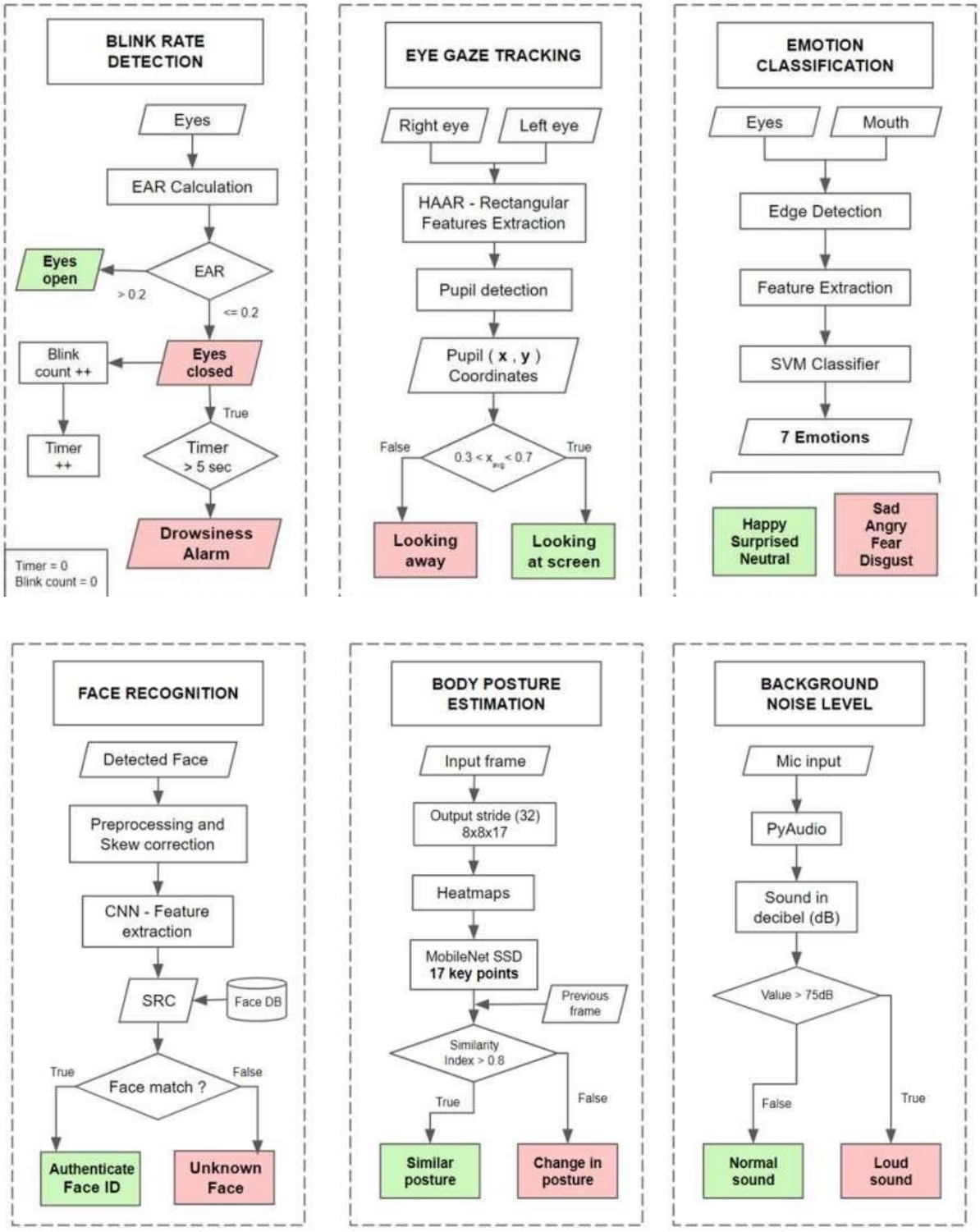


Fig 6.2:Flow Diagrams

6.2.1 Haar Cascade

Here we will work with face detection. Initially, the algorithm needs a lot of positive images (images of faces) and negative images (images without faces) to train the classifier. Then we need to extract features from it. For this, haar features shown in below image are used. They are just like our convolutional kernel. Each feature is a single value obtained by subtracting sum of pixels under white rectangle from sum of pixels under black rectangle.

Now all possible sizes and locations of each kernel is used to calculate plenty of features. (Just imagine how much computation it needs? Even a 24x24 window results over 160000 features). For each feature calculation, we need to find sum of pixels under white and black rectangles. To solve this, they introduced the integral images. It simplifies calculation of sum of pixels, how large may be the number of pixels, to an operation involving just four pixels. Nice, isn't it? It makes things super-fast.

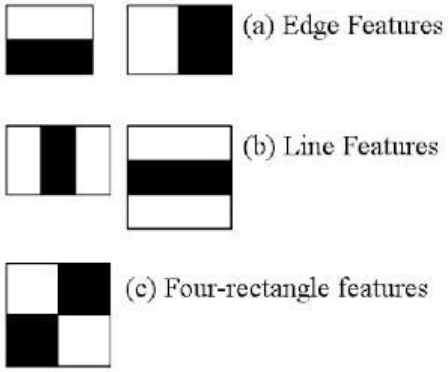


Fig 6.3 Haar Features extraction

But among all these features we calculated, most of them are irrelevant. For example, consider the image below. Top row shows two good features. The first feature selected seems to focus on the property that the region of the eyes is often darker than the region of the nose and cheeks. The second feature selected relies on the property that the eyes are darker than the bridge of the nose. But the same windows applying on cheeks or any other place is irrelevant. So how do we select the best features out of 160000+ features? It is achieved by Adaboost.

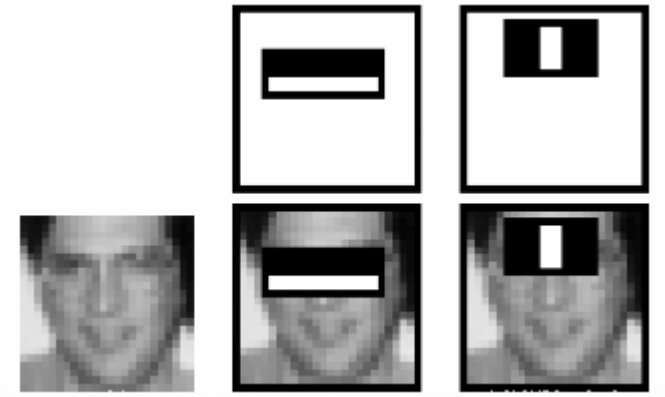


Fig 6.4 Feature Mapping

For this, we apply each and every feature on all the training images. For each feature, it finds the

best threshold which will classify the faces to positive and negative. But obviously, there will be errors or misclassifications. We select the features with minimum error rate, which means they are the features that best classifies the face and non-face images. (The process is not as simple as this. Each image is given an equal weight in the beginning. After each classification, weights of misclassified images are increased. Then again same process is done. New error rates are calculated. Also new weights. The process is continued until required accuracy or error rate is achieved or required number of features are found).

Final classifier is a weighted sum of these weak classifiers. It is called weak because it alone can't classify the image, but together with others forms a strong classifier. The paper says even 200 features provide detection with 95% accuracy. Their final setup had around 6000 features. (Imagine a reduction from 160000+ features to 6000 features. That is a big gain).

6.2.2 LBPH

Local Binary Patterns Histogram algorithm. It is based on local binary operator. It is widely used in facial recognition due to its computational simplicity and discriminative power.

The steps involved to achieve this are:

- creating dataset
- face acquisition
- feature extraction
- Classification

The LBPH algorithm is a part of opencv.

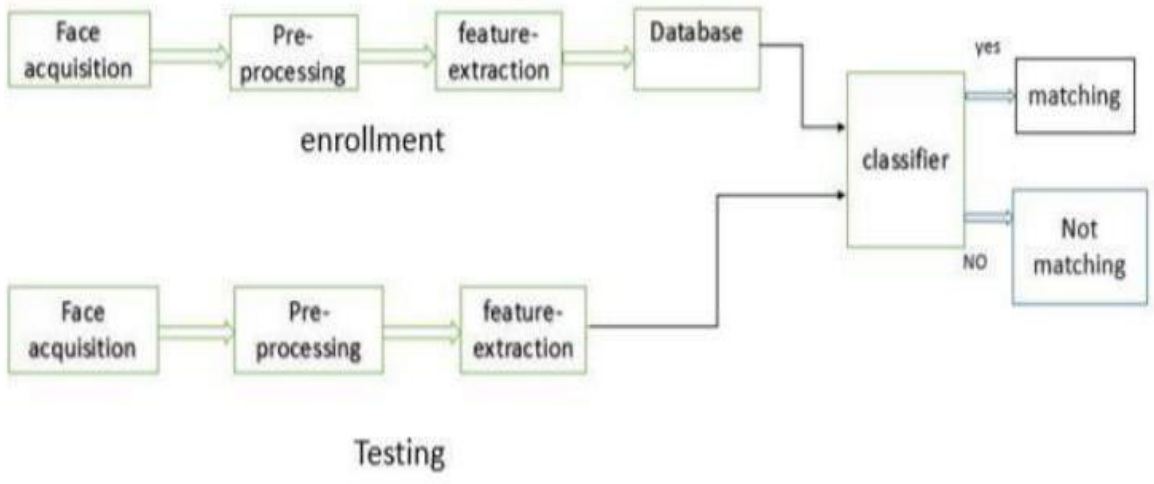


Fig 6.5 LBPH architecture

STEPS:

- Suppose we have an image having dimentions N x M.
- We divide it into regions of same height and width resulting in m x m dimension for every region.

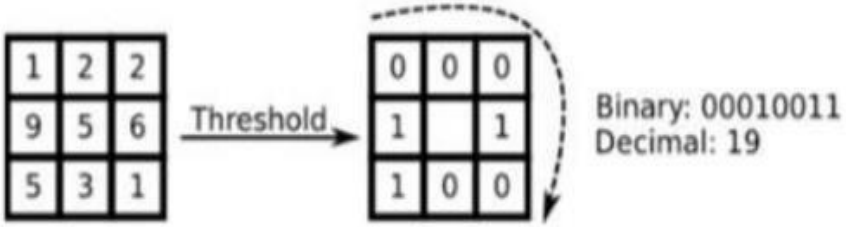
- Local binary operator is used for every region. The LBP operator is defined in window of 3x3.

$$LBP(x_c, y_c) = \sum_{p=0}^{P-1} 2^p s(i_p - i_c)$$

- here '(Xc,Yc)' is central pixel with intensity 'Ic'. And 'In' being the intensity of the neighbor pixel.
- Using median pixel value as threshold, it compares a pixel to its 8 closest pixels using this function.

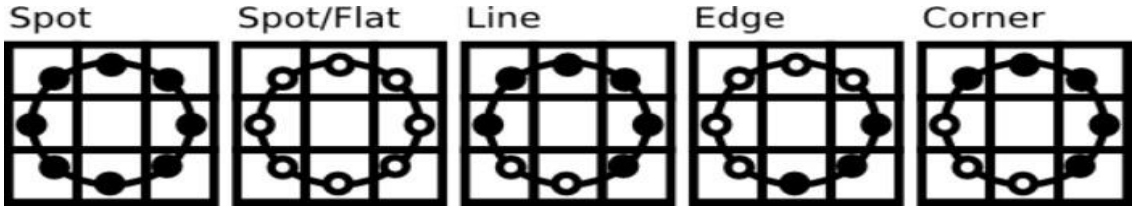
$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- If the value of neighbor is greater than or equal to the central value it is set as 1 otherwise it is set as 0.
- Thus, we obtain a total of 8 binary values from the 8 neighbors.
- After combining these values we get a 8 bit binary number which is translated to decimal number for our convenience.
- This decimal number is called the pixel LBP value and its range is 0-255.



- Later it was noted that a fixed neighborhood fails to encode details varying in scale .The algorithm was improved to use different number of radius and neighbors , now it was known as circular LBP.

- The idea here is to align an arbitrary number of neighbors on a circle with a variable radius. This way the following neighborhoods are captured:



- For a given point (Xc,Yc) the position of the neighbor (Xp,Yp), p belonging to P can be calculated by:

$$x_p = x_c + R\cos(\frac{2\pi p}{P})$$
$$y_p = y_c - R\sin(\frac{2\pi p}{P})$$

- here R is radius of the circle and P is the number of sample points.
- If a points coordinate on the circle doesn't correspond to image coordinates, it get's interpolated generally by bilinear interpolation:

$$f(x,y) \approx \begin{bmatrix} 1-x & x \end{bmatrix} \begin{bmatrix} f(0,0) & f(0,1) \\ f(1,0) & f(1,1) \end{bmatrix} \begin{bmatrix} 1-y \\ y \end{bmatrix}$$

- The LBP operator is robust against monotonic gray scale transformations.

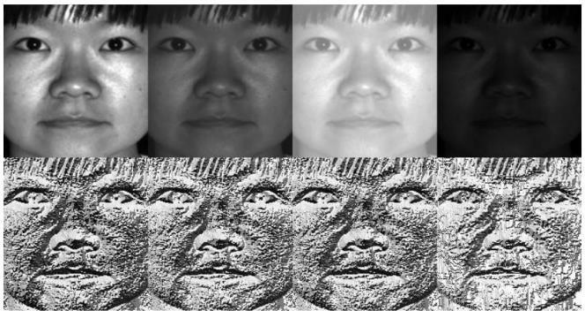
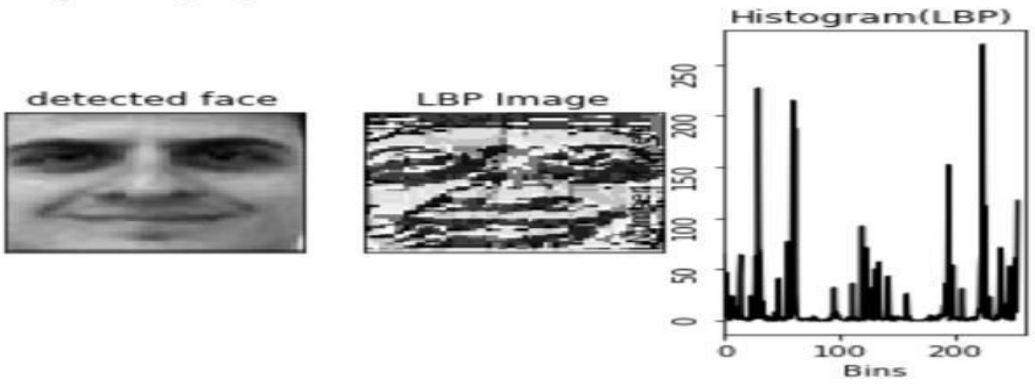


Fig 6.6:Monotonic image

- After the generation of LBP value histogram of the region is created by counting the number of similar LBP values in the region.
- After creation of histogram for each region all the histograms are merged to form a single histogram and this is known as feature vector of the image.
- Now we compare the histograms of the test image and the images in the database and then we return the image with the closest histogram.

$$d(a,b) = \sqrt{\sum_{i=1}^{n} |a_i - b_i|^2}$$

- (This can be done using many techniques like euclidean distance, chi-square, absolute value etc)
- The Euclidean distance is calculated by comparing the test image features with features stored in the dataset. The minimum distance between test and original image gives the matching rate.
- As an output we get an ID of the image from the database if the test image is recognized.



6.2.3 CNN

OpenPose also uses CNN as its main architecture. It consists of a VGG-19 convolutional network that is used to extract patterns and representations from the given input. The output from the VGG-19 goes into two branches of convolutional networks.

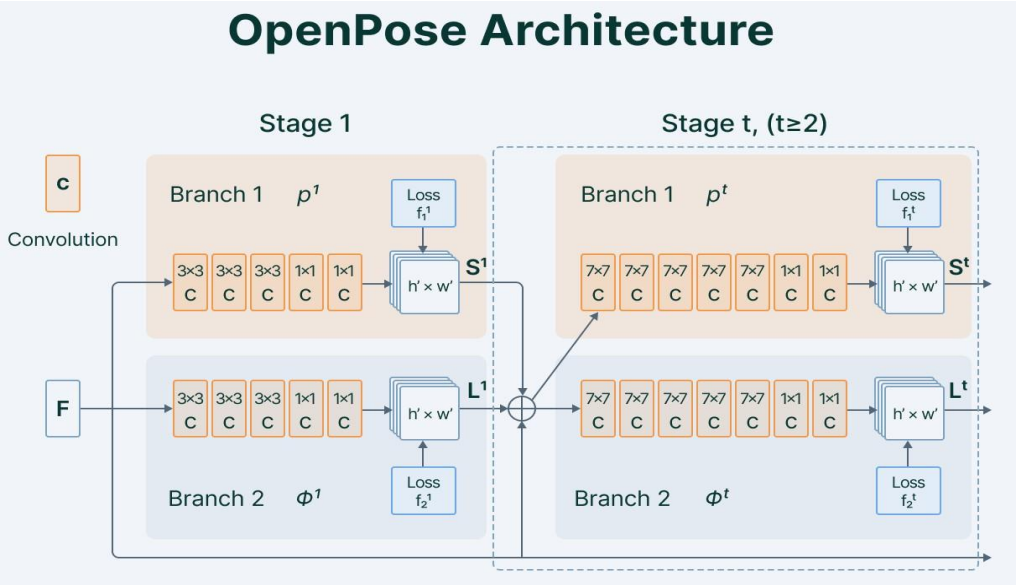


Fig 6.7: Open Pose Architecture

CHAPTER 7

PROJECT PLANING

Project planning is a discipline addressing how to complete a project in certain timeframe,usually with defined stages and designated resources.

7.2 Overview

During the execution phase, the project team develops the product or service and presents final product to customer.
There are 3 execution phases

- **Build Deliverables:**

We have completed the first two phases of this project. We begins by reviewing therequirements of the project.

- **Monitor and Control:**

This is the part of the execution phase because we must follow several steps to ensure the project meets the requirements such as

4. **Time management**: During the time management process, we controls the amount of time our team members spend working on each activity and monitor the stakeholders.

5. **Change management**: if a change in the project's scope of work we must formally request the change and get the changes approved.

6. **Risk Management**: This involves handling the potential disadvantages anddrawbacks that may arise during a project.

- **Review**: This helps you document the results of our project review, at the end of the execution project phase.

    - This is conducted at the end of the initiation, planning and execution phases within a project.
    - Project is currently delivering to schedule.
    - Risks have been controlled and mitigated.

7.1 Gantt Chart

Table7.1 Gantt Chart

	Sep 20	Oct 20	Nov 10	Dec 10	Feb 10	Mar 30	Apr 15	Apr 30	May 10	May 15
Planning phase										
Literature survey										
Analysis phase										
Design phase										
Implementation										
Testing										
Deployment										
Documentation Report										

CHAPTER 8

CONCLUSION

we have implemented a system to tackle the issues involved in online education using five parameters. We used the face recognition model to verify the student attending the online class. We used the other five parameters - blink rate, eye gaze, emotion, posture, and noise level to calculate the attention level of the student throughout the lecture. Since this involves real-time processing, we have implemented and used lightweight models to reduce the processing time. We visualize the scores in the form of a live graph and generate automated reports. The feedback generated can be used for: 1) Evaluating student performance 2) Improving teaching standards 3) Preventing malpractice during online examinations As a part of future works, we can improve our system's performance further by training our models using more data. Also, the same attention tracking mechanism can be further optimized to simultaneously work with multiple subjects in a classroom using video footage from the CCTV cameras. Moreover, we have used human observed attention scores as ground truth-values as we currently do not have any dataset for measuring the attention span during online lectures. A standard dataset can help to evaluate the system's performance more reliably.

BIBLIOGRAPHY

[1] M. Ramzan, H. U. Khan, S. M. Awan, A. Ismail, M. Ilyas, and A. Mahmood, "A Survey on State- of-the- Art Drowsiness Detection Techniques", in IEEE Access, vol. 7, pp. 61904-61919, 2019.

[2] Zhen Zhang, Yuan Zhang, "An Experimental Study on the Influence of Environmental Noise on Student'sAttention", EuroNoise 2018 conference.

[3] S. Sawhney, K. Kacker, S. Jain, S. N. Singh, and R. Garg, "Real-time Smart Attendance … System using Face Recognition Techniques", 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2019,
pp. 522-525.

[4] Eric-Juwei Cheng, Kuang-Pen Chou, Shantanu Rajora, Bo-Hao Jin, M. Tanveer, Chin-Teng Lin,Ku-Young Young, Wen-Chieh Lin, Mukesh Prasad,"Deep Sparse Representation Classifier for facial recognition and detection system".

[5] D. Yang, Abeer Alsadoona, P.W.C. Prasad*a, A.K. Singh, A. Elchouemi, "An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment".

[6] Bixler, R., D'Mello,"Automatic gaze-based user-independent detection of mind wandering during computerized reading". User Model User-Adap Inter 26, 33–68 2016.

[7] Liping Shen, Minjuan Wang, and Ruimin Shen, "Affective e-Learning: Using "Emotional" Data to Improve Learning in a Pervasive Learning Environment", International Forum of Educational Technology & Society 2009.

[8] C. Morimoto, D. Koons, A. Amir, and M. Flickner, "Pupil detection and tracking using multiple light sources," Image Vis. Comput, vol. 18, no. 4, pp. 331–336, 2008.