

# FACE RECOGNITION WEB APP USING OPENCV

A

*Project Report*

*Submitted*

*in partial fulfillment*

*for the award of the Degree of*

*Bachelor of Technology*

**In**

**Information Technology**



**Supervisor**

**Mrs. Nirmala Sharma**

Assistant Professor

**Ms. Parul Chouhan**

Assistant Professor

**Submitted By**

Neeraj Arora 14EUCIT019

Jagdish Verma 14EUCIT013

Nikhil Gurjar 14EUCIT020

**Department of Computer Science & Engineering**  
**Rajasthan Technical University, Kota**  
**Session (2017-18)**



## CERTIFICATE

---

This is to certify that **Neeraj Arora , Jagdish Verma and Nikhil Gurjar** of VIII Semester, B. Tech (Information Technology) “2017-2018”, has completed a project titled “FACE RECOGNITION WEB APP USING OPENCV” in partial fulfillment for the award of the degree of Bachelor of Technology under Rajasthan Technical University.

**Dr. Vikas Panthi**

**Assistant Professor**

**(Deptt. Of Computer Science & Engg.)**

## **CANDIDATE’S DECLARATION**

---

We hereby declare that the work, which is being presented in the Project, entitled **“FACE RECOGNITION WEB APP USING OPENCV”** in fulfillment for the award of Degree of **“Bachelor of Technology”** in Department of Computer Science & Engineering with Specialization in Information Technology, and submitted to the **Department of Computer Science & Engineering**, Rajasthan Technical University is a record of my own investigations carried under the Guidance of Dr. Vikas Panthi, Department of Computer Science & Engineering.

We have not submitted the matter presented in this Report anywhere for the award of any other Degree.

**Jagdish Verma      14EUCIT013**

**Nikhil Gurjar      14EUCIT020**

**Neeraj Arora      14EUCIT019**

**Date :**



## ACKNOWLEDGEMENT

---

Thank you GOD for giving me strength and ability to understand, learn and complete this report.

We have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals and teachers. We would like to extend my sincere thanks to all of them. The special environment at University College of Engineering, Rajasthan Technical University, Kota that always support educational activities, facilitated our work on this project.

We highly indebted to Dr. Vikas Panthi for their guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in completing the project.

We acknowledge the support, the encouragement , extended for this study by our  
Project Head Dr. Nirmala Sharma and Mrs. Parul Chouhan.

We would like to express our gratitude towards our parents & faculties of CS department for their kind co-operation and encouragement which help us in completion of this project. Our thanks and appreciations also go to our colleague in developing the project and people who have willingly helped us out with their abilities.

**Neeraj Arora 14EUCIT019**

**Jagdish Verma 14EUCIT013**

**Nikhil Gurjar 14EUCIT020**

## **ABSTRACT**

---

Face recognition project we propose a new locality preserving projections based approach called as LPP is implemented for mapping images in to subspace for analysis. This approach is different from other approaches like principle component analysis and linear discriminant analysis approaches. Which works on limited analysis on image analysis? In this new method it considers local information and obtains facial subspace which will detect most of the facial manifolds. This system works on Laplacian faces technology which will eliminate basic unwanted variations like light, facial expressions.

# TABLE OF CONTENTS

---

	Page no.
<b>1 INTRODUCTION</b>	<b>1</b>
1.1 Project Specification.....	1
1.2 Features of OpenCV Library.....	1
1.3 OpenCV Library Modules.....	2
1.4 Image Recognition.....	3
<b>2 REQUIREMENT ANALYSIS</b>	<b>5</b>
2.1 Software Requirements.....	5
2.1.1 Python.....	5
2.1.2 JAVA.....	6
2.1.3 Apache Tomcat Server.....	7
2.1.4 OpenCV Library.....	7
2.1.5 MySQL Database.....	7
2.1.6 Netbeans.....	8
2.1.7 Machine Learning.....	9
<b>3 SYSTEM OVERVIEW &amp; DESIGN</b>	<b>10</b>
<b>4 IMPLEMENTATION</b>	<b>16</b>
4.1 WORKING.....	16

4.1.1	User Section.....	16
4.1.2	Admin Section.....	19
<b>5</b>	<b>TESTING AND EVALUATION</b>	<b>24</b>
5.1	Performance Graph.....	24
5.2	LBPH Accuracy.....	25
<b>6</b>	<b>CONCLUSION AND FUTURE SCOPE</b>	<b>26</b>
6.1	Conclusion.....	26
6.2	Future Work.....	27
	<b>REFERENCES.....</b>	<b>28</b>

## **INTRODUCTION**

---

### **1.1PROJECT SPECIFICATION**

Facial recognition is a biometric software application capable of uniquely identifying or verifying a person by comparing and analyzing patterns based on the person's facial contours. Facial recognition is mostly used for security purposes, though there is increasing interest in other areas of use.

The Animetrics Face Recognition API can be used to detect human faces in pictures. Information on facial features or landmarks is returned as coordinates on the image.. Animetrics Face Recognition will also detect and return the orientation, or pose of faces along 3 axes.

Objective of our web App is to create a website in which face recognition is done using OpenCV library.The web page lets you create id on the website and let you upload pictures in the groups on website. Uploaded faces compares new photos to the existing collection to find a match.

We are creating a web app for a college. In this student can signup and upload Pictures by acting anonymous in various groups. Alumnis of the college can see their photos after a long time. There's a admin of this web app who can see details of each person in the images which are uploaded by users with help of face detection and recognition using OpenCV library.

Computer Vision can be defined as a discipline that explains how to reconstruct, interrupt, and understand a 3D scene from its 2D images, in terms of the properties of the structure present in the scene. It deals with modeling and replicating human vision using computer software and hardware.Image processing deals with image-to-image transformation. The input and output of image processing are both images.

Computer vision is the construction of explicit, meaningful descriptions of physical objects from their image. The output of computer vision is a description or an interpretation of structures in 3D scene.

### **1.2 Features of OpenCV Library**

1. Read and Write Images.
2. Detection of faces and its features.
3. Detection of shapes like Circle,rectangle etc in a image.
- EX. Detection of coin in images.
4. Text recognition in images. e.g Reading Number Plates/



5. Modifying image quality and colors e.g Instagram, CamScanner.

6. Developing Augmented reality apps.

### **1.3 OpenCV Library Modules**

Following are the main library modules of the OpenCV library.

#### **Core Functionality**

This module covers the basic data structures such as Scalar, Point, Range, etc., that are used to build OpenCV applications. In addition to these, it also includes the multidimensional array Mat, which is used to store the images. In the Java library of OpenCV, this module is included as a package with the name org.opencv.core.

#### **Image Processing**

This module covers various image processing operations such as image filtering, geometrical image transformations, color space conversion, histograms, etc. In the Java library of OpenCV, this module is included as a package with the name org.opencv.imgproc.

#### **Video**

This module covers the video analysis concepts such as motion estimation, background subtraction, and object tracking. In the Java library of OpenCV, this module is included as a package with the name org.opencv.video.

#### **Video I/O**

This module explains the video capturing and video codecs using OpenCV library. In the Java library of OpenCV, this module is included as a package with the name org.opencv.videoio.

#### **calib3d**

This module includes algorithms regarding basic multiple-view geometry algorithms, single and stereo camera calibration, object pose estimation, stereo correspondence and elements of 3D reconstruction. In the Java library of OpenCV, this module is included as a package with the name org.opencv.calib3d.

#### **features2d**

This module includes the concepts of feature detection and description. In the Java library of OpenCV, this module is included as a package with the name org.opencv.features2d.

#### **Objdetect**

This module includes the detection of objects and instances of the predefined classes such as faces, eyes, mugs, people, cars, etc. In the Java library of OpenCV, this module is included as a package with the name org.opencv.objdetect.

## **Highgui**

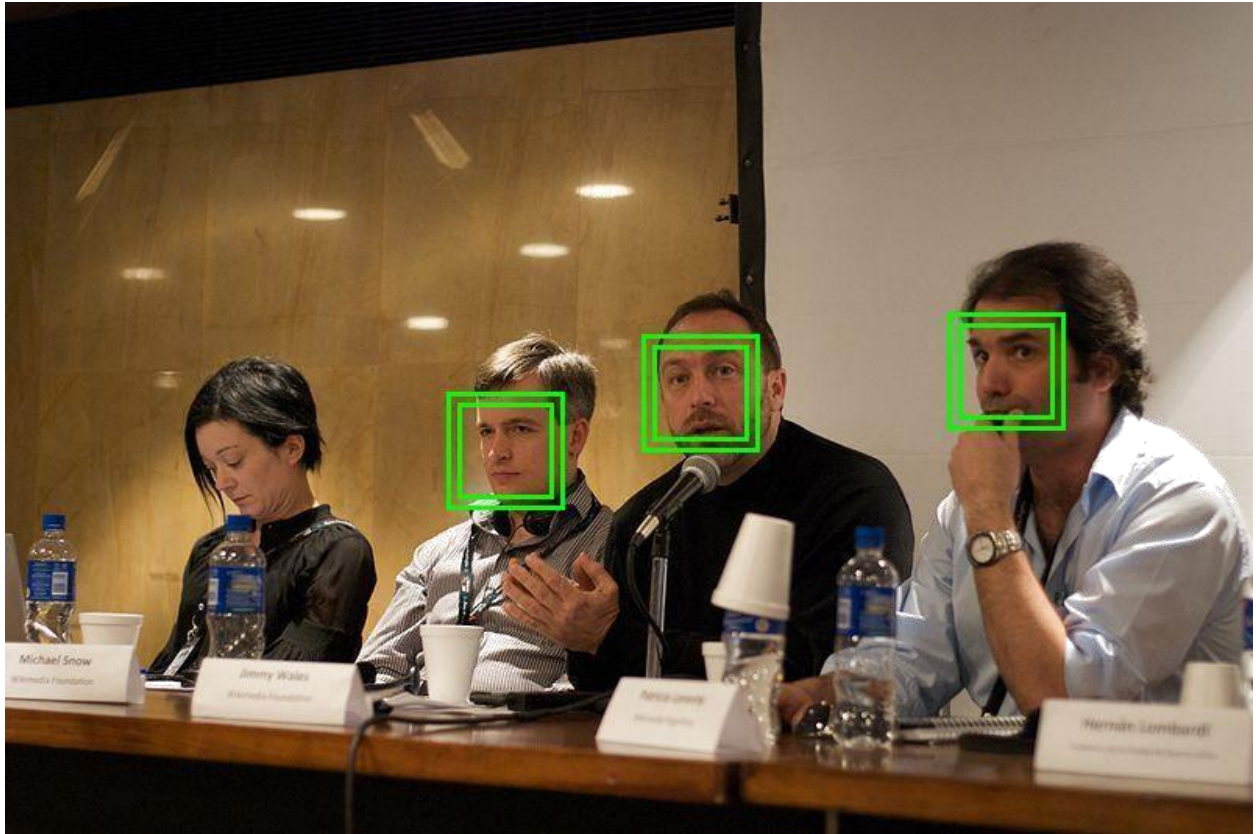
This is an easy-to-use interface with simple UI capabilities. In the Java library of OpenCV, the features of this module is included in two different packages namely, `org.opencv.imgcodecs` and `org.opencv.videoio`.

### **1.4 Image Recognition**

Image recognition is used to perform a large number of machine-based visual tasks, such as labeling the content of images with meta-tags, performing image content search and guiding autonomous robots, self-driving cars and accident avoidance systems.

While human and animal brains recognize objects with ease, computers have difficulty with the task. Software for image recognition requires deep machine learning. Performance is best on convolutional neural net processors as the specific task otherwise requires massive amounts of power for its compute-intensive nature. Image recognition algorithms can function by use of comparative 3D models, appearances from different angles using edge detection or by components. Image recognition algorithms are often trained on millions of pre-labeled pictures with guided computer learning.

Current and future applications of image recognition include smart photo libraries, targeted advertising, the interactivity of media, accessibility for the visually impaired and enhanced research capabilities. Google, Facebook, Microsoft, Apple and Pinterest are among the many companies that are investing significant resources and research into image recognition and related applications. Privacy concerns over image recognition and similar technologies are controversial as these companies can pull a large volume of data from user photos uploaded to their social media platforms.



To compare models, we examine how often the model fails to predict the correct answer as one of their top 5 guesses -- termed "top-5 error rate". AlexNet achieved by setting a top-5 error rate of 15.3% on the 2012 validation data set; Inception (GoogLeNet) achieved 6.67%; BN-Inceptionv2 achieved 4.9%; Inception-v3 reaches 3.46%.

## REQUIREMENT ANALYSIS

---

The chapter is intended to be accessible for enthusiasts, engineers, and data scientists at all skill levels. The only skills that you will need are a basic understanding of Python and enough knowledge of the command line to setup a project.

### 2.1 Software Requirements

The software requirements in this project include:

- Python
- JAVA
- Apache Tomcat Server
- OpenCV library
- MySQL database
- Netbeans
- Machine Learning

#### 2.1.1 Python:

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. ... Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. Python is used for creating backbone structure. Python is intended to be a highly readable language. It is designed to have an uncluttered visual layout, it uses whitespace indentation, rather than curly braces or keywords. Python has a large standard library, commonly cited as one of Python's greatest strengths. Python features a dynamic type

system and automatic memory management. It supports multiple programming paradigms, including object-oriented, imperative, functional and procedural, and has a large and comprehensive standard library.

### 2.1.2 JAVA

**Java** is a general-purpose computer-programming language **that** is concurrent, class-based, object-oriented,<sup>[15]</sup> and specifically designed to have as few implementation dependencies as possible. It is intended to let application developers "write once, run anywhere" (WORA),<sup>[16]</sup> meaning that compiled Java code can run on all platforms that support Java without the need for recompilation.<sup>[17]</sup> Java applications are typically compiled to bytecode that can run on any Java virtual machine (JVM) regardless of computer architecture. As of 2016, Java is one of the most popular programming languages in use,<sup>[18]</sup><sup>[19]</sup><sup>[20]</sup><sup>[21]</sup> particularly for client-server web applications, with a reported 9 million developers.<sup>[22]</sup> Java was originally developed by James Gosling at Sun Microsystems (which has since been acquired by Oracle Corporation) and released in 1995 as a core component of Sun Microsystems' Java platform. The language derives much of its syntax from C and C++, but it has fewer low-level facilities than either of them. The original and reference implementation Java compilers, virtual machines, and class libraries were originally released by Sun under proprietary licenses. As of May 2007, in compliance with the specifications of the Java Community Process, Sun relicensed most of its Java technologies under the GNU General Public License. Others have also developed alternative implementations of these Sun technologies, such as the GNU Compiler for Java (bytecode compiler), GNU Classpath (standard libraries), and IcedTea-Web (browser plugin for applets).

#### **What is JavaServer Pages?**

JavaServer Pages (JSP) is a technology for developing Webpages that supports dynamic content. This helps developers insert java code in HTML pages by making use of special JSP tags, most of which start with `<%` and end with `%>`.

A JavaServer Pages component is a type of Java servlet that is designed to fulfill the role of a user interface for a Java web application. Web developers write JSPs as text files that combine HTML or XHTML code, XML elements, and embedded JSP actions and commands.

Using JSP, you can collect input from users through Webpage forms, present records from a database or another source, and create Webpages dynamically.

JSP tags can be used for a variety of purposes, such as retrieving information from a database or registering user preferences, accessing JavaBeans components, passing control between pages, and sharing information between requests, pages etc.

### **2.1.3 Apache Tomcat Server**

Apache Tomcat, often referred to as Tomcat Server, is an open-source Java Servlet Container developed by the Apache Software Foundation (ASF). Tomcat implements several Java EE specifications including Java Servlet, JavaServer Pages (JSP), Java EL, and WebSocket, and provides a "pure Java" HTTPweb serverenvironment in which Javacode can run.

Tomcat is developed and maintained by an open community of developers under the auspices of the Apache Software Foundation, released under the Apache License 2.0 license, and is open-source software.

### **2.1.4 OpenCV Library**

OpenCV (Open Source Computer Vision Library) is released under a BSD license and hence it's free for both academic and commercial use. It has C++, Python and Java interfaces and supports Windows, Linux, Mac OS, iOS and Android. OpenCV was designed for computational efficiency and with a strong focus on real-time applications. Written in optimized C/C++, the library can take advantage of multi-core processing. Enabled with OpenCL, it can take advantage of the hardware acceleration of the underlying heterogeneous compute platform.

Adopted all around the world, OpenCV has more than 47 thousand people of user community and estimated number of downloads exceeding 14 million. Usage ranges from interactive art, to mines inspection, stitching maps on the web or through advanced robotics.

### **2.1.5 MySQL Database**

MySQL is a freely available open source Relational Database Management System (RDBMS) that uses Structured Query Language (SQL). SQL is the most popular language for adding, accessing and managing content in a database. ... MySQL is an essential part of almost every open source PHP application.

### **2.1.6 Netbeans**

NetBeans is an open-source integrated development environment (IDE) for developing with Java, PHP, C++, and other programming languages. NetBeans is also referred to as a platform of modular components used for developing Java desktop applications.

### 2.1.7 Machine Learning

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

Some machine learning methods

Machine learning algorithms are often categorized as supervised or unsupervised.

- Supervised machine learning algorithms can apply what has been learned in the past to new data using labeled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly.
- In contrast, unsupervised machine learning algorithms are used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data.
- Semi-supervised machine learning algorithms fall somewhere in between supervised and unsupervised learning, since they use both labeled and unlabeled data for training – typically a small amount of labeled data and a large amount of unlabeled data. The systems that use this method are able to considerably improve learning accuracy. Usually, semi-supervised learning is chosen when the acquired labeled data requires skilled and relevant resources in order to train it / learn from it. Otherwise, acquiring unlabeled data generally doesn't require additional resources.
- Reinforcement machine learning algorithms is a learning method that interacts with its environment by producing actions and discovers errors or rewards. Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning. This method allows machines and software agents to automatically determine the ideal behavior within a specific context in order to maximize its performance. Simple reward

feedback is required for the agent to learn which action is best; this is known as the reinforcement signal.

Machine learning enables analysis of massive quantities of data. While it generally delivers faster, more accurate results in order to identify profitable opportunities or dangerous risks, it may also require additional time and resources to train it properly. Combining machine learning with AI and cognitive technologies can make it even more effective in processing large volumes of information.



## SYSTEM OVERVIEW & DESIGN

---

Design is the first step in the development phase for any techniques and principles for the purpose of defining a device, a process or system in sufficient detail to permit its physical realization. Once the software requirements have been analyzed and specified the software design involves three technical activities design, coding, implementation and testing that are required to build and verify the software. The design activities are of main importance in this phase, because in this activity, decisions ultimately affecting the success of the software implementation and its ease of maintenance are made..

The system uses information extraction technique to parse job descriptions and resumes, and it gets information such as skills, job titles and education background. The information is used to create the models of job openings and job seekers. A domain specific ontology is used to construct the knowledge base, which includes the taxonomies that support resume-job matching. Our system follows the three tier architecture . First tier consist of GUI, Processing block and the Database.

**GUI:** The GUI(Graphical User Interface) in our project deals with the interface for the user where the user will submit his resumes and job description in any format(pdf, doc, docx,ect.).The GUI provides a platform for the user to communicate with the database. It acts as a connector as well as communicator which connects the database and helps in transfer of data between the GUI and the database.

**Processing block:** Processing block is the block where the actual processing of our project is done. This block connects the gui to the database i.e. it acts as a connector as well as communicator which connects the database and helps in transfer of data between the gui and the database. Its main function is to take input from resumes, convert them to text and parse it to store the information in database. After storing this information this system will give output using web application.

**Database:** Database tier is the tier used for the storage of data. This tier contains all the data that is need for the processing of the whole project. The data in this tier is related to the student information gathered form his/her resumes.

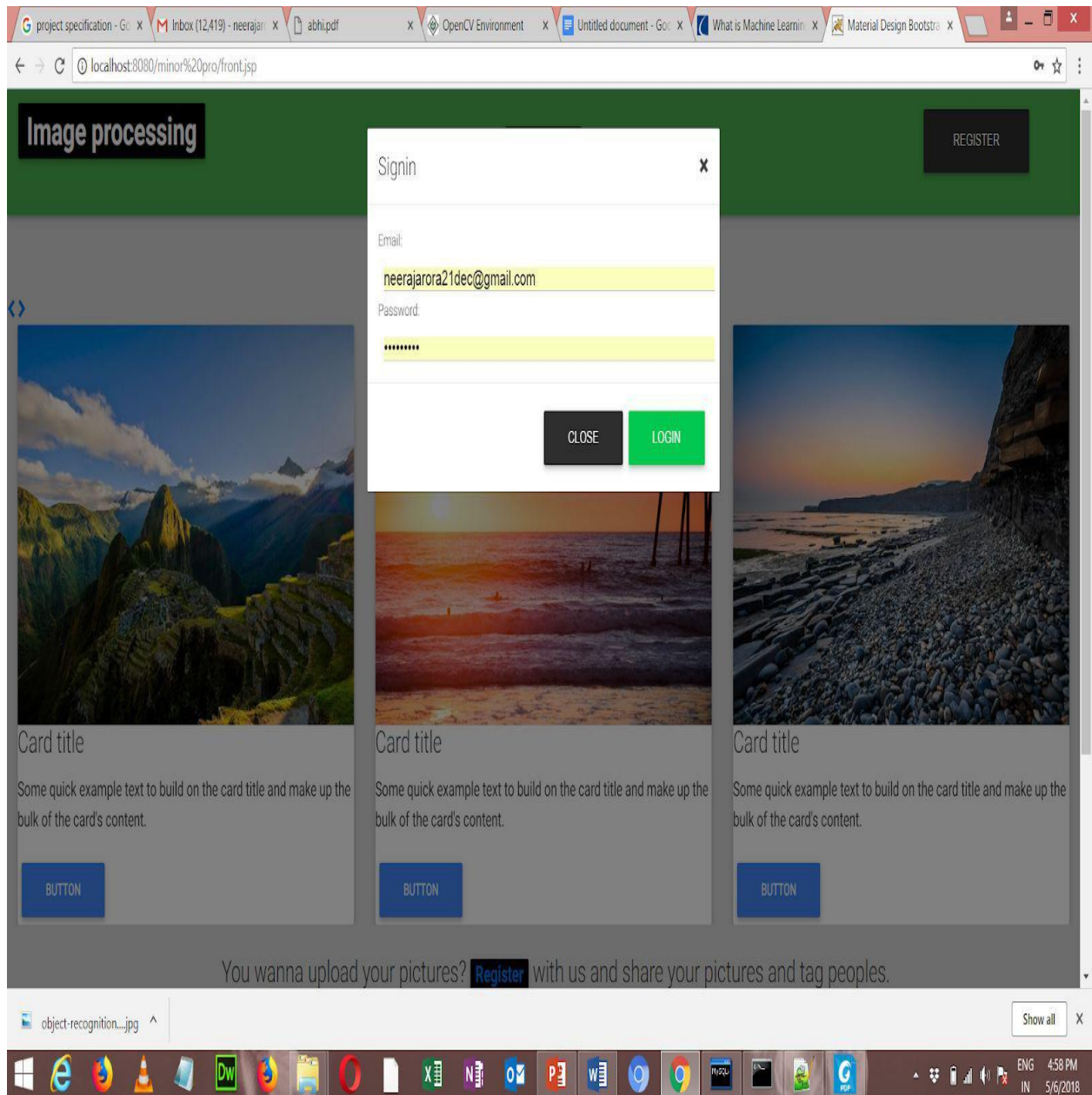


Fig. Home page of Web App

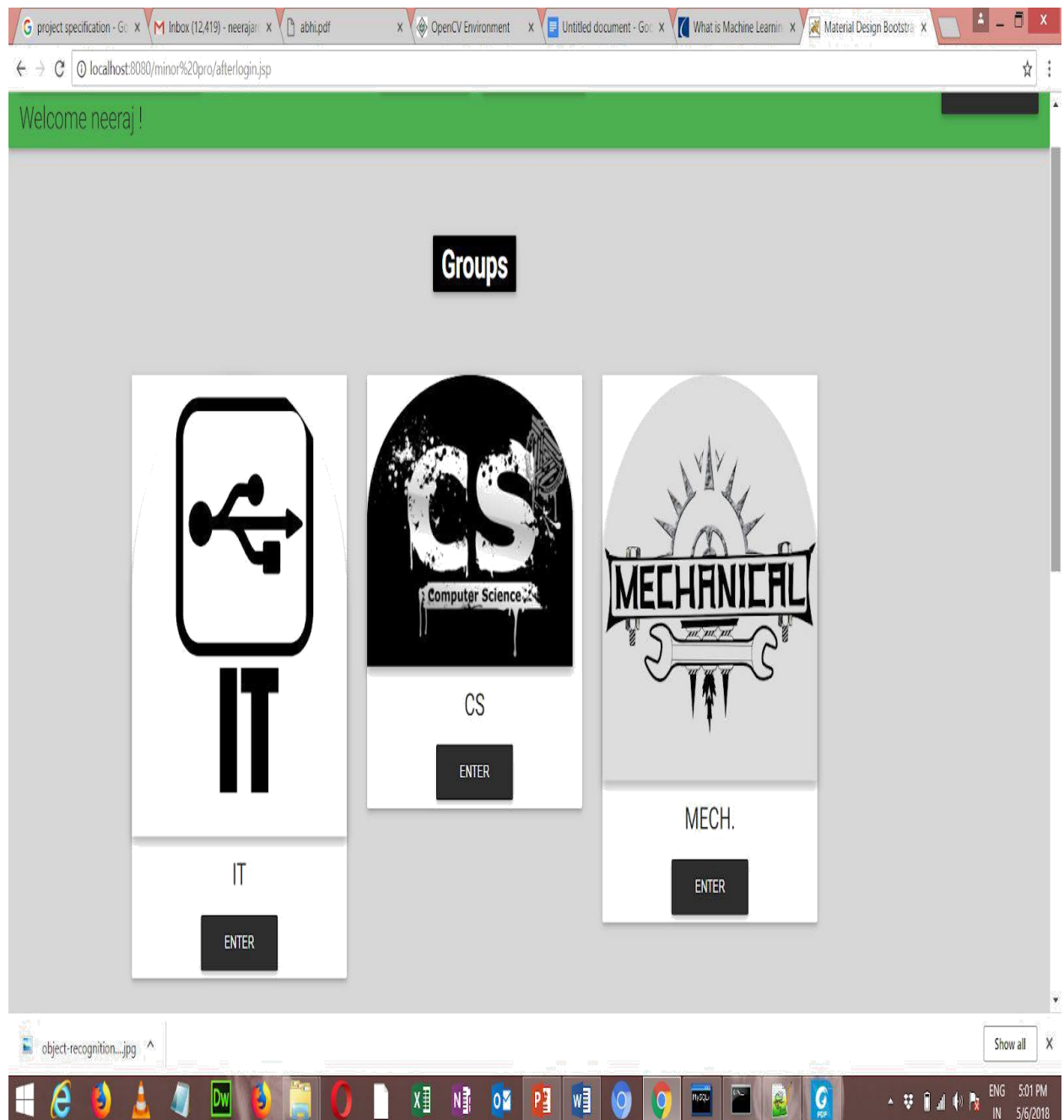


Fig. After Login page

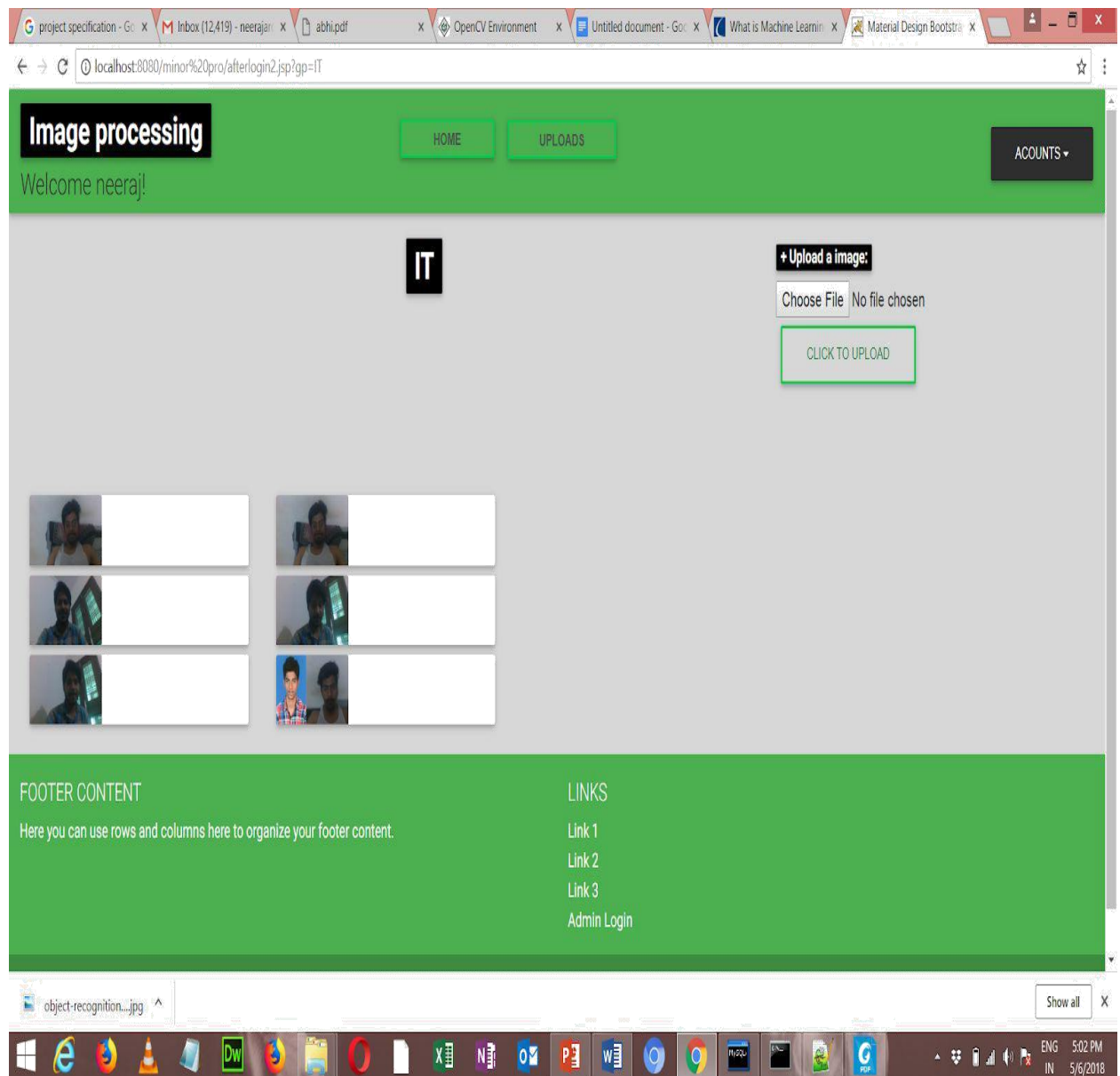


Fig. Upload Section

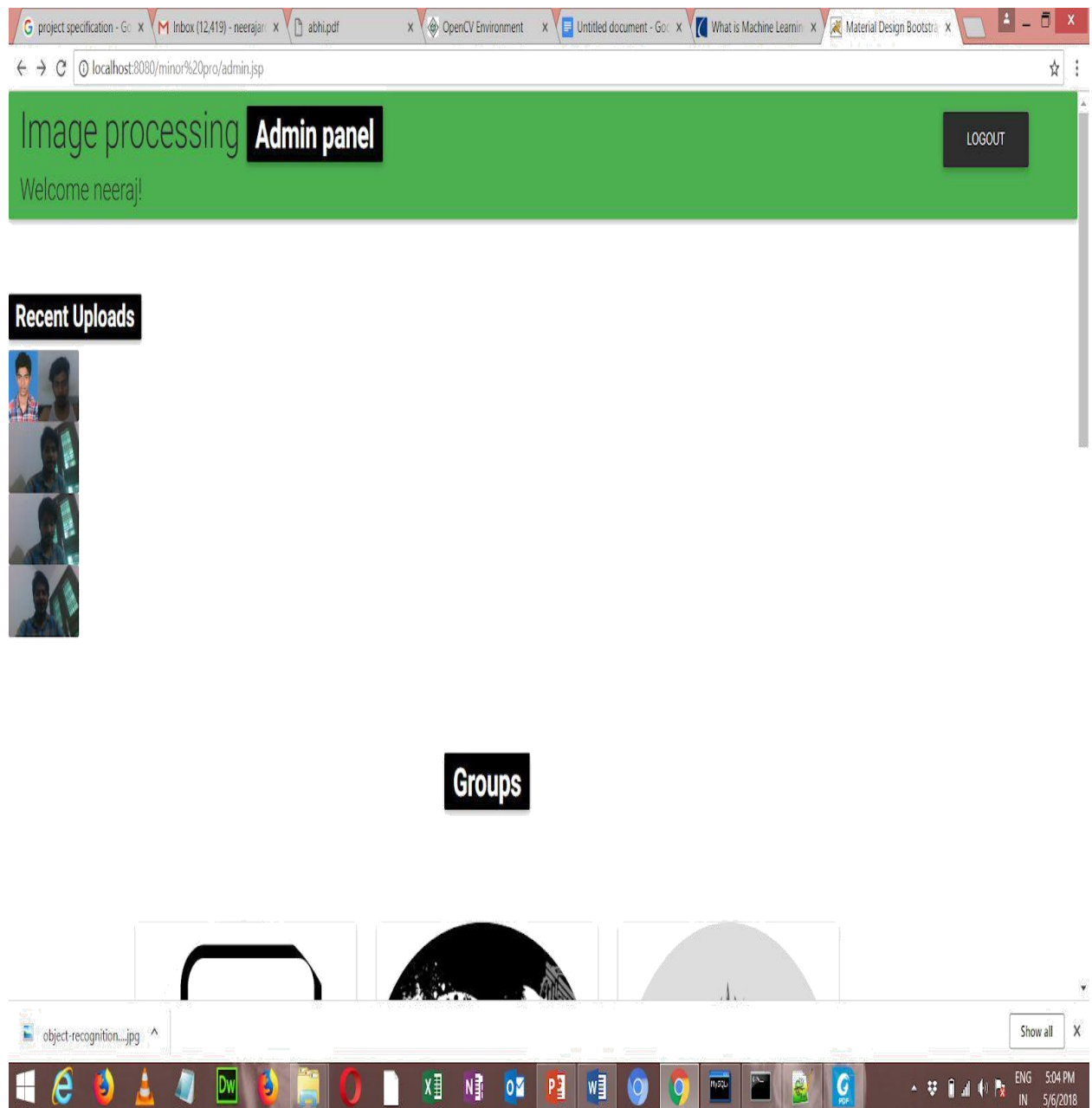


Fig. Admin Section

### **3.2 Overview of how website works?**

#### User Section

1. Register on the website.
2. Login into website.
3. Select a group.
4. Upload photos.

#### Admin Section

- 1.Login as admin.
- 2.Recognize faces of anonymous users with OpenCV library automatically.

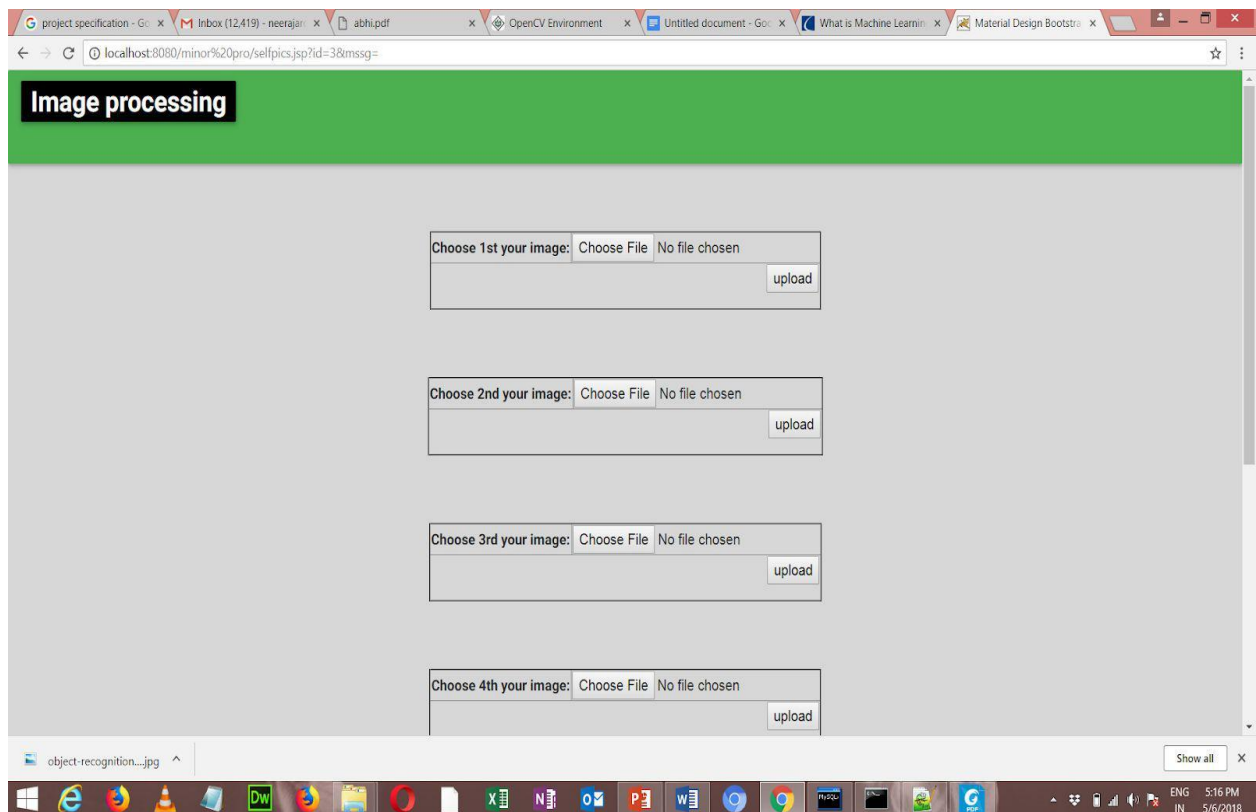
## IMPLEMENTATION

---

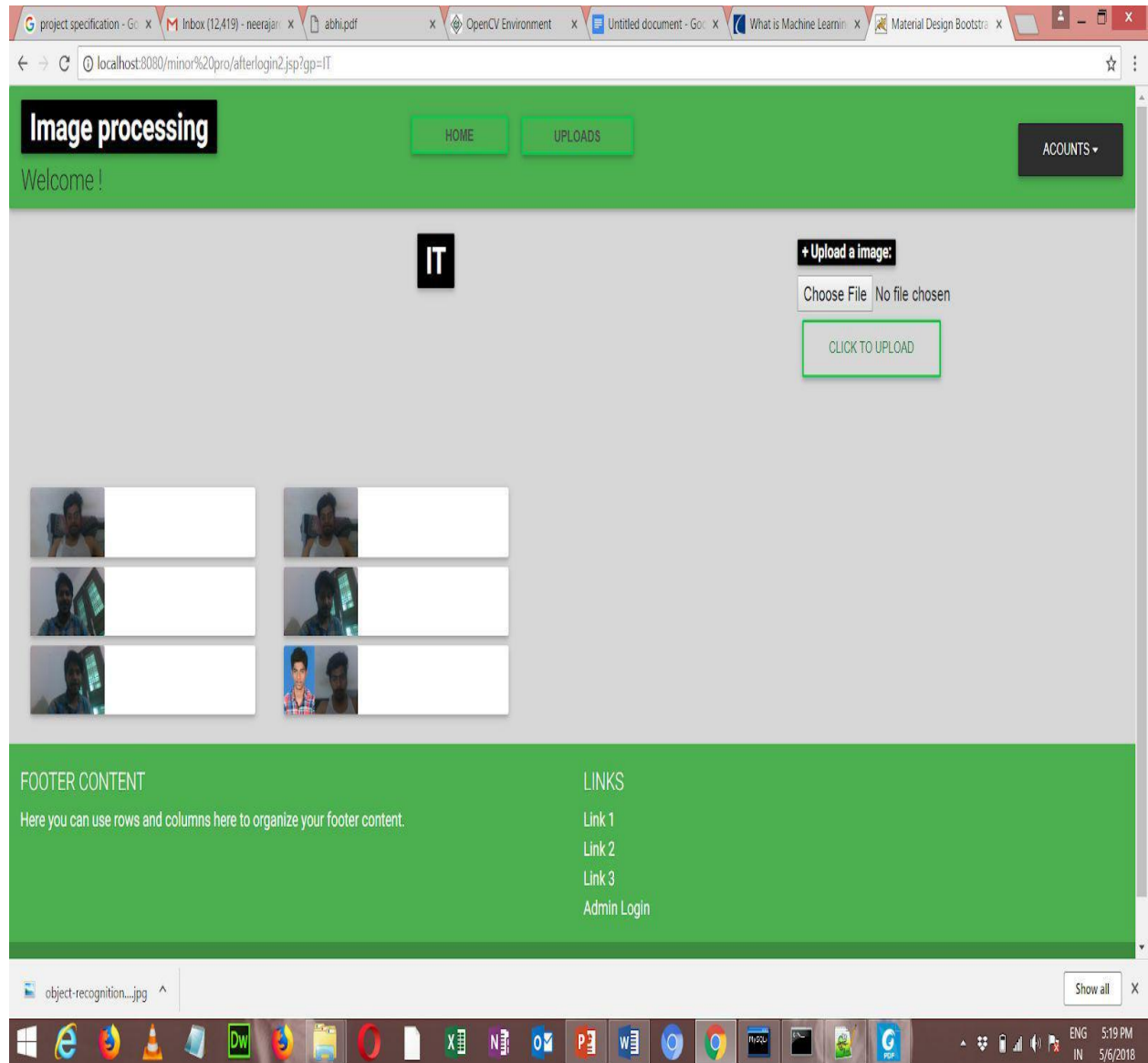
### 4.1 WORKING

#### 4.1.1 User Section

1. Go to registration page by clicking register button on top right of the home page.
2. Fill your entries and upload your 5 photos in which your face is there only.
3. Now go to login page after your registration.



4. After login now you can choose any group in which you wanna upload pictures.



Here Suppose we choose IT group and now we can upload pictures and we can also tag peoples if we want to tag in ou photos once we upload the theres a tag option then our webapp automatically recognzine faces from database with the help of machine learning and we can send tag request to person once they approve tag ,tag is made on that photo.



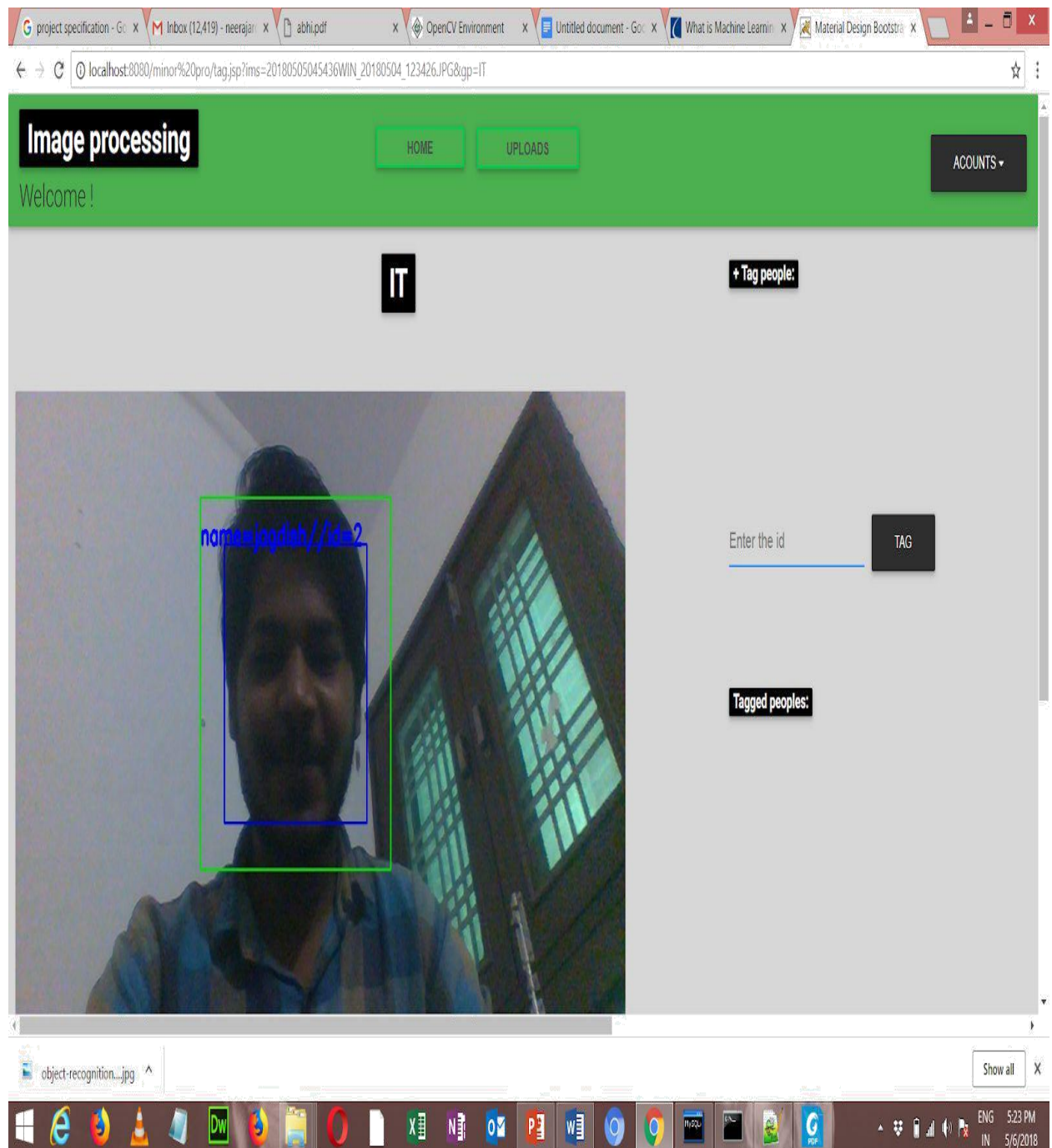
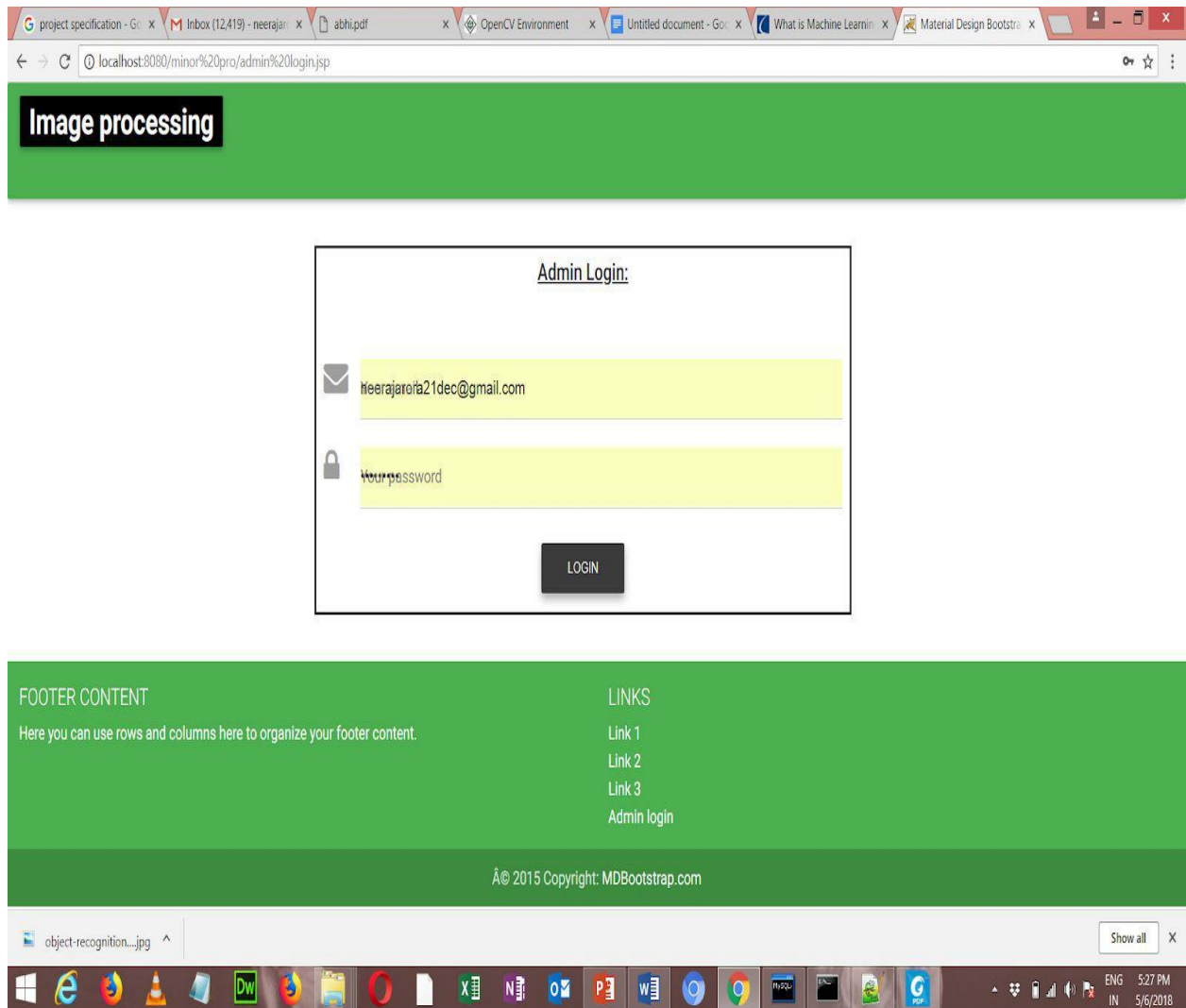


Fig. Tag page on uploaded photo

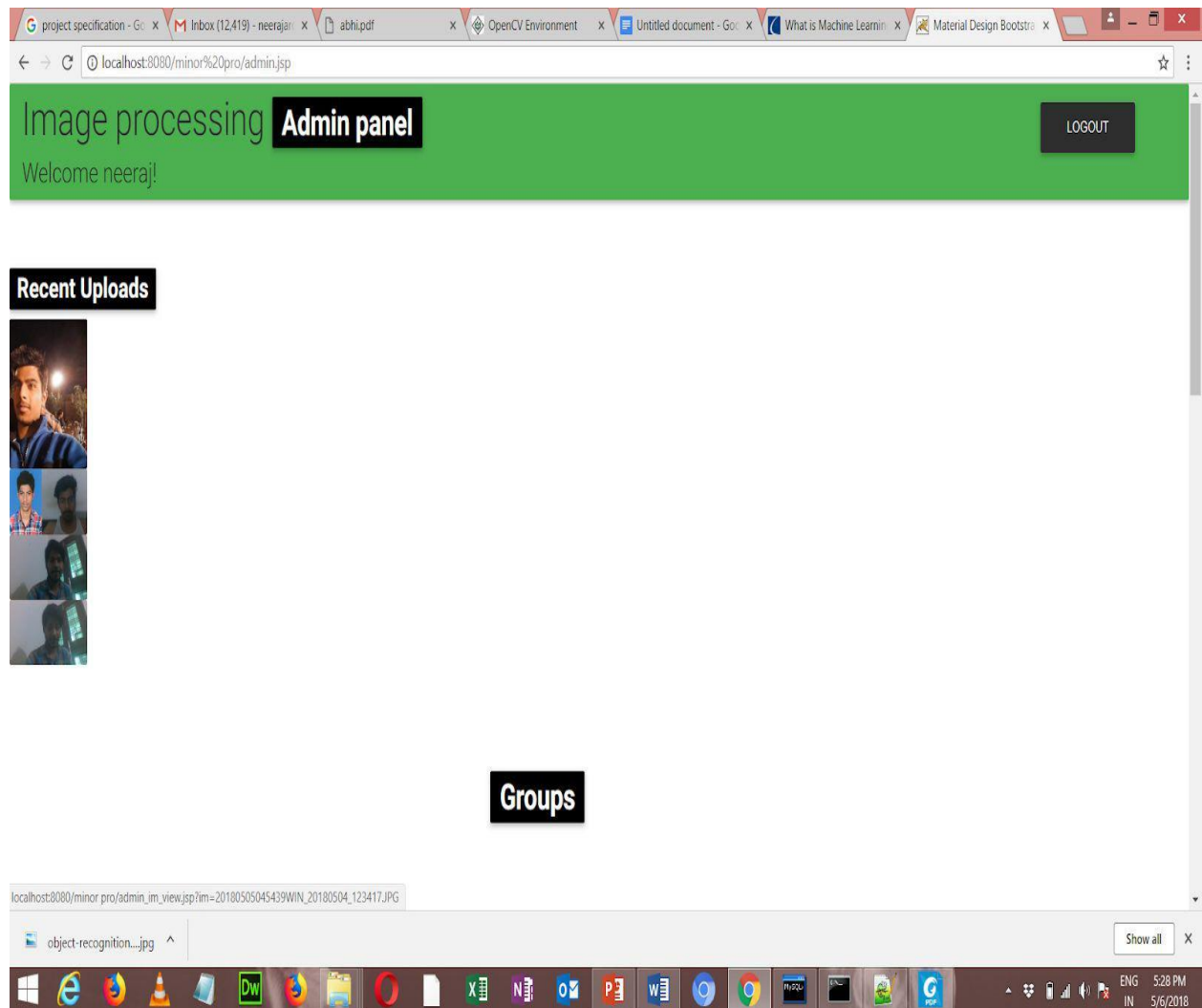
## 5.1.2 Admin Section

1. Admin of the website are registered by directly filling entries on database.
2. Login as admin .



Fig, Admin Login

### 3. Check out the users uploaded pictures.



Here Admin can see the recent uploaded photos .he can select any of the photos for checking details of any face int the photo.Admin can also choose groups to check out the photos uploaded in the groups.

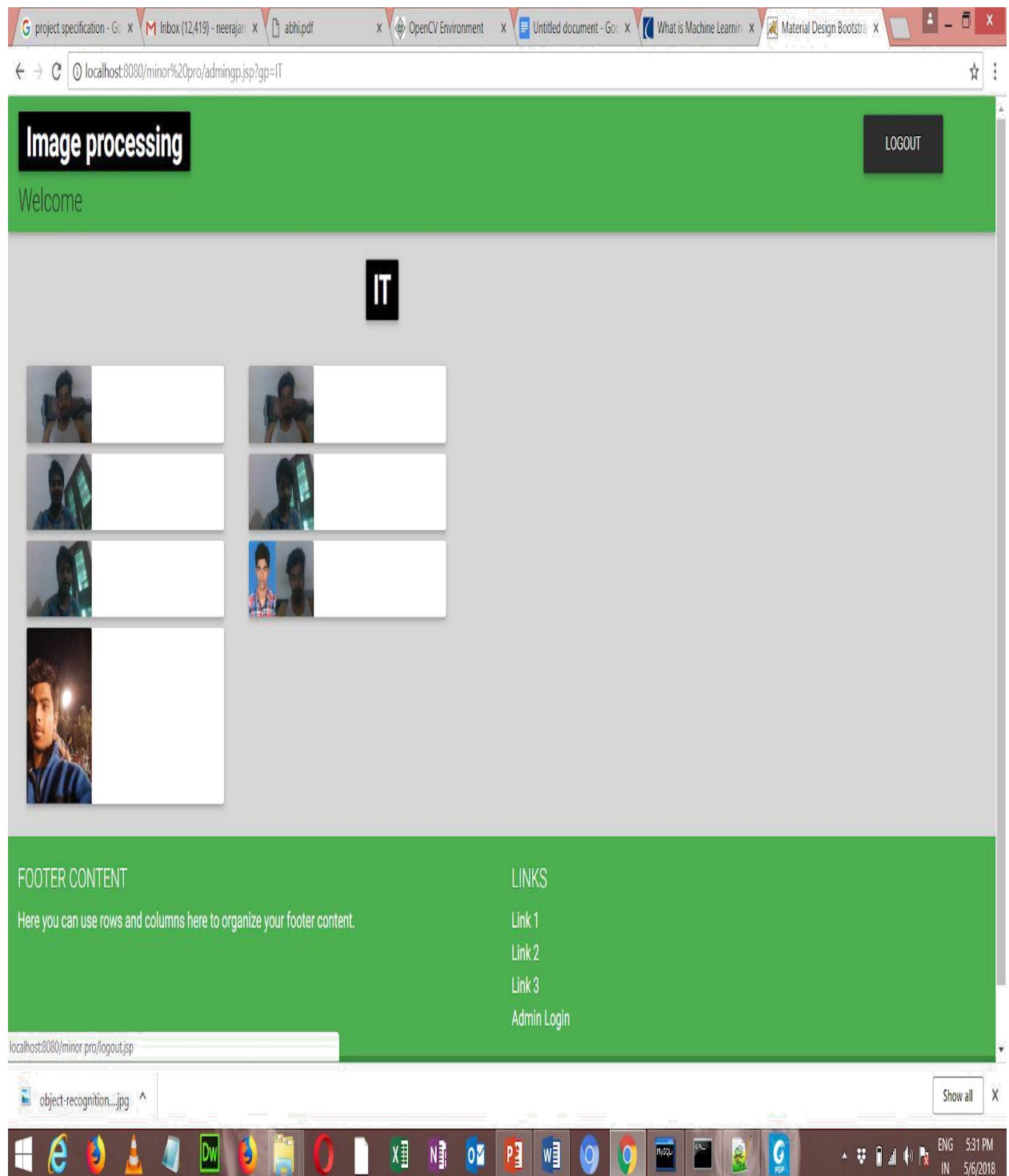
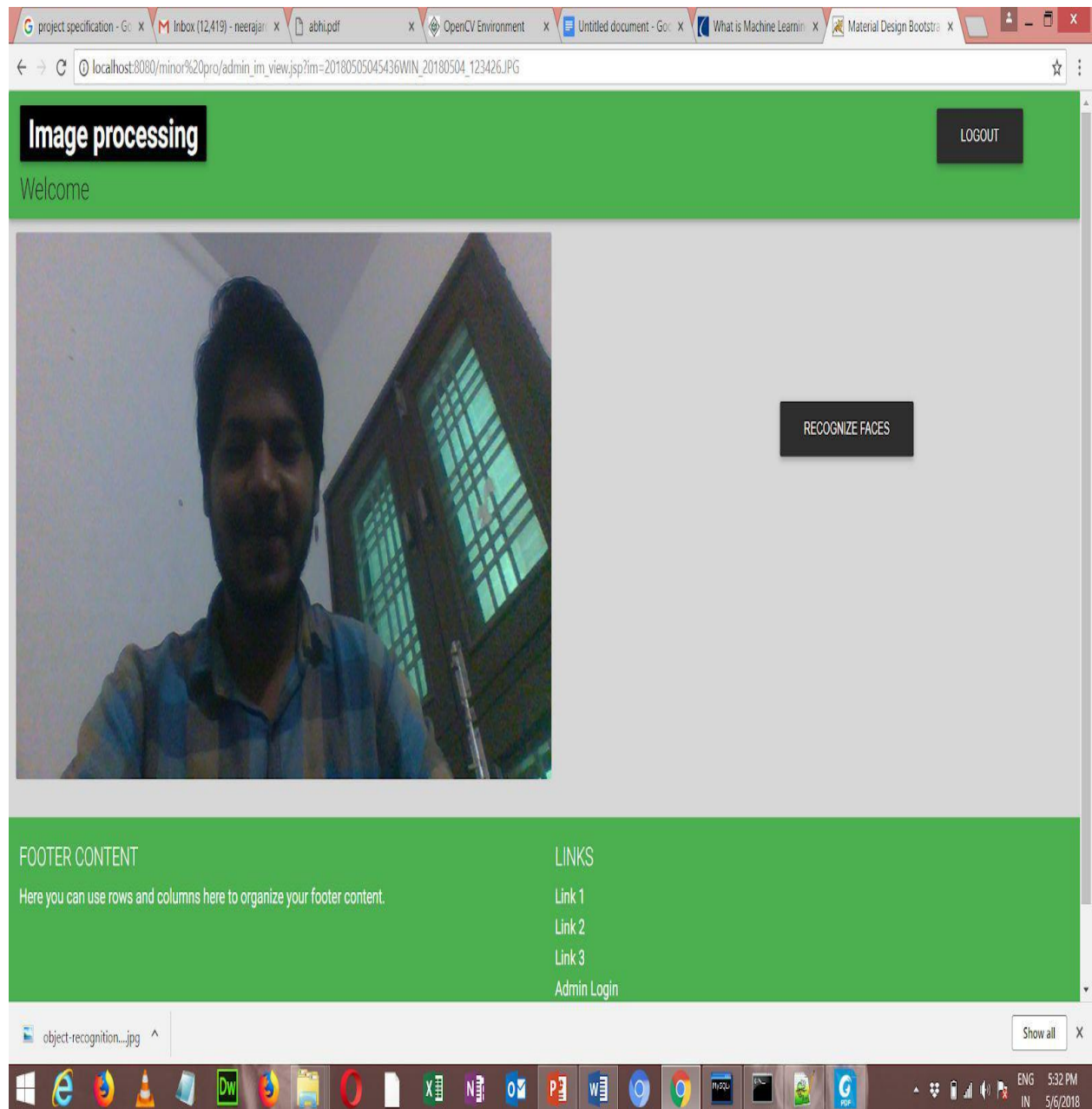
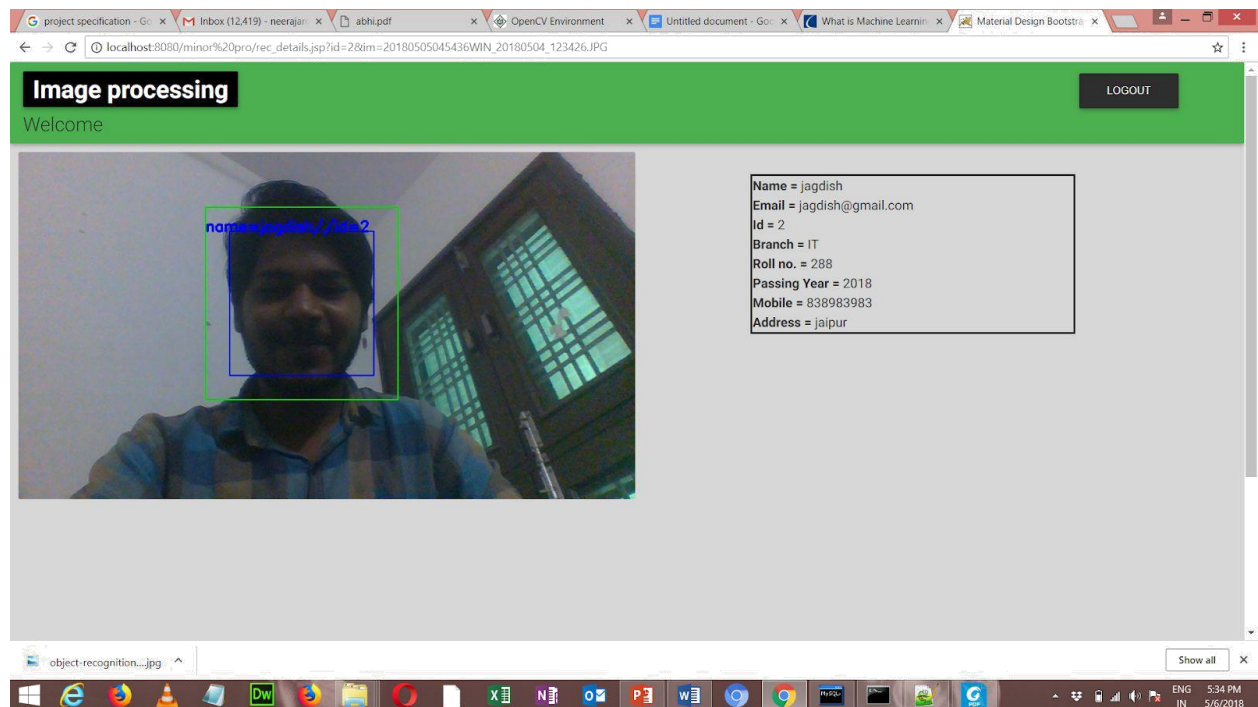


Fig. Admin Chosed Group

#### 4. Recognize and detect faces of various users with the help of OpenCV library.



Admin click on recognize faces button to detect face and fetch all information about the face. Here now where OpenCV comes to role then admin run a python script to recognize faces.



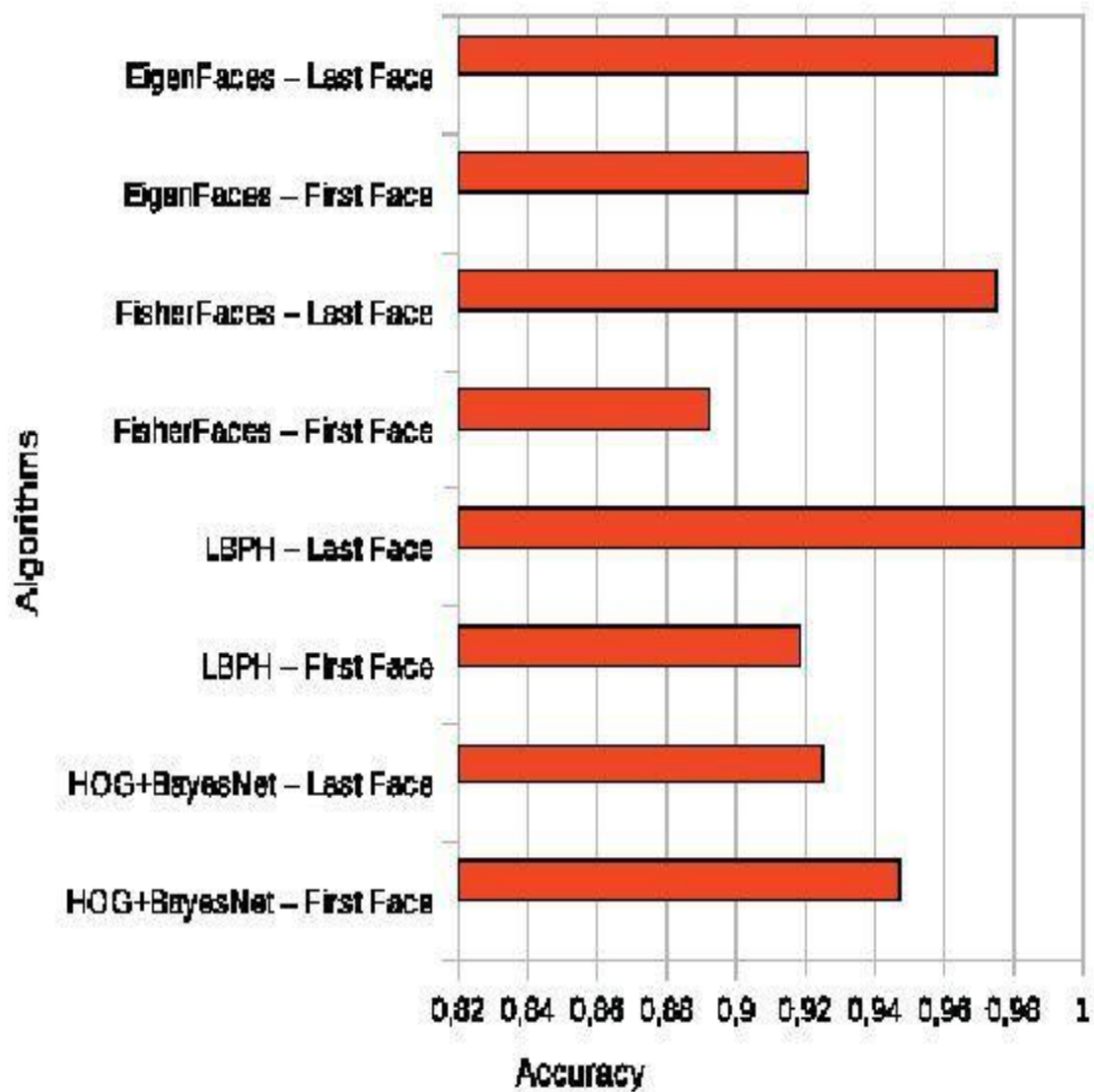
In this way no matter how old photo is but admin can easily recognize faces with the help of OpenCV machine learning library.



## TESTING AND EVALUATION

---

### 5.1 Performance Graph



## 5.2 LBPH Accuracy

In LBPH each images is analyzed independently, while the eigenfaces method looks at the dataset as a whole. The LBPH method is somewhat simpler, in the sense that we characterize each image in the dataset locally; and when a new unknown image is provided, we perform the same analysis on it and compare the result to each of the images in the dataset. The way which we analyze the images is by characterizing the local patterns in each location in the image.

Since you have already used Eigenfaces, you can try using LBPH method for face recognition, it will probably work better in different environments and light conditions, however, it will depend on your training and testing data sets. You will need around 10 different images of this person's face in order to be able to recognize him/her.

In case your willing to go beyond OpenCV, there is a lot of research around these areas and the latest results are quite promising. These methods use deep neural networks or convolutional neural networks.



## **CONCLUSION AND FUTURE SCOPE**

---

### **6.1 Conclusion**

Face recognition is one of the several techniques for recognising people.

There are several methods that can be used for that purpose. Some of the most common are using PCA or eigenfaces. Though there are other new techniques more simple to understand use and implement but also with very good performance. The ARENA algorithm is one of those algorithms. As we show ARENA has very good performance and is a very accurate especially if we use a feedforward neural network.

Face recognition technology has come a long way in the last twenty years. Today, machines are able to automatically verify identity information for secure transactions, for surveillance and security tasks, and for access control to buildings. These applications usually work in controlled environments and recognition algorithms that can take advantage of the environmental constraints to obtain high recognition accuracy. However, next generation face recognition systems are going to have widespread application in smart environments, where computers and machines are more like helpful assistants. A major factor of that evolution is the use of neural networks in face recognition. A different field of science that also is very fast becoming more and more efficient, popular and helpful to other applications.

The combination of these two fields of science manage to achieve the goal of computers to be able to reliably identify nearby people in a manner that fits naturally within the pattern of normal human interactions. "They must not require special interactions and must conform to human intuitions about when recognition is likely." [13] This implies that future smart environments should use the same modalities as humans, and have approximately the same limitations. "These goals now appear in reach however, substantial research remains to be done in making person recognition technology work reliably, in widely varying conditions using information from single or multiple modalities.

## 6.2 Future Work

Face recognition systems used today work very well under constrained conditions, although all systems work much better with frontal mug-shot images and constant lighting. All current face recognition algorithms fail under the vastly varying conditions under which humans need to and are able to identify other people. Next generation person recognition systems will need to recognize people in real-time and in much less constrained situations.

We believe that identification systems that are robust in natural environments, in the presence of noise and illumination changes, cannot rely on a single modality, so that fusion with other modalities is essential (see Figure 5). Technology used in smart environments has to be unobtrusive and allow users to act freely. Wearable systems in particular require their sensing technology to be small, low powered and easily integrable with the user's clothing. Considering all the requirements, identification systems that use face recognition and speaker identification seem to us to have the most potential for wide-spread application.

Cameras and microphones today are very small, light-weight and have been successfully integrated with wearable systems. Audio and video based recognition systems have the critical advantage that they use the modalities humans use for recognition. Finally, researchers are beginning to demonstrate that unobtrusive audio-and-video based person identification systems can achieve high recognition rates without requiring the user to be in highly controlled environments.

## References

- [1] M. B. Blaschko and C. H. Lampert. Learning to localize objects with structured output regression. In *Computer Vision—ECCV 2008*, pages 2–15. Springer, 2008.
- [2] L. Bourdev and J. Malik. Poselets: Body part detectors trained using 3d human pose annotations. In *International Conference on Computer Vision (ICCV)*, 2009.
- [3] H. Cai, Q. Wu, T. Corradi, and P. Hall. The crossdepiction problem: Computer vision algorithms for recognising objects in artwork and in photographs. *arXiv preprint arXiv:1505.00110*, 2015.
- [4] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.
- [5] T. Dean, M. Ruzon, M. Segal, J. Shlens, S. Vijayanarasimhan, J. Yagnik, et al. Fast, accurate detection of 100,000 object classes on a single machine. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1814–1821. IEEE, 2013.
- [6] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. *arXiv preprint arXiv:1310.1531*, 2013.
- [7] J. Dong, Q. Chen, S. Yan, and A. Yuille. Towards unified object detection and semantic segmentation. In *Computer Vision—ECCV 2014*, pages 299–314. Springer, 2014.
- [8] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov. Scalable object detection using deep neural networks. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 2155–2162. IEEE, 2014. 46
- [9] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual object classes challenge: A retrospective. *International Journal of Computer Vision*, 111(1):98–136, Jan. 2015.