**Criminal Identification System**

By

**Ehsanul Karim**

Roll: 1907039

&

**Ankon Chowdhury**

Roll: 1907048



**Supervisor:**

Dr. Sk. Md. Masudul Ahsan

Professor

Dept. of Computer Science and Engineering Signature

Khulna University of Engineering & Technology

**Department of Computer Science and Engineering**

**Khulna University of Engineering & Technology**

**Khulna 9203, Bangladesh**

**30th November 2023**

**Acknowledgement**

We extend our heartfelt gratitude to Allah, the Most Merciful and Powerful, for granting us the talents and capabilities to successfully complete the Criminal Identification System project. We are thankful for the invaluable guidance and supervision provided by Dr. Sk. Md. Masudul Ahsan, supervisor of this project. His expertise and knowledge in the field of computer science were very useful in the successful completion of the project. His constructive criticism, insightful suggestions, and constant encouragement helped us to overcome the challenges that we faced during the development of the system.

We would also like to thank our seniors, who inspired us to implement different ideas and suggestions throughout this project. We would also like to thank our friends for their association and cooperation with our project. Last but not least, our deepest gratitude is expressed to our parents, whose moral support and encouragement gave us the strength to complete this project.

**Ehsanul Karim and Ankon Chowdhury**

**Abstract**

The police force serves as a crucial public servant, and in their pursuit of criminals, they often employ various methods, including the utilization of sketch artists who create facial depictions based on victims’ descriptions. Subsequently, these sketches are compared with a criminal database to identify potential matches. Regrettably, these procedures are time consuming, and in the realm of law enforcement, time is of paramount importance. Again, the current online law enforcement website doesn’t extend to all criminal activities. Thus, to address these issues, the proposed system seeks to enhance efficiency by incorporating deep-learning models to capture facial features with a moderate degree of precision. The resulting data is then stored in a local database, facilitating expedited searches for potential matches, improving both time efficiency and accuracy. The web application has also been designed to handle a wide variety of criminal activities and provide better navigation of the system. The proposed models use a technique called Convolutional Neural Network (CNN) and have an accuracy of close to 88%. Mainly because the data set images consisted of good lightning conditions and a white background. A total of nearly 17,000 images were used to train the models, with each model having about 50 million trainable parameters. Facial features are extracted through various models but the base pre-trained model that was used for this system was Resnet50. The web application was also built with Django as the backend technology, as it provided lots of built in features, and incorporating the models was smoother due to the similar language environments.

**Contents**

|  |  |  |
| --- | --- | --- |
|  |  | **Page** |
| Acknowledgement |  | ii |
| Abstract |  | iii |
| Contents |  | iv |
| List of Tables |  | vi |
| List of Figures |  | vii |

|  |  |  |
| --- | --- | --- |
| **1** | **Introduction** | 1 |
|  | 1.1 Background | 1 |
|  | 1.2 Objectives | 1 |
|  | 1.3 Scopes | 2 |
|  | 1.4 Unfamiliarity of the problem | 2 |
|  | 1.5 Project planning | 2 |
| **2** | **Related Work** | 4 |
|  | 2.1 Existing solutions | 4 |
|  | 2.2 Limitation in existing solutions | 4 |
| **3** | **System Design** | 4 |
|  | 3.1 Analysis of the system  3.1.1 Data Flow Diagram  3.1.2 Use case Diagram | 4  4  5 |
|  | 3.2 System architecture  3.2.1 Class Diagram  3.2.2 ER Diagram | 7  7  7 |
|  | 3.3 Tools used | 7 |
|  | 3.3.1 Django | 7 |
|  | 3.3.2 Jupyter Notebook  3.3.3 VS Code  3.3.4 SQLite | 7  7  10 |
| **4** | **Project Implementation** | 10 |
|  | 4.1 System implementation  4.1.1 Web Implementation  4.1.2 Model Implementation  4.1.3 Model Architecture  4.1.4 Dataset | 10  10  16  16  18 |
|  | 4.2 Morality or Ethical issues | 19 |
|  | 4.3 Socio-economic impact and sustainability | 20 |
|  | 4.4 Financial analyses and budget | 20 |
| **5** | **Conclusions** | 21 |
|  | 5.1 Conclusion and challenges faced | 21 |
|  | 5.2 Future work | 22 |
|  | **References** | 22 |

**List of Tables**

|  |  |  |
| --- | --- | --- |
| **Table No.** | **Description** | **Page** |
| 4.1 | Statistics of individual model | 17 |
| 4.2 | Financial analysis and budget of the project | 20 |

**List of Figures**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Description** | **Page** |
| 1.1 | Gantt Chart of the project | 3 |
| 3.1  3.2  3.3  3.4  4.1  4.2  4.3  4.4  4.5  4.6  4.7  4.8  4.9 | DFD diagram of the project  Use case diagram of the project  Class diagram of the project  ER diagram of the project  Welcome page of the web app  Sign-up page of the web app  Sign-in page of the web app  User dashboard and advanced search page of the web app  Case description of the charge sheet form  Admin dashboard and criminal record insertion page of  Super-admin dashboard page of the web app  Architecture of Resnet50 model  A snapshot of the dataset | 5  6  8  9  11  11  12  13  14  15  16  18  19 |

1. Introduction
   1. Background

The current operational model of our police station presents several challenges in the effective handling of criminal activities. Currently, the police station lacks an extensive online charge sheet and FIR filling system, as well as the reliance on a manual sketch artist for suspect identifications which poses significant drawbacks. The sketching process is time-consuming, often leading to delays in investigations. To address these challenges, the proposed system implements an innovative solution, an AI-powered online charge sheet, FIR filing, and a suspect identification web application. This system enables user to file charge sheets online, additionally, the system will leverage facial recognition technology to extract features from facial images of criminals and store them in a database, facilitating a more efficient and accurate matching process. When a user submits a case and describes the criminal, the system will utilize the stored data to generate potential matches, presenting the results to the user.

* 1. Objectives

The projects major purposes or objectives are to improve the current online system and add a new solution to the limitations.

* To implement facial feature extraction technology using deep learning technique known as CNN.
* To develop a user-friendly online platform for citizens to file charge sheets for various criminal activities and complete the FIR filing by the admins.
* To gain knowledge about the working procedures of law enforcement and system development procedures.
* To reduce the hassle of being physically present for the procedures of law enforcement and filing charge sheets.
  1. Scope

The scope of this system is constrained to the storage of First Information Reports (FIR) exclusively within the police database. Regrettably, its functionality does not extend to encompass the official proceedings within the courtroom. Additionally, it is imperative to note that the AI facial extraction feature is dependent on favorable lighting conditions and the quality of the captured image.

Considering these limitations, strong efforts have been dedicated to maximizing the system’s capabilities. It excels in the identification and extraction of features while also adeptly navigating the procedures associated with charge sheets and FIR documentation.

* 1. Unfamiliarity of the problem

The integration of facial recognition technology for suspect identification represents a relatively unexplored territory as it wasn’t covered in our curriculum yet. Again, shifting from traditional, in-person charge sheet filing to an online platform introduces unfamiliarity for both users and law enforcement personnel. The learning curve of Django as the backend of this web application was also very steep and unfamiliar.

* 1. Project planning

The project planning consists of requirement analysis, study of techniques used, improvement of UI/UX, implementation of the whole system, and other sections. The relevant details are shown in the following Figure 1.1. The figure contains a Gantt chart, which has details of the work flow with time duration to specifically explain the planning and flow of operations in a very smooth manner.

**Figure 1.1**: Gantt Chart of the project.

1. Related Works

There exists already an online web application solution which is known as onlinegd.com. However, it contains some limitations which are addressed here.

* 1. Gap in Existing Solutions

The current online law enforcement website, known as onlinegd.com, exclusively facilitates the online filing of charge sheets for lost items and missing persons. Unfortunately, it does not facilitate reporting a broader spectrum of criminal activities. Furthermore, the platform lacks regular updates regarding news and notices to provide information to the public effectively. In addition to these limitations, the website does not incorporate features enabling users to identify potential suspect descriptions. Moreover, there is an absence of technological tools that could assist users in identifying suspects.

1. System Design

The system was designed with the fact that it can be scalable and adaptable for the common people. The reason is that with newer technology, the system can be easily upgraded, as well as having user-friendly, robust and advanced technology.

* 1. Analysis of the system

The system is analyzed through a couple of diagrams to depict the procedures and interactions of various elements. It follows with DFD and Use Case diagram to exhibit the sequence of operations.

* + 1. Data Flow Diagram

The system’s operation procedure consists of multiple steps. Working with large amounts of data or sensitive information, requires an efficient understanding of the flow of the system. As such the movement of data is shown in the Data Flow Diagram (DFD), Figure 3.1.

Citizen

Super-admin

Request for emergency

Complain

of incident

Add new crime type

Add information

Add Police-Officer to a dedicated Thana

Police Officer

Add Criminal Records

Update Case Records

Process criminal Image

Show user some criminal images

.

**Figure 3.1:** DFD diagram of the project.

* + 1. Use Case Diagram

The system consists of a wide variety of roles. Thus, it requires a high-level view of the system’s functionalities and interaction between actors. The perspective of these actor is exhibited in the Figure 3.2.

Police

User

SuperAdmin

**Figure 3.2:** Use case diagram of the project.

* 1. System architecture

The construction of the project requires in depth knowledge of its architecture. The structure of the project is exhibited in a class diagram, illustrating the encapsulation and association of key classes, and the database, which stores information and associated relationships with it, is shown in an ER diagram.

* + 1. Class Diagram

The information and functionality of association and encapsulation of each individual class or building block is shown in the Figure 3.3 where the blocks resonate the required information.

* + 1. ER Diagram

The Entity-Relationship diagram for this project illustrates the key entities with their attributes and relationship among them in the Figure 3.4.

* 1. Tools used

A wide variety of tools were used to complete the system. Each tool played a significant role in the completion of the project.

* + 1. Django

A high-level python web framework, chosen for its robustness and commitment to the MVC (Model-View-Controller) architecture, providing an organized structure for efficient web development.

* + 1. Jupyter Notebook

Jupyter Notebook facilitates interactive development. Its versatility makes it an ideal choice for prototyping and refining deep learning techniques, such as Convolutional Neural Networks (CNN).

* + 1. VS Code

Visual Studio Code serves as a versatile code editor. Various extensions can be found to support both python and web development languages which increases coding productivity.

Admin Profile

Name: String

Email: String

Password: String

Allocated Thana: Address()

Search\_fir & criminal\_record()

View\_fir & criminal\_record()

Add\_criminal\_record()

Handle\_Request()

Criminal Profile

Criminal\_image: Blob

Criminal name: String

Criminal\_personal\_infos

Physicial Structure: Criminal Structure()

Calculate\_age()

Criminal Structure

Crimes: String

Arrest date: Date

Gender: String

Hair color: String

Hair length: String

Look age: String

Face shape: String

Facial hair: String

Height: Float

Weight: Float

Mark: String

Predict()

Load\_model()

Noti Panel

Title: String

Created\_at: Date

Time\_expired()

Destroy()

Request

Description: String

Complain no:

Status: String

Address

Division: String

District: String

Thana: String

Update\_dis()

Update\_thana()

User Profile

Name: String

Email: String

Password: String

PermanentAddress: Address()

File\_new\_complain()

Search\_View\_Self\_complain()

Complain

Victim info’s: User Profile()

Occurance Date: Date()

Occurance Place: Address()

Brief\_of\_incident: String()

Criminal\_structure\_details: CriminalStructure()

Check\_column\_filling()

View\_summery()

**Figure 3.3:** Class diagram of the project.

Notification

User

Complain

Admin

Criminal

Physical Structure

Address

Authentication

Has a

Assign

Arrest

Get

Submit

Has a

Signin

Suspect

Information

1

M

1

1

1

M

1

M

1

1

1

1

1

Happen

1

1

1

1

1

1

1

1

**Figure 3.4:** ER diagram of the project.

* + 1. SQLite

SQLite, a lightweight and structured relational database is chosen for simplicity and effective managing of the project’s data.

1. Project Implementation

This chapter will describe the steps and individual procedures that was taken to complete the system project. The system consists of collaborative and systemic process which is described in the following.

* 1. System implementation

The system being a web application, a website was designed using Django as the backend and basic HTML, CSS, JS as the frontend. The CNN technique was used for the model training, and a pre-trained model was collected for the task. The dataset for the task was also acquired through the web. The details of these procedures are as following.

* + 1. Web Implementation

The frontend design is done through the help of frameworks known as bootstrap and Scss. Additionally, CSS is also used for custom styling and JS to incorporate some dynamic features. The backend of the structure is completed by Django which uses python language. For database system, the built in SQLite system is used. Initially, the frontend of the pages was designed, which was then connected according to the Django framework. The SQLite database is chosen due it’s simplicity and easy to maintain structured form. The following exhibition of Figures will describe the web application pages for all the actors giving a glimpse of the criminal identification system.

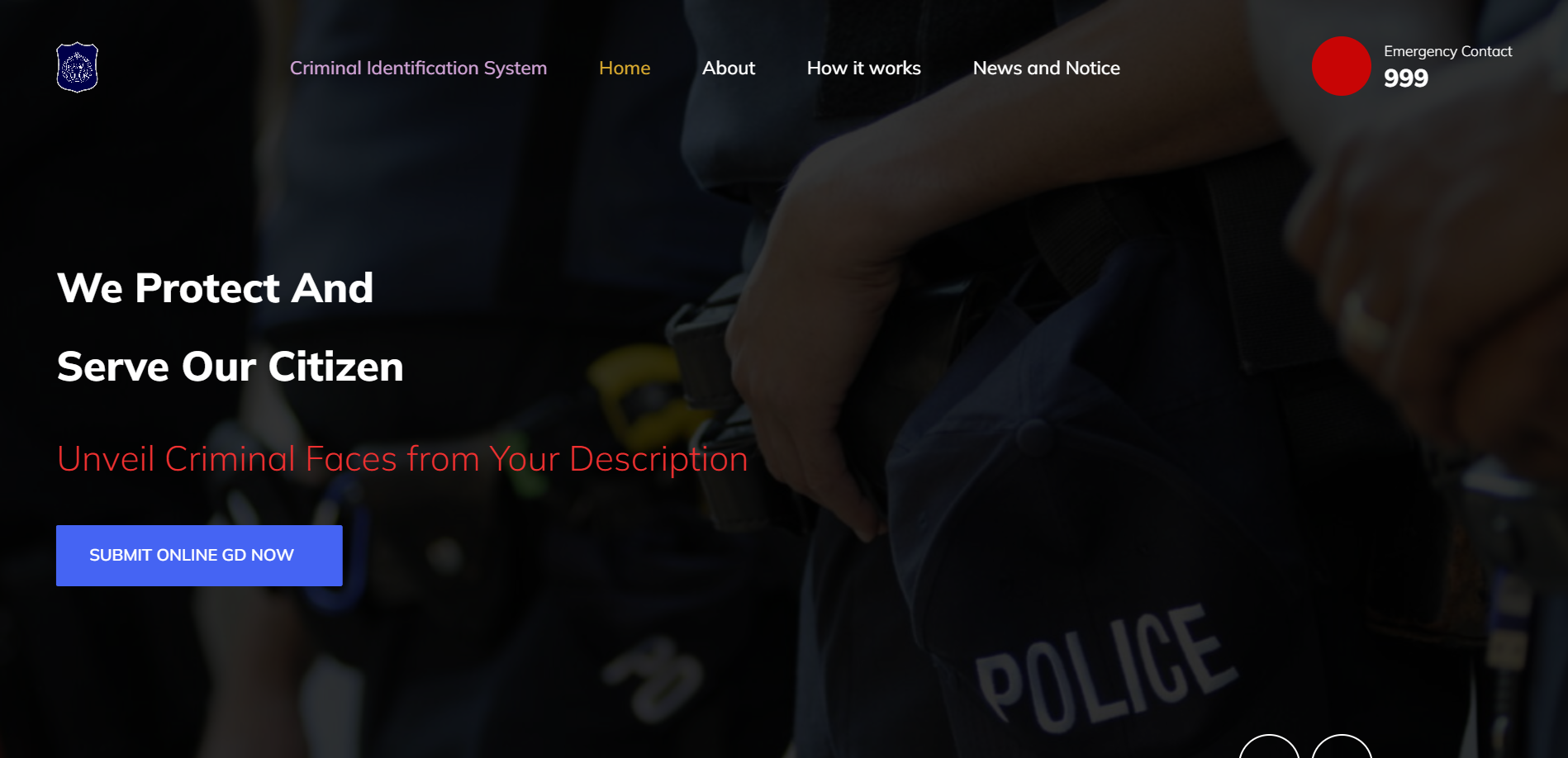


Figure 4.1: Welcome page of the web application.

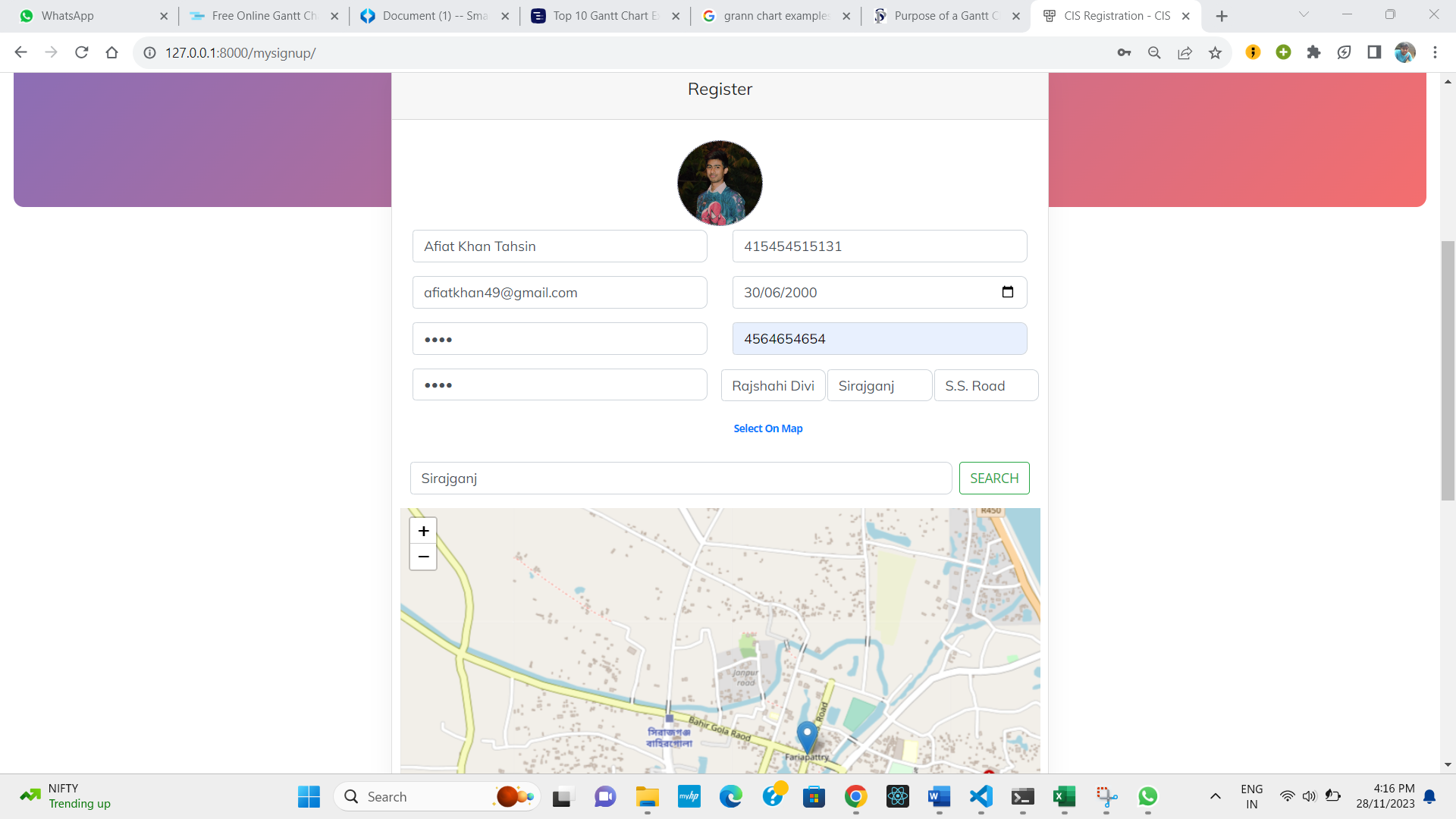
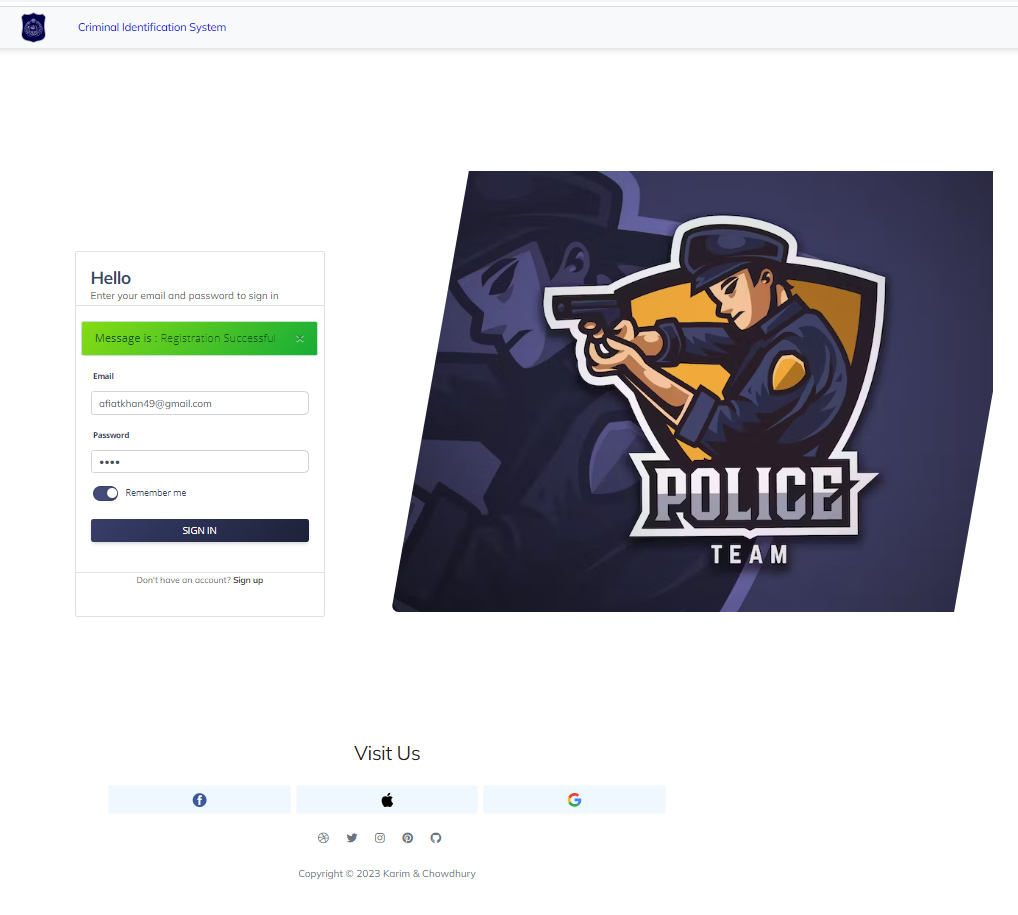
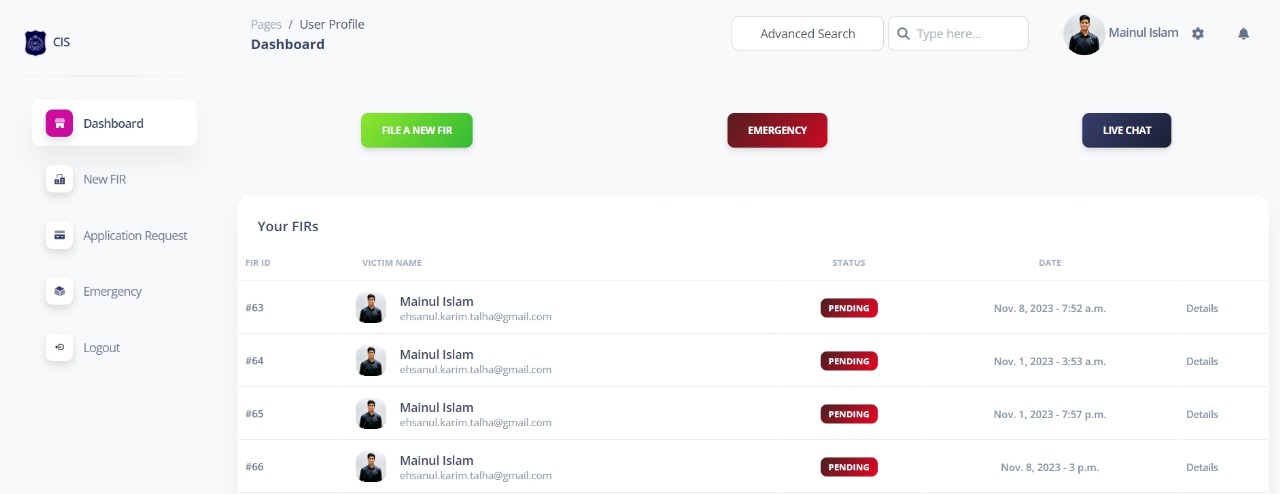


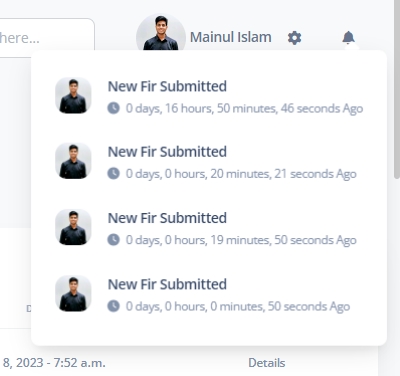
Figure 4.2: Sign-up page of the web application.

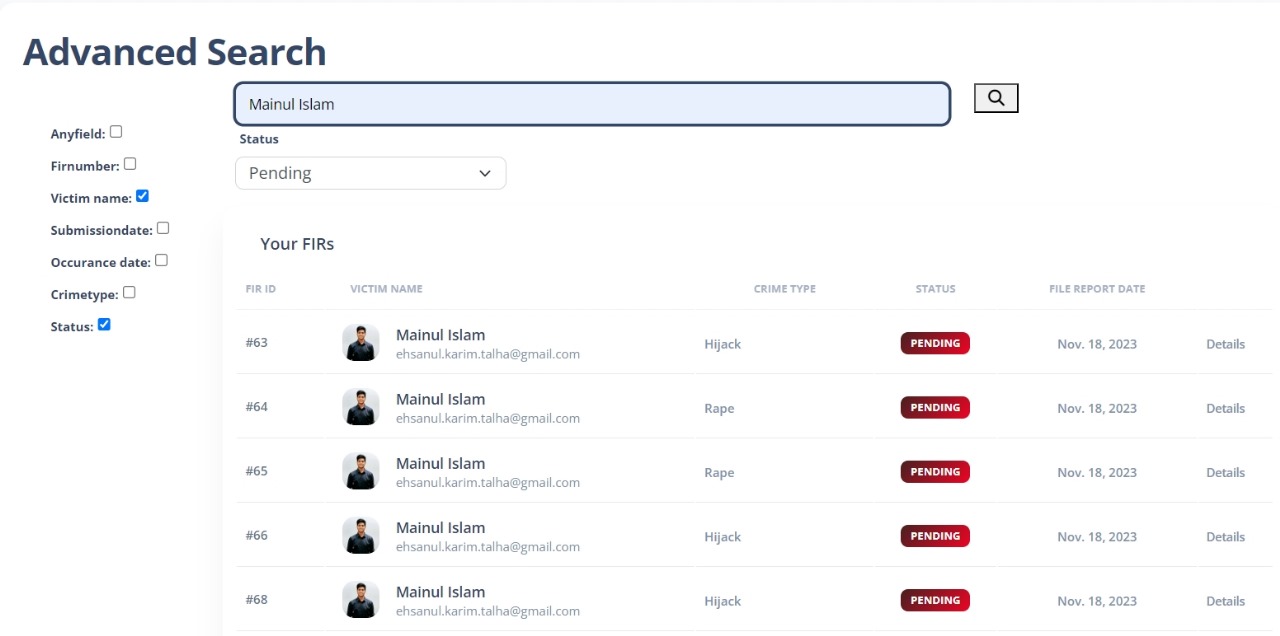


**Figure 4.3:** Sign-in page of the web application.

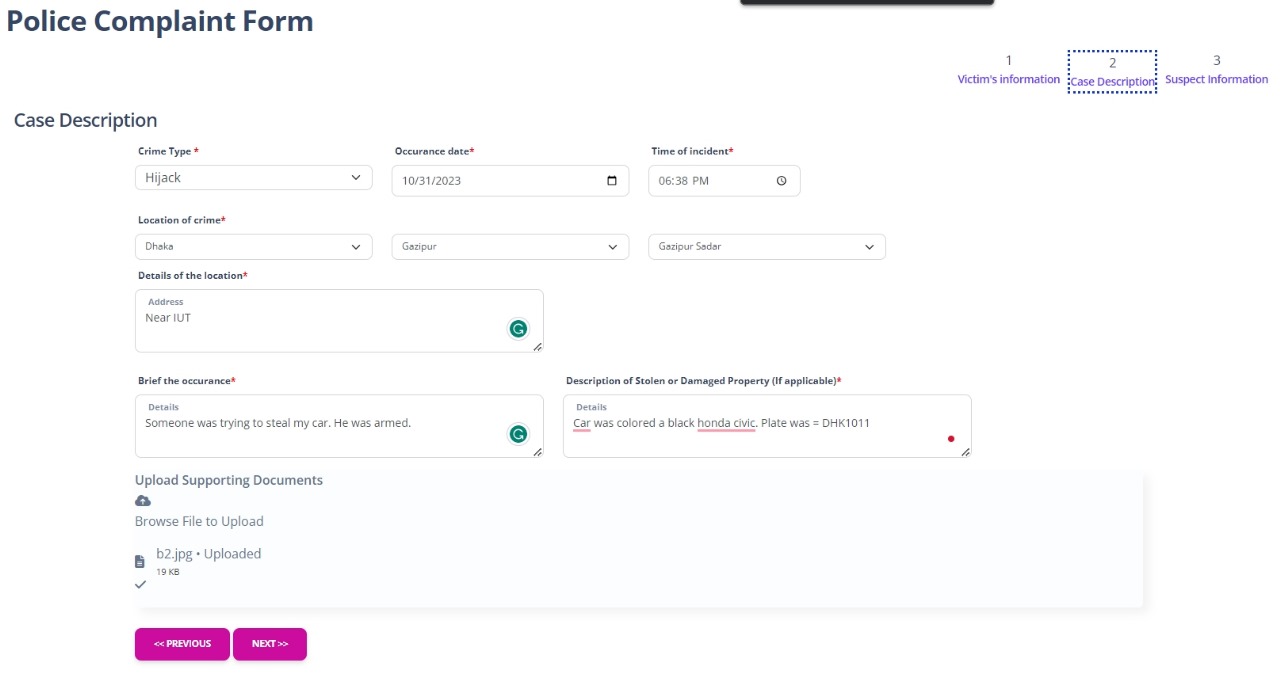
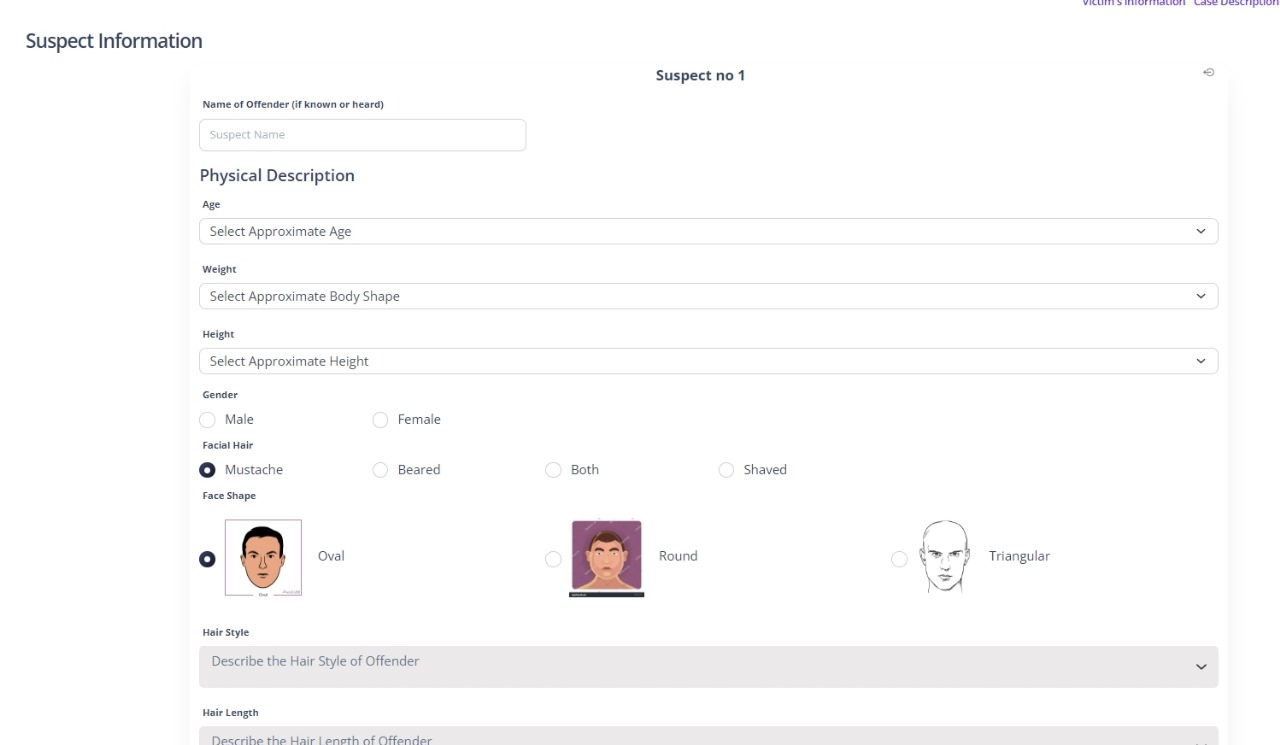
The welcome page greets the user when they first visit. The user has the option to either register for the website or login if they have previously set up an account. After entering the login credentials, the dashboard appears. The dashboard shows some general information exhibited in Figure 4.4. The user can have a comprehensive view of all of his filed charge sheets by using the advanced search. The notification section provides the most recent updates. To initiate the process of filing a new charge sheet, the user needs to select the charge sheet filing tab and enter the details of the case along with the suspect's description, as illustrated in Figure 4.5. Since law enforcement officers can also access the website, when logging in, they are presented with an additional dashboard, the admin dashboard, as shown in Figure 4.6. The criminal insertion format is also shown in Figure 4.6, which the administrator uses to add criminal records to the database. The Super-admin page, which can assign agents to the designated thana and submit messages with additional powers, is shown in Figure 4.7.



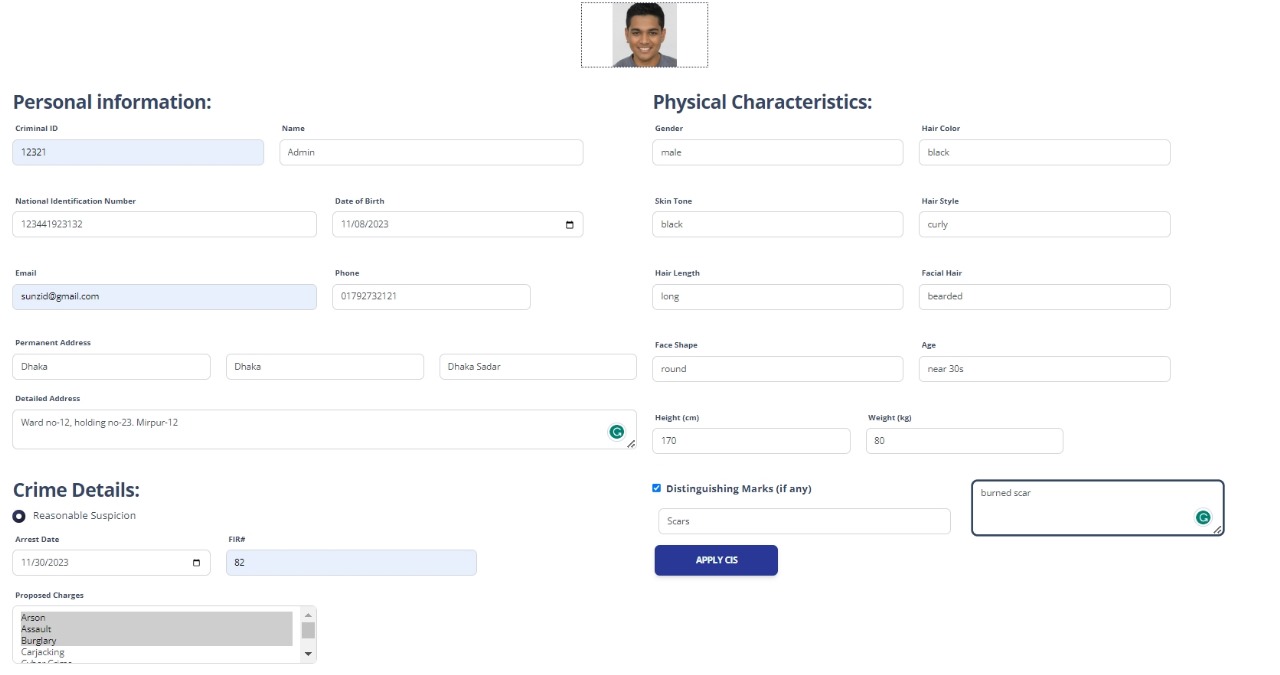
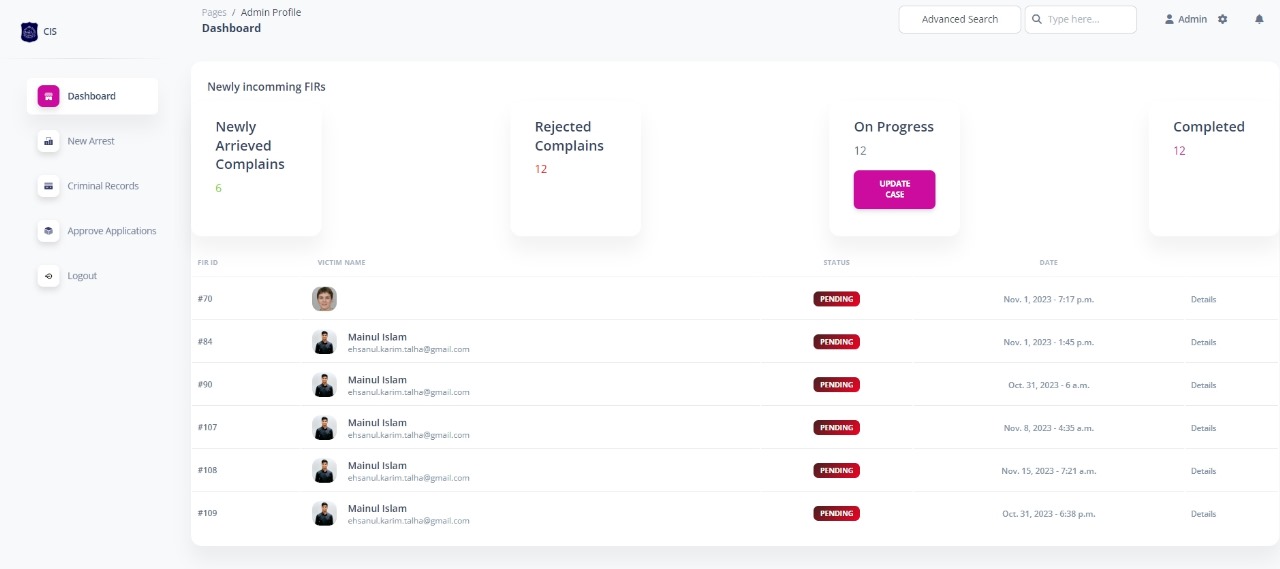




**Figure 4.4:** User dashboard and advanced search page of the web application.



**Figure 4.5:** Case description of the charge sheet form.



**Figure 4.6:** Admin dashboard and criminal record insertion page of the web app.

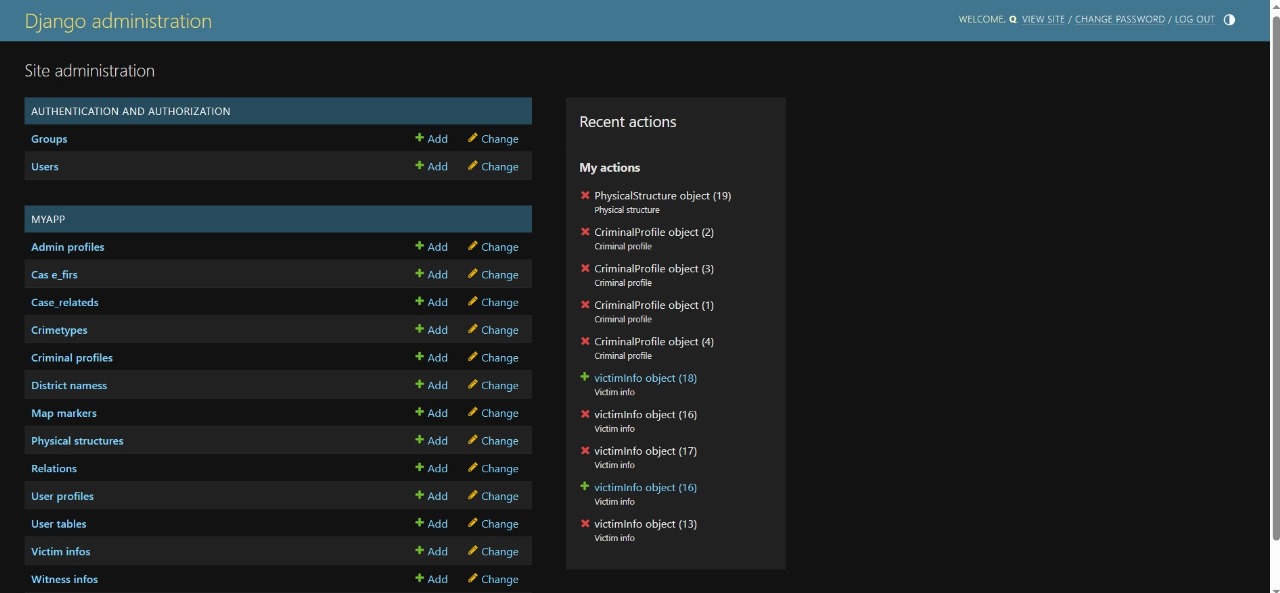


Figure 4.7: Super-admin dashboard page of the web app.

* + 1. Model Implementation

The model utilizes tensor-flow and the keras library of python as well as open-cv for image processing. The images are split between training and validation sets and pre-processed through 224 x 224 pixels, with both keras and open-cv libraries. For the extraction of hair information, a masking technique is used through the open-cv library as it exhibited better results. The Resnet50 model is utilized from the keras library, and on top of that, a custom fully connected layer is attached, which completes the CNN network for implementation. Graphs are created to acquire knowledge on accuracy and loss. These details assisted with the epoch number, which is close to 15 on average. As for the optimizer, Adam is utilized with a learning rate of 0.001 looking at the accuracy metric, thus achieving the finest accuracy.

* + 1. Model Architecture

The pre-trained model Resnet50 is a convolutional neural network (CNN) utilized in image classification. It introduces the concept of residual learning, where the blocks have shortcut connections bypassing one or more layers, allowing the model to learn residual functions rather than attempting to approximate the direct mapping. The model, consisting of 50 layers, is equipped with ReLU activation functions. Batch normalization is integrated to enhance stability and accelerate convergence; the global average pooling layer is also connected to reduce spatial dimensions. The final output of the model is connected to a custom layer consisting of Flatten and a couple of Dense layers. The intrinsic details of each individual model are in the following Table 4.1 and, the architecture of the base Resnet50 model is exhibited in Figure 4.8.

Table 4.1: Statistics of individual model

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model Name | Number of Parameters | Accuracy | Number of Class | Class Names | Number of Images in each Class |
| Gender | 51,381,762 | 95.6% | 2 | Female, Male | 2500 |
| Age-group | 103,287,811 | 85.5% | 3 | Near 30s, Near 40s, Near 50s | 2000 |
| Skin-tone | 51,381,762 | 88.2% | 3 | Black, Brown,  White | 1800 |
| Hair-color | 51,381,762 | 90.2% | 3 | Black, Blonde, Brown | 2000 |
| Hair-length | 51,381,762 | 84.2% | 3 | Long,  Medium,  Short | 1500 |
| Hair-type | 51,381,762 | 83.2% | 2 | Curly, Straight | 550 |

**Zero Padding**

**CONV**

**Batch Norm**

**ReLu**

**Max Pool**

**Conv Block**

**ID Block**

**Conv Block**

**ID Block**

**Avg Pool**

**Flattening**

**FC**

**Input**

**Output**

……

Stage 3

Stage 2

Stage 1

Figure 4.8: Architecture of Resnet50 model.

* + 1. Dataset

A dataset is very important for training a model. A total of six models were trained and, for each model, a wide number of images of faces were collected and used. A total of around 17,000 images were collected from the web for training and validation. Each model trained with a different number of images for its purpose whose details are given in the previous Table 4.1. The dataset consists of images of size 250 x 250 pixels, which are then resized to 224 x 224 pixels as it works best for the model. Images in the dataset are of good quality. With a noiseless background and better lighting conditions, ensuring a well-rounded environment for the training set.

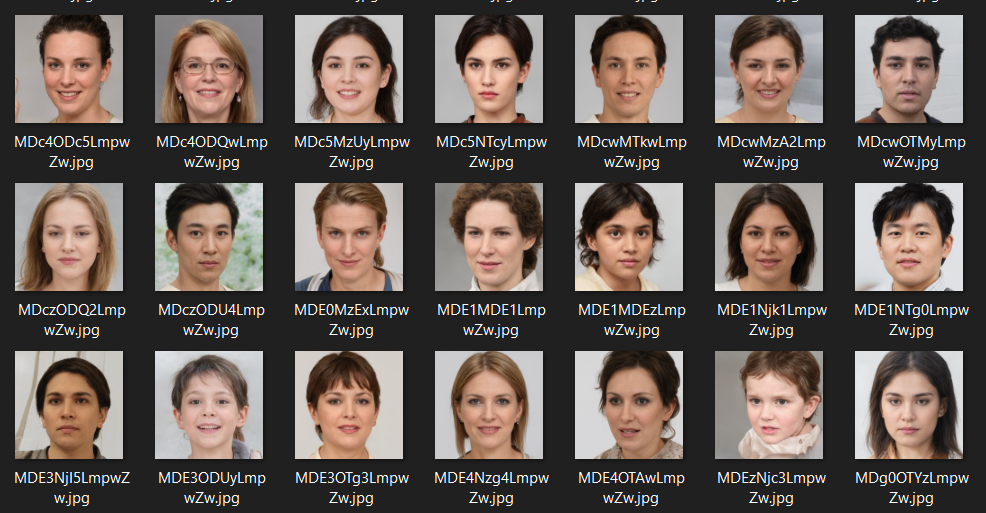


Figure 4.9: A snapshot of the dataset.

* 1. Morality or ethical issues

The project developed took a wide variety of sources. The dataset was collected from a website that generated and collected facial images of size 250 x 250 pixels. The base pre-trained model was also imported from the keras library, and its procedures were learned from the official documentation. The technique that was implied in the project was learned from various papers and books. The preferred sources are referred to in the reference section, but among them, the most notable one is deep-learning by python written, by Francois Chollet. The masking of images was done by the open-cv library of python. As for the web application, the Django documentation was followed to ensure seamless productivity. The implementation of database system and class diagrams were learned from previous academic curricula. The system also includes sensitive information about its users, but it is fully ensured that the information or data will be kept discrete from others. The data will not be available to the public, and full security of the data is confined within the system. The project upholds ethical standards, notably prioritizing user privacy.

* 1. Socio-economic impact and sustainability

The project encourages the community by empowering citizens to actively participate in the criminal reporting process. This leads to a reduction in criminal activities, which in turn provides the community with a safe and healthy environment. The system also has a good impact on the socio-economic status of society by reducing the hassle and time spent by citizens to file a charge sheet. It also provides a dynamic way for law enforcement to complete FIRs, allowing them to improve efficiency and accessibility. The system provides long term service as it can be adapted to newer technology, which makes it easier to scale and sustain. The system could potentially increase the amount of network traffic in the area. Regardless, the system holds a potential amount of significance for society and upholds its law and order. This system has a positive legal impact, ensuring that justice is accessible to a broader segment of the population.

* 1. Financial analysis and budget

Financial analysis and budget for this project involves the cost of development, equipment, implementation, maintenance and other relevant expenses. A general measurement of the budget has been estimated in following Table 4.2.

**Table 4.2:** Financial analysis and budget of the project

|  |  |
| --- | --- |
| **Types of cost** | **Budget(tk)** |
| Personnel Salaries | 1,00,000 – 1,50,000 |
| Hardware | 50,000 – 80,000 |
| Software | 10,000 – 40,000 |
| Data collection and processing | 10,000 – 30,000 |
| Development tools | 10,000 – 15,000 |
| Admin training | 10,000 – 20,000 |
| **Types of cost** | **Budget(tk)** |
| Project management tools | 5,000 – 12,000 |
| Miscellaneous | 5,000 – 8,000 |
| Total: | 2,00,000 – 3,50,000 |

The overall budget is characterized by minimal external expenditures, with majority of costs being self-funded and utilization of available software and technologies.

1. Conclusion

The development of the online charge sheet, FIR filing and criminal records management system represents a significant leap forward in leveraging technology to enhance the efficiency of law enforcement.

* 1. Conclusion and challenges faced

The system provides users with a seamless and accessible platform to file charge sheets online. Coupled with identification of criminals and an approval mechanism, it fluently incorporates the procedures of law enforcement. Admins also have extensive search capabilities, an efficient way of viewing and updating cases and criminal records, FIR filing, which ensures a dynamic investigative purpose. Thus, the system not only optimizes operational efficiency but also contributes to the evolution of modern, tech-driven law enforcement practices. Although the system was built successfully, there were many challenges or hurdles along the way. One of which was the accurate extraction of facial features, which required the training of various models on distinct datasets. The limitations of a large dataset further impacted the precision of the results. Another hurdle that was faced was in the domain of web technology, particularly in integrating the models as it was taking a long time to load them and addressing various UI/UX design issues. Nevertheless, the challenges were confronted directly, ultimately ensuring the project’s overall success.

* 1. Future Study

# As we overcame some of the challenges, there are still some fields left for improvement. The project can be extended to include additional legal procedures, such as incorporating courtroom processes and addressing a broader spectrum of criminal activities. Further study can be done to increase the number of datasets, particularly in each class where having more than 10,000 images can significantly improve the results of the model. A better understanding of the technology and time efficiency can lead to a well-rounded architectural model that can produce all feature extractions in one model. Further features can be implemented on the web application to scale the project for a wide variety of societies.

# References

[1] Alarifi, J. S., Goyal, M., Davison, A. K., Dancey, D., Khan, R., & Yap, M. H. (2017). Facial Skin Classification Using Convolutional Neural Networks (pp. 479–485). https://doi.org/10.1007/978-3-319-59876-5\_53

[2] LIEW, S. S., KHALIL-HANI, M., AHMAD RADZI, S., & BAKHTERI, R. (2016). Gender classification: a convolutional neural network approach. TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES, 24, 1248–1264. https://doi.org/10.3906/elk-1311-58

[3] Muhammad, U. R., Svanera, M., Leonardi, R., & Benini, S. (2018). Hair detection, segmentation, and hairstyle classification in the wild. Image and Vision Computing, 71, 25–37. https://doi.org/10.1016/j.imavis.2018.02.001

[4] Mustapha, M. F., Mohamad, N. M., Osman, G., & Ab Hamid, S. H. (2021). Age Group Classification using Convolutional Neural Network (CNN). Journal of Physics:ConferenceSeries,2084(1),012028.https://doi.org/10.1088/1742-6596/2084/1/012028

[5] Deep Learning with Python Book by Francois Chollet

[6] [Django documentation | Django documentation | Django (djangoproject.com)](https://docs.djangoproject.com/en/4.2/)

[7] [ResNet and ResNetV2 (keras.io)](https://keras.io/api/applications/resnet/#resnet50-function)

[8] [AI-Generated Faces: Diverse & Customizable | Generated.Photos](https://generated.photos/faces)