**Smart Criminal Identification system using Image Processing and Classification Machine learning Algorithms**

# Chapter 1: Problem Identification

**1.0 Introduction**

Zimbabwe as a developing nation, is experiencing a surge in serious crimes such as robbery, unlawful entry, theft, and the sexual abuse of minors (Zimbabwe National Statistics Agency, 2023). Law enforcement agencies are often overwhelmed by logistical limitations, lack of data-driven crime monitoring tools, and the absence of centralized criminal databases. This has resulted in a reactive approach to crime management, where police respond only after crimes have occurred, often with limited success in tracking or apprehending repeat offenders (Mawere & Chikowore, 2022). In contrast, developed countries have embraced artificial intelligence (AI), in facial recognition identification systems, biometric systems. These tools have proven to be efficient in tracking and managing crime occurrence more effectively. In United States and the United Kingdom there have integrated AI-powered surveillance cameras that automatically detect and flag known offenders through facial recognition, combined with real-time access to criminal history databases (Smith and Li, 2020). These systems also use predictive analytics to determine where crimes are likely to occur, allowing for proactive deployment of resources. In India, facial recognition is being used in major cities to identify repeat offenders, while South Africa has introduced AI-enhanced CCTV systems in high-crime areas such as Johannesburg and Cape Town (Raghavan, 2021). However, in many marginalized communities across the globe, such systems are either inaccessible or poorly integrated, often due to infrastructural, financial, or policy limitations. As a result, communities in these areas continue to suffer from high crime rates and low prosecution rates, with no systematic way of identifying or tracking serial offenders (Makoni, 2022).

Despite global technological advancements, Zimbabwe has not yet adopted a national, AI-driven crime monitoring and offender profiling system. There is currently no centralized platform that keeps a cumulative record of criminal offenses committed by individuals, especially those involved in recurrent or serious crimes such as child rape, armed robbery, and burglary. Furthermore, the Zimbabwe Republic Police (ZRP) lacks the capacity to instantly identify suspects from images, classify their criminal risk levels, or issue public alerts for high-risk individuals. This presents a critical research and implementation gap, highlighting the need for a digital solution that integrates AI, biometric data including facial recognition and crime severity grading to track and manage wanted persons across all regions, including underserved and marginalized communities (Ngwenya et al., 2022).​

The current study therefore proposes the development of Smart Criminal Identification system that uses Convolutional Neural Networks for image processing and classification algorithms for criminal profiling and tracking system tailored to Zimbabwe's context. This system aims to store detailed criminal histories, use facial recognition to match individuals with known records, and automatically classify offenders based on the severity and frequency of their crimes. This solution is designed to function effectively in Zimbabwe's marginalized areas, which often suffer from limited internet access and technological infrastructure. To overcome these challenges, the AI system will incorporate offline capabilities, allowing data collection and preliminary processing without constant internet connectivity. Periodic synchronization with central databases can occur when connectivity is available, ensuring that remote areas are not excluded from national crime monitoring efforts.

## Background Analysis

Zimbabwe's criminal justice framework operates through a dual system comprising formal courts and traditional community courts. The formal judiciary encompasses the Constitutional Court, Supreme Court, High Court, and Magistrates’ Courts, which handle serious criminal offenses such as murder, armed robbery, and sexual assault. These courts adhere to codified legal statutes, including the Criminal Law (Codification and Reform) Act [Chapter 9:23] and the Criminal Procedure and Evidence Act [Chapter 9:07], ensuring standardized legal procedures and evidence-based adjudication. ​

Conversely, traditional courts, presided over by chiefs, headmen, and village heads, function under the Customary Law and Local Courts Act [Chapter 7:05]. These courts address minor civil disputes and petty offenses within rural communities, emphasizing restorative justice and community harmony. However, the informal nature of these courts, reliance on oral traditions, and lack of standardized procedures have raised concerns about impartiality and human rights adherence. Instances of traditional leaders overstepping their jurisdiction, imposing excessive fines, and exhibiting gender biases have been documented, highlighting the need for oversight and reform (Blessing Hodzi, 2024). ​

The integration of Artificial Intelligence (AI) into Zimbabwe's criminal justice system presents an opportunity to enhance efficiency, accuracy, and accessibility. Globally, AI tools such as facial recognition, predictive policing algorithms, and AI-powered surveillance systems have been employed to assist in crime detection and prevention. In Zimbabwe, the city of Bulawayo has pioneered the adoption of AI in law enforcement through the implementation of the Smart City Surveillance System. This system utilizes AI-enhanced CCTV cameras equipped with facial recognition, license plate recognition, and behavioral analytics to monitor public spaces, identify suspects, and predict potential criminal activities. The initiative has led to a notable reduction in vehicle-related crimes within the Central Business District, demonstrating the potential of AI in enhancing public safety. ​

A pertinent case underscoring the challenges and potential of AI integration is the Ecobank Bulawayo Heist of October 2024. In this incident, armed robbers stole over US$4 million from a cash-in-transit vehicle outside the Ecobank branch in Bulawayo (…). The perpetrators, armed with AK-47 rifles, swiftly overpowered security personnel and fled the scene within minutes. Investigations revealed that the getaway vehicle's number plates were stolen, a tactic employed to mislead law enforcement. The Zimbabwe Republic Police, in collaboration with Interpol, initiated investigations, utilizing forensic accounting and digital forensics to trace the culprits. However, the absence of CCTV coverage on Fife Street, where the robbery occurred, impeded the investigation, highlighting the limitations of existing surveillance infrastructure (). ​

The Ecobank heist exemplifies the critical need for comprehensive AI integration in Zimbabwe's law enforcement strategies. While urban centers such as Bulawayo have begun adopting AI technologies, marginalized and rural areas remain underserved due to infrastructural limitations, lack of internet connectivity, and inadequate technological resources. Furthermore, the traditional courts operating in these regions lack the capacity to handle sophisticated criminal activities and often function without standardized records, making it challenging to track repeat offenders.​

This context reveals a significant research gap: the necessity for a Smart Criminal Identification System tailored to Zimbabwe's unique socio-legal landscape. Such a system would leverage AI technologies, including Convolutional Neural Networks for image processing and classification algorithms for offender profiling, to enhance criminal monitoring across both formal and traditional justice mechanisms. By incorporating offline capabilities, the system would address infrastructural challenges in marginalized areas, ensuring inclusivity and effectiveness in crime prevention and law enforcement. The integration of AI into Zimbabwe's criminal justice system holds the promise of bridging the divide between formal and traditional courts, enhancing efficiency, and ensuring equitable access to justice. The proposed Smart Criminal Identification System represents a strategic approach to modernizing law enforcement, addressing current limitations, and fostering a more robust and inclusive justice system.​

These insights are supported by (Ngwenya et al.,2022), who emphasize the potential of AI technologies such as facial recognition, AI-enhanced surveillance, and predictive analytics to transform policing in Zimbabwe. Their research highlights how AI can improve crime detection and prevention, especially when adapted to local contexts and infrastructural realities. They also underscore the importance of integrating AI within national development strategies like Zimbabwe's National Development Strategy 1 (NDS1), which advocates for the digitization and modernization of public services, including law enforcement. Furthermore, the study points out that implementing AI systems with offline capabilities can mitigate infrastructural challenges in remote areas, ensuring broader access to justice. The authors conclude that a well-designed AI-driven criminal identification system could bridge the gap between formal and traditional justice mechanisms, promoting a more inclusive and effective criminal justice system in Zimbabwe.

#### **Problem Statement**

## Zimbabwe is currently grappling with a significant increase in serious crimes, including robbery, unlawful entry, theft, and sexual abuse of minors. According to the Zimbabwe National Statistics Agency (ZIMSTAT), over 8,500 sexual abuse cases were reported in 2023, highlighting the severity of the issue. Despite these alarming figures, law enforcement agencies face numerous challenges, such as logistical constraints, the absence of centralized criminal databases, and a lack of data-driven crime monitoring tools. This has resulted in a predominantly reactive approach to crime management, where police respond only after crimes have occurred, often with limited success in tracking or apprehending repeat offenders.​ The Zimbabwe Republic Police (ZRP) lacks the capacity to instantly identify suspects from images, classify their criminal risk levels, or issue public alerts for high-risk individuals (Chikowore, 2023). This gap is particularly pronounced in marginalized and rural communities, where infrastructural limitations and inadequate technological resources hinder effective law enforcement. Traditional courts in these regions often operate without standardized records, making it challenging to track repeat offenders and manage serious crimes.​ In order to address these challenges, there is a critical need for the development of a Smart Criminal Identification System tailored to Zimbabwe's unique socio-legal landscape. Such a system would leverage AI technologies, including Convolutional Neural Networks for image processing and classification algorithms for offender profiling, to enhance criminal monitoring across both formal and traditional justice mechanisms. Therefore, by incorporating offline capabilities, the system would function effectively in areas with limited internet access, ensuring inclusivity and effectiveness in crime prevention and law enforcement. This integration of AI into Zimbabwe's criminal justice system holds the promise of bridging the divide between formal and traditional courts, enhancing efficiency, and ensuring equitable access to justice. ​

### Aims

### The aim of this project is to design and implement a Smart Criminal Identification System that

### employ image and classification algorithms for Masvingo Central Police Station.

**1.4.1 Objectives**

## To build a dataset containing at least 1,000 facial images of convicts and wanted persons at Masvingo Central Police Station, annotated with relevant metadata within two months.

## To implement a Convolutional Neural Network (CNN) model for facial recognition, achieving at least 85% accuracy within three months after dataset completion.

## To develop a crime classification algorithm that categorizes offenses into defined grades (e.g., Grade A for Murder, Grade B for Theft) based on Standing Orders Volume 1 & 2, with 95% classification accuracy within one month.

## To test and evaluate the complete system using a new dataset of at least 200 images, achieving an F1-score of at least 80%, within one month of full integration

**1.5 Tools Used**

* Convolutional Neural Networks (CNN)
* Crime Classification Algorithms
* OpenCV for Image Preprocessing
* CSV files
* Scikit-learn for Model Evaluation

## Methodology

## This research explores the integration of Artificial Intelligence (AI) to develop a Smart Criminal Identification System tailored to the needs of Masvingo Central Police Station in Zimbabwe. The approach combines both primary and secondary data sources, where facial images of at least 1,000 convicts and wanted persons will be gathered from police records, alongside relevant metadata such as criminal history and crime type. Secondary data, including critical legal frameworks like the Standing Order Volumes 1 and 2 and the Zimbabwe Criminal Law (Codification and Reform) Act, will provide a basis for categorizing crimes, such as Grade A for serious offenses like murder and Grade B for less severe crimes such as theft. Using a purposive sampling technique, this study ensures a diverse representation of crime categories, age groups, gender, and facial characteristics. To ensure the quality and accuracy of the model, image preprocessing techniques such as normalization and augmentation will be applied. The core of the system will leverage Convolutional Neural Networks (CNN) for face recognition, with machine learning algorithms like Random Forest or Support Vector Machines (SVM) used to classify crimes based on severity. The proposed system aims to offer real-time identification and classification of offenders. Evaluation will involve testing the system with a separate dataset of 200 images, measuring performance through key metrics like precision, recall, F1-score, and ROC analysis. Additionally, feedback from police officers will be incorporated to assess the practical utility and effectiveness of the system. This approach not only promises to enhance the operational capacity of law enforcement in Zimbabwe but also presents an innovative application of AI in the criminal justice system.

* 1. **Conclusion**

## This chapter outlines the proposed research aimed at developing a Smart Criminal Identification System for Masvingo Central Police Station in Zimbabwe, leveraging Artificial Intelligence (AI) techniques to enhance crime detection and offender tracking. The proposed system integrates Convolutional Neural Networks (CNN) for facial recognition and supervised machine learning algorithms to classify crimes based on their severity. The research adopts an experimental design, with data collection involving primary sources, such as facial images of convicts and wanted individuals, and secondary sources, including legal frameworks such as the Standing Order Volumes and the Zimbabwe Criminal Law Act. The proposed system seeks to address gaps in the current policing infrastructure, especially in marginalized areas, by enabling real-time identification and crime categorization. The chapter also outlines the proposed methodology, data collection process, and evaluation metrics, emphasizing the potential of AI to improve law enforcement effectiveness in Zimbabwe.

**Chapter 2: Literature Review and Feasibility Study**

**2.1 Literature Review**

This chapter critically examines existing scholarly contributions related to the use of image processing and machine learning (ML) algorithms in smart criminal identification systems. It presents a comparative analysis between developments in high-income countries and the emerging landscape within the Southern African Development Community (SADC), with a specific focus on Zimbabwe. The chapter further highlights contextual challenges, ethical considerations, and research gaps that this study seeks to address. In high-income countries, criminal identification has been significantly transformed by the integration of image processing and ML techniques, especially through the adoption of advanced deep learning models. Technologies such as Convolutional Neural Networks (CNNs) have demonstrated remarkable performance in facial recognition, video surveillance, and biometric data matching (Parkhi, Vedaldi, & Zisserman, 2015; Schroff, Kalenichenko, & Philbin, 2015).

**2.2 What has been done in developed countries**

The United States’ Next Generation Identification (NGI) system, managed by the Federal Bureau of Investigation, combines biometric modalities such as facial, fingerprint, and iris recognition within a centralized AI-driven infrastructure (Jain, Ross, & Nandakumar, 2011). Similarly, the UK’s law enforcement agencies have tested real-time facial recognition technology in public spaces, linking it with national criminal databases for proactive surveillance (Babuta, Oswald, & Sullivan, 2020). Deep learning models such as FaceNet and DeepID3 achieve high precision when trained on large-scale datasets like MS-Celeb-1M, Labeled Faces in the Wild (LFW), and VGGFace2 (Schroff et al., 2015; Taigman, Yang, Ranzato, & Wolf, 2014). These systems not only support retrospective investigations but are increasingly applied in real-time surveillance contexts through edge computing and cloud-based architectures (Zhao, Jia, & Liu, 2019). Nonetheless, there are growing concerns about algorithmic bias and fairness in facial recognition technologies. Buolamwini and Gebru (2018) exposed significant racial and gender disparities in commercial facial analysis algorithms, with error rates as high as 34% for dark-skinned women, compared to less than 1% for light-skinned men. Such findings have ignited global debates on ethical AI, leading to increased efforts in algorithmic auditing and transparency (Raji & Buolamwini, 2019).

**2.3 What has been done in developing countries**

South Africa is at the forefront of AI-driven law enforcement in the region. Research conducted by the Council for Scientific and Industrial Research (CSIR) and universities such as the University of Pretoria has demonstrated the feasibility of low-cost facial recognition systems based on MobileNet and SqueezeNet, particularly when enhanced by histogram equalization and noise filtering (Mbatha, Moodley, & Twala, 2021). Despite their technical success, these models face challenges when applied to African populations, due to the lack of local facial datasets. As a result, efforts have emerged to build regionally inclusive datasets—such as the AfricanFaceSet project designed to improve model generalizability and accuracy (Abdulrahman & Adewumi, 2020). In Zimbabwe, scholarly engagement with smart criminal identification is in an early developmental phase. Universities such as the University of Zimbabwe and the Harare Institute of Technology have initiated pilot projects that apply transfer learning techniques on pre-trained CNNs (e.g., EfficientNet, ResNet18) using small-scale datasets curated from local surveillance systems (Chikobvu & Muchabaiwa, 2022).

**2.4 What has been done in Zimbabwe?**

Zimbabwean researchers have explored the use of offline facial recognition models on mobile devices, an approach that addresses local limitations in internet connectivity and computational power. Image preprocessing methods, such as contrast-limited adaptive histogram equalization (CLAHE) and face alignment, have been shown to significantly enhance model performance in low-quality imaging environments (Moyo & Sibanda, 2023). However, practical implementation remains hindered by the absence of a centralized biometric database, outdated surveillance infrastructure, and weak legal frameworks to regulate data privacy and algorithmic accountability (Goredema, 2020). There is also minimal public discourse on the ethical implications of AI-powered surveillance, creating potential risks for misuse in politically sensitive environments. The SADC Regional AI and Digital Innovation Strategy (2023–2027) outlines public safety and smart surveillance as key focus areas for regional digital transformation. Cross-border initiatives, such as the Smart Policing Project piloted in Botswana and under review in Zimbabwe, propose the use of facial recognition-enabled mobile scanners for real-time suspect verification. These initiatives are supported by international development organizations and seek to promote inclusive, ethical, and secure AI applications in the justice sector (SADC Secretariat, 2023). Open-source platforms such as TensorFlow Lite and OpenCV have been instrumental in democratizing access to AI tools in the region, enabling researchers and developers in Zimbabwe to experiment with real-time image processing and ML classification even with limited hardware (Hove, 2022). However, sustainability and scalability remain contingent on state investment, public-private partnerships, and regional policy harmonization.

**2.5 Synthesis and Research Gap**

While the literature indicates substantial global progress in smart criminal identification, the SADC region—and Zimbabwe in particular—remains under-researched and under-resourced. Existing studies often rely on imported datasets and architectures, leading to contextual misalignment. Moreover, legal and ethical frameworks lag behind technological development, posing serious risks for surveillance abuse and data exploitation.This study addresses these gaps by proposing a contextually adapted smart identification model that leverages lightweight machine learning architectures and local image datasets. It also advocates for a rights-based framework for deployment, integrating technical efficacy with ethical governance.

**2.6 Conclusion**

Smart criminal identification systems using image processing and machine learning offer transformative potential for law enforcement worldwide. However, their development and application remain highly uneven. While countries in the Global North advance toward predictive policing and behavioral analytics, nations like Zimbabwe are still building the foundational structures necessary for effective and ethical adoption. As such, this research contributes not only to technological innovation but also to the regional discourse on AI governance, justice, and digital sovereignty.

## 2.7 Feasibility Study: Smart Criminal Identification System for Zimbabwe

This feasibility study examines the viability of developing and implementing a smart criminal identification system in Zimbabwe, utilizing image processing and classification machine learning algorithms.

**2.7.1 Project Goal:**

To assess the feasibility of a smart criminal identification system for Zimbabwe.

* **Objectives:**
  + Evaluate technical feasibility.
  + Assess operational feasibility.
  + Provide a project schedule.
  + Conduct a preliminary economic cost-benefit analysis.
  + Identify potential security threats and controls.

### 2.7.2 Project Description

* The proposed system will use image processing and machine learning to identify criminals from visual data (images and videos).
* Key components include:
  + Image and video capture devices
  + Data storage and management systems
  + Image processing software
  + Machine learning algorithms (CNNs, SVMs, etc.)
  + A database of criminal records
  + User interface for law enforcement personnel

### 2.7.3 Feasibility Analysis

#### 2.7.3.1 Technical Feasibility

* **Assessment:**
  + **Technology Availability:**
    - Image capture devices: Readily available (CCTV, mobile cameras).
    - Hardware: Servers, processing units are procurable.
    - Software: Open-source and commercial options exist for image processing and ML (e.g., OpenCV, TensorFlow, PyTorch).
  + **Technical Expertise:**
    - Availability of skilled personnel (data scientists, software engineers) in Zimbabwe or the potential for training and development.
  + **Data Availability:**
    - The availability and quality of criminal record data, including images and videos, within Zimbabwe.
  + **Infrastructure:**
    - Existing infrastructure, including power supply and network connectivity, to support system deployment and operation.
* **Findings:** Technically feasible, but requires investment in infrastructure and skills development.

#### 2.7.3.2 Operational Feasibility

* **Assessment:**
  + **System Integration:**
    - The ability to integrate the new system with existing law enforcement workflows and databases.
  + **User Acceptance:**
    - The willingness of law enforcement personnel to adopt and use the new system.
  + **Organizational Structure:**
    - The existing organizational structure and the changes needed to accommodate the system.
  + **Training Requirements:**
    - The extent of training required for law enforcement personnel to effectively use the system.
  + **Policy and Legal Frameworks:**
    - The need for new policies and legal frameworks to govern the use of the technology, ensuring compliance with privacy laws and human rights.
* **Findings:** Operationally feasible with proper change management, training, and policy development.

### 2.7.4 Project Schedule

Stages and Time Frames:

|  |  |  |
| --- | --- | --- |
| Stage | Description | Timeframe |
| Project Initiation | Project planning, resource allocation, and stakeholder engagement. | 1-2 Weeks |
| Requirements Gathering | Detailed analysis of user needs, system requirements, and data sources. | 2-3 Weeks |
| System Design | System architecture design, software development plan, and hardware specification. | 3-4 Weeks |
| System Development | Software development, hardware procurement, and system integration. | 6-12 Weeks |
| Testing and Deployment | System testing, pilot deployment, and user training. | 4-6 Weeks |
| Evaluation and Maintenance | System performance evaluation, ongoing maintenance, and updates. | Ongoing |

* **Total Estimated Time:** 16 - 27 weeks + Ongoing Maintenance

### 2.7.5 Economic Cost-Benefit Analysis

#### 2.7.5.1 Preliminary Budget

* **Cost Categories:**
  + **Hardware:** Servers, cameras, storage devices.
  + **Software:** Development, licenses, integration.
  + **Personnel:** Salaries for developers, project managers, trainers.
  + **Training:** Law enforcement personnel training.
  + **Data Acquisition:** Costs associated with gathering and pre-processing data.
  + **Infrastructure:** Network upgrades, power supply.
  + **Maintenance:** Ongoing system maintenance and support.
* **Estimated Costs:** A detailed budget will be developed in the next phase, but a preliminary estimate is provided below.

| **Item** | **Estimated Cost (USD)** |
| --- | --- |
| CCTV and Capture Equipment | $15,000 – $30,000 |
| Server and Storage Hardware | $10,000 |
| Software Development & Integration | $5,000 – $15,000 |
| Maintenance & Training | $2,000/year |

**Funding Sources**:

* Government Smart City initiatives
* NGO partnerships (e.g., UNDP, African Development Bank)
* Local government and community policing projects

**2.7.5.2 Benefits**

* **Direct Benefits:**
  + Reduced crime rates.
  + Faster criminal identification.
  + Improved law enforcement efficiency.
  + Reduced investigation costs.
  + Increased public safety.
* **Indirect Benefits:**
  + Deterrence of crime.
  + Enhanced public trust in law enforcement.
  + Attraction of investment and tourism.

#### 2.7.5.3 Costs

* **Direct Costs:**
  + System development and implementation costs.
  + Hardware and software procurement costs.
  + Training and personnel costs.
  + Maintenance and operational costs.
* **Indirect Costs:**
  + Potential impact on civil liberties.
  + Risk of system errors and biases.
  + Need for regulatory and legal frameworks.
  + Public concerns about privacy.

#### 2.7.5.4 Cost-Benefit Techniques/Methods

* **Methods:**
  + **Net Present Value (NPV):** Discounting future benefits and costs to present value.
  + **Cost-Benefit Ratio (CBR):** Comparing the present value of benefits to the present value of costs.
  + **Return on Investment (ROI):** Measuring the efficiency of the investment.
  + **Sensitivity Analysis:** Assessing how changes in key variables (e.g., crime rates, operating costs) affect the results.

#### 2.7.5.5 Recommendations

* Conduct a detailed cost-benefit analysis using the methods mentioned above.
* Prioritize investments that offer the highest return on investment and align with Zimbabwe's specific needs and priorities.
* Explore funding options from government, international organizations, and private sector.

### 2.7.6 Threats to Security and Controls

* **Threats:**
  + **Unauthorized Access:**
    - Risk of unauthorized access to the system and criminal database.
  + **Data Breaches:**
    - Potential for data breaches and leakage of sensitive information.
  + **System Tampering:**
    - Risk of system tampering or manipulation of data.
  + **Cyberattacks:**
    - Vulnerability to cyberattacks, including malware and denial-of-service attacks.
  + **Data Bias and Inaccuracy:**
    - The system may produce biased or inaccurate results if the training data is flawed or if the algorithms are not properly designed and validated.
* **Controls:**
  + **Access Control:**
    - Implement strict access control measures, including strong passwords, multi-factor authentication, and role-based access control.
  + **Data Encryption:**
    - Encrypt sensitive data both in transit and at rest.
  + **Security Audits:**
    - Conduct regular security audits and vulnerability assessments.
  + **Intrusion Detection:**
    - Implement intrusion detection and prevention systems.
  + **Data Integrity:**
    - Ensure data integrity through hashing, digital signatures, and regular backups.
  + **Security Policies:**
    - Develop and enforce comprehensive security policies and procedures.
  + **Regular Updates:**
    - Keep all software and systems up to date with the latest security patches.
  + **Bias Mitigation:**
    - Employ techniques to mitigate bias in the data and algorithms, and conduct rigorous testing to ensure fairness and accuracy.

# Chapter 3: Methodology

**3.0 Introduction**

This chapter presents the methodological framework used in the development of the Smart Criminal Identification System for Masvingo Central Police Station. The aim of the system is to enhance the efficiency of law enforcement by using artificial intelligence to identify suspects based on facial features from images and classify them according to crime severity. The methodology includes the research design, data acquisition, preprocessing, model training, system development and testing.

**3.1 Research Methodology**

The study makes use of a Design Science Approach as the core methodology, which emphasizes on the iterative development and evaluation of a technological artifact (a functional system). This approach is appropriate as it allows for the creation of an AI-based identification system that addresses specific criminal justice challenges and is usable by the end users. In addition to the Design Science Approach, the study applies an experimental methodology for testing and validating the machine learning models. The effectiveness of the models is evaluated using established performance metrics such as accuracy and F1-score. The methodology also incorporates quantitative analysis through statistical evaluation of the models performance metrics.

**3.2 Data Collection**

**3.2.1 Primary Data**

Primary data were collected from archived criminal records at Masvingo Central Police Station. The dataset comprises facial images of known convicts and suspects along with metadata including name, crime type, crime severity grade, date of offense and location. Ethical clearance was obtained for use of this data and all personally identifiable information was anonymized such as the convicts’ names. This data was used to train the models.

**3.2.2 Secondary Data**

Secondary data sources included legal texts such as the Criminal Law (Codification and Reform) Act [Chapter 9:23], Standing Orders Volumes 1 and 2, and previous research on facial recognition systems in policing. These sources informed the crime categorization scheme and the legal compliance requirements of the system.

**3.3 Dataset Preparation**

Images were processed using OpenCV to ensure consistency and compatibility with the Convolutional Neural Network (CNN). The following preprocessing steps were applied:

* Face detection using Haar cascade classifiers
* Cropping and resizing
* Normalization to scale pixel values between 0 and 1
* Converting to gray scale (since CNN learns well from grayscale)
* Data augmentation through flipping, rotation, and converting to gray scale to enhance model generalization

Each image was linked to its metadata via a CSV file, and the dataset was split into training (70%), validation (20%), and testing (10%) sets.

**3.4 System Architecture and Tools**

The system is composed of three main components:

* 1. Frontend – A React.js web application that allows officers to upload facial images and receive results.
  2. Backend – A Flask-based REST API responsible for image processing, facial recognition, classification, and database communication.
  3. Database – A local SQLite database stores processed records, user activity, and metadata. The system supports offline operations with synchronization functionality for later syncing with a central server.

**3.5 Model Development**

**3.5.1 Facial Recognition (CNN)**

A Convolutional Neural Network (CNN) was developed using TensorFlow to detect and recognize faces. The model consists of several convolutional and pooling layers followed by fully connected layers and it outputs a face embedding vector that is matched against existing records using cosine similarity. The model was trained on the preprocessed dataset and optimized using categorical cross-entropy loss and the Adam optimizer.

**3.5.2 Crime Classification**

A Random Forest classifier, implemented in Scikit-learn, was used to categorize crimes into three severity grades:

* + - * Grade A: Serious crimes (e.g., murder, armed robbery)
      * Grade B: Moderate crimes (e.g., theft, assault)
      * Grade C: Minor crimes (e.g., public disturbance, trespassing)

Classification accuracy was evaluated using 5-fold cross-validation and confusion matrix analysis.

**3.6 System Integration**

The system components were developed independently and then integrated into a functional pipeline:

* + 1. The React.js frontend provides an interface for image uploads.
    2. The image is sent to the Flask backend API, which preprocesses the image and performs facial recognition.
    3. If a match is found, the system retrieves the associated metadata and crime severity grade from the SQLite database.
    4. The result is returned as a JSON object and rendered on the dashboard for the officer’s review.
    5. If the criminal is not found they are stored with their associated crime in the database

The architecture supports offline-first deployment, with synchronization to a central database once connectivity is available.

**3.7 Evaluation and Testing**

**3.7.1 Quantitative Evaluation**

The system was evaluated using a separate dataset of 200 images to test recognition and classification performance. The following metrics were recorded:

* + - * Accuracy
      * Precision
      * Recall
      * F1-Score
      * ROC-AUC

Target thresholds:

* + - * Facial recognition accuracy ≥ 85%
      * Classification F1-score ≥ 80%

**3.7.2 Qualitative Evaluation**

The system was deployed on a test machine and evaluated by a small group of officers. Users were asked to:

* + - * Upload images of known suspects
      * Interpret the output
      * Rate system usability on a 5-point Likert scale
      * Provide feedback on interface design and speed

**3.7.3 Test Plan**

The test plan included two phases:

* + 1. Model Testing – Using unseen data to evaluate CNN and classifier performance.
    2. System Testing – Full pipeline testing, from frontend upload to result display. Errors were logged, and performance metrics such as API response time and user interface rendering were captured.

**3.8 Ethical Considerations**

* + - * **Data Privacy and Consent**: All criminal records were anonymized, and consent was obtained for academic use.
      * **Bias Mitigation**: Care was taken to ensure that the dataset was balanced across gender, age and location.
      * **Security**: User access is restricted via authentication. All sensitive data is encrypted at rest and in transit.

**3.9 Conclusion**

This chapter has outlined the methodology used to build and evaluate the Smart Criminal Identification System. The use of Design Science and Experimental approaches has enabled the development of a practical and functional solution suited to the Zimbabwe’s law enforcement context. The next chapter presents the system’s architectural and technical design in the form of diagrams and system models.

# Chapter 4: System Design

**4.0 Introduction**

This chapter presents the system design of the Smart Criminal Identification System. It covers functional modeling through data flow diagrams, system modeling using UML diagrams, data modeling, database structure, interface mockups, and security mechanisms. These elements form the blueprint for the system’s implementation and ensure functionality, usability, and security in deployment.

**4.1 Functional Modeling**

**4.1.1 Context Diagram (Level 0 DFD)**

At the highest level, the system interacts with three external entities: the police officer, the administrator, and the criminal database. The officer uploads a suspect image and receives potential matches. The administrator manages users and verifies uncertain cases. The system queries the criminal database for known offenders.

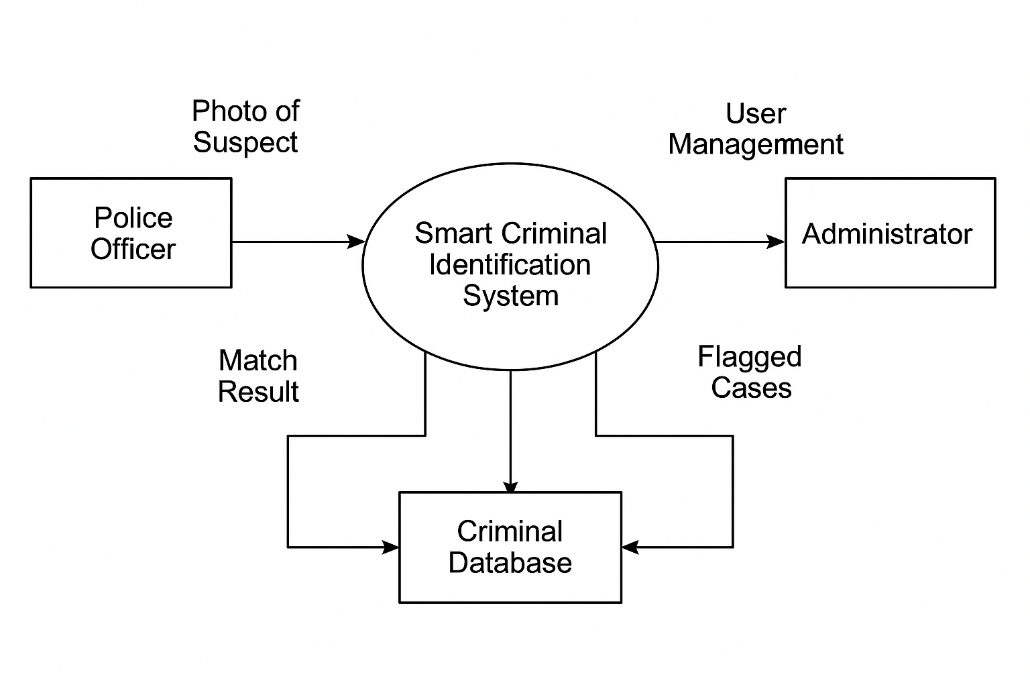
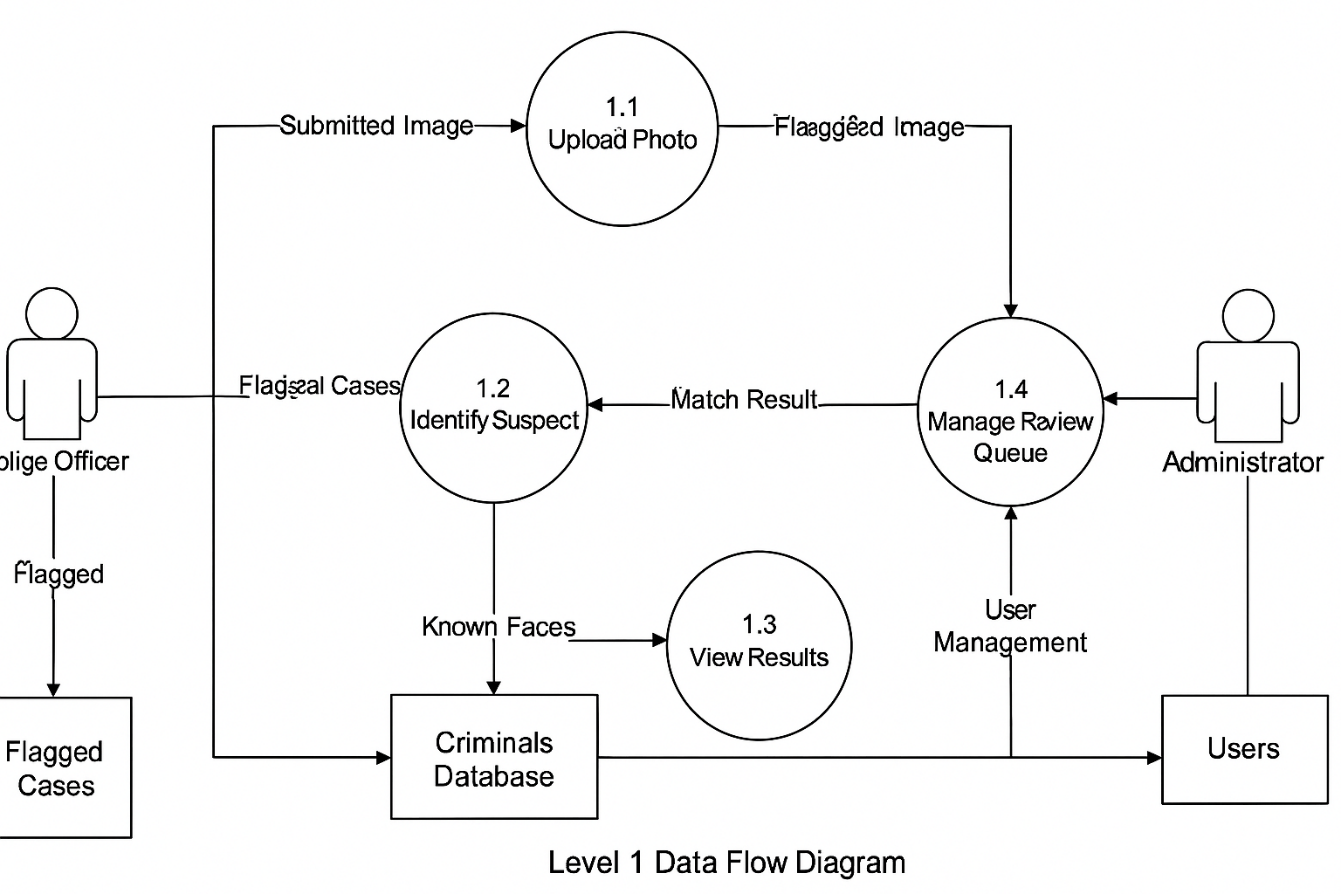


Figure 1Level 0 DFD

**4.1.2 Level 1 Data Flow Diagram (DFD)**

The Level 1 DFD breaks down the main system into subprocesses such as: image upload, preprocessing and recognition, classification, database query, and result display. Each function communicates through data flows and contributes to the overall identification and classification process.



**4.2 System Modeling**

**4.2.1 Use Case Diagram**

Actors include the police officer and the administrator. Key use cases are login, upload photo, view result, add new criminal record, and manage user accounts (admin only).

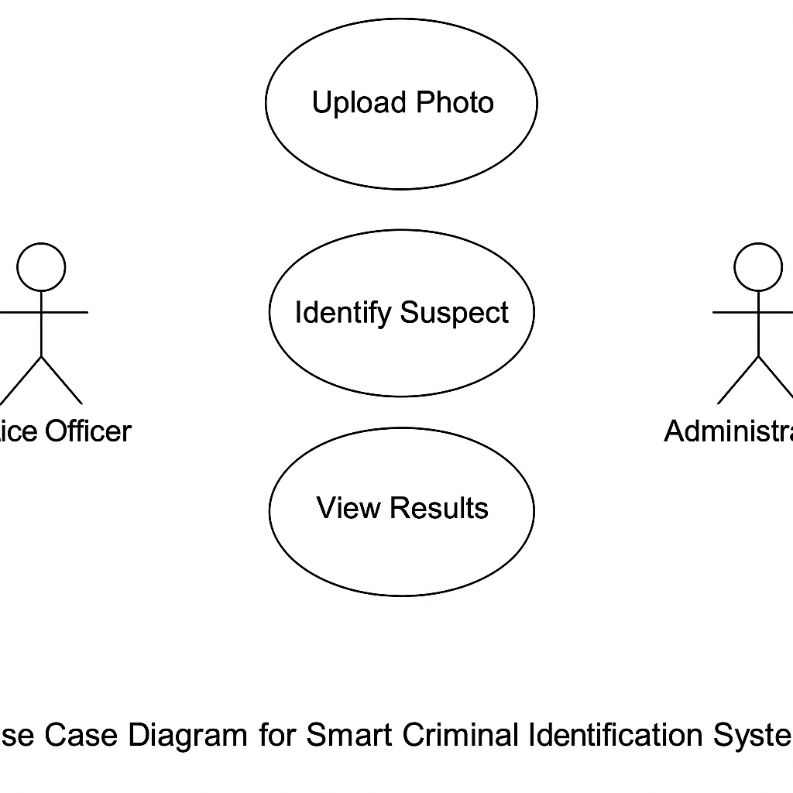


Figure 2Use case diagram

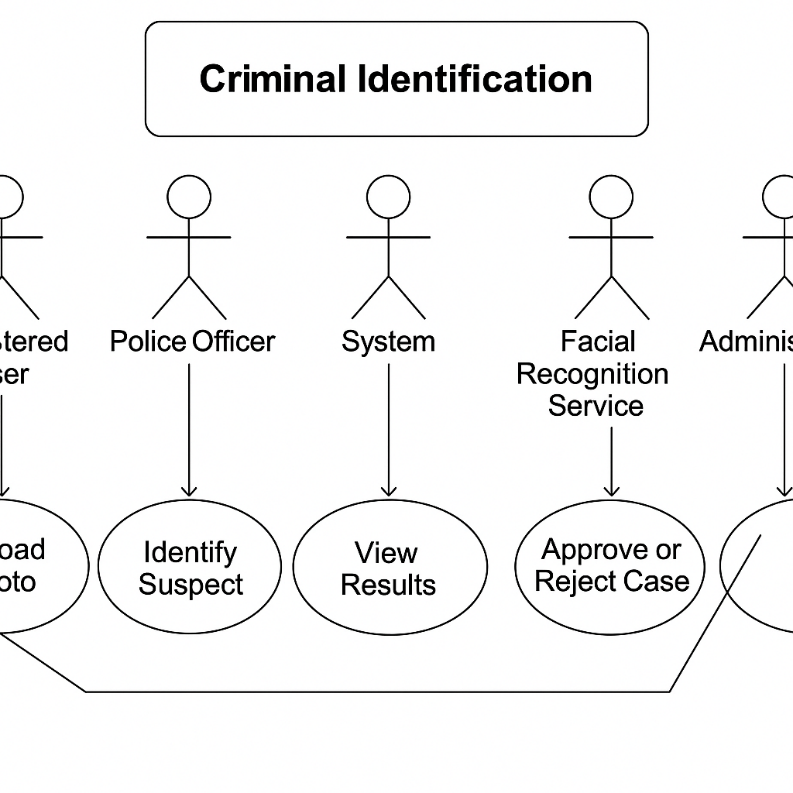


Figure 3Expanded Use case diagram

**4.2.2 Sequence Diagram**

The sequence begins with the user logging in and uploading a photo. The backend receives the image, processes it, and uses the CNN model to identify faces. A query is made to the database, and if a match is found, relevant data is returned to the frontend for display.

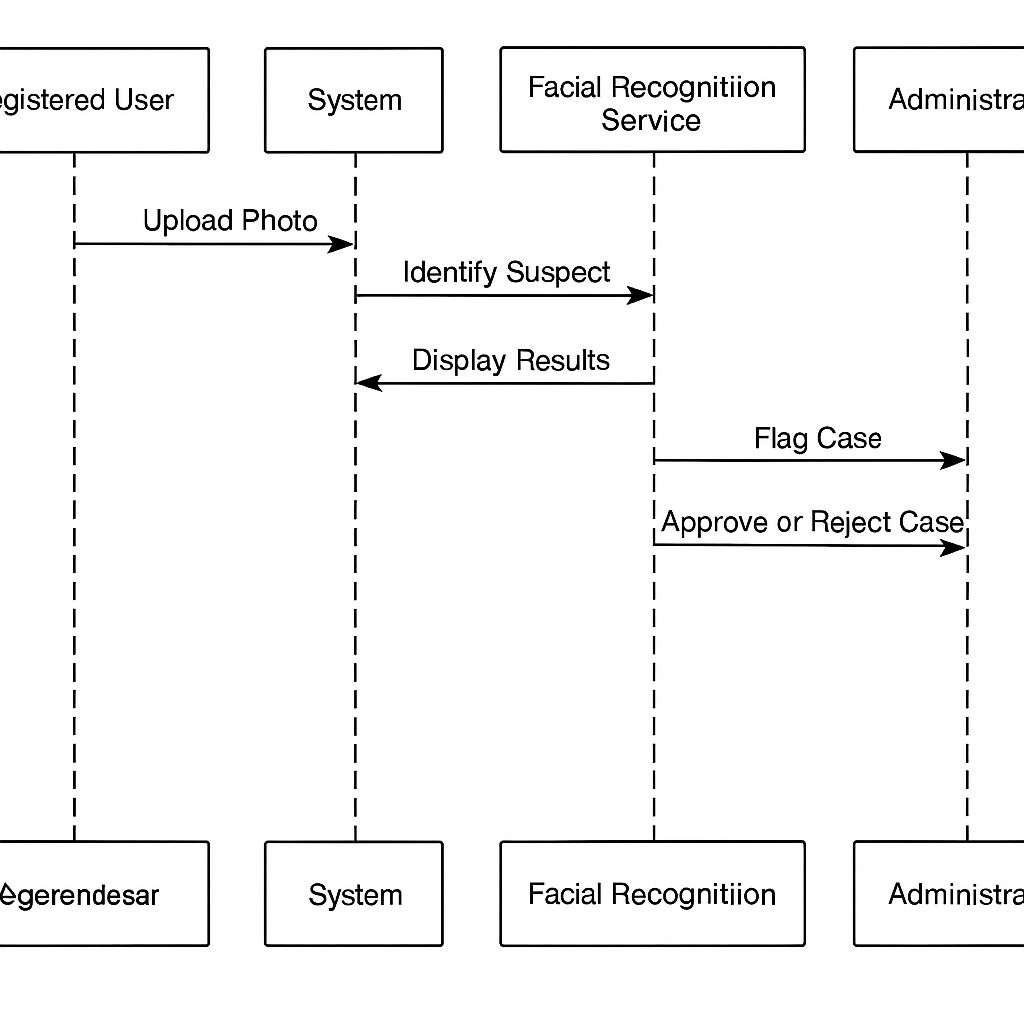


Figure 4UML Sequence Diagram

**4.2.3 State Diagram**

A suspect profile moves through states such as: new entry → unverified → verified → flagged → archived, depending on recognition and administrative review outcomes.

**4.2.4 Activity Diagram**

A police officer logs in, uploads a photo, receives results, and can confirm or flag a match. An admin later reviews flagged entries.

**4.2.5 Architectural Design**

The system is based on a 3-tier architecture: (1) Frontend developed with React.js; (2) Backend using Flask API for preprocessing and model inference; (3) SQLite database for local storage with optional PostgreSQL syncing. RESTful API calls ensure modularity and scalability.

**4.2.6 Component Design / Module Functions**

* + - * Upload Module: Accepts and validates image files.
      * Preprocessing: Detects and normalizes facial features using OpenCV.
      * Recognition Engine: Matches input face against stored embeddings.
      * Classifier: Categorizes crimes by severity.
      * Database Interface: Manages CRUD operations.
      * Dashboard: Presents recognition outcomes.
      * Admin Panel: Manages reviews and users.

**4.3 Data Modeling**

**4.3.1 Entity Relationship Diagram (ERD)**

Entities include users, criminals, upload\_results, and review\_queue. The users table is linked to uploads and reviews. Upload\_results connects users with criminals and stores recognition logs. The review\_queue stores flagged cases pending admin review.

**4.3.2 Class Diagram**

Classes include Criminal, User, UploadResult, and ReviewQueue. Each class contains relevant attributes and methods for creating, reading, and updating system records. Inheritance and relationships define interactions between user actions and system data.

**4.4 Database Design**

The system database includes four main tables: users, criminals, upload\_results, and review\_queue.

* + - * users(user\_id, username, password\_hash, role, created\_at, status)
      * criminals(criminal\_id, full\_name, photo\_path, crime\_type, crime\_grade, location, arrest\_date, is\_repeat\_offender)
      * upload\_results(upload\_id, uploaded\_by, image\_path, match\_confidence, matched\_criminal\_id, status, upload\_time)
      * review\_queue(review\_id, upload\_id, flagged\_by, reason, review\_status, reviewed\_by, review\_time)

This schema ensures normalization and traceability of all activities in the system.

**4.5 Interface Design**

The system has two main dashboards tailored for officers and administrators.  
  
Police Officer Dashboard:

* + - * Login screen for user authentication
      * Upload page for image submission
      * Result viewer for match details, confidence, and crime info
      * Search page to retrieve records by ID or name

Admin Dashboard:

* + - * Includes all officer features
      * Additional 'Pending Reviews' page to handle flagged uploads
      * User Management Panel to create and manage user accounts

**4.6 Security Design**

Security is ensured through:

* + - * Authentication: Username and hashed password with role-based access
      * Input validation: For all file uploads and queries
      * Access Control: Admin-only access to sensitive functions
      * Data Encryption: Sensitive data encrypted in transit (HTTPS) and hashed where necessary
      * Review Mechanism: Manual review of low-confidence matches to prevent false positives

**4.7 Conclusion**

This chapter has provided the complete system design of the Smart Criminal Identification System. Through functional modeling, architectural planning, database structuring, and user interface specification, the chapter establishes the foundation for the next stage of system development and implementation.