Pension Safe: Ensuring Secure Pension Access Through Biometric Authentication

Software Design and Requirement Specification

Session: 2021 – 2025

Submitted by:

Hammad Shahbaz	2021-CS-607
Zain Manzoor	2021-CS-608
Tooba Eman	2021-CS-612

Supervised by:

Dr. Farah Adeeba

Department of Computer Science, New Campus

University of Engineering and Technology Lahore, Pakistan

Contents

List of Figures

List of Tables

Chapter 1

Requirement Specifications

1.1 Functional Requirements

In developing a robust pensioner verification system, a comprehensive set of functional requirements has been defined.

The functional requirements are grouped into four categories: Business requirements, administrative requirements, user requirements and system requirements. Each category is assigned a unique identifier,such as B for business, A for administrative, U for user and S for system requirements. The requirements are then numbered within each category, starting with 01.

Following list describe the main functions the system will do:

- **FR-B-01**: The system must enable and disable super admin.
- **FR-A-01**: The system must share login ID new password with super admin.
- **FR-A-02**: The system must allow super admin to change its password.
- **FR-A-03**: The system must provide forget password functionality to reset password for all users.
- **FR-B-02**: The system must allow super admin to add data entry operators by adding their personal info.
- **FR-B-03**: The system must allow super admin to manage data entry operators.
- **FR-B-04**: The system must allow super admin to remove the data of deadly pensioner and data entry operator.
- **FR-A-04**: The super admin must share login ID and password to data entry operators.

- **FR-A-05**: The super admin must allow data entry operators to change its password.
- **FR-A-06**: The super admin must allow data entry operators to registers pensioners.
- **FR-U-01**: The system must allow data entry operators to enter credentials of pensioners.
- **FR-U-02**: The system shall enable a data entry operator to conduct interviews with pensioners by asking them a set of questions and capturing their responses through video recordings.
- **FR-U-03**: The system shall enable data entry operators to upload videos of pensioners during the registration phase.
- **FR-A-07**: The data entry operators must share login ID and password with Pensioners.
- **FR-U-04**: The data entry operator must allow Pensioners to change its password.
- **FR-U-05**: The system shall provide tutorials that guide pensioners on how to verify their identity.
- **FR-U-06**: The system shall include a dialogue manager that engages pensioners by asking five questions.
- **FR-U-07**: The system shall facilitate seamless interaction in Urdu to accommodate the language preferences of pensioners.
- **FR-S-01**: The system must allow pensioners to record their own videos.
- **FR-S-02**: The system must start recording video by pressing an icon.
- **FR-S-03**: The system shall allow users to change the camera direction (by double-tapping or using an icon).
- **FR-S-04**: The system shall support the collection of video data to aid in development and testing phases.
- **FR-S-06**: The system shall have the capability to save recorded videos.
- **FR-S-07**: The system shall automatically create unique folders for each pensioner to store their recorded videos.
- **FR-S-08**: The system shall be able to convert recorded .mp4 files to .wav format with a 16kHz sampling rate and mono channel.
- **FR-S-09**: The system shall process voice samples (.wav files) to convert speech into text using automatic speech recognition (ASR) technology.

- **FR-S-10**: The system shall verify that the pensioner has answered the specific question being asked.
- **FR-B-05**: The system shall compare audio signals with the trained model and analyze voice features such as tone and frequency to confirm the pensioner's identity.
- **FR-S-11**: The system shall analyze video frames to detect unique facial features.
- **FR-S-12**: The system shall recognize pensioners even if their appearance changes.
- **FR-S-13**: The system shall confirm the pensioner's identity by recognizing live facial features rather than static images.
- **FR-B-06**: The system shall validate the identity of pensioners in a single streamline process that includes facial recognition, voice recognition, and speech recognition..
- **FR-B-07**: The system shall maintain a record of the number of times a pensioner verifies their identity.

1.2 Non-functional Requirements

Complementing the functional requirements, the non-functional requirements define the qualitative attributes and performance expectations of the pensioner verification system. Non-functional requirements outline how the system accomplishes its tasks beyond the core functionality. These aspects focus on enhancing the user experience, system reliability, scalability, performance and support.

The non-functional requirements are grouped into four categories: usability, reliability, scalability, performance. Each category is assigned a unique identifier, such as U for usability, R for reliability, S for scalability and P for performance. The requirements are then numbered within each category, starting with 01.

- **NR-U-01:** The user interface shall have an intuitive design, ensuring easy navigation and user-friendly interactions.
- **NR-U-02:** Color schemes shall be visually appealing and consistent across the application.
- **NR-U-04:** The system shall support interaction exclusively in Urdu to accommodate the language of pensioners.

- **NR-R-01:** The system shall maintain high availability, ensuring it operates 24/7 without significant downtime.
- **NR-R-02:** In case of any disruptions, the system shall have mechanisms to quickly recover and resume normal operation.
- **NR-R-03:** The system shall ensure data integrity during verification processes and storage.
- **NR-S-01:** The system architecture shall be designed to accommodate growing demands without sacrificing performance.
- **NR-S-02:** The system shall support future enhancements and the addition of new biometric modalities.
- **NR-P-01:** The system shall verify a pensioner's identity within 90 to 100 seconds.
- **NR-P-02:** The system shall support concurrent users without performance degradation.
- **NR-P-03:** Actions such as facial recognition, voice recognition, and speech recognition shall occur within acceptable time frames.

Chapter 2

Design Specifications

2.1 Detailed Literature Review

Pensioner verification system introduce various tasks, the study attempts to realize core tasks with face and speaker recognition technology, significant research is happening for this purpose. In this section, we begin by discussing the most prevalent methodologies for visual and audio recognition, followed by the summary of existing methods.

Kong et al[1] in 2024, conducted a study that utilized advanced machine learning models to enhance facial and image recognition technologies. The research made use of datasets such as Face2Gene, Elife, and DermNet, and employed models including K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Deep Convolutional Neural Networks (DCNN), VGG16, and 3DResNet-50. The study achieved a high accuracy of 93.72% across various tasks. Limitations were noted in terms of the reliability of the models, completeness of the dataset, and the models' ability to generalize across different scenarios.

Ohi and Gavrilova[2] in 2024 engaged in a study aimed at improving speaker recognition using datasets such as VoxCeleb1, LibriSpeech, and Musan. The research employed LVDNet, a deep learning model, and achieved an accuracy of 94.98%. A significant limitation of the study was that the model could not be utilized for speaker diarization problems, which are essential for separating and identifying multiple speakers in a conversation.

ALANSA RI[3] in 2023, conducted a study focused on improving face recognition technologies using the MS1MV2 and MS1MV3 datasets. The research employed GhostFaceNets, a lightweight neural network model, and achieved an impressive accuracy of 99.73%. Despite the high accuracy, the study highlighted a key limitation: the model may not generalize well to all face recognition scenarios, particularly in more diverse or uncontrolled environments.

Hangaragi [4] in 2023 conducted a study focused on facial recognition using the LFW and

BU3DFE datasets. The research utilized Face Mesh and MediaPie models, achieving an accuracy of 94.23%. However accuracy was reduced when faces were captured from certain angles, affecting the overall reliability in real-world scenarios

Trau-Margalit[5] in 2023 conducted a study on eye-tracking using the Tobii Eye Tracker dataset. The research employed models such as Generalized Correlation Analysis (GCA) and Multilevel Modeling (MLM), yielding significant results. The study faced limitations related to sample size, potential bias, and the scope of the research, which may have impacted the generalizability of the findings.

Tkauc[6] in 2022 undertook a study focused on recognizing spoken names using the Kperson website and spoken names datasets. The research applied Amazon Rekognition and Google Cloud Speech-to-Text (GCST), achieving an accuracy of 73.1%. Despite these results, the study faced challenges related to language barriers and background noise, which impacted the performance of the models. As Isaac Asimov once said, "The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom." These issues reflect how technology, while advancing, still struggles with real-world complexities like linguistic diversity and environmental noise.

Fahn et al [7] in 2022 conducted a study focused on the development of humanized care robots to address the growing need for elderly care due to the aging population and shortage of caregivers. The research centered on improving four key tasks in image and speech recognition technology: human pose recognition, facial expression recognition, eye gazing recognition, and Chinese speech recognition. Fahn used models such as Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and Continuous Bag of Words (CBOW) to enhance these tasks. The study achieved accuracies of 74.1% in pose recognition, 87.8% in facial expression recognition, and 89.0% in speech recognition.Further improvements are needed, especially in minimizing interference and refining recognition techniques to better suit elderly care environments

Asha et al.[8] in 2022 carried out a study on deepfake detection using datasets such as MMDFD, FakeAVCeleb, and DFDC. The research employed a combination of VGG16 and MFCC (Mel Frequency Cepstral Coefficients) models, achieving an accuracy of 92%. However, the study highlighted limitations in deepfake detection, which could affect the model's robustness in detecting manipulations across more sophisticated deepfakes.

Teoh[9] in 2021 carried out a study aimed at enhancing image recognition using high-quality image datasets. The research applied Haar Cascade, OpenCV, and Deep Neural Networks (DNN) to achieve an accuracy of 91.7%. However, the performance of the models was found to be affected by factors such as distance from the camera and low-light conditions, limiting their effectiveness in real-world applications where these variables are common.

Li et al[10] in 2021 performed a study focused on speaker recognition using datasets such as Aishell1 and VoxCeleb1. The research utilized Hybrid Deep Neural Networks (DNNs) and achieved a 26% reduction in Equal Error Rate (EER). A limitation in that it did not incorporate a deep speaker embedding learning approach, which may have further improved recognition performance.

Ohi et al. [11] in 2021 conducted a study on speaker recognition using datasets such as NIST SRE, TIMIT, and LibriSpeech. The research utilized SincNet, SpeakerGAN, and i-vector models, achieving an accuracy of 95%. However, the study identified that noisy input significantly impacted the performance of the models, reducing their effectiveness in real-world noisy environments.

Tabassum[12] in 2020 undertook a study focusing on facial expression recognition using datasets such as Grimace, CK+, FERET, and ORL. The research utilized Principal Component Analysis (PCA), Discrete Wavelet Transform (DWT), and Linear Discriminant Analysis (LDA) models, achieving an accuracy of 95.44%. Key limitation identified in the study was the reliance on DWT across all methods, which may have restricted the overall versatility of the models in handling different types of input data.

Table 2.1 presents research work in the areas of face, speaker, and speech recognition.:

TABLE 2.1: Research work on facial and speaker recognition models

Year	**Authors**	**Dataset**	**Models**	**Accuracy**	**Limitations**
2024	Kong et al [1]	Face2Gene, Elife, DermNet, 3DResNet-50	KNN, SVM, DCNN, VGG16, ResNet	93.72%	Reliability, dataset completeness, generalization
2024	Ohi and Gavrilova [2]	VoxCeleb1, LibriSpeech, Musan	LVDNet	94.98%	Can't be used for speaker diarization problems

Year	Authors	Dataset	Models	Accuracy	Limitations
2024	Li et al [10]	Aishell1, VoxCeleb1	Hybrid DNNs	26% EER reduction	No deep speaker embedding learning approach
2023	Hangaragi [4]	LFW, BU3DFE	Face Mesh, MediaPie	94.23%	Limited accuracy under certain angles
2023	ALANSA RI [3]	MS1MV2, MS1MV3	GhostFaceNets	99.73%	May not generalize to all face recognition scenarios
2023	Trau-Margalit [13]	Tobii Eye Tracker	GCA, MLM	Significant results	Sample size, bias, scope limitations
2022	Fahn et al [7]	Elderly body images, Chinese speech	LSTM, CNN, CBOW	74.1%, 87.8%, 89.0%	Improving techniques, minimizing interference
2022	Tkauc [6]	Kperson website, spoken names	Amazon Rekognition, GCST	73.1%	Issues with language barriers, noise
2022	Asha et al [8]	MMDFD, FakeAVCeleb, DFDC	VGG16+MFCC	92%	Language support, deepfake detection
2021	Saxena and Varshney [5]	RSR2015, 202,599 images	DeepID2, FaceNet, GMM, CNN	95.43%, 95.92%, 87%	Eye-based recognition of masked users
2021	Teoh [9]	High-quality images	Haar Cascade, OpenCV, DNN	91.7%	Affected by distance, low light
2021	Ohi et al [11]	NIST SRE, TIMIT, LibriSpeech	SincNet, SpeakerGAN, i-vector	95%	Noisy input affects performance

Year	Authors	Dataset	Models	Accuracy	Limitations
2021	Park [14]	LibriSpeech 100-860	LSTM, RNN-T, CNN encoders	WER improved to 3.9%, 8.8%	Minimal effect on LibriLight
2020	Tabassum [12]	Grimace, CK+, FERET, ORL	PCA, DWT, LDA	95.44%	Relying on DWT for all methods
2019	David et al [15]	VoxCeleb-1+2, SITW	TDNN, ResNet43+LDE	Min error 1.50%, 1.55%	Data dependent, complexity with deeper architectures

The need of pensioner verification system is evident from the prevalence weaknesses of current methods. Existing research highlights vulnerabilities in verification systems and underscores the necessity for robust authentication mechanisms. They are limited scalability due to reliance on traditional image methods and potential interference with multithreading in speech recognition. And most importantly locally, there is a notable absence of systems utilizing both face and voice recognition for pensioner verification. This gap poses risks for pensioners and emphasizes the urgent need for a modernized system to enhance authentication.

2.2 Proposed Methodology

The proposed architecture diagram is shown in Figure 2.1:

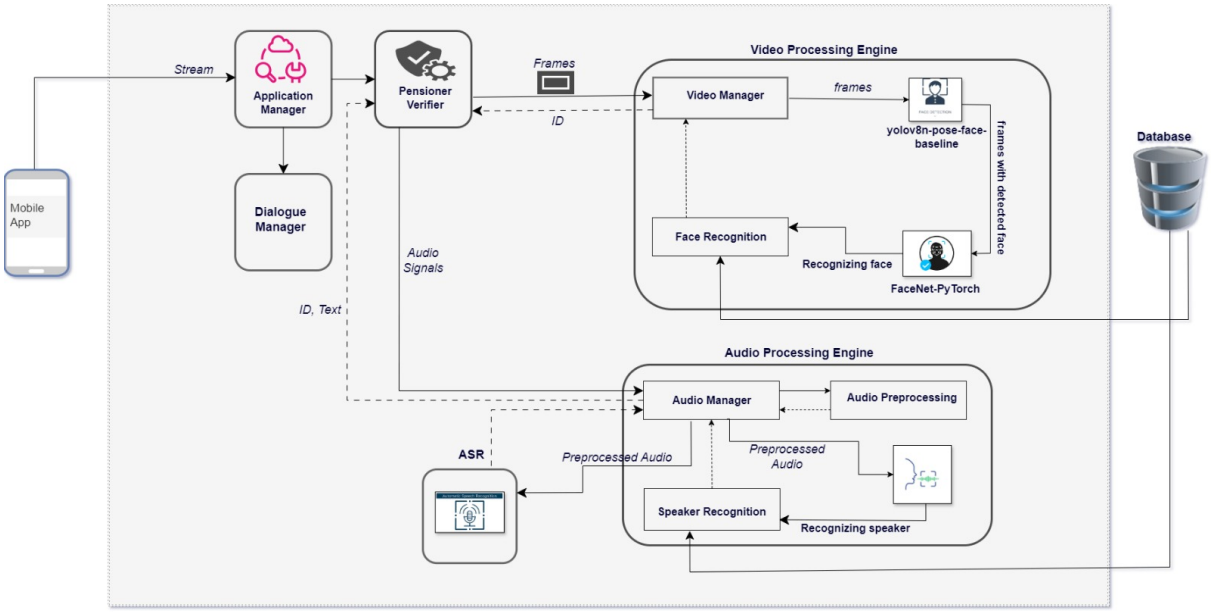

FIGURE 2.1: Architecture Diagram

Major components of app are:

- **Mobile App:** The Mobile App captures streaming data, which is then transferred to the Application Manager for further processing.
- **Dialogue Manager:** This component prompts questions for the pensioner to answer during the verification process.
- **Application Manager:** It collaborates with the Dialogue Manager to ask the user questions and receives their responses. It then sends the streaming input to the Pensioner Verifier for analysis.
- **Pensioner Verifier:** Receives streaming input from the Application Manager and forwards frames to the Video Manager and audio signals to the Audio Manager. It processes outputs from both modules and determines the Final Output.
- **Video Processing Engine:** The Video Manager sends frames to the Face Recognition module, which compares them with the database to generate facial recognition outputs. These outputs are then sent back to the Video Manager and forwarded to the Pensioner Verifier.

- **Audio Processing Engine:** The Audio Manager sends incoming audio data to the Audio Preprocessing module. Preprocessed audio is then forwarded to both the Automatic Speech Recognition (ASR) and Speaker Recognition modules. ASR outputs are sent back to the Audio Manager, while Speaker Recognition results, after comparison with the database, are also sent to the Audio Manager. Finally, the Audio Manager forwards outputs from both modules to the Pensioner Verifier for further analysis.
- **Evaluation:** Testing validates facial and speaker recognition accuracy, assesses module integration, and ensures dialogue manager effectiveness across diverse scenarios, enhancing system reliability and performance.

As the Pensioner Verifier app relies on face and speaker recognition models for its functionality, the following Section 2.2.1 is a detailed explanation of how these models are developed and how they work.

2.2.1 Model Preparation

In our pensioner verification app, we use face recognition and speaker recognition models. These models work together to verify the user's identity through both face and voice, ensuring secure and accurate verification.

Following section provide an overview of the techniques employed to gather and analyze data effectively.

1)- Face Recognition Model:

This section outlines the methodology and procedures for preparing the face recognition model, including data design, data collection, data preprocessing, model training, model testing, model improvement, and model deployment to ensure robust and accurate facial recognition performance.

- **Data Design:**

  A dataset of 35 to 50 images per pensioner will be collected, ensuring diversity by capturing different facial expressions and orientations.

  The images will be taken under various conditions, such as eyes open/closed, looking at the camera or in different directions, and mouth open/closed.

  The dataset will also include images with different lighting conditions (bright and dark) and varied backgrounds to improve the model's robustness.

- **Data Collection:**

  A camera will be used to capture high-quality images of each pensioner in diverse scenarios to simulate real-world variations in facial expressions and lighting.

  The collected images will be organized into separate folders for each individual, properly labeled for training and testing purposes.

- **Data Preprocessing:**

  The images will be converted to the correct format using the Python PIL library.

  The images will be resized to standard dimensions (160x160) and normalized with the values used by FaceNet (mean=[0.5, 0.5, 0.5] and std=[0.5, 0.5, 0.5]).

  The pre-processed images will be transformed into tensors suitable for input into the FaceNet model.

- **Model Training:** FaceNet, with pretrained weights from the VGGFace2 dataset, will be utilized to generate embeddings from the facial images.

  Embeddings will be created for each pensioner, representing unique feature vectors for facial recognition.

  The embeddings and corresponding labels for each person will be stored in .npy files for later use in recognition.

  The diverse dataset of pensioner images will ensure high accuracy during training, allowing the model to generalize across various real-world conditions.

- **Model Testing:**

  The model will be tested using a Flask API, where the pensioner's video will be uploaded for processing.

  During testing, YOLOv8n-pose-face-baseline will be used for face detection, and FaceNet will be applied for facial recognition based on the pre-trained embeddings.

  The Euclidean distance between the detected face embeddings and the stored embeddings will be calculated to verify the pensioner's identity.

  Additional features like eye aspect ratio (EAR) for blink detection and mouth aspect ratio (MAR) for lip movement detection will be calculated to ensure liveness detection, preventing spoofing.

- **Model Improvement:**

  Liveness detection through blink and lip movement will improve the robustness of the system, reducing the chances of false acceptance or rejection.

If the recognition confidence falls below the threshold, additional data will be collected to fine-tune the model for improved accuracy.

- **Model Deployment:** The model will be deployed as a Flask API, where pensioners can upload their video, and the system will return a verification result.

  Video frames will be processed in real-time for face detection and recognition, with results displayed in the form of bounding boxes, landmarks, and recognition labels.

  The API will support real-time tasks managed through threading and queuing to efficiently handle multiple requests.

2)- Speaker Recognition Model:

This section outlines the methodology and procedures for developing the voice recognition model, which will encompass data design, data collection, segmentation of audio files, data preprocessing, generating train and test data, model training, and model testing to ensure high accuracy and reliability in voice recognition.

- **Data Design:**

  A structured set of verification questions will be prepared, primarily focusing on personal and career-related information. These questions will be crafted to prompt natural and varied responses, ensuring the collection of diverse speech patterns.

- **Data Collection:**

  Each pensioner will be recorded in a series of 1-minute video clips, resulting in a total of 10 minutes of video footage per individual. Rather than recording separate voice samples, the voice data will be extracted directly from the recorded video clips. These videos will capture participants answering a series of questions, providing the necessary speech patterns for training the system's voice recognition capabilities.

- **Data Preprocessing:**

  Segmentation of Audio Files: Audio files will be split into smaller segments based on criteria such as silence detection or duration thresholds, and the segmented files will be saved accordingly.

  Preprocessing: Tasks like renaming and format conversion will be applied to ensure the audio files meet quality standards, including minimum duration requirements. Data will be further refined and organized by moving non-waste files to new directories, and audio duration's will be calculated.

Generating Train and Test Data: The preprocessed files will be organized into training and testing sets for machine learning. This step will ensure proper distribution of data for effective model training and evaluation.

- **Model Training:**

  I-vectors will be extracted from the segmented audio files, and a speaker recognition model will be trained. This process will involve feature extraction, i-vector normalization, and model training using tools like Kaldi, ultimately resulting in a trained speaker recognition model.

- **Model Testing:**

  Identification Process: The speaker identification process will be run on the test data, where the trained model will be loaded, the test data processed, and identification results produced.

  Model Evaluation: The model's performance will be evaluated by testing it on segmented audio files or features. This step will output the model's performance metrics and help identify areas for improvement.

- **Model Improvement:**

  Model improvement will happen after testing, especially if results are inconsistent. To fix this, we will provide more voice data to retrain the model, helping it perform better. After retraining, we will adjust the threshold to ensure the model makes accurate decisions. This process will help the model give more reliable and stable results.

- **Model Deployment:**

  Model deployment will occur as a Flask API. When a pensioner verifies themselves through a video, the audio in .wav format will be extracted from the video and sent to the model. The model will process this audio and return the pensioner's ID and the text from the speech.

2.3 Data Collection Techniques

The journey of data collection for the Pensioner Verification System will be a meticulously planned and executed process. The primary goal will be to gather comprehensive video and audio data from individuals aged 65 to 80, serving as the foundation for developing and testing the system. This process will involve several carefully designed steps to ensure the accuracy and diversity of the dataset.

- **Preliminary Discussions and Question Design:** We will initiate the data collection process by engaging in detailed discussions and conducting thorough research. This will allow us to prepare a structured set of verification questions, primarily focusing on personal and career-related information. These questions will be crafted to prompt natural and varied responses, ensuring the collection of diverse speech patterns and facial expressions.

- **Video Recording:**

  Each participant will be recorded in a series of 1-minute video clips, totaling 10 minutes of video footage per individual. The video recordings will capture participants as they answer the prepared questions, ensuring that the spoken content is in Urdu. Special attention will be given to the participants' facial expressions, with instructions to change expressions and face angles during the recording, mimicking real-world verification scenarios. These variations will help improve the robustness of the facial recognition model.

- **Pose and Expression Variations:**

  To simulate the conditions encountered in real-world verification processes, participants will be encouraged to alter their facial poses and angles during the recording. This step will be crucial to train the model on diverse facial expressions and orientations, enhancing the system's ability to perform under different circumstances.

- **Voice Samples Extraction:**

  Voice data will play a crucial role in building a comprehensive dataset. Rather than recording separate voice samples, we will extract the voice data directly from the recorded video clips. These videos will capture participants answering a series of questions, providing the necessary speech patterns for training the system's voice recognition capabilities. This integrated approach will allow us to collect both visual and auditory data simultaneously, ensuring that the voice samples correspond naturally with facial expressions and body language.

- **Environment and Background Noise Management:**

  As the Pensioner Verification System is designed for indoor use, we will focus on maintaining an environment where background noise is controlled to a manageable level. While some background noise will be present to mimic realistic indoor settings, care will be taken to ensure that the noise does not significantly affect the quality of the audio. The system will be optimized to handle the typical background sounds encountered in indoor environments, balancing the need for realism without overwhelming the voice recognition functionality.

- **Participant Selection:**

  A total of 60 participants will be selected for the data collection process, comprising 20 women and 40 men, all aged between 65 to 80 years. This age group will be specifically chosen to cater to the system's target audience—retired pensioners. We will also ensure that none of the participants wear glasses during the recording, as this could interfere with accurate facial recognition.

- **Diversity and Realism in Data:**

  he collected data will aim to reflect realistic verification environments, ensuring that the model can generalize well. By incorporating variations in poses, expressions, speech patterns, and background conditions, we will build a diverse and representative dataset likely to enhance the accuracy and robustness of the Pensioner Verification System.

2.4 Experimental Design

The experimental design for the pensioner verification system will involve collecting data from 60 individuals including 40 men and 20 women aged from 65 to 80 years. The dataset will be split into two parts: data from 40 participants will be used to train the system, while the remaining 20 participants' data will be reserved for testing the system's accuracy and reliability.

Testing will be conducted in an indoor environment with moderate background noise, such as the sound of a fan, to simulate real-world conditions. To ensure the effectiveness of the face recognition component, pensioners will be instructed not to wear glasses or masks. Additionally, clear speech will be required for accurate speaker and speech recognition, ensuring optimal performance of the system under these controlled conditions.

2.5 WireFrames

This section presents the wireframes designed for the application, illustrating the user interface and interaction flow. These wireframes provide a visual representation of the application's layout and functionality, guiding the development and ensuring a user-friendly experience.

2.5.1 Pensioner Enrollment App

The wireframes below illustrate the user interface for the pensioner enrollment process, showcasing screens for login, home, initializing enrollment process and enrollment process.

2.5.1.1 Login(Data Entry Operator)

The login wireframe for data entry operator is shown in Figure 2.2:

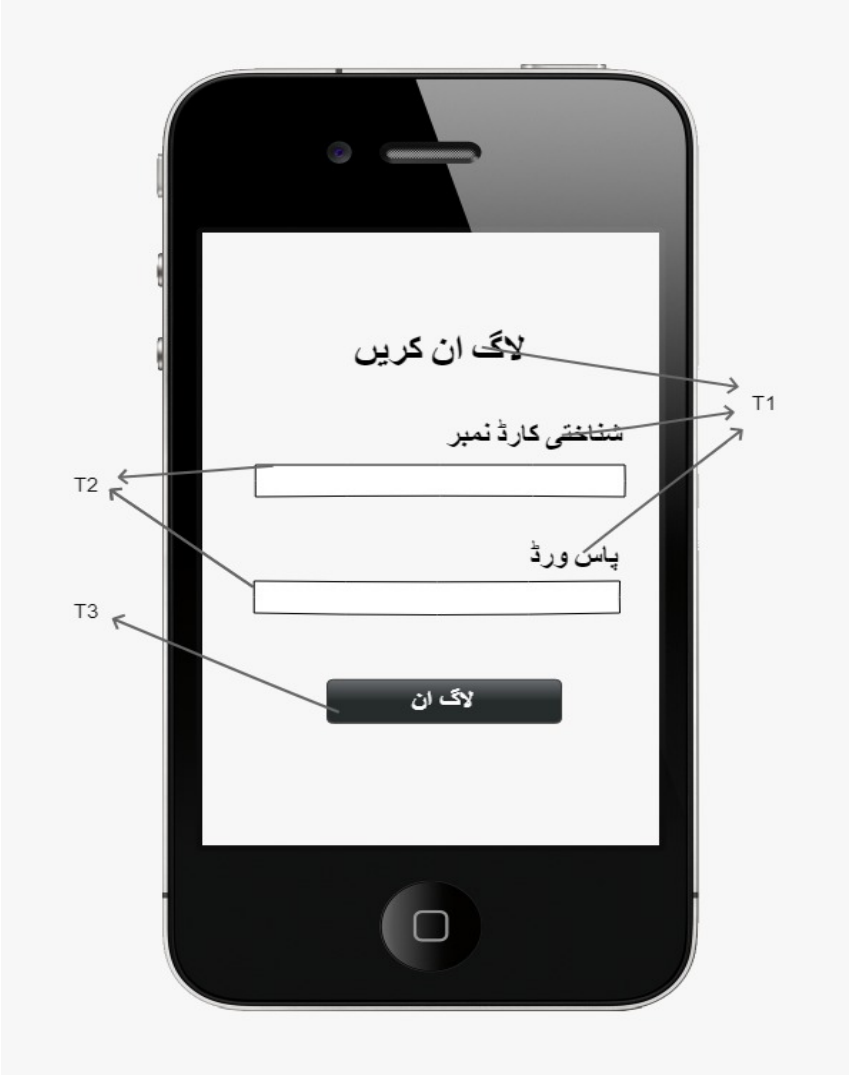

FIGURE 2.2: Login(Data Entry Operator)

TABLE 2.3: Login (Data Entry Operator)

Sr.No	Description
T1	These are the labels
T2	These are the text area to enter data entry operator credentials
T3	It's a button for login

2.5.1.2 Home Screen(Data Entry Operator)

The home screen wireframe for data entry operator is shown in Figure 2.3:

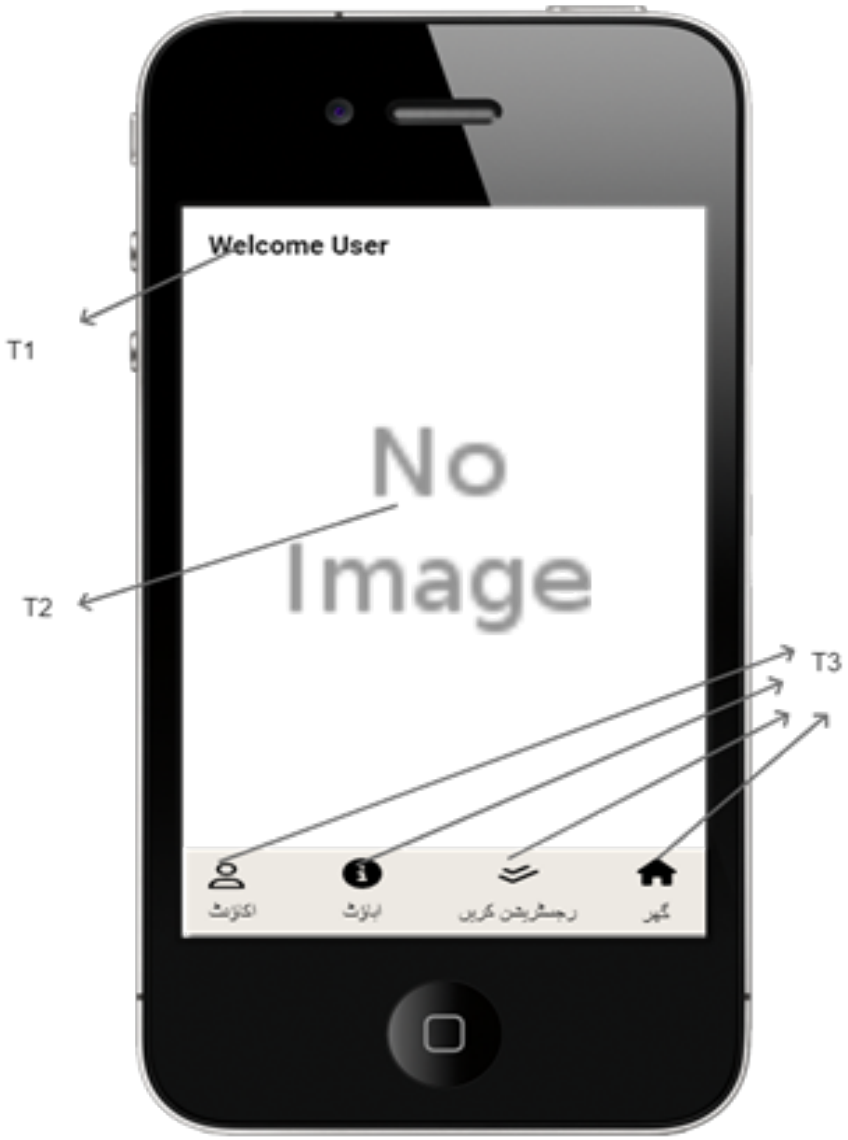

FIGURE 2.3: Home Screen(Data Entry Operator)

TABLE 2.4: Home Screen(Data Entry Opertor)

Sr.No	Description
T1	It is a welcome message for data entry operator
T2	It's the animation of scanning face
T3	These are icons to navigate to other screens

2.5.1.3 Initialize Enrollment

The initialize enrollment wireframe is shown in Figure 2.4:

FIGURE 2.4: Initialize Enrollment

TABLE 2.5: Initialize Enrollment

Sr.No	Description
T1	It is a welcome message for data entry operator
T2	It's a button to start enrollment process
T3	These are icons to navigate to other screens

2.5.1.4 Enrollment Process

The enrollment process wireframe is shown in Figure 2.5:

FIGURE 2.5: Enrollment Process

TABLE 2.6: Enrollment Process

Sr.No	Description
T1	These are the random questions asked by app to pensioners
T2	It's a flip icon to rotate camera
T3	It's a camera open in app
T4	It's recording screen icon
T5	It's a button to end enrollment process
T6	It's a button to get change question in registration process
T7	It's a timer for recording s screen

2.5.2 Pensioner Verification App

The following wireframes depict the verification process, highlighting key screens for login, home, initialization of verification process and face and voice recognition.

2.5.2.1 Login(Pensioner)

The login wireframe for pensioner is shown in Figure 2.6:

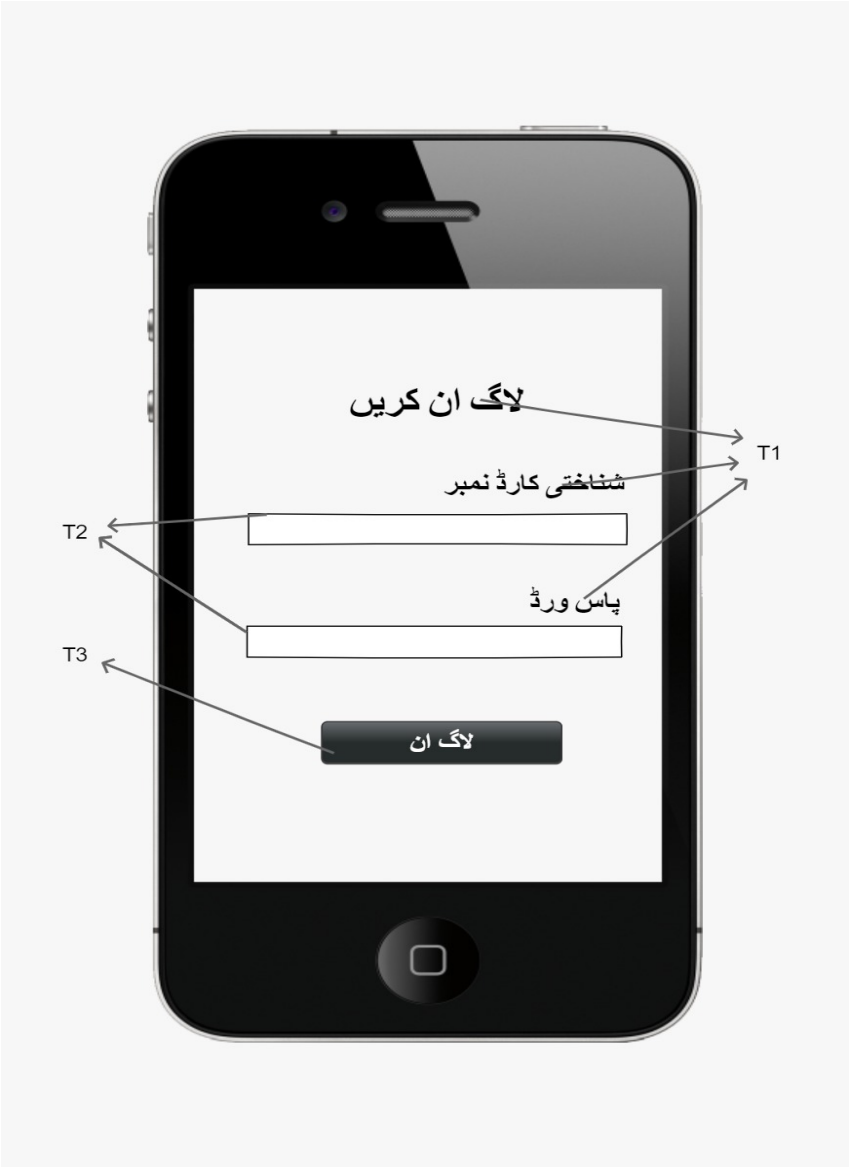

FIGURE 2.6: Login(Pensioner)

TABLE 2.7: Login(Pensioner)

Sr.No	Description
T1	These are the labels
T2	These are the text area to enter pensioner credentials
T3	It's a button for login

2.5.2.2 Home Screen(Pensioner)

The home screen wireframe for pensioner is shown in Figure 2.7:

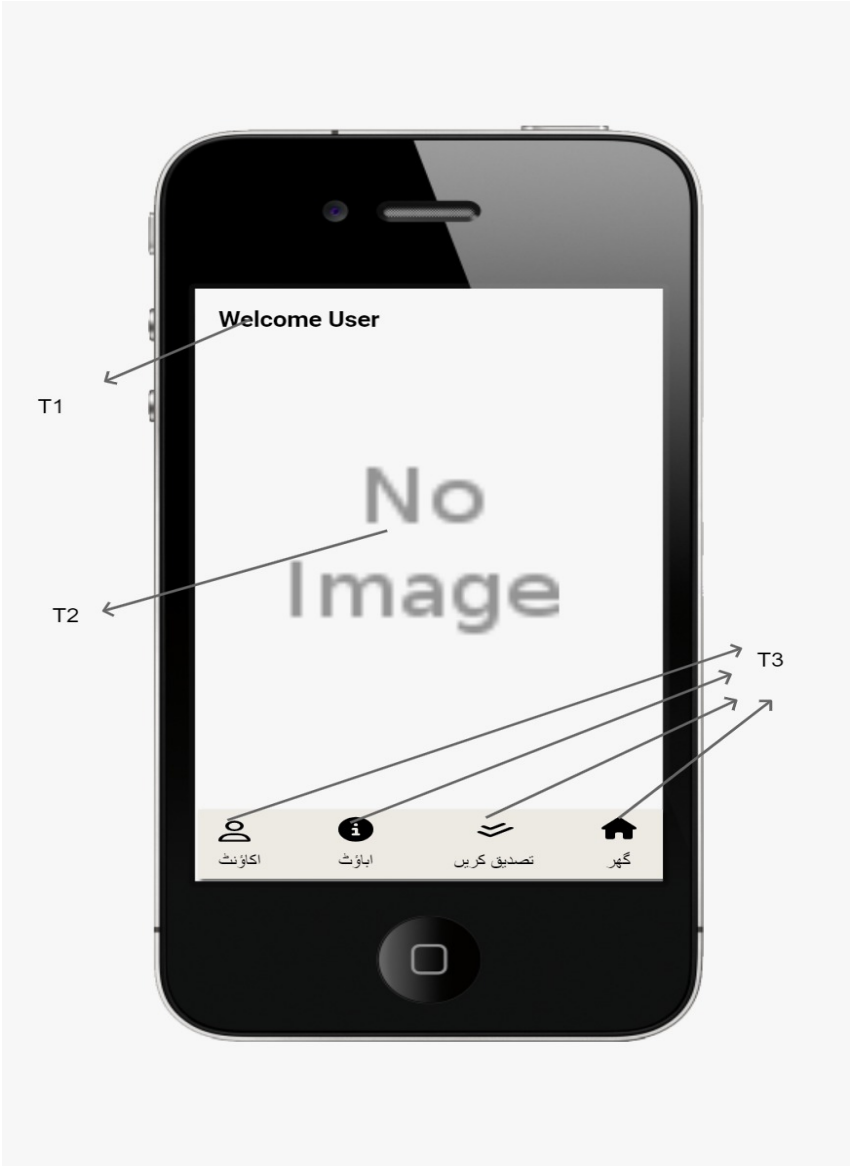

FIGURE 2.7: Home Screen(Pensioner)

TABLE 2.8: Home Screen(Pensoner)

Sr.No	Description
T1	It is a welcome message for pensioner
T2	It's the animation of scanning a face
T3	These are icons to navigate to other screens

2.5.2.3 Initialize Verification:

The initialize verification wireframe is shown in Figure 2.8:

FIGURE 2.8: Initialize Verification

TABLE 2.9: Initialize Verification

Sr.No	Description
T1	It is a welcome message for pensioner
T2	It's a button to start verification process
T3	These are icons to navigate to other screens

2.5.2.4 Verification Process

The verification process wireframe is shown in Figure 2.9:

FIGURE 2.9: Verification Process

TABLE 2.10: Verification Process

Sr.No	Description
T1	These are the random questions asked by app to pensioners
T2	It's a flip icon to rotate camera
T3	It's a camera open in app
T4	It's recording screen icon
T5	It's a button to end verification process

2.6 Use Case Diagram

The use case diagram is shown in Figure 2.10, visualizes the interactions between super admin, data entry operators, pensioners and the system outlining the key functionalities and goals.

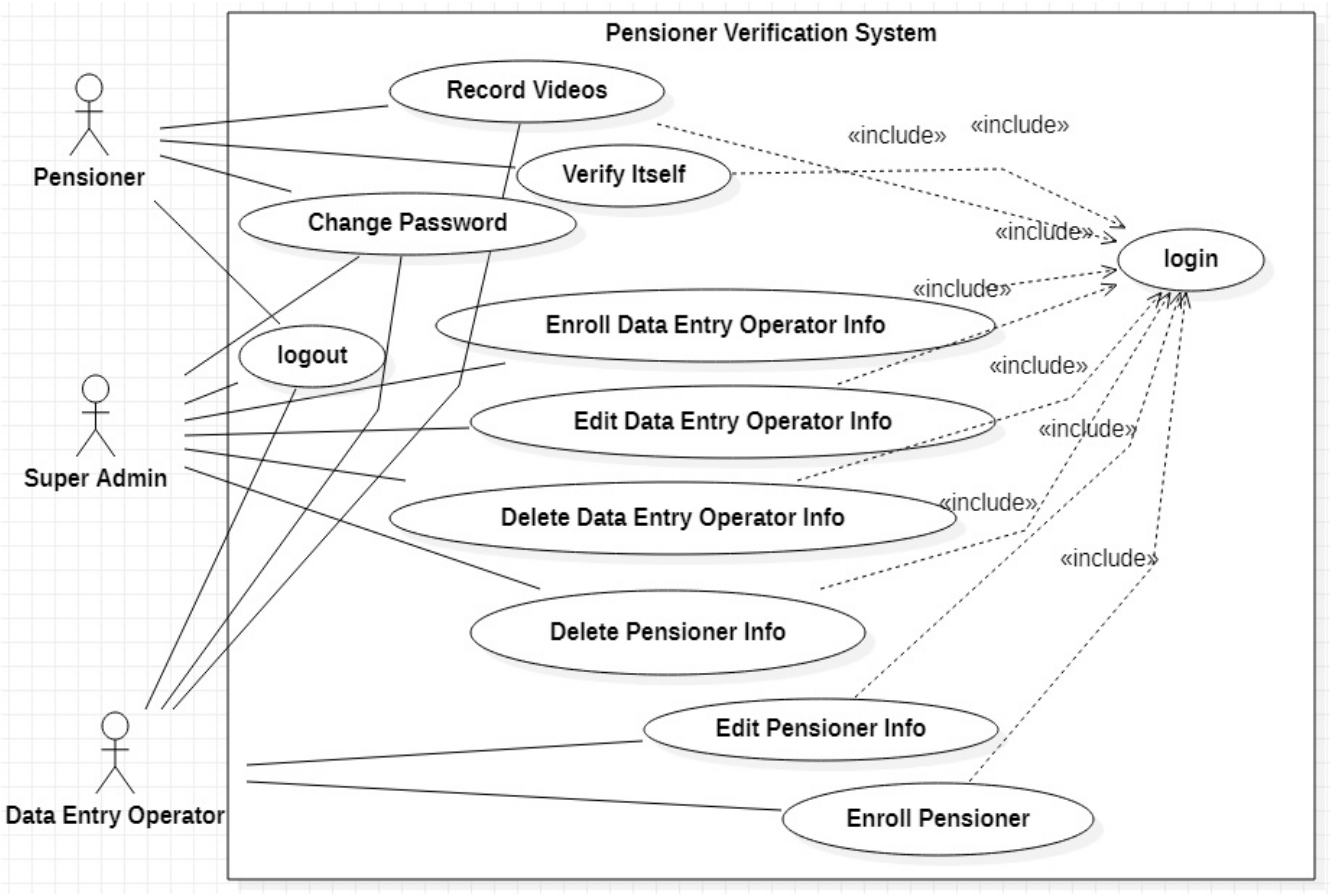

FIGURE 2.10: Use Case Diagram

Table 2.11 outlines the use cases associated with the Pensioner Verification System. Each use case describes a specific interaction that users (pensioners, super admins, and data entry operators) have with the system, detailing their roles and functionalities.

TABLE 2.11: Use Cases

Use Case Name	**Use Case ID**	**Use Case Description**
Login (Pensioner)	UC-1	Pensioner logs into the system (see Table 2.12)
Record Videos	UC-2	Pensioner records verification videos (see Table 2.13)
Verify Itself	UC-3	Pensioner completes identity verification (see Table 2.14)
Logout (Pensioner)	UC-4	Pensioner logs out of the system (see Table 2.15)
Change Password (Pensioner)	UC-5	Pensioner changes their password (see Table 2.16)
Login (Super Admin)	UC-6	Super Admin logs into the system (see Table 2.17)
Enrol Data Entry Operator	UC-7	Super Admin enrols a new data entry operator (see Table 2.18)
Edit Data Entry Operator Info	UC-8	Super Admin edits data entry operator info (see Table 2.19)
Delete Data Entry Operator Info	UC-9	Super Admin deletes data entry operator info (see Table 2.20)
Delete Pensioner Info	UC-10	Super Admin deletes pensioner info (see Table 2.21)
Change Password (Super Admin)	UC-11	Super Admin changes their password (see Table 2.22)
Logout (Super Admin)	UC-12	Super Admin logs out of the system (see Table 2.23)
Login (Data Entry Operator)	UC-13	Data Entry Operator logs into the system (see Table 2.24)
Enrol Pensioner	UC-14	Data Entry Operator enrols a new pensioner (see Table 2.25)

Use Case Name	**Use Case ID**	**Use Case Description**
Edit Pensioner Info	UC-15	Data Entry Operator edits pensioner info (see Table 2.26)
Change Password	UC-16	Data Entry Operator changes their password (see Table 2.27)
Logout (Data Entry Operator)	UC-17	Data Entry Operator logs out of the system (see Table 2.28)

The tables below detail the descriptions and specifications for each use case in the Pensioner Verification System, outlining interactions for pensioners, super admins, and data entry operators.

TABLE 2.12: UC-1: Login(Pensioner)

Use Case ID	UC-1
Name	Login
Primary Actor	Pensioner
Goal	As a pensioner, I want to login to the system.
Trigger	When the pensioner clicks on the login button, a message will display.
Pre-Condition	• Pensioner must have an account. • Pensioner must enter CNIC and password.
Post-Condition	System will navigate to the homepage.
Basic Flow	1. Pensioner enters CNIC. 2. Pensioner enters password. 3. Pensioner clicks on the login button. 4. After successfully logging in, the system shall navigate to the next page.

Alternate Flow	Not Available
Exceptions	• Wrong pensioner details: Pensioner must enter his/her correct information. • Not registered: Pensioner must have an account.
Qualities	• Pensioner logged into the system. • System navigates to the homepage successfully logged in.

TABLE 2.13: UC-2: Record Videos

Use Case ID	UC-2
Name	Record Videos
Primary Actor	Pensioner
Goal	As a pensioner, I want to record videos of my answers asked by the system.
Trigger	When the pensioner clicks on the record icon, the recording will start.
Pre-Condition	• Pensioner must be logged in to the system. • Pensioner must provide required permissions.
Post-Condition	System will record videos of the pensioner.
Basic Flow	1. Pensioner logged in to the system. 2. On the verification page, the pensioner can record videos. 3. After successfully recording, the video will be saved.

Alternate Flow	Not Available
Exceptions	• Camera not working: Pensioner has to start the session again.
Qualities	• Pensioner can record videos. • The videos will be saved.

TABLE 2.14: UC-3: Verify Itself

Use Case ID	UC-3
Name	Verify Itself
Primary Actor	Pensioner
Goal	As a pensioner, I want to verify myself to get the next pension.
Trigger	When the pensioner records all videos, verification starts.
Pre-Condition	• Pensioner must be logged in to the system. • Pensioner must record videos for all questions asked by the system.
Post-Condition	System will save videos of the pensioner for verification.
Basic Flow	1. Pensioner logged in to the system. 2. On the verification page, the pensioner records videos. 3. After successfully recording videos, the video will save and send for verification.
Alternate Flow	Not Available

Exceptions	• Internet issue: Try troubleshooting internet issues.
Qualities	• Pensioner can verify itself.

TABLE 2.15: UC-4: Logout(Pensioner)

Use Case ID	UC-4
Name	Logout
Primary Actor	Pensioner
Goal	As a pensioner, I want to log out from the system.
Trigger	When the pensioner clicks on the logout button, a message will display.
Pre-Condition	• Pensioner must be logged in.
Post-Condition	System will navigate to the login page.
Basic Flow	1. Pensioner logged in to the system. 2. Pensioner clicks on the logout button. 3. System will navigate to the login page.
Alternate Flow	Not Available
Exceptions	• Not signed in: Pensioner must log in to the system.

Qualities	• Pensioner can log out from the system.

TABLE 2.16: UC-5: Change Password(Pensioner)

Use Case ID	UC-5
Name	Change Password
Primary Actor	Pensioner
Goal	As a pensioner, I want to change my password.
Trigger	When the pensioner clicks on the change password button, a message will display.
Pre-Condition	• Pensioner must be logged in to the system.
Post-Condition	Password will be changed.
Basic Flow	1. Pensioner logged in to the system. 2. Pensioner goes to the account section. 3. After that, the pensioner can change the password.
Alternate Flow	Not Available
Exceptions	• Server down: Pensioner must wait for server response.
Qualities	• Password will be changed.

TABLE 2.17: UC-6: Login(Super Admin)

Use Case ID	UC-6
Name	Login
Primary Actor	Super Admin
Goal	As a super admin, I want to log in to the system.
Trigger	When the super admin clicks on the login button, a message will display.
Pre-Condition	• Super Admin must have an account. • Super Admin must enter CNIC and password.
Post-Condition	System will navigate to the homepage.
Basic Flow	1. Super Admin enters CNIC. 2. Super Admin enters password. 3. Super Admin selects role. 4. Super Admin clicks on the login button. 5. After successfully logging in, the system shall navigate to the next page.
Alternate Flow	Not Available
Exceptions	• Wrong super admin details: Super Admin must enter his/her correct information.
Qualities	• Super Admin logged in to the system. • System will navigate to the homepage successfully logged in.

TABLE 2.18: UC-7: Enroll Data Entry Operator

Use Case ID	UC-7
Name	Enrol Data Entry Operator
Primary Actor	Super Admin
Goal	As a super admin, I want to register data entry operators.
Trigger	When the super admin clicks on the register button, a message will display.
Pre-Condition	• Super Admin must be logged in to the system. • Super Admin must enter data entry details (name, CNIC, password).
Post-Condition	A new data entry operator will be registered in the system.
Basic Flow	1. Super Admin logged in to the system. 2. Super Admin enters details for the data entry operator. 3. After that, the new data entry operator will be registered in the system.
Alternate Flow	Not Available
Exceptions	• Wrong credentials for data entry operator: Super Admin must enter correct information.
Qualities	• New data entry operator will be registered in the system.

TABLE 2.19: UC-8: Edit Data Entry Operator Info

Use Case ID	UC-8
Name	Edit Data Entry Operator Info
Primary Actor	Super Admin
Goal	As a super admin, I want to edit data entry operator information.
Trigger	When the super admin clicks on the edit button, a message will display.
Pre-Condition	• Super Admin must be logged in to the system. • Super Admin must edit data entry details (name, CNIC, password).
Post-Condition	Data entry operator details will be updated.
Basic Flow	1. Super Admin logged in to the system. 2. Super Admin edits details for the data entry operator. 3. After that, the data entry operator details will be updated.
Alternate Flow	Not Available
Exceptions	• Wrong credentials for data entry operator: Super Admin must enter correct information.
Qualities	• Data entry operator details will be updated in the system.

TABLE 2.20: UC-9: Delete Data Entry Operator Info

Use Case ID	UC-9
Name	Delete Data Entry Operator Info

Primary Actor	Super Admin
Goal	As a super admin, I want to delete data entry operator information.
Trigger	When the super admin clicks on the delete button, a message will display.
Pre-Condition	• Super Admin must be logged in to the system.
Post-Condition	Data entry operator details will be deleted.
Basic Flow	1. Super Admin logged in to the system. 2. Super Admin views all data entry operators. 3. Super Admin can delete any data entry operator.
Alternate Flow	Not Available
Exceptions	• Server down: Wait for the server response.
Qualities	• Data entry operator details will be deleted in the system.

TABLE 2.21: UC-10: Delete Pensioner Info

Use Case ID	UC-10
Name	Delete Pensioner Info
Primary Actor	Super Admin
Goal	As a super admin, I want to delete pensioner information.
Trigger	When the super admin clicks on the delete button, a message will display.

Pre-Condition	• Super Admin must be logged in to the system.
Post-Condition	Pensioner details will be deleted.
Basic Flow	1. Super Admin logged in to the system. 2. Super Admin views all pensioners. 3. Super Admin can delete any pensioner.
Alternate Flow	Not Available
Exceptions	• Server down: Wait for the server response.
Qualities	• Pensioner details will be deleted from the system.

TABLE 2.22: UC-11: Change Password(Super Admin)

Use Case ID	UC-11
Name	Change Password
Primary Actor	Super Admin
Goal	As a super admin, I want to change my password.
Trigger	When the super admin clicks on the change password button, a message will display.
Pre-Condition	• Super Admin must be logged in to the system.
Post-Condition	Password will be changed.

Basic Flow	1. Super Admin logged in to the system. 2. Super Admin goes to the account section. 3. After that, the super admin can change the password.
Alternate Flow	Not Available
Exceptions	• Server down: Super Admin must wait for server response.
Qualities	• Password will be changed.

TABLE 2.23: UC-12: Logout(Super Admin)

Use Case ID	UC-12
Name	Logout
Primary Actor	Super Admin
Goal	As a super admin, I want to log out from the system.
Trigger	When the super admin clicks on the log out button, a message will display.
Pre-Condition	• Super Admin must be logged in.
Post-Condition	System will navigate to the login page.

Basic Flow	1. Super Admin logged in to the system. 2. Super Admin clicks on the log out button. 3. System will navigate to the login page.
Alternate Flow	Not Available
Exceptions	• Not signed in: Super Admin must log in to the system.
Qualities	• User session will be ended.

TABLE 2.24: UC-13: Login(Data Entry Operator)

Use Case ID	UC-13
Name	Login
Primary Actor	Data Entry Operator
Goal	As a data entry operator I want to login to the system.
Trigger	when data entry operator clicks on log in button a message will display.
Pre-Condition	• Data Entry Operator must have an account. • Data Entry Operator enter cnic and password.
Post-Condition	System will navigate to the homepage.

Basic Flow	1. Data Entry Operator enter cnic. 2. Data Entry Operator enter password. 3. Data Entry Operator selects role. 4. Data Entry Operator clicks on login button. 5. After successfully login in system shall navigate to the next page.
Alternate Flow	Not Available
Exceptions	• Wrong Data Entry Operator details: Data Entry Operator must enter his/her correct information.
Qualities	• Data Entry Operator min logged in to the system. • System will navigate to the homepage successfully logged in to the system.

TABLE 2.25: UC-14: Enroll Pensioner

Use Case ID	UC-13
Name	Enrol Pensioner
Primary Actor	Data Entry Operator
Goal	As a data entry operator, I want to register pensioners.
Trigger	When the data entry operator clicks on the register button, a message will display.

Pre-Condition	• Data entry operator must be logged in to the system. • Data entry operator must enter details (name, CNIC, password).
Post-Condition	New pensioners will be registered in the system.
Basic Flow	1. Data entry operator logged in to the system. 2. Data entry operator enters details for the pensioners. 3. After that, the new pensioners will be registered in the system.
Alternate Flow	Not Available
Exceptions	• Wrong credentials for pensioners: Data entry operator must enter correct information.
Qualities	• New pensioners will be registered in the system.

TABLE 2.26: UC-15: Edit Pensioner Info

Use Case ID	UC-14
Name	Edit Pensioner Info
Primary Actor	Data Entry Operator
Goal	As a data entry operator, I want to edit pensioner information.
Trigger	When the data entry operator clicks on the edit button, a message will display.

Pre-Condition	• Data entry operator must be logged in to the system. • Data entry operator must edit details (name, CNIC, password).
Post-Condition	Pensioner details will be updated.
Basic Flow	1. Data entry operator logged in to the system. 2. Data entry operator edits details for pensioners. 3. After that, pensioner details will be updated.
Alternate Flow	Not Available
Exceptions	• Wrong credentials for pensioners: Data entry operator must enter correct information.
Qualities	• Pensioner details will be updated in the system.

TABLE 2.27: UC-16: Change Password(Data Entry Operator)

Use Case ID	UC-15
Name	Change Password
Primary Actor	Data Entry Operator
Goal	As a data entry operator, I want to change my password.
Trigger	When the data entry operator clicks on the change password button, a message will display.
Pre-Condition	• Data entry operator must be logged in to the system.

Post-Condition	Password will be changed.
Basic Flow	1. Data entry operator logged in to the system. 2. Data entry operator goes to the account section. 3. After that, the data entry operator can change the password.
Alternate Flow	Not Available
Exceptions	• Server down: Data entry operator must wait for server response.
Qualities	• Password will be changed.

TABLE 2.28: UC-17: Logout(Data Entry Operator)

Use Case ID	UC-16
Name	Logout
Primary Actor	Data Entry Operator
Goal	As a data entry operator, I want to log out from the system.
Trigger	When the data entry operator clicks on the log out button, a message will display.
Pre-Condition	• Data Entry Operator must be logged in.
Post-Condition	System will navigate to the login page.

Basic Flow	1. Data Entry Operator logged in to the system. 2. Data Entry Operator clicks on the log out button. 3. System will navigate to the login page.
Alternate Flow	Not Available
Exceptions	• Not signed in: Data Entry Operator must log in to the system.
Qualities	• Data Entry Operator can log out from the system.

2.7 Class Diagram

Figure 2.11 presents a class diagram that include three main modules of the Pensioner Verification System: the mobile app (beige), the face recognition module (light green), and the speaker recognition module (blue). The mobile app manages user interactions and session recordings. The face recognition module handles facial detection, liveness checks, and video processing, while the speaker recognition module processes and verifies audio inputs for voice identification. Together, these modules ensure accurate pensioner verification.

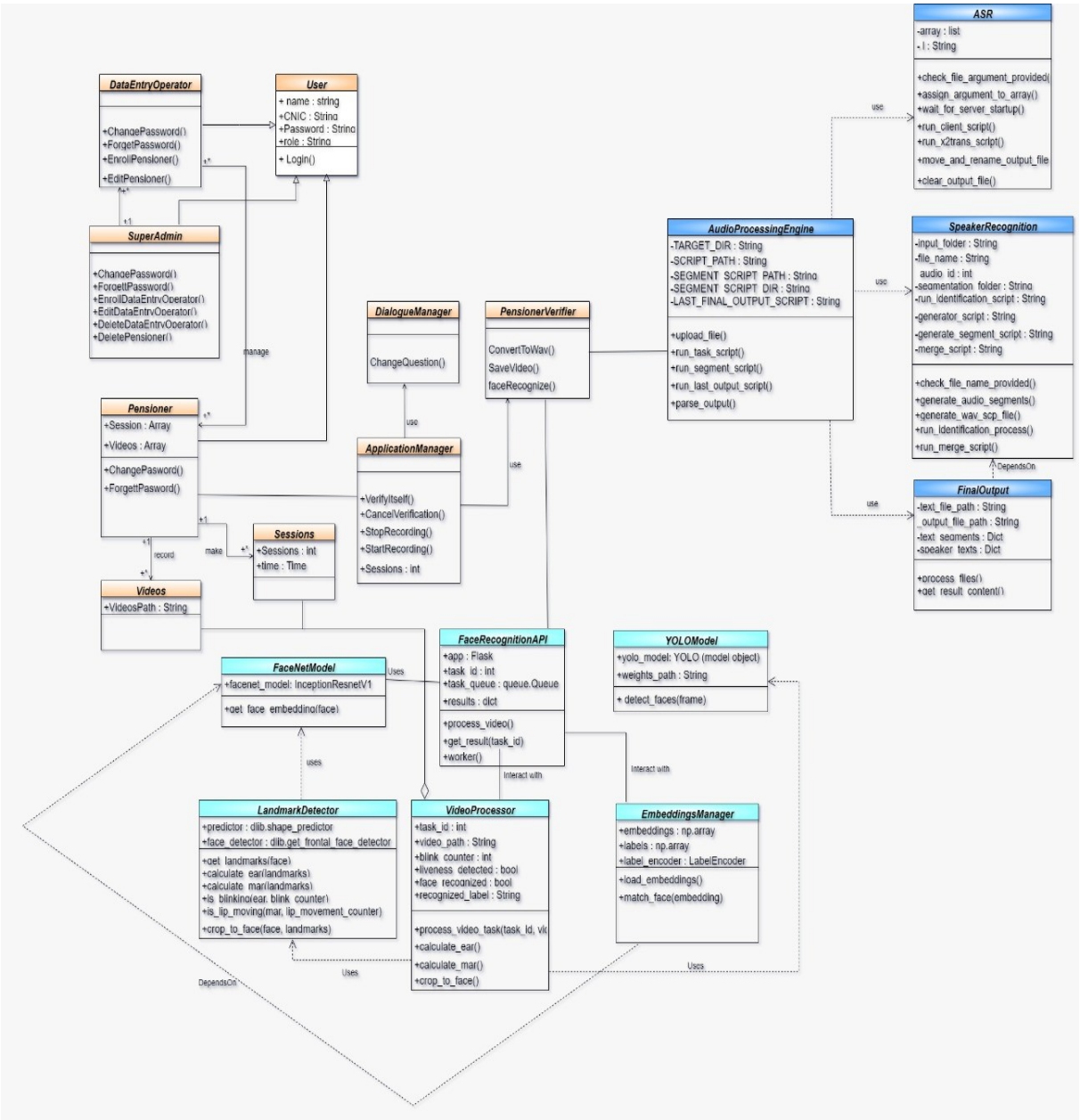

FIGURE 2.11: Class Diagram

2.8 Sequence Diagram

Sequence diagrams shows the flow of interactions between objects in proposed apps over time, showing the sequence of messages exchanged to complete a specific process.

2.8.1 Enrollment Diagram

The sequence diagram of proposed enrollment process is shown in Figure 2.12:

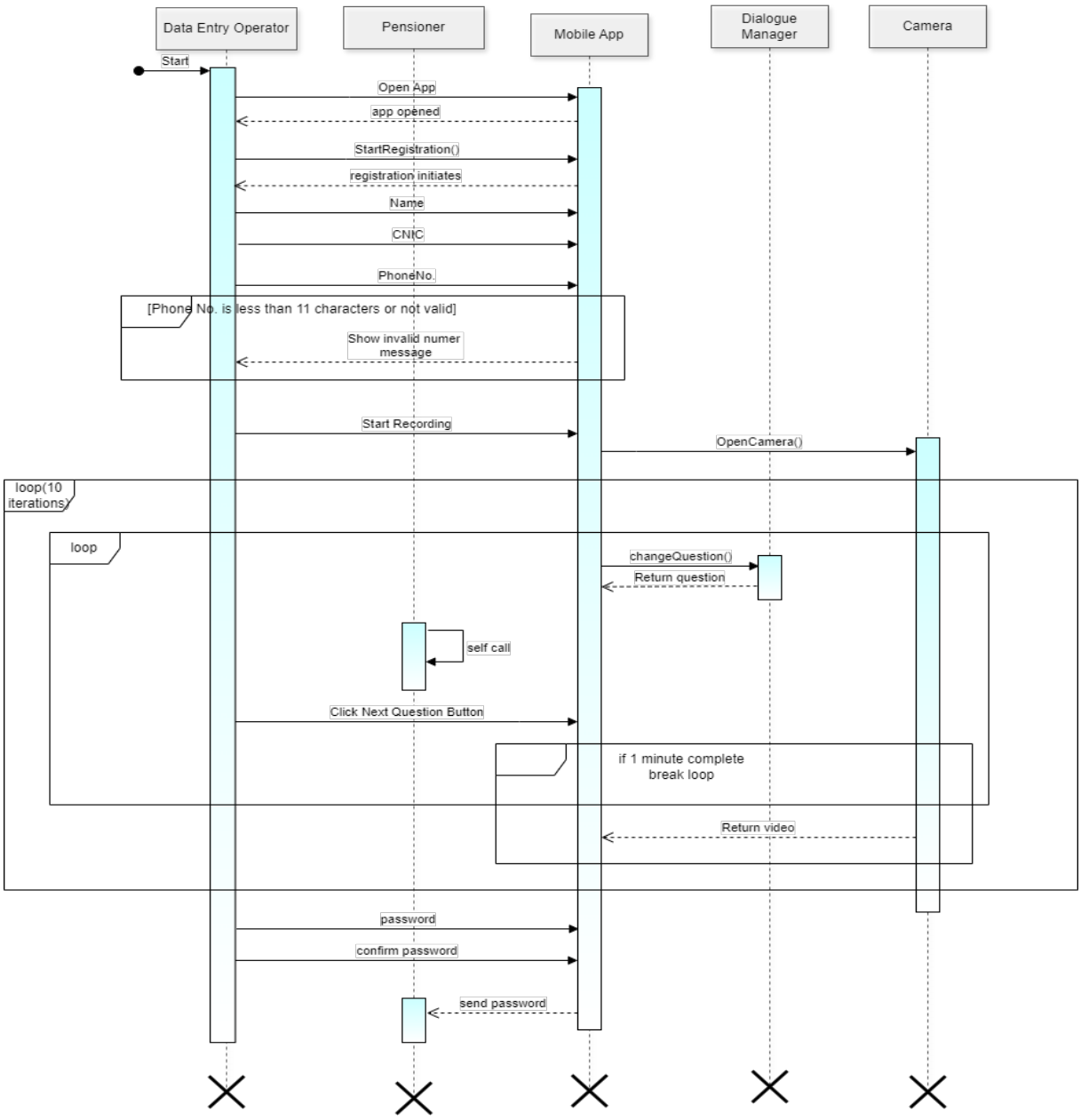

FIGURE 2.12: Enrollment Sequence Diagram

2.8.2 Verification Diagram

The sequence diagram of proposed verification process is shown in Figure 2.13:

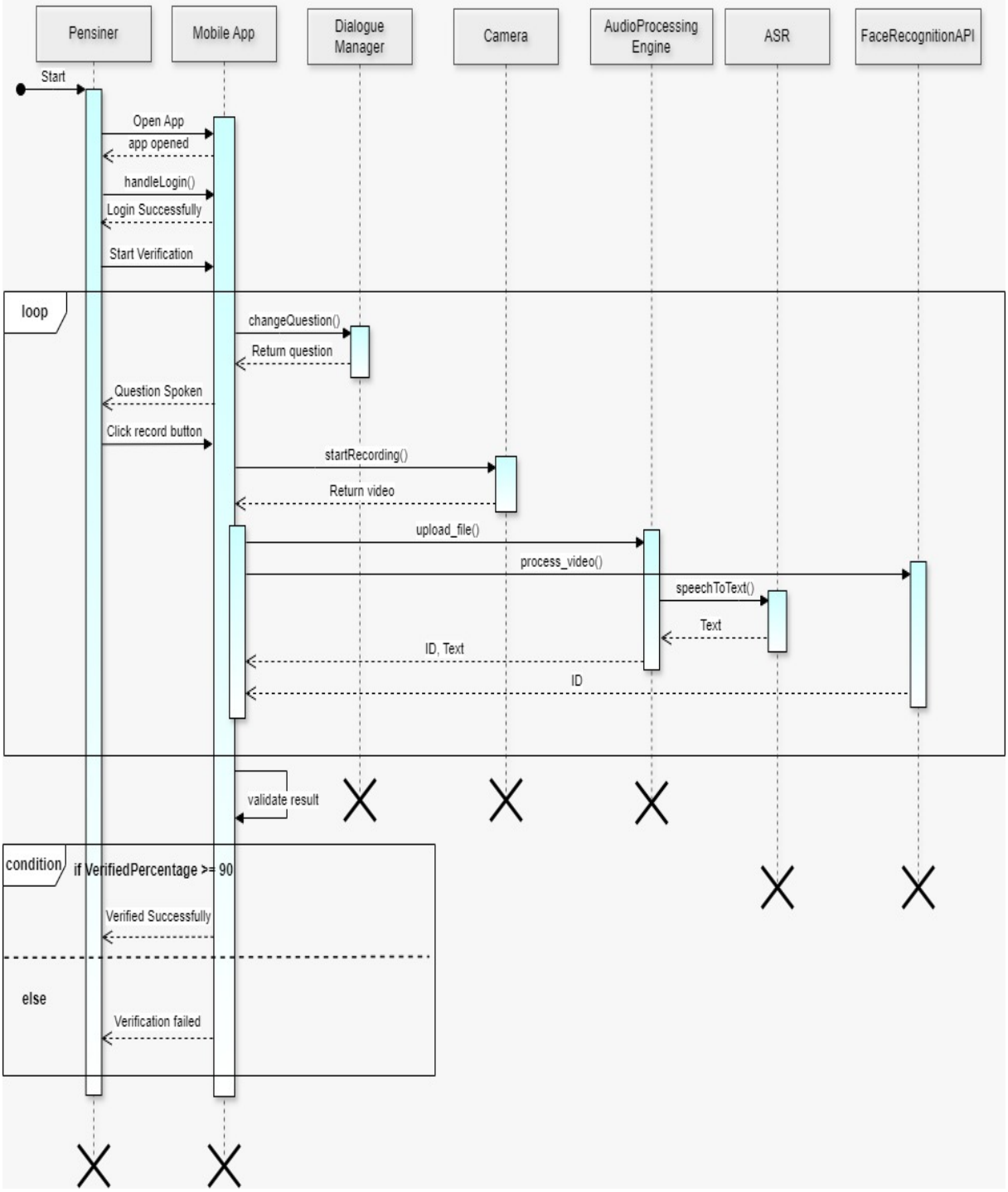

FIGURE 2.13: Verification Sequence Diagram

2.9 Entity Relationship Diagram

The entity-relationship (ER) diagram for the pensioner verification mobile app outlines the relationships between core database tables, including Pensioners, DataEntryOperator, Sessions, Videos, and SuperAdmin. It captures how pensioner data is stored and linked to video sessions for verification purposes, with administrative roles managing the overall process.

Figure 2.14 ensures efficient data organization and retrieval for the verification system.

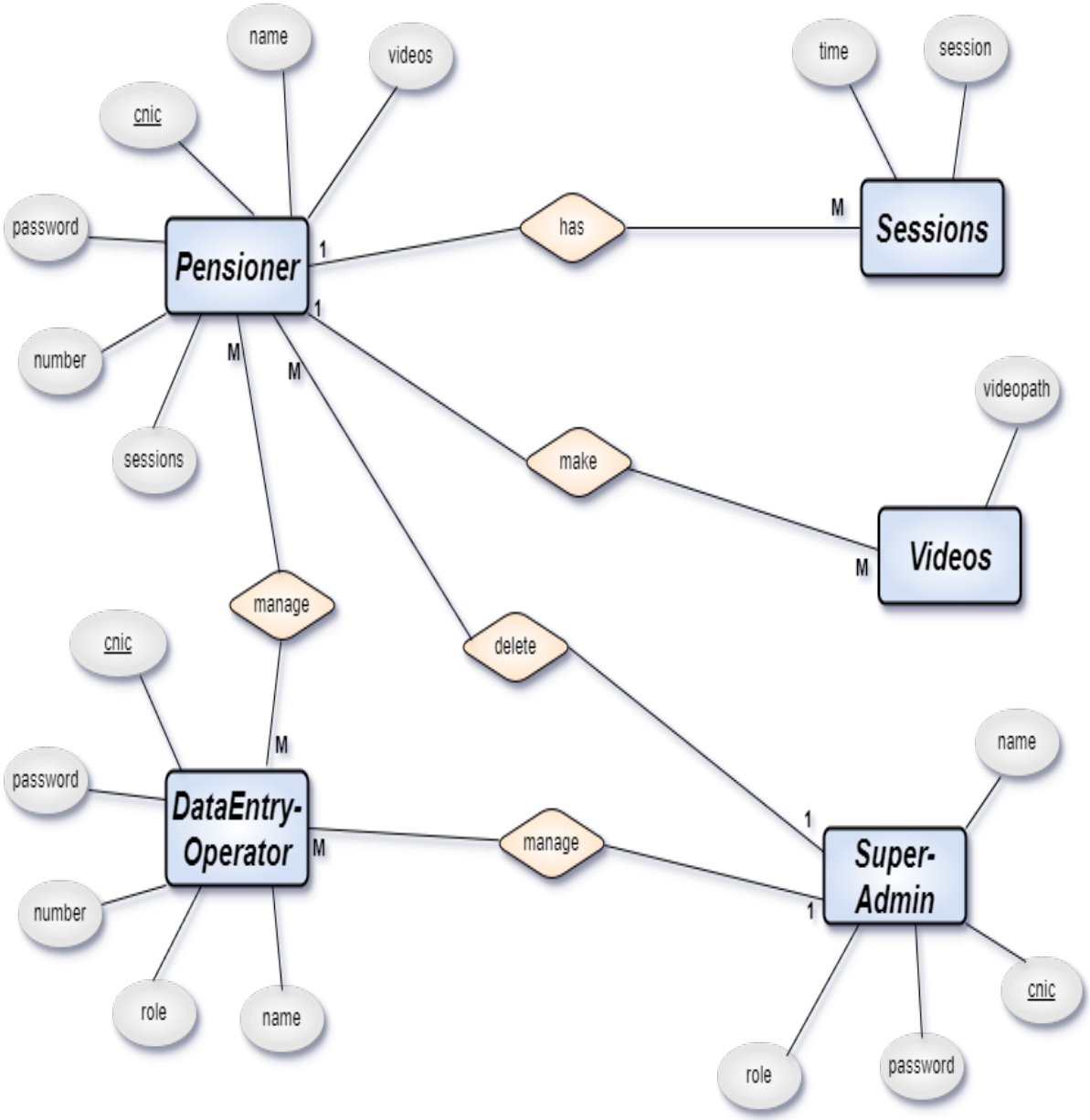

FIGURE 2.14: Entity Relationship Diagram

References

[1] Xinru Kong, Ziyue Wang, Jie Sun, Xianghua Qi, Qianhui Qiu, and Xiao Ding. Facial recognition for disease diagnosis using a deep learning convolutional neural network: a systematic review and meta-analysis. *Postgraduate Medical Journal*, 2024.

[2] Abu Quwsar Ohi and Marina L. Gavrilova. Self-supervised open-set speaker recognition with laguerre–voronoi descriptors. *Sensors*, 24(6):1996, 2024.

[3] Mohamad Alansari, Oussama Abdul Hay, Sajid Javed, Abdulhadi Shoufan, Yahya Zweiri, and Naoufel Werghi. Ghostfacenets: Lightweight face recognition model from cheap operations. *IEEE Access*, 11:35429–35446, 2023.

[4] Shivalila Hangaragi Tripty Singh and Neelima N. Face detection and recognition using face mesh and deep neural network. *Procedia Computer Science*, pages 741–749, 2023.

[5] Navya Saxena and Devina Varshney. Smart home security solutions using facial authentication and speaker recognition through artificial neural networks. *International Journal of Cognitive Computing in Engineering*, 2:154–164, 2021.

[6] Nathalie Tkauc, Thao Tran, Kevin Hernandez-Diaz, and Fernando AlonsoFernandez. Cloud-based face and speech recognition for access control applications. In *2020 IEEE Conference on Communications and Network Security (CNS)*, pages 1–8. IEEE, 2020.

[7] Chin-Shyurng Fahn, Szu-Chieh Chen, Po-Yuan Wu, Tsung-Lan Chu, Cheng-Hung Li, Deng-Quan Hsu, Hsiu-Hung Wang, and Hsiu-Min Tsai. Image and speech recognition technology in the development of an elderly care robot: Practical issues review and improvement strategies. *Healthcare*, 10(11):2252, 2022.

[8] Irene Amerini and Varun G. Menon. D-fence layer: an ensemble framework for comprehensive deepfake detection. *Multimedia Tools and Applications*, pages 1–24, 2024.

[9] KH Teoh, RC Ismail, and SZM Naziri. Face recognition and identification using deep learning approach. *Journal of Physics: Conference Series*, 2021.

[10] Gayathri Girish, Deepak Mishra, and Subrahamanian KS Moosath. Utilising energy function and variational inference training for learning a graph neural network architecture. *Machine Learning*, 113(3):1219–1241, 2024.

[11] Abu Quwsar Ohi, Muhammad Firoz Mridha, Md Abdul Hamid, and Muhammad Mostafa Monowar. Deep speaker recognition: Process, progress, and challenges. *IEEE Access*, 9:89619–89643, 2021.

[12] Fahima Tabassum, Md Imdadul Islam, Risala Tasin Khan, and M. Ruhul Amin. Human face recognition with combination of dwt and machine learning. *Journal of King Saud University-Computer and Information Sciences*, 34(3):546–556, 2022.

[13] Avital Trau-Margalit, Leah Fostick, Tami Harel-Arbeli, and Rachel Nissanholtz-Gannot. Speech recognition in noise task among children and young-adults: a pupillometry study. *Original Research*, 2023.

[14] Daniel S. Park, Yu Zhang, Ye Jia, Wei Han, Chung-Cheng Chiu, Bo Li, Yonghui Wu, and Quoc V. Le. Improved noisy student training for automatic speech recognition. *arXiv preprint*, 2020. arXiv:2005.09629.

[15] David Snyder, Jesús Villalba, Nanxin Chen, Daniel Povey, Gregory Sell, Najim Dehak, and Sanjeev Khudanpur. The jhu speaker recognition system for the voices 2019 challenge. In *INTERSPEECH*, pages 2468–2472, 2019.