COMP2043.GRP Final Group Report: AI-Powered Digital Signage for Targeted and Personalized Advertisement

TEAM2024.06

Yiwei LI (scyyl30, **20513831**)
Zelin XIA (scyzx5, **20513999**)
Xinna SU (scyxs4, **20514175**)
Ang LI (scyal8, **20514282**)
Yuyang ZHANG (scyyz26, **20514470**)

Supervised by Kian Ming Lim

School of Computer Science

University of Nottingham Ningbo China

Date: March 24, 2025

Contents

Chapter 1

Introduction

The AI digital signage system developed by our team this year is an innovative intelligent advertising solution that enhances user engagement through personalized content recommendations. Unlike traditional digital signage that displays static or carousel ads, our system utilizes artificial intelligence (AI) analysis to customize ads based on real-time audience attributes.

The structure of this report is as follows: **Chapter 2** provides a background introduction, modifications to the initial problem, and the existing system. **Chapter 3** introduces the research conducted before development, including analysis of existing advertising signs and market research. **Chapter 4** provides a detailed introduction to requirements engineering, with a focus on the needs of stakeholders. Next is the system design, which is detailed in **Chapter 5**. **Chapters 6** to **8** document the user interface, implementation, and testing methods used. **Chapter 9** discusses the problems encountered during project development and their solutions. The **final chapter** summarizes our work, reflects on it, and provides suggestions for the future, while the appendix includes supporting documents such as test cases, meeting minutes, and other technical details.

Chapter 2

Background

This chapter discusses the background and the difference between the initial system and the current system.

2.1 Initial Statement

This project aims to develop an AI-based digital signage system that can achieve personalized advertising recommendations through computer vision and large language models (LLM). Traditional digital signage ads often rely on preset content for playback and lack accurate audience analysis, resulting in poor advertising delivery results (Greco et al., 2020). The goal of this system is to use AI technology to make advertising delivery more intelligent, make real-time adjustments based on the characteristics and interests of the audience, and improve the relevance and interactivity of advertising.

The system leverages computer vision to detect audience information, including demographic characteristics such as age, gender, and emotions to analyze audience behavior and preferences. Computer vision technology will be used to analyze videos and images to identify the characteristics of live audiences and provide a basis for advertising recommendations. At the same time, LLM will automatically generate personalized advertising copy based on audience data, making advertising content more in line with audience interests and increasing user engagement. In terms of advertising delivery, the system will select the most suitable advertisements from the preset advertising library based on the

analysis results and play them in real time on the digital signage. At the same time, the system will provide a user-friendly management interface to make advertising delivery management and data monitoring more convenient.

2.2 Current System

While a degree of personalized advertising has been achieved in current AI-powered digital signage products, there are still some unresolved challenges. The main existing system issues identified after discussion with stakeholders are as follows.

- **Limited personalization** Few features of facial analysis technology lead to less accurate results in personalized pushes.
- **Ineffective or no advertisement text** Lack of personalized and attractive advertising messages due to the requirement of a large investment.
- **Real-Time performance tracking** Most digital signage solutions do not track user engagement in real-time. They lack the ability to analyze factors such as viewing duration, user gaze direction, or interaction feedback, making it difficult to optimize content based on audience behavior.
- **Lack of analysis of data** There is a lack of backend data for administrators and advertisers to analyze ad completion rates.
- **Data privacy concerns** Consumer privacy could be violated due to facial recognition and data analytics, leading to the possibility that some users concerned about privacy may refuse to use (Phelps et al., 2001).

These limitations result in low engagement rates, making it harder for advertisers to connect with their target audiences and measure the effectiveness of their campaigns.

2.3 Proposed Solution

The proposed digital signage system aims to address the identified challenges by meeting the specific requirements of the stakeholders. The key objectives are as follows.

- **Enhancing Personalization:** Improve the accuracy of personalized advertisements using advanced computer vision and AI technologies. In addition to age and gender, the user's ethnicity is also detected to refine the classification of the population. This ensures that advertisements are effectively tailored to diverse audience groups, enhancing relevance and engagement.

- **Create effective advertisement messages** Instead of using static ad messages, the system leverages large language models (LLM) to generate personalized and context-sensitive advertisements. The LLM adapts the message based on the detected demographics, emotions, and engagement level, making advertisements more compelling and relevant.

- **Eye-Tracking Model for Real-Time Engagement Analysis** Integrate an eye-tracking model to analyze user viewing time during the advertisement playing. This allows to measure how long users engage with each advertisement, providing valuable insights into ad effectiveness.

- **Comprehensive Performance Dashboard** A dashboard is implemented to track and analyze viewing duration and ad completion rates intuitively. The system calculates various completion rate metrics to enable advertisers to adjust content strategies for higher engagement and better conversion rates.

- **Ensuring Data Privacy** The system employs real-time facial recognition and analysis without storing or transmitting identifiable user data. This is to ensure the compliance with privacy regulations and addresses stakeholder concerns about data security.

These objectives are designed to meet the needs of all stakeholders, including advertisers and end-users. By aligning the system's features with these requirements, the project aims to deliver a reliable, efficient, and user-friendly AI-powered digital signage solution.

Chapter 3

Literature Research

This chapter provides a comprehensive examination of the theoretical and technical foundations relevant to this project, evaluating the effectiveness of AI-powered digital signage.

3.1 Background Information

With the advent of digital technology, the traditional outdoor advertising market is facing a new transformation (Taylor, 2010; Truong et al., 2010; Taylor, 2015). The media landscape is changing rapidly due to digitalization, blurring the boundaries and fields of different advertising media, thereby accelerating the application and development of digital convergence in digital outdoor advertising (Choi et al., 2012; Franke and Taylor, 2017; Taylor and Sarkees, 2016). According to Ranvnik and Solina, digital signage as a new advertising platform, has received much attention from advertisers recently Ravnik and Solina (2013). In the latest report from a global market research company, the global digital signage market is expected to reach **$3.23** billion by 2020, up from **$1.53** billion in 2014. With this rapid growth, digital signage has become the "fourth screen" in the advertising media environment after TV, Internet, and mobile devices (Kim and Shim, 2015).

Digital signage is fundamentally an autonomous digital display network designed to deliver targeted audiovisual content (Hong, 2014). Modern digital signage is mainly for public or internal information display, advertising, branding, and improving customer experience (Krumm, 2011; Lundström, 2008; Schaeffler, 2008; Müller et al., 2011). Digital

signage has changed the way advertisements are delivered. Digital signage displays have the advantage over static signs because they can display the multimedia content such as images, animations, video and audio. The content can be adapted in real time to a different context and audience (Bauer and Spiekermann, 2011).
From a stakeholder perspective, digital signage provides substantial benefits:

- **For consumers** Digital signage enhances viewer engagement by tailoring advertisements to individual characteristics and preferences, making content more appealing and relevant (Greco et al., 2020; van Reijmersdal et al., 2017).

- **For retailers** In-store digital signage can increase customer traffic and sales(Burke, 2006, 2009). Moreover, digital signage can reduce expected execution time and costs through personalized advertising(van Reijmersdal et al., 2017).

- **For malls** The shopping environment, emotions and approach behaviors can be effectively improved (Dennis et al., 2010). Digital signage enhances the shopping environment, influences consumer emotions, and promotes positive approach behaviors. Additionally, it improves the overall image of shopping malls, fostering a more favorable shopping atmosphere (Dennis et al., 2010; Newman et al., 2010).

Despite these advantages, existing digital signage solutions still face several limitations, such as a lack of personalization, limited interactivity, and challenges in measuring user engagement. A comparison of current digital signage technologies (see Table 3.1) provides a comparative analysis of current digital signage technologies, identifying areas for further innovation and improvement.

Table 3.1: Comparison of Digital Signage Solutions and the Proposed System

Feature	Static Billboards	Basic Digital Signage	AI-driven Digital Signage	Proposed Solution
Adaptability	✗	✗	✓	✓
Personalization	✗	✗	✓	✓
Ads Prompt Generation	✗	✗	✗	✓
Advertisement Completion Guarantee	✓	✓	✗	✓
Privacy Protection	✓	✓	✗	✓
User Feedback Integration	✗	✗	✗	✓

3.2 Technical Research

The AI-driven digital signage system integrates multiple datasets and machine learning models to construct an intelligent, interactive advertising platform. This project uses **UTKFace** and **FER2013** datasets, combined with **YOLOv8**, **Llama3.2**, **ResNet50** and

shape_predictor_68_face_landmarks.dat models, each component performs specialized functions, collaboratively contributing to the system's effectiveness.

In the workflow, **YOLOv8** first quickly detects and accurately locates the person, and circles the approximate position of the face, providing a basis for subsequent processing. Next, **ResNet50**, trained on the **UTKFace** and **FER2013** datasets, performs age, gender, and facial expression recognition, extracting nuanced user characteristics. Based on these insights, **Llama3.2** generates personalized advertising content tailored to the user's profile, which is then displayed in real time. Concurrently, the

shape_predictor_68_face_landmarks.dattracks the user gaze duration, and transmit it to the dashboard for analytics.

This synergistic combination of models enables a highly adaptive and personalized digital signage solution, significantly enhancing user experience and showcasing the innovation and efficiency of the proposed system.

3.2.1 Dataset

- UTKFace

  The UTKFace dataset is a comprehensive facial dataset encompassing a wide age range, spanning from 0 to 116 years. It comprises over 20,000 face images, each annotated with age, gender, and ethnicity. The dataset exhibits significant diversity in terms of pose, facial expressions, lighting conditions, occlusions, and image resolution, making it valuable for various facial analysis tasks (Zhang et al., 2017).

- Fer2013

  The FER2013 dataset contains 48×48 pixel grayscale facial images, where each face has been automatically aligned to ensure consistent position and proportion. It is designed for facial expression recognition and classifies faces into seven emotion categories (Lagwankar, 2021). However, in our project, the system will focus on detecting four emotions: neutral, happy, sad, and angry.

3.2.2 Model

- YOLOv8

  YOLOv8's architecture is optimized for real-time object detection, making it ideal for applications that require fast and accurate object identification in images (Alshammari and Alshammari, 2024).

- Llama3.2

  The Llama 3.2 multilingual large language model (LLM) suite consists of pretrained and instruction-tuned generative models, designed for text input and output (Nguyen et al., 2024).

- ResNet50

  ResNet50 is a powerful image classification model that can be trained on large datasets and achieve state-of-the-art results.

- shape_predictor_68_face_landmarks.dat

  It contains the trained weights and architecture of a deep learning model designed

to predict 68 specific facial landmarks (Saad et al., 2024).

3.3 Market Analysis

There are some digital signage solutions in the market, all capable of dynamically displaying the advertisements videos. However, their solutions have significant limitations. Two representative products are compared, **BrightSign** and **NEC Analytics Learning Platform(ALP)**, to highlight their strengths and weaknesses.

- **BrightSign**

  The product is powered by their own system BrightOS, which would provide the signage a suitable environment display the advertisement videos. Based on the BrightOS system, it may create a screen wall and make a multiple-screen advertisement. According to the data showing on BrightSign official website, over 2.5 million digital signage were sold and more than 130 countries installed BrightOS. However, it still limited in time audience analysis and personalized content delivery.

- **NEC ALP**

  NEC ALP is an AI-Driven digital signage that develops AI-based analytics to automatically create recommendations for target customers based on their age and gender. However, it cannot generate highly personalized advertisement messages based on user characteristics. Additionally, has mentioned legal frameworks differ across regions, raising concerns that facial recognition technology may violate user privacy.

Chapter 4

Requirements Engineering

Requirements engineering is the systematic process of eliciting, analyzing, specifying, validating, and managing the requirements of a system to ensure it aligns with stakeholder needs and project objectives (Macaulay, 2012). This chapter delineates the requirements engineering process for the AI-powered digital signage system, ensuring a rigorous and methodical approach to system development.

4.1 Requirements Elicitation

Requirements elicitation serves as the foundational step in understanding stakeholder expectations and ensuring the project aligns with academic and ethical standards. The following methods were employed:

- **Academic Literature Review**

  Relevant literature was reviewed to explore methodologies for delivering targeted and personalized advertisements. This analysis provided insights into best practices, ethical concerns, and emerging trends in AI-driven digital advertising systems.

- **Stakeholder Meetings**

  Structured meetings were conducted with the supervisor to validate initial requirements and facilitate iterative updates based on feedback. This ensured alignment with stakeholder expectations and project objectives.

  Comparative Analysis Existing AI-driven digital signage solutions were exam-

ined to identify best practices, usability considerations, and potential limitations, ensuring that the proposed system incorporates state-of-the-art features.

4.2 Requirements Specification

The system requirements are categorized into functional and non-functional requirements. Functional requirements define the system's core capabilities in addressing user needs, while non-functional requirements delineate performance constraints and operational standards.

4.2.1 Functional Requirements

1. **User**
    - Engage with the digital signage to view personalized advertisement videos.
    - Receive real-time demographic and emotional state analysis based on system detection.
    - The system will track user engagement through eye-tracking models to measure user engagement through advertisement watch time and completion rate.
2. **System Administrator**
    - Access a marketing dashboard to monitor advertising performance providing key indicators, including viewer count, watch duration, and advertisement completion rate.

4.2.2 Non-Functional Requirements

1. **Computer Vision (CV) Module**
    - Detect and capture facial features of approaching users.
    - Classify the demographic attributes (age, gender, ethnicity).
    - Analyze the user's emotional state (happy, angry, sad, neutral).
2. **LLM Module**

- Utilize outputs from the CV module to formulate tailored advertising prompts.
- Generate personalized advertising messages using an integrated LLM.

3. **Advertisement(AD) Pool**

- Maintain a database for storing the advertisement videos.
- Retain metadata, including corresponding advertisement briefs.

4. **Analytical Module**

- Employ an eye-tracking model to detect viewer engagement, including watch time and advertisement completion rate.
- Store and retrieve formatted feedback data from the user.
- Provide a dashboard for system administrators to monitor marketing effectiveness.

5. **Privacy Protection**

- In compliance with GDPR guidelines (ITGP, PTIP, 2017), all captured facial images will not be saved following analysis to ensure data protection and user privacy.

6. **Performance**

- Process user data and deliver ads within 10 seconds to ensure timeliness.

7. **User Interface**

- Display personalized advertisement messages at the bottom of the advertisement video.
- Present demographic analysis messages in the secondary screen.

8. **Maintainability**

- The codebase should meet key software engineering development standards for later maintenance.

4.3 Requirements Validation

Throughout the project lifecycle, iterative refinements were made based on stakeholder feedback. The following modifications were incorporated to enhance system efficacy and adherence to academic and ethical considerations:

1. Equipment purchase requests were removed to optimize project expenses and focus on core features.

2. System usage by minors under the age of 17 was prohibited to uphold ethical standards.

3. Vocal and touch interaction features were omitted, as per supervisor recommendations during the project bidding stage.

4. The prerequisite for user agreement to ethical terms before face detection was removed to simplify user interaction.

5. The CV Module initially expected 10 face images input per person and output analysis result with the highest probability. Changed CV Module to expect only 1 image input to simplify workflow.

6. Face detection was paused during advertisement playback to reduce computational overhead.

7. Replaced QR code-based feedback collection with an eye-tracking model for engagement assessment. After user testing, few users have the willingness to take the initiative to scan the code and fill in a long list of suggestions for this product.

8. Simplified the matching module for improved efficiency.

9. Implemented a structured mechanism for tracking requirement changes and their impact on project objectives.

These modifications were validated against project constraints, feasibility, and ethical considerations, ensuring a rigorous and transparent requirements engineering process.

Chapter 5

Design

5.1 Pseudocode

Pseudocode is a tool for designing program logic and computational algorithms that limits the programmer's flexibility by specifying the algorithm in what is essentially a high-level language(Davis, 2019). Click here to visit the Pseudocode in Appendix.

5.2 Diagram Design

5.2.1 Use Case Diagram

Use case diagrams are utilized to capture the requirements of a system, focusing primarily on the design aspects of the system. During the analysis of a system to identify its functionalities, use cases are developed and the corresponding actors are determined Mule et al. (2015).

For the proposed AI-powered digital signage system, the use case diagram highlights its main functions and the interactions between two primary actors: users and system administrators.

Key Information

- This project introduces an AI-Powered Digital Signage System for targeted and personalized advertisements. Users can interact with the system by viewing personalized advertisements and provide feedback to improve and personalization.

- The system uses the CV Module to analyze real-time demographics and emotion data. In addition, LLM Module is employed generate targeted messages to for personalized advertisements.

- The system does not store user facial data in any database, ensuring full compliance with privacy regulations.

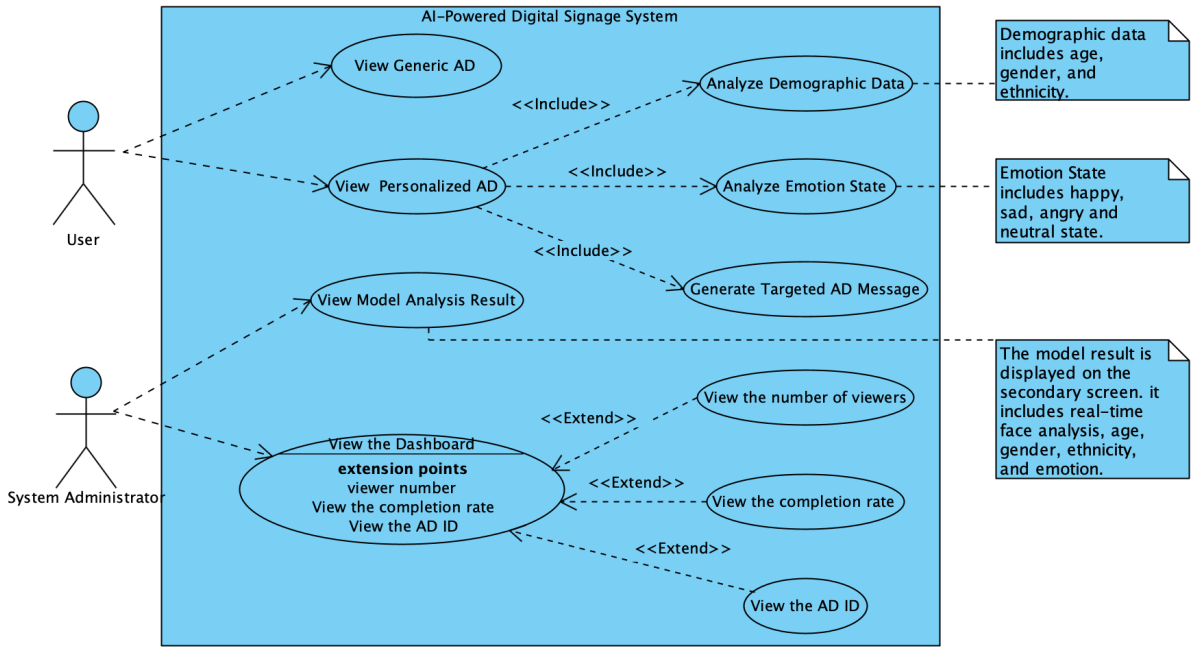


Figure 5.1: Use case Diagram

Description

Actors:

- User: The primary users who interact with AI-powered digital signage will view generic ads and personalized ads.

- System Administrator: Technical users are responsible for monitoring and managing the system in the dashboard, including tracking viewer numbers, ad id and analyzing completion rates.

Use Cases:

1. **View Generic AD:** When no faces are detected or when faces are detected but LLM has not yet finished generating text, the users will view generic ads. In both cases, the default ad rotation continues to ensure uninterrupted content display.

2. **View Personalized AD:** The user stands in front of the AI-powered digital signage, where their demographic data are captured and analyzed for demographic information (age, gender, and ethnicity), along with their emotional state (happy, sad, angry, or neutral). Based on this analysis, the system generates a targeted advertising message that is tailored to the user's demographics and current emotions. The user then views the personalized advertisement displayed on the screen, enhancing engagement and relevance.

3. **View Model Analysis Result:** The system displays facial demographic information on the secondary screen, such as age, gender, and ethnicity, as well as emotions. This information is visible to the system administrator as part of the interaction.

4. **View the Dashboard:** The System Administrator can navigate to the dashboard, where they can view key metrics such as the number of viewers, ad id and ad completion rates. The dashboard provides real-time information on user engagement, allowing the administrator to monitor system performance. By analyzing viewer data and completion rates, the administrator can optimize content and system functionality.

5.2.2 Sequence Diagram

Sequence diagrams define interactions between objects in sequential order, bridging use cases and system design. They help developers to detail object behavior, help business staff understand workflows, and refine requirements into formal designs Unhelkar (2005).

View Generic AD

Key Information

- This sequence diagram focuses on how the system determines whether to play a generic ad or a personalized ad based on the YOLO confidence score and the LLM text generation state.

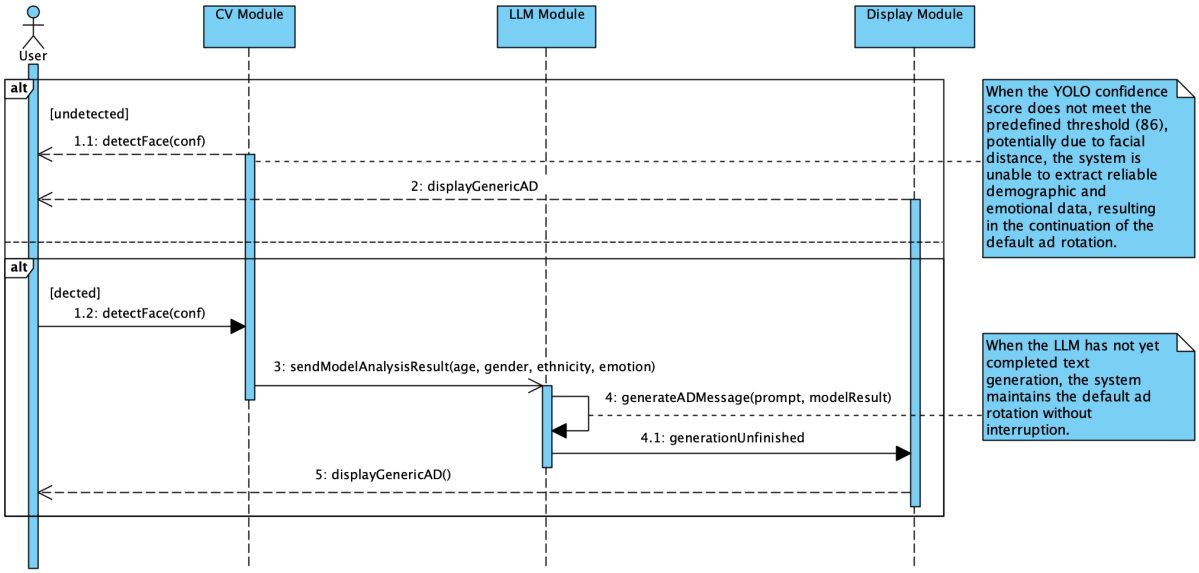


Figure 5.2: Sequence Diagram 1: View Generic AD

Description

- **CV (Computer Vision) Module:** The CV module is responsible for analyzing facial features of the face, such as age, gender, ethnicity, and emotion. The CV module sends the analysis results to the LLM module and the matching module. It does not participate directly in ad generation or matching logic.
- **LLM (Large Language) Module:** The LLM module generates personalized advertising messages based on the model analysis results.
- **Display Module:** This module is responsible for displaying ads to users, displaying a generic ad when no face is detected or when the LLM has not yet completed generating text. Once demographic data and emotion analysis are successfully processed, a personalized ad is displayed.

View Personalized AD

Key Information

- The sequence diagram shows the functions of the system through the collaboration between modules. The loosely coupled design method is used to enable each module to work independently and interact through clear data for efficient integration.

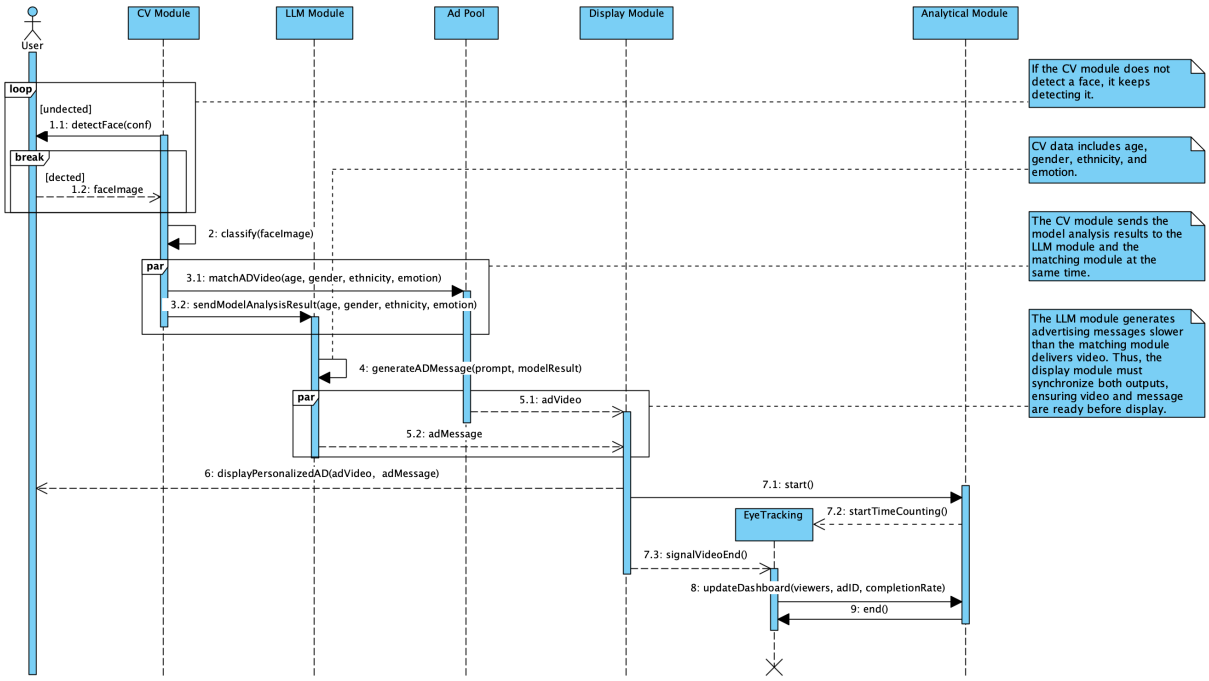


Figure 5.3: Sequence Diagram 2: View Personalized AD

Description

- **CV (Computer Vision) Module:** The CV module is responsible for detecting faces and analyzing facial features such as age, gender, ethnicity, and emotion. If no face is detected, the module continues scanning until a face is found. Once a face is detected, the module sends the analysis results simultaneously to the LLM Module. It does not directly participate in ad generation or matching logic but provides crucial user demographic and emotional data for personalization.

- **LLM (Large Language) Module:** The LLM module generates personalized advertising messages based on demographic data and emotion analysis provided by

the Computer Vision (CV) module. However, due to the LLM's relatively slow text generation process, synchronization is required to ensure the seamless display of the final personalized advertisement.

- **Display Module:** The Display Module is responsible for displaying advertisements to the user. It ensures that both the matched ad video and the generated ad message are synchronized before display. If no face is detected or the LLM has not yet completed text generation, the system remains in the ad rotating state, continuously displaying a predefined sequence of advertisements from the ad pool to ensure uninterrupted screen utilization.

- **Analytical Module:** The Analytical Module collects feedback from users and updates the system with viewer engagement data. It records key metrics such as viewer count and ad completion rate. This module integrates eye-tracking data to analyze user attention and engagement levels. The data collected are processed into a dashboard that provides insights into the effectiveness of the advertisement and user behavior.

View the Dashboard

Key Information

- This sequence diagram illustrates the content of the dashboard in the use case diagram.

Description

- **Authorization:** The system administrator logs in by using a username and password, and the process is allowed to proceed upon successful authentication.

- **Analytical Module:** After successful authentication, the Analytical Module processes dashboard requests by retrieving advertising-related data from the Database, which systematically stores key performance metrics, including ad ID, viewer count,

and ad completion rate. Leveraging this data, the Analytical Module conducts in-depth analysis to derive comprehensive insights into ad performance and audience engagement.

- **Dashboard:** The analytical module responds with dashboard data, which includes additional optional metrics such as the number of viewers ad id and ad completion rate.

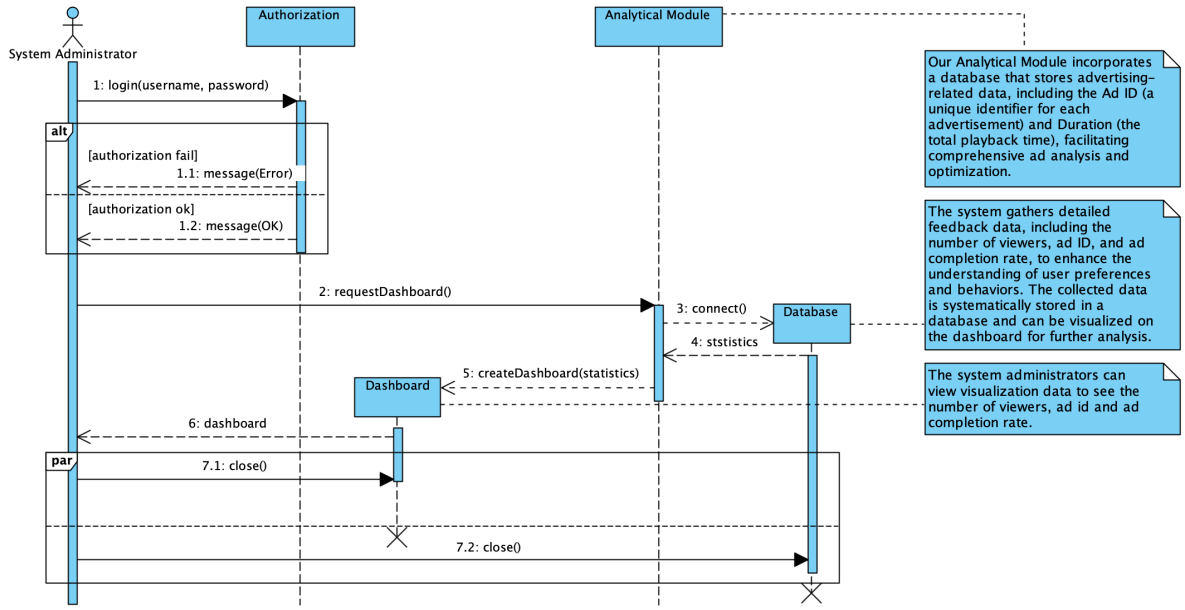


Figure 5.4: Sequence Diagram 3: View the Dashboard

View Model Analysis Result

Key Information

- The sequence diagram shows the process of the system administrator viewing model analysis results.

Description

- **Display Module:** The system administrator requests the display module to view the model analysis result.

- **CV Module:** The display module requests the CV module to view the analysis result. The CV module displays the analysis result to display module, including the face image, age, gender, ethnicity, and emotion.

- **Display Module:** The display module displays the analysis results from the CV module, including age, gender, race, and emotion, to the system administrator.

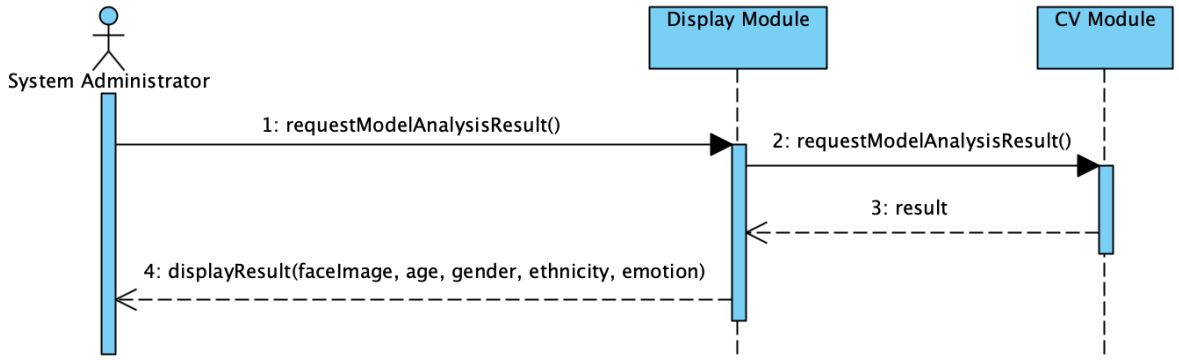
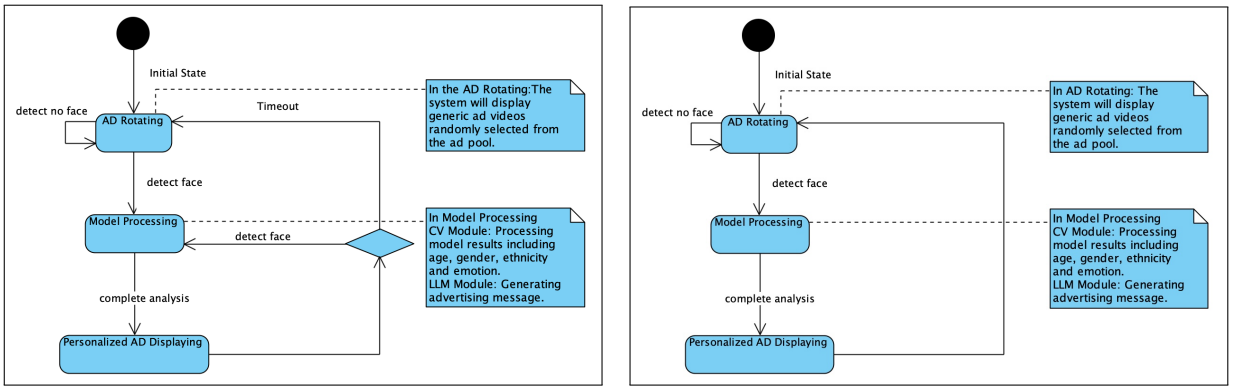


Figure 5.5: Sequence Diagram 4: View Model Analysis Result

5.2.3 State Diagram

Key Information

- The State Diagram shows transitions between various states in the system.



(a) State Diagram: Ideal (b) State Diagram: Realistic

Figure 5.6: State Diagram

Note: The state diagram includes two states due to the LLM's slow output generation, necessitating a compromise to accommodate its processing constraints.

Description

- **AD Rotating:** The system starts in the AD Rotating state. In this state, it continuously cycles through advertisements from the ad pool while detecting faces in real time.
- **Model Processing:** Upon detecting a face, the system enters the Model Processing state for analysis.
- **Personalized AD Displaying:** Once the analysis is completed, the system enters the Personalized AD Displaying state, showing personalized advertisements to the users.

Chapter 6

User Interface

This chapter provides a comprehensive exploration of the system's user interface, the backend processes associated with each operational state, and the dashboard.

6.1 User Interface Per State

This section illustrates the user interface for each state and its interactions in the State Diagram. In the following figures, sub-figure (a) represents the User Screen, where advertisements are displayed to the user, while sub-figure (b) represents the System Backend, where system states are monitored by the system administrator.

State: AD Rotating

When no face is detected by the CV Module or when the LLM has not yet completed generating text, the system remains in the AD rotating state. The User Screen continuously displays a predefined sequence of advertisements from the AD pool, ensuring uninterrupted screen utilization. Meanwhile, the CV Module persistently monitors for facial presence while ads are playing, allowing for the real-time transition to personalized advertising once a face is detected and the LLM-generated content is ready.

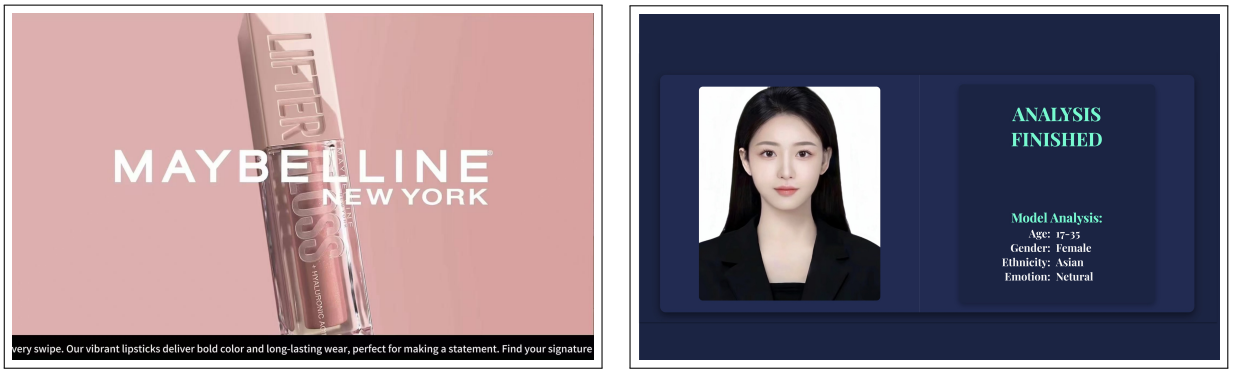
State: Personalized AD Displaying

When the CV Module detects any face, the system begins the model process. Once the model process is complete and the advertisement rotation ends, the system matches the appropriate video from the ad pool based on the user's demographic information.

(a) User Screen (b) Backend

Figure 6.1: AD Rotating

Simultaneously, the personalized content generated by LLM is displayed at the bottom of the video. In the back-end, the analysis results are intuitively displayed, including information such as age, gender, ethnicity, and emotions. During the personalized ad segment, the eye-tracking system records the viewer's view time and stores the data in the database. The dashboard then retrieves these data for further analysis and optimization.



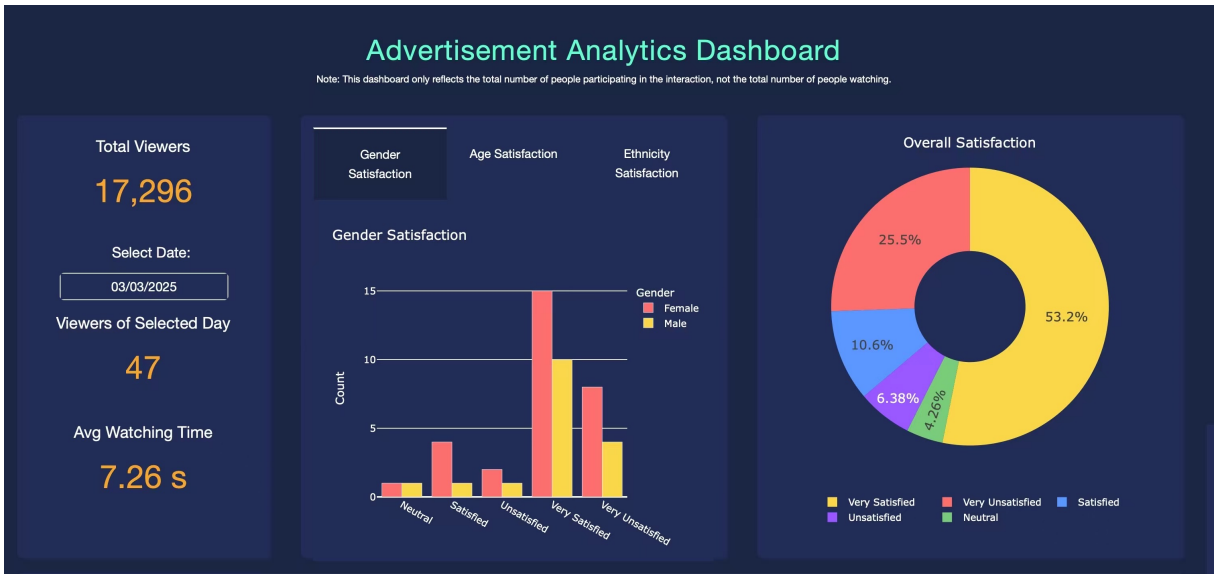
(a) User Screen (b) Backend

Figure 6.2: Personalized AD Displaying

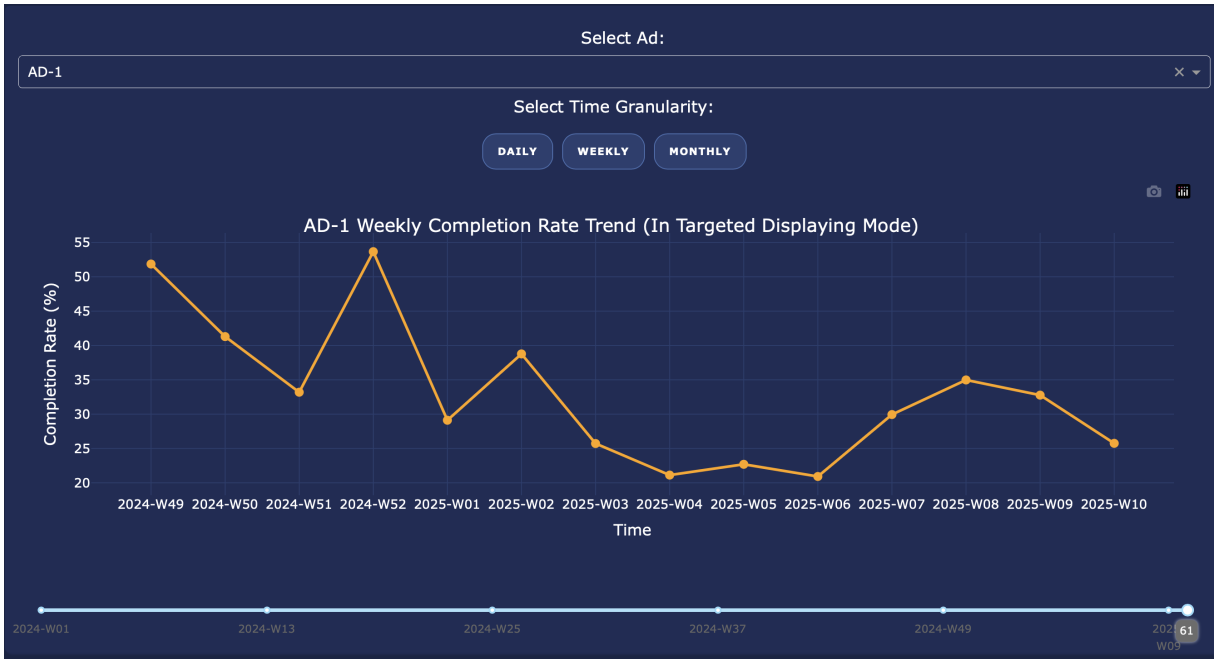
6.2 Dashboard

The dashboard, created by the Analytical Module, enables the system administrator to review advertising effectiveness, including the number of viewers and ad completion rate. It provides comprehensive insights through viewer engagement metrics (total audience reach, daily participation, and average viewing duration), and user satisfaction analytics (incorporating overall sentiment, as well as demographic segmentation based on gender,

age, and ethnicity). Furthermore, the dashboard offers adjustable temporal granularity (daily, weekly, and monthly), facilitating detailed trend assessments.



(a) Dashboard 1



(b) Dashboard 2

Figure 6.3: Dashboard

Chapter 7

Implementation

This chapter outlines the progress made in system implementation and prototypes development.

7.1 Key Technology Stack

- Python – Base programming language
- PyTorch – Machine learning framework
- OpenCV – Image Capture tool
- YOLOv11 – Face detection
- ResNet18 – Backbone for emotion classification task
- ResNet50 – Backbone for demographic data classification task
- Llama-3.2-1b-Instruct – Pretrained LLM for personalized advertisement message generation

7.2 Iterative Development Process

7.2.1 Initial Research

Description

At the beginning of this project, the team decided to perform a comprehensive review of publicly available datasets for demographic and emotional analysis. Additionally, the implementation of the Computer Vision (CV) module has been prioritized as the first

phase of development.

Achievement

Team members chose the open-source dataset UTKFace for the training of the demographic model and FER2013 for the training of the emotional analysis model.

Feedback

The datasets were deemed sufficient. Stakeholders suggested parallel development of the CV and LLM modules.

Improvement

The team decided to prioritize the development of the Computer Vision (CV) module, while concurrently assigning one team member to investigate pre-trained large language models (LLM).

7.2.2 CV Module Development

Description

The CV Module needs to perform tasks on:

- face detection
- demographic classification
- emotion analysis

Literature research was conducted to identify the best approaches for training these models.

Achievement

1. YOLOv8 is leveraged to perform the face detection due to its effectiveness and efficiency. However, it occasionally misclassified incomplete faces which caused negative

effects on downstream tasks.

2. ResNet is selected to be the backbone for both demographic and emotion classification due to its accuracy.

3. All models have been trained using Adam Optimizer with learning rate of 0.0001 together with a weight decay of $1e-5$.

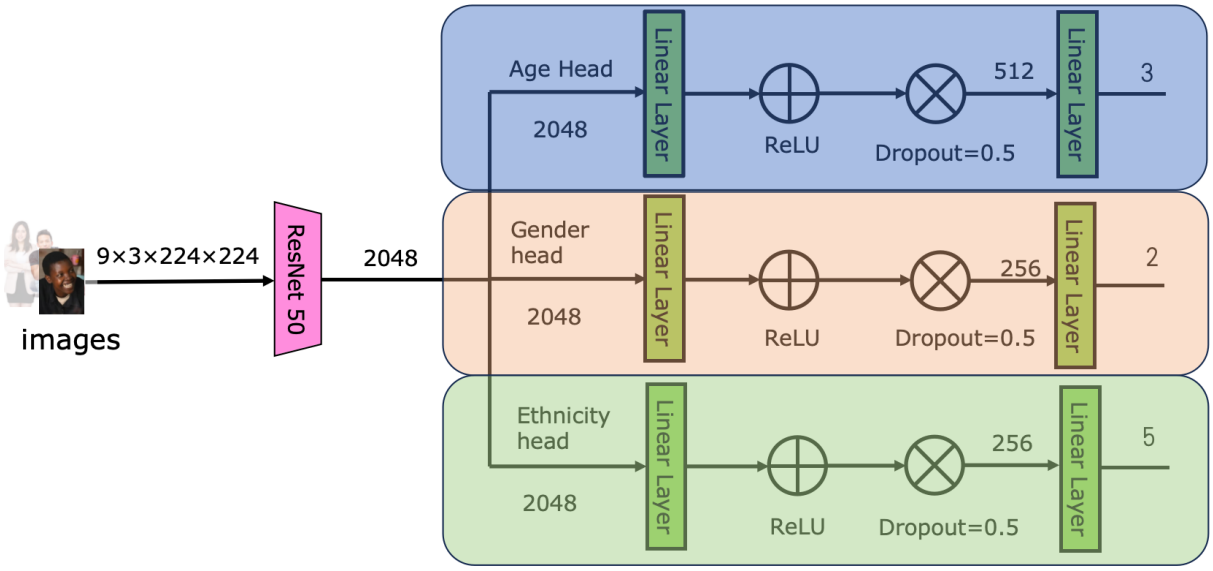


Figure 7.1: Proposed ResNet50 backbone model training on UTKFaces for classification on demographic data. The age head classifies for "17-30", "30-55", "55+". The ethnicity head classifies for "Asia", "India", "Middle East", "White" and "Black". Gender heads classifies for "men" and "women".

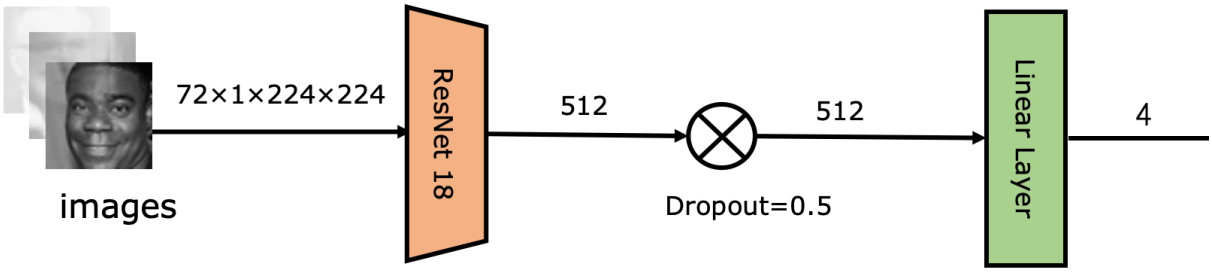


Figure 7.2: Proposed ResNet18 backbone model training on FER2013 for emotional analysis. The model classifies for "happy", "angry", "neutral", "sad".

Feedback

Stakeholder suggested replacing YOLOv8 with a simpler model if system performance was hindered. In addition, ResNet is suitable for the classification tasks. The team members

were also encouraged to explore ways to improve model accuracy.

Improvement

1. Upgraded to YOLOv11 for faster and more accurate face detection. Set confidence threshold to 0.86 to ensure full-face detection.

2. Data augmentation techniques (random cutting, color jitter, random flip, randomly adjusted crop, random rotation) are applied on both datasets.

3. UTKFace has more than 20,000 images while FER2013 has only about 2,000 images. After changing ResNet50 to ResNet18 as backbone and adding dropout layer on emotion analysis task to handle identified over-fitting problem started from epoch 8, the accuracy on validation set is increased from 61.43% to 66.52%. The team also tried a VGG type model as a backbone but it is less effective.

4. In FER2013, the emotions were divided into 7 classes, but there were few opportunities for the emergence of "fear", "surprise" and "disgust", so the team decided to reduce it to 4 classes (happy, angry, neutral, sad). The accuracy has increased to 69.04%.

Accuracy of each model (acc@1)	**RenNet50**	**Resnet18**	**VGG50**
UTKFace	80.13%	73.68%	78.92%
FER2013	61.43%	69.04%	66.82%

Table 7.1: Accuracies of Different Models and Datasets

7.2.3 Advertising Database & Pool System

Description

The integrated advertising system combines a relational database with a dynamic advertisement (ad) pool to manage targeted advertising effectively. The database consists of three core tables: *demographics*, *ads*, and *viewers*. The *demographics* table stores viewer characteristics such as age group, gender, and ethnicity. The *ads* table records advertisement details, including content, descriptions, target audience (linked via demographics

ID), priority weights, product names, duration, and core metadata such as video names and product specifications (e.g., background and weight coefficients). The *viewers* table logs interactions between viewers and advertisements, capturing view time, visit date, and associated demographic and ad IDs. Additionally, an ad pool stores advertisement videos and accompanying background information, supporting dynamic ad management. The system maintains an Ad Matching Matrix, ensuring a minimum of three pre-optimized video candidates per demographic cluster, and employs a tag-based classification system to categorize ads by audience group, industry type, and engagement level.

Achievement

The advertisements database and pool system provide a structured, scalable framework for personalized advertising. Implemented with hybrid storage—SQL tables for demographics, ads, and viewers, alongside a video pool—the system ensures efficient data retrieval and integrity through optimized indexing and foreign key constraints. A 3:1 ad-to-demographic redundancy ratio enhances matching flexibility. The ad pool was successfully populated with curated videos and mock data, seamlessly integrating with the Computer Vision (CV) Module for dynamic filtering based on recognized demographic attributes. The tag-based classification improves targeting precision, while engagement data (viewership metrics such as time, date, and duration) supports comprehensive analytics and data-driven optimization.

Feedback

Stakeholders praised the system's efficiency in demographic-based targeting and engagement analysis but suggested further enhancements. Recommendations included sourcing additional copyright-free ad videos from platforms like YouTube and refining categorization to align with CV Module outputs. Additional feedback highlighted the potential for advanced indexing, query caching, and real-time processing to boost performance, alongside integrating machine learning for predictive analytics and audience segmentation.

Improvement

The system was enhanced by incorporating additional copyright-free videos into the ad pool, sourced from reputable platforms, and implementing a structured categorization system mapped to CV Module outputs for automated, precise ad selection. Database optimizations, including advanced indexing, query caching, and partitioning, improved retrieval efficiency, while archiving and compression reduced redundancy. Functionally, real-time processing and advanced user behavior analytics were added to enhance targeting precision. Plans to integrate machine learning for personalized recommendations and predictive analytics aim to further optimize engagement and effectiveness, ensuring relevance and maximizing impact across demographic clusters.

7.2.4 LLM Module Development

Description

The LLM Module is designed to generate personalized advertising messages tailored to selected advertisements from the AD Pool, utilizing demographic and emotional data identified by the CV module. The proposed system synergizes computer vision-derived demographic/emotional cues with a LLM (deployed via HuggingFace Transformers) to dynamically synthesize contextually relevant promotional messages. Our architecture innovatively embeds real-time emotional as soft prompts into the LLM's attention mechanism, enabling joint optimization of linguistic coherence and affective alignment.

Prompt Structure

To enable systematic comparative testing, our team developed a Python-based prompt generator that automatically constructs structured LLM inputs from demographic parameters (race, gender, age range, emotion). The parametric transformation is formally defined as:

$$\mathcal{P}(r, g, a_r, e) = \text{“Target: } \underbrace{r}_{\text{race}} \underbrace{g}_{\text{gender}}\text{, aged } \underbrace{a_r}_{\text{age range}}\text{, feeling } \underbrace{e}_{\text{emotion}}\text{.”}$$

where (r, g, a_r, e) respectively denote categorical variables for race, gender, age range (e.g., "17-35"), and emotional state.

Final result

In the final version of the system, the LLM module enables population-driven AD personalization through an optimized processing pipeline. After receiving the demographic information transmitted by the CV module, the system automatically generates targeted advertising text by associating demographic characteristics, including age groups and population density indicators, with predefined marketing strategies. The AI-driven personalization process completes in about 15 seconds (all systems are powered by the Nvidia GeForce MX350), meeting the operational requirements of real-time adaptive content generation.

7.2.5 Analytical Module

Description

The effectiveness of the selected ads is evaluated by analyzing user engagement through advanced eye-tracking technology. This approach provides objective metrics on viewer attention and interaction with the advertisements.

Improvement

The way of collecting feedback from qr code in analytical module has been deleted, because it does not conform to hci design principles. After user testing, few users are willing to actively scan the code and fill in a long list of suggestions for this product. Instead, use only eyetracking models to capture the duration of a user's watch, measure the completion rate to reflect the performance of the AD, and visually visualize it in the dashboard.

Final result

An eye-tracking model has been integrated into the system to measure user engagement during the playback of personalized advertisement videos. The model detects and records

the duration of eye contact with the screen, allowing for the calculation of completion rates for each ad. These metrics are then visualized in a dashboard, providing intuitive insights into the viewing effectiveness of the advertisements. This method replaces the previously considered QR code-based questionnaire approach, which was abandoned due to low user participation rates during testing, as it did not align with HCI design principles.

7.2.6 Dashboard

Description

The Dashboard serves as the central monitoring and analytics platform for the AI-powered digital signage system. It provides real-time insights into advertisement performance, user engagement, and demographic distribution. Built using Flask and Dash visualization libraries, the Dashboard retrieves data from the Analytical Module and presents key metrics through an interactive web interface. The system continuously updates engagement statistics, enabling administrators to assess advertising effectiveness dynamically.

Achievements

The dashboard successfully integrates data from the database and the eye-tracking module, providing a structured and insightful visualization of advertisements performance. It can display key statistics such as total viewers, today's viewers, and ad completion rate. In addition, it provides a breakdown of audience attributes and advertisemnets completion rates, including the distribution of age, gender, race, and completion rate. Finally, ad performance can be tracked to show the engagement of each ad. The dashboard runs as a web-based application accessible to administrators. It uses asynchronous updates to dynamically obtain data from the backend, ensuring a smooth user experience.

7.2.7 Integration

Description

The integration phase aimed to combine the Computer Vision (CV) Module, Large Language Model (LLM) Module, and Analytical Module into a unified application. This was

achieved using the Flask web framework, which connects these modules into a seamless end-to-end system. The workflow begins with the CV Module detecting faces and extracting demographic data from video input. The LLM Module then generates personalized advertisement text based on this data. Finally, the Analytical Module, incorporating eye-tracking technology, analyzes user engagement by measuring viewing duration and ad completion rates. The resulting system delivers a cohesive experience accessible via a web interface.

Achievement

The integration effort successfully unified the CV Module, LLM Module, and Analytical Module into a single, functional application. The system operates through a state machine with two primary states:

- **AdRotating**: Displays generic advertisements while actively detecting faces using the CV Module.
- **PersonalizedADDisplaying**: Shows tailored ads generated by the LLM Module based on detected demographics, with the Analytical Module tracking user engagement via eye-tracking.

The CV Module processes video frames to identify faces and demographics, feeding this data into the LLM Module for ad generation. The Analytical Module, enhanced by eye-tracking, calculates total watch time and updates a database with engagement metrics. The Flask application provides three key interfaces:

- **User Screen**: Displays ads to the viewer.
- **Secondary Screen**: Shows supplementary information or controls.
- **Dashboard**: Presents engagement analytics in real-time.

These interfaces are accessible through a web browser, with threading used to manage concurrent tasks such as video capture, state transitions, and eye-tracking, ensuring smooth operation.

Final Result

The integrated system has been fully implemented, achieving a cohesive framework for personalized advertisement delivery. The state machine enables seamless transitions between displaying generic ads and personalized ads, with the CV Module successfully detecting faces and demographics from video input. The LLM Module generates tailored ad content based on this data, while the Analytical Module's eye-tracking functionality captures total watch time and updates the database accordingly. The Flask-based web interfaces—the User Screen, Secondary Screen, and Dashboard—operate concurrently, supported by multi-threaded execution, providing a stable and responsive system. This successful integration demonstrates the system's ability to deliver targeted advertisements and track engagement effectively.

7.3 Team Management Tools

1. **Feishu**
    - Manage personnel by establishing different chat groups to make it easier for team members to communicate respective tasks.
    - Record and summarize members' work in stages through shared documents.
    - File transfer is convenient and will not be automatically cleaned.
2. **GitHub**
    - Version control tool used in system developing.
    - GitHub link: https://github.com/RockYYY888/ai-digital-signage

Chapter 8

Test

8.1 Eyetracking Test

Eyetracking test							
ID	Test Scenario	Test Function	Test File	Test Input	Expected Output	Test Output	Test Result
1	Accuracy Test	Whether eyetracking module can record watch time and demographic ID	eyetrack.py & yolov8.py	Watched For 5.35s	5.35	4.96	Pass
2				Watched For 14.67s	14.67	14.2	Pass
3				Watched For 10.88s	10.88	11.23	Pass
4				Watched For 9.44s	9.44	8.64	Pass
5				Watched For 13.16s	13.16	13.47	Pass

Figure 8.1: Test output result of eyetrack.py

Figure 8.1 illustrates the table of eye-tracking test cases. The test propose is shown on the third column from left to right (Test Function). Test input, expected output and test results are also provided. All tests meet the expected result (test inputs are based on manual timing, small errors may be unavoidable).

8.2 LLM Test

LLM Test							
ID	TestScenario	Test Function	Test File	Test Input	Expected Output	Test Output	Test Result
1	Compatibility test*	Input a tuple from CV module, output the correct ad text	LLM/LLM.py	('17-35', 'Female', 'Asian', 'happy')	videos/0041.mp4 Hi high heels are more than just shoes for young women - they're a symbol of confidence, elegance, and personal style.	Output as Expected	Pass
2				('35-50', 'Male', 'White', 'sad')	videos/0012.mp4 A well-fitted suit embodies comfort & professionalism; perfect for formal events, business meetings, family gatherings & special occasions.	Output as Expected	Pass
3				('35-50', 'Female', 'Black', 'sad')	videos/0028.mp4 Feeling down? Let our lipstick add color & comfort with its soft shades, style & confidence that suits you from timeless reds to bold pinks.	Output as Expected	Pass
4				('50+', 'Female', 'Other', 'angry')	videos/0046.mp4 Indulge in our organic vegetables grown using natural farming methods without synthetic pesticides, fertilizers & GMs. Fresh, high nutritionally complete, reduced exposure to toxins, promoting peace of mind.	Output as Expected	Pass
5				('50+', 'Male', 'Indian', 'sad')	videos/0008.mp4 Feeling down? Our stylish sunglasses protect your eyes from harsh UV rays, reduce glare & improve visibility with clear comfort & bright colors.	Output as Expected	Pass

Figure 8.2: Test output result of LLM.py

Figure 8.2 demonstrates test cases of LLM.py. This test is to show whether the LLM model can generate the correct AD text with specified test input tuples. Expected output and test output are provided. All test cases have passed and have the expected result.

8.3 Dashboard Test

DashBoard Test							
ID	Test Scenario	Test Function	Test File	Test Input	Epected Output	Test Output	Test Result
1	Compatibility test*	Select a specific date, ad ID, age gender and ethnicity see if the dashboard can update the correct data	dashboard .py	(03/03/2025, AD-1, gender Completion rate, daily)	The dashboard would select data of 03/03/2025 from database, the bar chart would show the gender completion rate and line chart below will show the ad-1 completion rate daily	update as expected	Pass
2				(25/02/2025, AD-53, ethnicity Completion rate, weekly)	The dashboard would select data of 25/02/2025 from database, the bar chart would show the ethnicity completion rate and line chart below will show the ad-53 completion rate weekly	update as expected	Pass
3				(02/02/2025, AD-23, gender Completion rate, monthly)	The dashboard would select data of 02/02/2025 from database, the bar chart would show the gender completion rate and line chart below will show the ad-23 completion rate monthly	update as expected	Pass
4				(18/02/2025, AD-63, age Completion rate, monthly)	The dashboard would select data of 18/02/2025 from database, the bar chart would show the age completion rate and line chart below will show the ad-63 completion rate monthly	update as expected	Pass
5				(22/02/2025, AD-14, ethnicity Completion rate, daily)	The dashboard would select data of 22/02/2025 from database, the bar chart would show the ethnicity completion rate and line chart below will show the ad-14 completion rate daily	update as expected	Pass

Figure 8.3: Test output result of dashboard

Figure 8.3 shows test cases of dashboard.py. This test is to show whether the dashboard can update the new data in the database and deliver new data to the dashboard screen. Input data shown as above(a tuple contains data, AD-ID, different completion rate and sorted by different standards). Test output is same as expected. All test cases have passed.

8.4 CV test

CV Tests							
ID	Test scenario	Test Function	Test File	Test Input	Expected Output	Test Output	Test Result
1	Accuracy test	demographic and emotional recognation	CV/yolov8.py		(17-35, Male, India, Happy)	(17-35, Male, India, Happy)	Pass
2					(17-35, Female, Black, Netural)	(17-35, Female, Black, Netural)	Pass
3					(17-35, Female, White, Happy)	(17-35, Female, White, Happy)	Pass
4					(35-50, Male, White, Netural)	(35-50, Male, White, Netural)	Pass
5					(50+, Female, Asian, Happy)	(50+, Female, Asian, Happy)	Pass

Figure 8.4: Test output result of CV.py

Figure 8.4 shows test cases table of CV.py. This test is to show whether the CV module can produce an accurate demographic and emotional prediction. Input data shown as above(a tuple contains age, gender race and emotion). The test output is the same as expected. All test cases have passed.

8.5 State test

State Test --- Main Screen							
ID	Test Scenario	Test Function	Test File	Test Input	Expected Output	Test Output	Test Result
1	Logical Test	Test the state machine to see if it can switch main screen into different states	state.py	No Input	Ad rotating state, ad would play randomly	Output as expected	Pass
2				('17-35', 'Female', 'Asian', 'happy')	Ad rotating state will switch into personalized ad displaying state	Output as expected	Pass
3				('35-50', 'Male', 'White', 'sad') then no input	After finishing displaying personalized ad, the state machine will switch into ad rotating instantly	Output as expected	Pass
4				('35-50', 'Female', 'Black', 'sad') then ('50+', 'Female', 'Other', 'angry')	After personalized ad displaying state, return to ad rotating, then play personalized ad again	Output as expected	Pass
5				('50+', 'Male', 'Indian', 'sad')	Ad rotating state will switch into personalized ad displaying state	Output as expected	Pass

Figure 8.5: Test output result of state—user screen

Figure 8.5 shows the table of test cases of user screen. This test is to show whether the state logic can switch correctly. All tests have passed.

State Test---Secondary Screen							
ID	Test Scenario	Test Function	Test File	Test Input	Expected Output	Test Output	Test Result
1	Logical Test	Test the state machine to see if it can switch main screen into different states	state.py	No Input	Stay in ad rotating screen	Output as expected	Pass
2				('17-35', 'Female', 'Asian', 'happy')	received CV data, show user's photo, age gender race and emotion	Output as expected	Pass
3				('35-50', 'Male', 'White', 'sad') then no input	received CV data, show user's photo, age gender race and emotion then return to ad rotating screen	Output as expected	Pass
4				('35-50', 'Female', 'Black', 'sad') then ('50+', 'Female', 'Other', 'angry')	received CV data, show user's photo, age gender race and emotion then return to ad rotating screen and then back to received CV data state	Output as expected	Pass
5				('50+', 'Male', 'Indian', 'sad')	received CV data, show user's photo, age gender race and emotion	Output as expected	Pass

Figure 8.6: Test output result of state—secondary screen

Figure 8.6 shows the table of secondary screen test cases. This test is to show whether the state logic can switch correctly. All tests have passed.

8.6 User test

Questionnaire	
Are you satisfied with the recommended advertising content overall?	How do you rate the accuracy of advertising signs based on age and gender recommendations?
Do you think the content on the advertising sign meets your interests and needs?	Do you think the recommended subtitles align well with the advertising video?
Do you think the AD content is relevant to your age group?	Do you find our digital signage system user-friendly and easy to use? Please share your experience, such as whether functions are easy to find and interactions meet your expectations.
Do you feel that your gender is taken into account in the advertising content?	
Do you feel that the advertisements recommended by the advertising signs match your personal characteristics?	What are your suggestions for improving the signage system?

Figure 8.7: Question in the in the questionnaire

Figure 8.7 This table contains the questions in the user questionnaire. This questionnaire aims at a diverse population with a randomized selection of all age groups, gender, and ethnicities.

Chapter 9

Manual

AI-Powered Digital Signage

This repo is for [P2024-08] AI-Powered Digital Signage for Targeted and Personalized Advertisement.

Introduction

This project aims to develop an AI digital signage capable of collecting users' facial and emotional data via a camera, enabling the recommendation of personalized advertisements tailored to individual preferences, thereby enhancing the fluidity and comfort of the shopping experience for users.

How To Get Started

Configure the conda environment.

1. Create new conda environment:

Open the terminal on Mac or Anaconda Prompt on Windows and run the following command to create a new environment:

```
conda create python=3.9
```

2. Activate and deactivate the Environment:

After creating the environment, activate it using the command:

- Activate on MacOS/Linux/Windows:

```
conda activate <project>
```

- Deactivate on MacOS/Linux/Windows:

```
conda deactivate <project>
```

Create a .env file in the project root directory.

Enter

```
HF_TOKEN=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
```

in the file and save it.

This step is intended to provide the token as an environment variable (You may need VPN for the first time).

How To Use

Start up the project.

We have provided multiple entrance for user to execute the project

- For Windows users: The user only needs to execute the pre-packaged .exe file to run the program.
- For Mac users: Mac users only need to run the corresponding executable file to launch the program.
- To run the source code: If you want to view our source code and execute it, please run the following command:

```
python state.py
```

You may need to install the required dependencies, execute the following command to install them:

```
pip install -r requirements.txt
```

After initializing the project, it will ask user to input a number (0 or 1), 0 represents the default camera, and 1 represents the external camera.

```
(base) liang@liangdeMacBook-Air ai-digital-signage % python state.py
Using device: cpu
cpu
Start up. Today's date is: 2025-03-24
Input camera index:
```

It will open a port to display the frontend components of the project.

```
 * Serving Flask app 'state'
 * Debug mode: off
 * Running on http://127.0.0.1:5000
Press CTRL+C to quit
```

The system acquires access to the camera resource, it will immediately initiate face detection and retrieve facial and emotional data.

```
0: 384x640 1 face, 303.1ms
Speed: 16.6ms preprocess, 303.1ms inference, 22.1ms postprocess per image at shape
    (1, 3, 384, 640)
[CV] Predicted Demographics: ('17-35', 'Male', 'Asian', 'sad')
[CV] Putting to detected_face_queue...
[State] LLM Processing: Generating ad text.
```

- Activate on MacOS/Linux/Windows

```
conda activate <project>
```

- Deactivate on MacOS/Linux/Windows

```
conda deactivate <project>
```

Three browser windows will show up as below, dashboard, secondary screen and user screen:

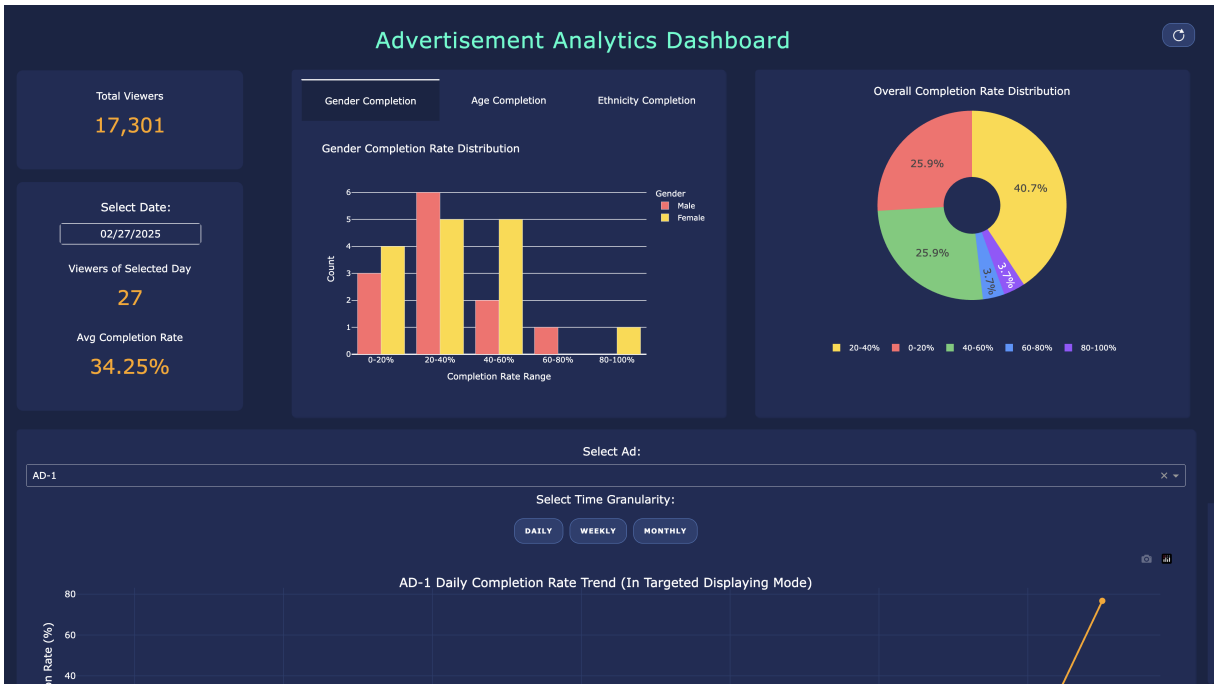


Figure 9.1: Dashboard

If you have any confusion about the project, please contact us: scyal8@nottingham.edu.cn

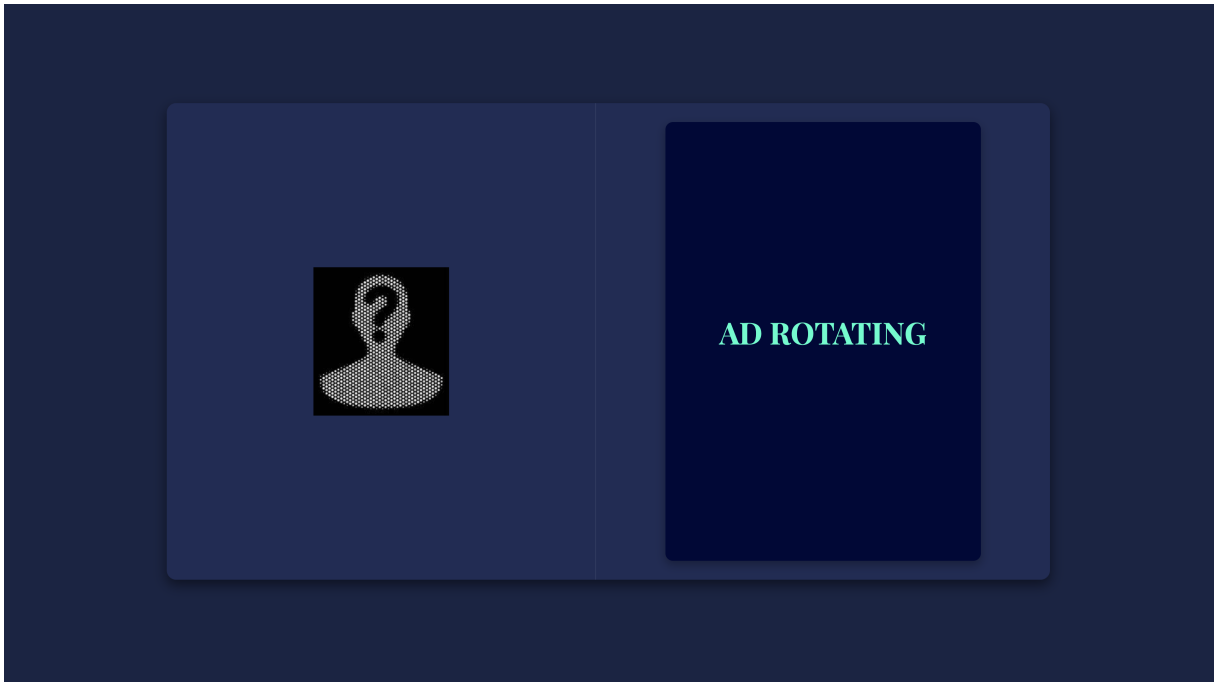


Figure 9.2: secondary screen(ad rotating state

Figure 9.3: secondary screen(target video state)

Figure 9.4: ad rotating displaying

Figure 9.5: targeted video displaying

9.0.1 Appendix

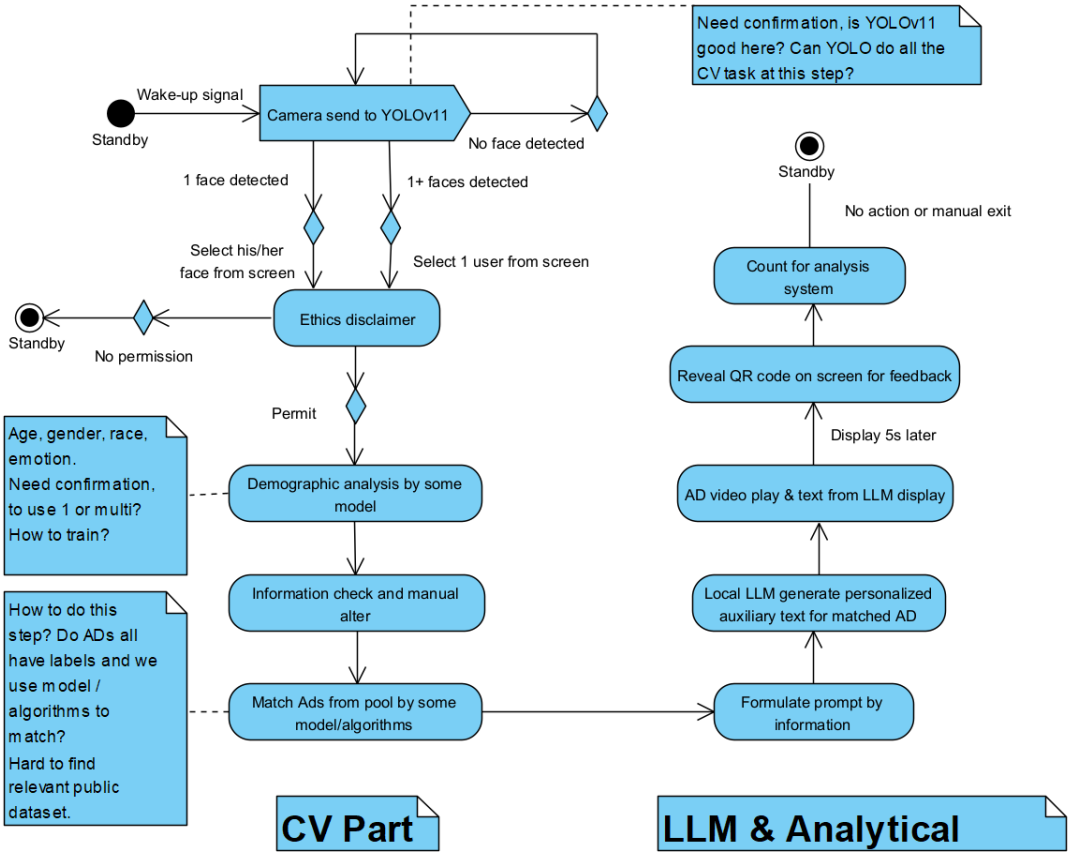


Figure 9.6: Activity Diagram For The Project

Chapter 10

Problem Encountered

10.1 Technical

Complex Multi-Module Code Synchronization

- **Problem**

  The system integrates multiple modules, such as the Large Language Model (LLM), Computer Vision (CV), and Eye Tracking, making synchronization challenging. Managing dependencies and ensuring smooth interaction between these components is difficult due to the need for concurrent execution, precise timing in data exchange, and consistent state management. Without proper coordination, issues like race conditions and data inconsistency may occur.

- **Remedial action**

  To address these challenges, several synchronization mechanisms were introduced:

  - **Queues**: Thread-safe queues (e.g., `detected_face_queue`, `ad_text_queue`) are employed to facilitate reliable data transfer between modules, ensuring that information such as detected faces or generated ad text is passed without loss or corruption.
  - **Events**: Threading events (e.g., `face_detection_active`, `eye_tracking_active`) are used to control the activation and deactivation of modules, enabling precise timing and coordination of their operations.

- **Locks**: Locks (e.g., `state_lock`, `watching_lock`) protect shared resources and critical sections, preventing concurrent access conflicts and ensuring atomic updates to system states.

These mechanisms work together to ensure seamless interaction between modules, maintaining data consistency and enabling smooth state transitions throughout the system.

Long LLM Processing Time

- **Problem**

  The LLM takes a long time to generate advertising messages, leading to potential delays in ad display.

- **Remedial Action**

  After weighing the impact of output quality and output speed on the project, the project team decided to modify the LLM parameters to slightly reduce the output quality of LLM and increase the output speed of LLM to provide users with a smoother experience.

Eye Tracking Accuracy Issues

- **Problem**

  The eye tracking system may incorrectly assign watch time from two individuals to the same user, raising concerns about accuracy in tracking and identification.

- **Remedial Action**

  The team set a confidence threshold of 0.86 for YOLO, increasing the likelihood of capturing full facial images.

Adaptation Issues Of Responsive Design

- **Problem**

  The variation in computer screen sizes may lead to differences in the rendering of the same HTML content, affecting both layout and dimensions.

- **Remedial Action**

  To mitigate variations in HTML rendering across different computer screen sizes, we implement full-screen mode to maintain the integrity of both layout and text presentation.

Secondary Screen Refresh Issues

- **Problem**

  Because the project does not store any information about the user, the secondary screen page cannot be manually refreshed when displaying the user's face captured by the camera, and refreshing will cause the captured face photo to disappear.

- **Remedial Action**

  Given the system's foundational constraint of not retaining user information, this issue is recognized as an inherent design trade-off rather than a technical flaw. The prioritization of user privacy precludes persistent storage of the captured image, rendering manual refresh functionality unfeasible within the current framework.

Model Recognition Accuracy Issues

- **Problem**

  The model's performance in discerning demographic details, notably age, exhibits suboptimal accuracy.

- **Remedial Action**

  Since the process of model recognition of population information is affected by factors such as light and camera frame rate, the project has not yet found a solution, which remain ongoing challenges for the project.

10.2 Interpersonal

Increased Workload Due to Team Member Transfer

- **Problem**

  One team member transferred to another university, leading to an increased work-

load for the remaining members. This affected both task quality and overall efficiency.

- **Remedial action**

  Last semester, we divided the work among ourselves and held weekly meetings. However, this semester, we have 3-4 meetings per week, each lasting four hours, to compensate for the workload of an additional person.

Mac-Windows Compatibility Issues

- **Problem**

  Team members use different operating systems (Mac and Windows), which caused compatibility issues, particularly in running and debugging the software.

- **Remedial action**

  Due to compatibility issues with the CV module, the existing code may not run on Mac. To address this, we modified certain parts to ensure Mac compatibility and added relevant comments to facilitate execution on any system.

Skill Gaps Among Team Members

- **Problem**

  Some members had limited experience with AI-related frameworks (e.g., deep learning models and computer vision), requiring additional time for learning and adaptation, which slowed down the project.

- **Remedial action**

  Experienced members take priority on tasks related to AI frameworks, such as deep learning models and computer vision, while other members make up for skill gaps by contributing in other areas or learning relevant knowledge using various tools.

10.3 Management

Inefficient Online Communication

- **Problem**

  Task allocation and discussions were primarily conducted online, which resulted in ineffective communication and unclear task assignments, slowing down development progress.

- **Remedial action**

  All the original online communications have been changed to offline communications. As mentioned before, all the problems between us will be discussed in the weekly informal meetings.

Inadequate Time Allocation in Interim Report

- **Problem**

  In the interim report, the estimated completion time was too short, creating a heavy workload for the team.

- **Remedial action**

  For the final report, we proactively addressed this issue by developing a detailed time plan in advance.

10.4 General Improvement

To improve team performance, collaboration, we proposed to apply two main methods:

Peer Review

To evaluate our team's work, the team used a structured peer review approach to assess the work, where team members check each other's work and provide feedback and suggestions to improve the quality of the deliverables. This process highlights the strengths of teamwork and mutual support, points out areas for improvement, and provides guidance for individual and team development throughout the project.

Reflexive Thinking

In addition to peer review, we have reflection sessions at each meeting to promote reflection on what the team had worked on and the decision-making process. At the end of each project phase, the team collectively discussed what went well, challenges encountered, and possible improvements for the next phase. These can be based on individual expertise for effective task allocation and good communication practices, as well as areas for improvement. By addressing these issues, the team implemented strategies such as setting clearer goals for meetings, which led to significant improvements in efficiency and productivity. Reflective thinking fostered a culture of continuous learning and adaptation, ensuring continued progress throughout the project.

Chapter 11

Summary

11.1 Summary

This project focuses on developing an AI-powered digital signage system that enhances advertising effectiveness through computer vision (CV) and machine learning (ML) techniques. The system analyzes user demographics to provide a highly adaptive user experience while ensuring compliance with ethical and privacy regulations. Additionally, by integrating emotion recognition into digital signage advertising, which is a rarely explored dimension, we enhanced the system's ability to adapt to real-time audience dynamics, making it more responsive and effective. Furthermore, the system incorporates eye-tracking technology to measure user engagement by analyzing viewing duration and ad completion rates, enabling data-driven adjustments to advertisement strategies. A Dashboard provides real-time analytics, visualizing key engagement metrics, demographic insights, and advertisement performance. This integrated monitoring platform ensures administrators can track the system's effectiveness and make informed optimizations.

Bibliography

Alshammari, A. and Alshammari, M. E. (2024), 'Emotional facial expression detection using yolov8', *Engineering, Technology Applied Science Research* **14**(5), 16619–16623.

Bauer, C. and Spiekermann, S. (2011), Conceptualizing context for pervasive advertising, *in* H. J. Müller, F. Alt and D. Michelis, eds, 'Pervasive Advertising', Springer, pp. 159–183.

Burke, R. R. (2006), The third wave of marketing intelligence, *in* M. Krafft and M. Mantrala, eds, 'Retailing in the 21st Century: Current and Future Trends', Springer, pp. 159–171.

Burke, R. R. (2009), 'Behavioral effects of digital signage', *Journal of Advertising Research* **49**, 180–185.

Choi, M. S., Kang, M. C. and Yang, S. B. (2012), 'An investigation of affecting factors on consumers' perceived value and attitude towards advertising in smart signage', *Journal of Information Technology Applications and Management* **19**(4), 115–135.

Davis, W. S. (2019), *The Information System Consultant's Handbook*, CRC Press.

Dennis, C., Newman, A., Michon, R., Brakus, J. J. and Wright, L. T. (2010), 'The mediating effects of perception and emotion: digital signage in mall atmospherics', *Journal of Retailing and Consumer Services* **17**, 205–215.

Franke, G. R. and Taylor, C. R. (2017), 'Public perceptions of billboards: A meta-analysis', *Journal of Advertising* **46**(3), 395–410.

Greco, A., Saggese, A. and Vento, M. (2020), Digital signage by real-time gender recognition from face images, *in* '2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT', pp. 309–313.

Hong, J. (2014), 'A study on activation plan of digital signage: Focusing on its status and legislative system', *KCA Trends & Perspectives* **76**, 5–36.

ITGP, PTIP (2017), *EU General Data Protection Regulation (GDPR): An Implementation and Compliance Guide*, IT Governance Ltd, Ely. Available from: ProQuest Ebook Central.

Kim, S. J. and Shim, S. W. (2015), 'The media effect on user context of digital signage focusing on the display ad of convenient store', *The Korean Journal of Advertising and Public Relations* **17**(2), 104–137.

Krumm, J. (2011), 'Ubiquitous advertising: the killer application for the 21st century', *IEEE Pervasive Computing* **10**, 66–73.

Lagwankar, A. A. (2021), Face recognition and emotion identification, Ms thesis.

Lundström, L. I. (2008), *Digital Signage Broadcasting: Content Management and Distribution Techniques*, Focal Press, Oxford, GB.

Macaulay, L. A. (2012), *Requirements Engineering*, Springer Science & Business Media.

Mule, S., Waykar, Y. and Mahavidyalaya, S. V. (2015), 'Role of use case diagram in s/w development', *International Journal of Management and Economics* .

Müller, H. J., Alt, F. and Michelis, D. (2011), Pervasive advertising, *in* H. J. Müller, F. Alt and D. Michelis, eds, 'Pervasive Advertising', Springer, pp. 1–29.

Newman, A., Dennis, C., Wright, L. T. and King, T. (2010), 'Shoppers' experiences of digital signage-a cross-national qualitative study', *JDCTA* **4**, 50–57.

Nguyen, X.-P., Aljunied, S. M., Joty, S. and Bing, L. (2024), 'Democratizing llms for low-resource languages by leveraging their english dominant abilities with linguistically-

diverse prompts'.

URL: *https://arxiv.org/abs/2306.11372*

Phelps, J. E., D'Souza, G. and Nowak, G. J. (2001), 'Antecedents and consequences of consumer privacy concerns: An empirical investigation', *Journal of Interactive Marketing* **15**(4), 2–17.

Ravnik, R. and Solina, F. (2013), 'Audience measurement of digital signage: Quantitative study in real-world environment using computer vision', *Interacting with Computers* **25**(3), 218–228.

Saad, F. H., Farook, T. H., Ahmed, S., Zhao, Y., Liao, Z., Verjans, J. W. and Dudley, J. (2024), 'Facial and mandibular landmark tracking with habitual head posture estimation using linear and fiducial markers', *Healthcare Technology Letters* **11**, 21–30.

Schaeffler, J. (2008), *Digital Signage: Software, Networks, Advertising, and Displays: A Primer for Understanding the Business*, Focal Press.

Taylor, C. R. (2010), 'Editorial. innovation in traditional media: Keeping our 'eyes on' an innovative new measurement system for out of home advertising', *International Journal of Advertising* **29**(4), 521–525.

Taylor, C. R. (2015), 'Editorial. creating win-win situations via advertising: New developments in digital out-of-home advertising', *International Journal of Advertising* **34**(2), 177–180.

Taylor, C. R. and Sarkees, M. E. (2016), 'Do bans on illuminated on-premises signs matter? balancing environmental impact with the impact on businesses', *International Journal of Advertising* **35**(1), 61–73.

Truong, Y., McColl, R. and Kitchen, P. (2010), 'Practitioners' perceptions of advertising strategies for digital media', *International Journal of Advertising* **29**(5), 709–725.

Unhelkar, B. (2005), *Verification and Validation for Quality of UML 2.0 Models*, Wiley, United States. Available at: `https://www.google.com/books/edition/`

Verification_and_Validation_for_Quality/OHyhLEJ96uIC?hl=zh-CN&gbpv=0 [Accessed 28-Nov.-2024].

van Reijmersdal, E. A., Rozendaal, E., Smink, N., van Noort, G. and Buijzen, M. (2017), 'Processes and effects of targeted online advertising among children', *International Journal of Advertising* **36**(3), 396–414.

Zhang, Z., Song, Y. and Qi, H. (2017), Title of the paper, *in* 'Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)', pp. 5810–5818.

Appendix A

LSEPI

Intellectual Property

Our project has created intellectual property (IP), including the source code for various components such as the ad pool, computer vision (CV), dashboard, data integration, eye-tracking, and large language model (LLM). These assets are owned by our team. Currently, there aree no plans to commercially exploit the entire project, but future commercialization is possible. Should commercialization occur, we will protect the IP and consider licensing or selling the technology. We have used open-source libraries like PyTorch ResNet and LLaMA 3 under the BSD 3-Clause License and the LLaMA 3.2 Community License. We comply with their terms, retaining copyright notices, and documenting any modifications. The IP is owned by the team, and future distribution may occur under open-source licenses, adhering to legal and licensing requirements.

Research Ethics

Our project uses computer vision to collect demographic data through facial recognition. We adhere to the University of Nottingham's research ethics requirements, ensuring informed consent and voluntary participation. Facial recognition occurs only when users are in close proximity and have opted in. No personal data is stored, and only real-time processing is used for detected faces. We protect privacy by ensuring that the system is

restricted to users aged 17 and above. Ethical considerations were continuously reviewed, and the system was designed to minimize privacy risks. Our project follows the University's guidelines, ensures informed consent, and guarantees privacy by not storing any personal data.

Data protection

The project collects facial data for real-time processing without storing it. This data is used exclusively for personalized content delivery, ensuring compliance with privacy protection standards and safeguarding users' data rights.

Other Legislation

The project adheres to relevant legislation, including:

1. **Equality Act (2010)**: We ensure the system is accessible to all users, including those with disabilities.
2. **Computer Misuse Act (1990) and Counter-Terrorism and Security Act (2015)**: We implement security measures to protect against unauthorized access.
3. **Online Safety Bill (UK)**: The system enforces age restrictions (17+) to ensure the safety of vulnerable groups.
4. **Privacy and Electronic Communications Regulations (2003)**: The system does not use cookies or tracking mechanisms without user consent, ensuring transparency and privacy.

Broader Ethical and Social Considerations

Beneficiaries and Impact: Businesses and consumers benefit from personalized advertising. Businesses achieve better targeting and ROI, while consumers receive personalized recommendations. The system also enhances retail efficiency.

Unintended Consequences: Personalized advertising risks over-targeting, bias reinforcement, and inaccurate personalization targeting.

Affected Stakeholders: The work may directly impact users who interact with the signage and indirectly affect organizations deploying the system. To include stakeholder perspectives, we gathered feedback from target users during testing, and adjusted our interface and model behaviour accordingly.

Accessibility and Inclusiveness: We have aimed to ensure the system is accessible to a broad audience by designing a simple, intuitive interface and considering potential limitations faced by users with visual or mobility impairments.

Societal Inequality: Our system helps reduce social inequality by providing accessible design and intuitive interfaces for all users, including those with visual or mobility impairments. This ensures the system can be used by a diverse audience.

Environmental Considerations: The project minimizes its environmental impact by using local data processing rather than relying on cloud infrastructure, reducing energy consumption and carbon footprint.

Justification and Value: The project responsibly uses AI to improve public-facing technology, aligns with Responsible Research and Innovation (RRI) principles, and prioritizes societal benefit by addressing ethical, legal, and environmental considerations.

Appendix B

Meeting Minutes

Week 3.1

Meeting Overview

- **Time:** 2024.10.14 12:00 PM
- **Chairperson & Secretary:** Kian Ming Lim & Xinna Su
- **Present Members:** Supervisor & All Team Members

Meeting Content

1. **Brief Self-Introduction**
2. **Discuss Project Purpose.**
3. **Project Outline**
    - **Data Collection:** Open-source dataset (Asian version).
    - **Computer Vision:** Automated recognition.
    - **LLM:** Online API for language model integration.
    - **Advertising Delivery:** Demo online video for testing.
    - **Interaction & Optimization:** User interaction and system optimization.
    - **Analytics:** Dashboard for analytics and optimization.

4. **Additional Information**
    - Confirm necessity of mentioned functions.
    - Ethics form submission (Deadline: Before 10.31).
    - Complete functionalities before interim report.

To-Do List

1. Confirm purchase list – Zhitong GUO, Xinna SU
2. Choose dataset – Yiwei LI, Ang LI
3. Complete ethics form – Yuyang ZHANG, Zelin XIA
4. Identify learning needs and plan – Yiwei Zelin Ang

Week 4.1

Meeting Overview

- **Time:** 2024.10.21 12:00 PM
- **Chairperson & Secretary:** Yiwei Li & Xinna Su
- **Present Members:** Supervisor & All Team Members

Meeting Content

1. **Review Minutes**
2. **Recap and outline remaining tasks.**
3. **Model Presentation**
    - Software development.
    - Recognition of age, gender, and emotion.
    - Eye-tracking for engagement.
    - Dashboard for data analysis.

- Focus on CV and LLM.

4. **Equipment Adjustment:** Borrow a big screen instead of buying one.

To-Do List

1. Device Request – Xinna SU
2. LaTeX Notes – Xinna SU, Zelin XIA
3. Website Development – Zhitong GUO, Ang LI
4. YOLO Model – Zelin XIA, Yiwei LI
5. Feedback System – Yuyang ZHANG, Xinna SU
6. Diagram Check – Yiwei LI

Week 5.1

Meeting Overview

- **Time:** 2024.10.28 12:00 PM
- **Chairperson & Secretary:** Zhitong GUO & Zelin Xia
- **Present Members:** Supervisor & All Team Members

Meeting Content

1. **Completed & submitted the Project site.**
2. **Learnd and utilized a pretrained face detection YOLOv11 model.**
3. **Show better model we found and outline the product development plan:**
    - Additional features for gender and age recognition.
    - Potential for racial recognition.
4. **Website review and recommendations:**
    - Streamline content to highlight key attributes.
    - Add images and rearrange them for better aesthetics.
    - Deploy the updated website to the server.

To-Do List

1. Complete & Submit the site – Ang LI, Zhitong GUO, Yuyang ZHANG, Yiwei LI
2. Train & learn the YOLO model – Zelin XIA, Yiwei LI, Ang LI
3. Find suitable ads for different people. – Yuyang ZHANG
4. Starting doing the interim report – Xinna SU, Zhitong GUO, Yuyang ZHANG
5. Recheck the criteria and requirements in the handbook – Yiwei LI, Yuyang ZHANG, Xinna SU

Week 6.1

Meeting Overview

- **Time:** 2024.11.04 12:00 PM
- **Chairperson & Secretary:** Xinna SU& Zhitong GUO
- **Present Members:** Supervisor & All Team Members

Meeting Content

1. **Completed & Submitted the Project site**
2. **Learned and utilized a pretrained face detection YOLOv1 model.**
3. **Trained a Resnet18 for demographic data classification with UTK dataset:**
    - **Gender:** Male/Female
    - **Age Group:** 16- (Young)/17-45 (Young)/46+ (Senior)
    - **Ethnicity:** White/Black/Asian/Indian/Other
4. **Find suitable ads (10s) for different people:**
    - Find reasonable advertisements for each group.

To-Do List

1. Continue with the interim report –Xinna SU, Zhitong GUO
2. Emotional Analysis model training and integrating –Ang Li, Zelin Xia
3. Laptop runnable LLM exploring & prompt engine – Ang Li, Yiwei Li
4. Look for articles on age division & consumption customs –Yuyang ZHANG, Xinna SU
5. Keep training CV part –Yiwei LI
6. UML diagram –Zelin XIA

Week 7.1

Meeting Overview

- **Time:** 2024.11.11 12:00 PM
- **Chairperson & Secretary:** Zelin XIA & Yiwei LI
- **Present Members:** Supervisor & All Team Members

Meeting Content

1. **Update CV models**
    - Adjusted all CV models to process only one picture at a time, reducing the input size from 10 pictures to 1 for better performance and efficiency.
2. **Improving User Experience**
    - Revised the video display logic to ensure a smoother and more seamless experience for users.

To-Do List

1. Continue with the interim report – Xinna SU, Zhitong GUO
2. Training the CV part model – Ang LI, Yiwei LI
3. Prompt engineering – Zelin XIA
4. Fix the AD folder in the new age group – Yuyang ZHANG
5. Set up GitHub – Ang LI, Yiwei LI

Week 8.1

Meeting Overview

- **Time:** 2024.11.18 12:00 PM
- **Chairperson & Secretary:** Yuyang Zhang & Ang LI
- **Present Members:** Supervisor & All Team Members

Meeting Content

1. **Combine LLM MOdel**
    - Combine LLM models and demographic-emotion model and produce it.
2. **Automatic Search**
    - LLM Model will search advertisements from advertisement pool automatically
3. **Softw Engineering**
    - Very Important for the project.
4. **Interim Report**
    - Our model may show in the report.

To-Do List

1. Complete the interim report – Xinna SU, Zhitong GUO
2. Implement a matching algorithm for demographic-based ad selection – Yiwei LI, Ang LI
3. Integrate LLM with CV data and matched text files – Zelin XIA, Yuyang ZHANG
4. Combine and debug code for parts 2 and 3 – Yiwei LI, Ang LI
5. Submit the final code to the Git repository.

Week 9.1

Meeting Overview

- **Time:** 2024.11.28 15:00 PM
- **Chairperson & Secretary:** Yuyang Zhang & Ang LI
- **Present Members:** Supervisor & All Team Members

Meeting Content

1. **Interim Report**
    - The interim report was completed and sent to the supervisor for comments.
2. **Comments for interim report**
    - Correct the interim report according to the marking criteria provided.
3. **Comments for Background and Research**
    - Identify and highlight how the actual needs of the market and the shortcomings of existing technology drive the development of the project.
    - Strengthen arguments with references to industry statistics and literature.
4. **Lack of statistics**
    - Develop a comparison table that lists the functions of different types of signage.
5. **Include additional diagrams**
    - Ensure the following diagrams are included:
        - Use Case Diagram: To showcase user interaction with the system.
        - Activity Diagram: To visualize the workflow of key processes.
        - Sequence Diagram: To detail the interactions between system modules.

To-Do List

1. Continue developing our interim report. – Xinna SU, Zhitong GUO
2. Initial plan after interim report DDL.
3. Add thread control that lets the CV model wait for LLM model to deliver the result, then do the next detection.

Appendix C

User Evaluation Questionnaire

This is the evaluation questionnaire.

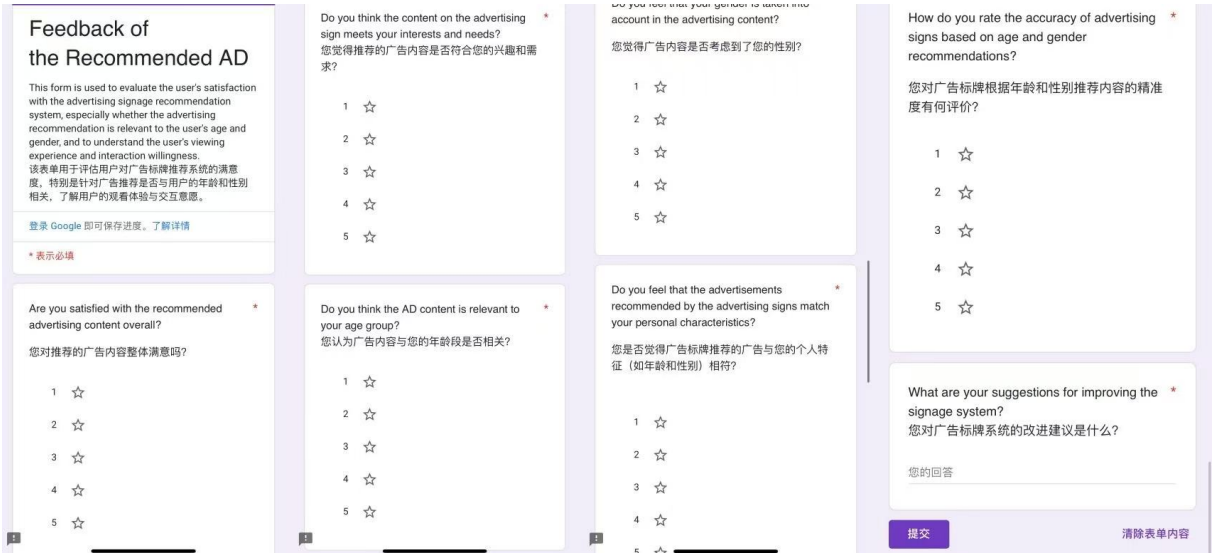


Figure C.1: The contents of questionnaire

Appendix D

Pseudocode

```
# Pseudocode for Face Detection, Model Inference, and Text Generation Process
Begin Program
While (Video Stream Active):
    # Capture a frame from the webcam
    frame = webcam.capture()

    # Wait for previous detection and processing tasks to complete
    Wait for detection_done_event

    # Perform face detection using YOLOv11
    faces = YOLO.detect_faces(frame)

    # If faces are detected in the frame
    If faces detected:
        # Crop the detected face from the frame
        face_region = crop(frame, face)

        # Predict demographic features (age, gender, race)
        demographic_predictions = DemographicModel.predict(face_region)

        # Predict the emotion of the detected face
        emotion_predictions = EmotionModel.predict(face_region)

        # Combine the demographic and emotion predictions
        combined_predictions = Combine(demographic_predictions,
emotion_predictions)

        # Initiate text generation for personalized advertisement based on
predictions
        Start new_thread(generate_personalized_text, combined_predictions)

        # Wait until the advertisement text generation is complete
        Wait for text_generation_done_event

        # Output the combined predictions and the generated advertisement
message
        output(combined_predictions, personalized_message)

        # Display the advertising message/video
        display(advertising_video)

    # Allow processing of the next frame
    continue
End Program
```

Figure D.1: Pseudocode

Appendix E

Timeline

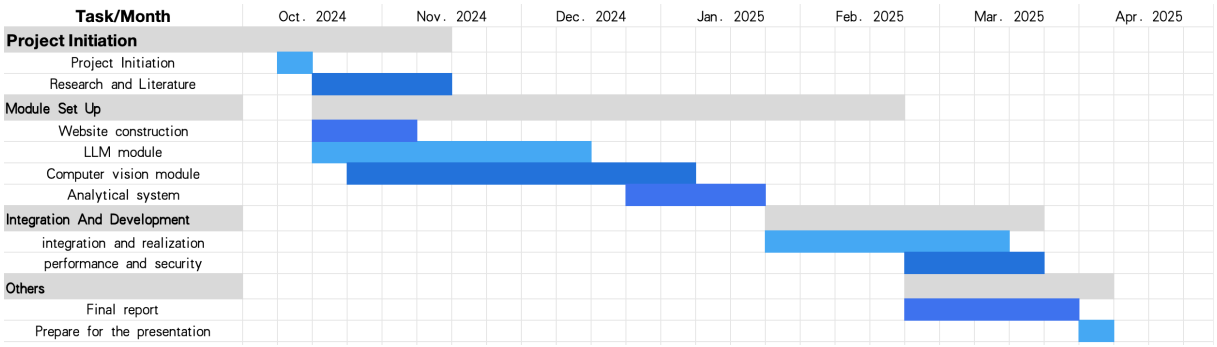
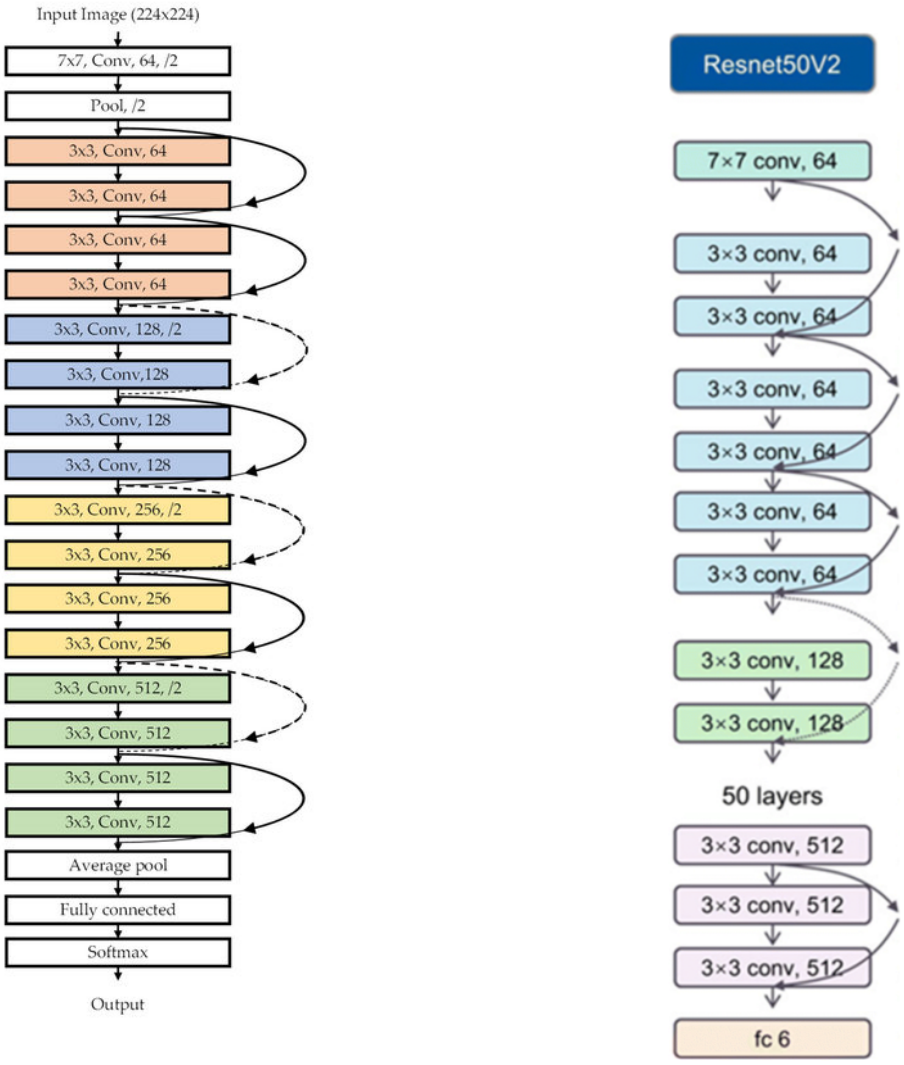


Figure E.1: Timeline

This shows the project time table for the recent months. Project Initiation concluded in November 2024, followed by the commencement of Research and Literature and Module Set Up. Research and Literature ran concurrently with the initial Module Set Up, which included Website construction and the LLM module, both ending in December 2024. Subsequently, the Computer vision module was developed until January 2025, and the Analytical system was completed by the end of February 2025. The Integration And Development phase, encompassing integration and realization, primarily took place in February and March 2025, with a focus on performance and security extending into April 2025. The team is scheduled to finalize the Final report and Prepare for the presentation in April 2025.

Appendix F

Backbone Structure



(a) The structure of ResNet18 (b) The structure of ResNet50

Figure F.1: ResNet50 is more capable for larger dataset and complex task while increasing risk of overfitting compared to ResNet18.