成都理工大学中英合作办学
CDUT Sino-British Collaborative Education

OXFORD BROOKES UNIVERSITY

UNDERGRADUATE PROJECT PROPOSAL

Project Title:	**Facial Recognition Attendance System**
Surname:	Qin
First Name:	**Hong Sheng**
Student Number:	**202118010213**
Supervisor Name:	**Joojo Walker**
Module Code:	**CHC 6096**
Module Name:	**Project**
Date Submitted:	**2024/11/4**

Table of Contents

1 Introduction

1.1 Background

Facial recognition technology has been extensively applied as a prominent biometric solution in security, surveillance, and personal identification[1]. This technology leverages computer vision and machine algorithms to recognize and verify individuals based on their unique facial features[2]. One domain where facial recognition can be particularly beneficial is attendance tracking, as it can provide a more efficient and reliable alternative to traditional methods like manual sign-in sheets or ID card scans. Existing attendance tracking methods have imperfections including time-consuming procedures, susceptibility to human error, and vulnerability to dishonest activities[3]. Therefore, by integrating facial recognition technology, the proposed Facial Recognition Attendance System (FRAS) aims to provide a more accurate, secure, and convenient way to monitor and record attendance, benefiting both the administrators and the attendees. The FRAS mainly combine computer vision and deep learning models to actualize robust and reliable identification of individuals even in the presence of variations in lighting, poses, and occlusions[4]. The remainder of this report is structured as follows: Section 2 provides a summary of existing approaches and related literature on facial recognition and attendance tracking systems. In Section 3, the recognition approach and implementation of FRAS are briefly elaborated. Section 4 covers project schedules, plans and deliverables. Ultimately, the cited articles will be shown in Section 5.

1.2 Aim

The aim of this project is to design and develop a Facial Recognition Attendance System that capable of efficient and reliable tracking and recording attendance in various scenario.

1.3 Objectives

The objectives are depicted as follows:

i. Conduct review on facial recognition and attendance tracking methodologies.
ii. Develop facial recognition model based on deep learning techniques.
iii. Ascertain and utilize appropriate datasets for training and evaluation.
iv. Implement and integrate model into unified system to optimize user experience.
v. Measure FRAS prototype through rigorous testing and documentation, and provide recommendations for future enhancements.

1.4 Project Overview

1.4.1 Scope

The purpose of this study is to develop a novel facial recognition model that can serve as the core component of the FRAS. The proposed model will focus on improving upon the accuracy, robustness, and computational efficiency of existing algorithms. The significance of this study lies in its potential to enhance the efficiency and reliability of attendance monitoring systems, benefiting multiple organizations. Meanwhile, this study will contribute to advancements in computer vision and machine learning research by exploring the integration of deep learning model to construct a computationally efficient facial recognition system.

1.4.2 Audience

The primary beneficiaries of this project will be organizations that demand efficient and reliable attendance tracking systems, such as educational institutions, workplaces and event venues. This project aims to provide a transformative attendance tracking solution which can benefit a wide range of organizations and their stakeholders, including administrators, employees, and attendees.

2 Background Review

2.1 Summary of literature review

A. Earlier state-of-the-art method:

The fundamental face recognition attendance system originated from traditional manual attendance tracking, representing one of the first attempts to digitize attendance management[5]. This system utilized fundamental image processing algorithms for face detection and matching, though it faced significant limitations in accuracy and environmental adaptability. Following this, the smart database-integrated system emerged as an enhancement, incorporating basic digital record management with SQL databases[6]. While this advancement improved data organization, the system still relied on rudimentary face matching algorithms and struggled with scalability issues.

B. Machine learning-based techniques:

The evolution continued with the development of multi-modal verification systems, which addressed growing security concerns in remote attendance tracking[7]. This approach innovatively combined facial recognition with geo-location verification, utilizing traditional machine learning algorithms alongside GPS integration. Parallel to this, research into synthetic data generation emerged as a solution to data scarcity problems[8]. This technique employed Generative Adversarial Networks (GANs) and conventional machine learning methods to create artificial training data, though challenges persisted in generating realistic and diverse facial datasets.

C. Latest state-of-the-art method using deep learning:

The comprehensive survey of 3D face recognition techniques represents the cutting edge of this technology's evolution[4], [9]. This research spans a decade of developments, documenting the transition from conventional methods to sophisticated deep learning approaches. The integration of Convolutional Neural Networks (CNNs) with 3D facial data marking has significantly advanced the field. This latest state-of-the-art approach not only addresses previous limitations in accuracy and reliability but also establishes new benchmarks for performance and sets directions for future research in facial recognition technology.

2.2 Annotated Bibliography

Essay Topic	Summary
Jing et al [4]	The essay examines a decade of 3D face recognition technology, analyzing both traditional and deep learning methods while comparing databases and performance metrics to provide researchers a comprehensive reference of current developments and future opportunities.
Patnaik et al [5]	The review discusses an automated attendance system using face recognition to improve accuracy, but lacks crucial details about implementation, performance metrics, and operational challenges.
Arjun et al [6]	The essay describes a face recognition-based attendance system with database integration, but omits essential information about technical architecture, implementation methods, and performance evaluation.
Singh et al[7]	The review explores an attendance system combining facial recognition and geo-location tracking, though it lacks technical details on architecture and performance metrics needed to fully evaluate its practical effectiveness.
Boutros et al [8]	The essay examines synthetic data generation for face recognition, highlighting its potential to address dataset limitations while acknowledging challenges in creating realistic and diverse synthetic faces.

Table 1: Annotated Bibliography

3 Methodology

3.1 Approach

The approach for a machine learning/deep learning-based project should focus on describing the core machine learning model to be employed. Briefly describe the mathematical basis, the algorithm details, and the optimization strategy, if applicable. Also, describe the datasets and data processing techniques to be used where relevant.

The project proposes a hybrid approach that combines convolutional neural networks (CNNs) and the Vision Transformer (ViT) architecture. Instead of facilitating raw image patches as input to the ViT, the hybrid model utilizes feature maps from a CNN. The patch embedding projection is applied to the CNN-derived patches, and the resulting sequence is fed into the Transformer encoder.

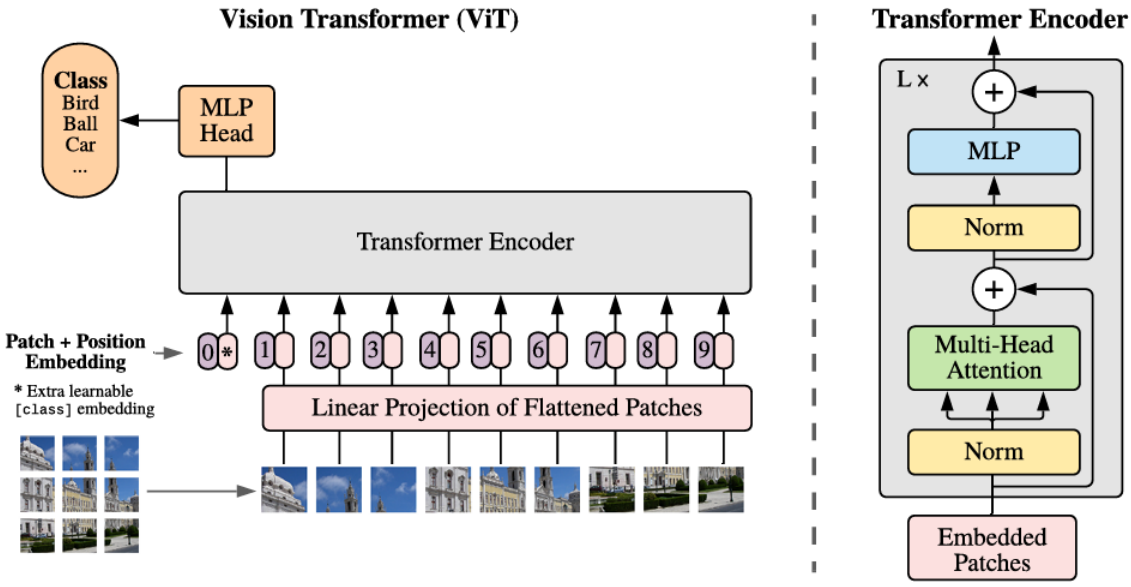

Figure 1 Overview of Vision Transformer with Encoder[10]

The hybrid design aims to leverage the inductive biases of CNNs including locality and translation equivariance, while also benefiting from the global attention mechanism of the Transformer. The approach experiment utilizes the output of different CNN stages as the input to the ViT component.

The hybrid model will be evaluated on datasets including ImageNet, ImageNet-21k, and JFT-300M. Meanwhile, deduplication will be employed when processing data to ensure fair evaluation.

3.2 Technology

The implementation tools are shown as follows:

Hardware	Software
System: Windows 11	Version: Tensorflow GPU @cuda 11.8
CPU: Intel core i7 11800H	IDE: Python 3.10.0
GPU: GeForce RTX 3060 16GB	
Memory: 16 GB	

Table 2: Development Infrastructures and Utilities

3.3 Version management plan

To efficiently manage and backup project code and document, the follwing will be utilized:

- Repository in GitHub, all code and documents will be uploaded.

  URL: RandyQin628/obu-project

4 Project Management

4.1 Activities

Objectives	Plans
Research on the facial recognition techniques in attendance systems	1.1 Research associative models thoroughly 1.2 Accomplish comparison table 1.3 Complete literature review
Develop facial recognition model based on deep learning techniques	2.1 Download necessary tools and IDE 2.2 Configuration setting 2.3 Comprehend each parameter and component 2.4 Design new model
Ascertain and utilize appropriate datasets for training and evaluation	3.1 Search and select dataset to be utilized 3.2 Initiate data preprocessing 3.3 Split into training, validation and test dataset 3.4 Model evaluation and prediction
Implement and integrate model into unified system to optimize user experience	4.1 Evaluate the integration degree 4.2 Test the stability and usability of the system with embeeded model 4.3 List out result of model weight and necessary syntax
Measure FRAS prototype through rigorous testing and documentation, and provide recommendations for future enhancements	5.1 Summarize the main features and innovation point 5.2 Demonstrate accomplishments 5.3 Recommend for future model enhancements

Table 3: Detailed Activities

4.2 Schedule

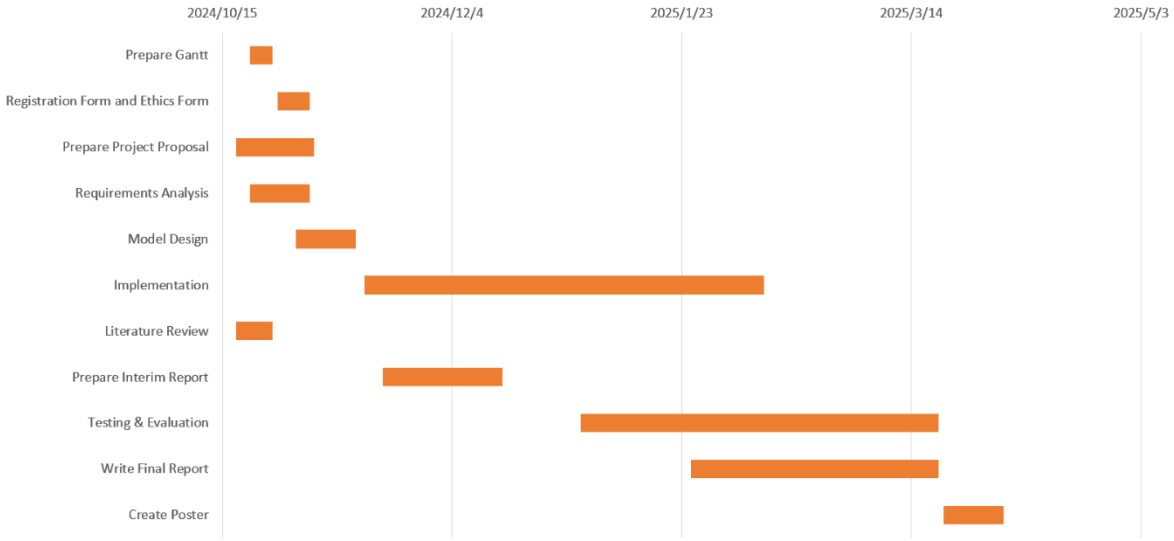

Figure 2: Schedule Plan

4.3 Data management plan

Git repositories are utilized to plan the whole documents and upload the latest code. The folder's contents will be updated in the future:

Git URL: [RandyQin628/obu-project]

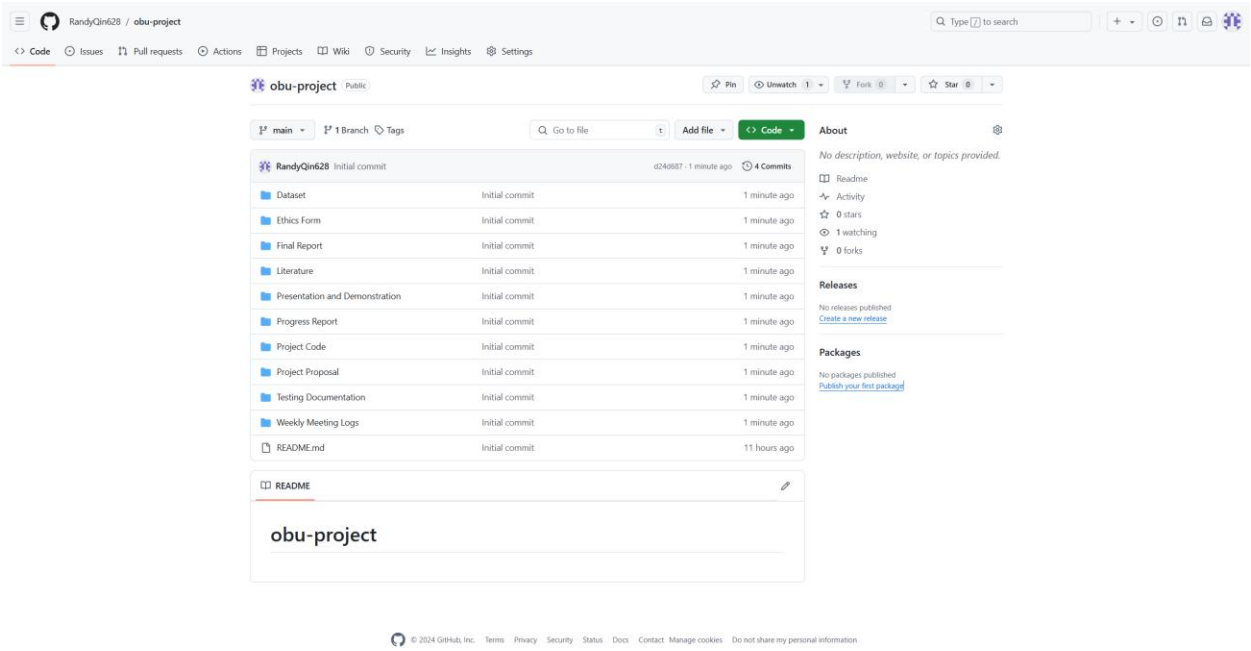

Figure 3: Repository Structure

4.4 Project Deliverables

To guarantee all the documents that must be submitted for assessment are clear, they are listed as follows:

- Project Log
- Interim Deliverables
- Detailed Project Proposal and Ethics Forms
- Progress Report
- Project Presentation and Demonstration
- Final Report
- Legal, Social, Ethical, Environmental and Professional Issues
- Dataset
- Project code
- Evaluation of the models

5 References

[1] Y. Guo, L. Zhang, Y. Hu, X. He, and J. Gao, 'MS-Celeb-1M: A Dataset and Benchmark for Large-Scale Face Recognition', in *Computer Vision – ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., Cham: Springer International Publishing, 2016, pp. 87–102. doi: 10.1007/978-3-319-46487-9_6.

[2] Z. Guo and Y. Fan, 'Sparse Representation for 3D Face Recognition', in *2013 Fourth World Congress on Software Engineering*, Dec. 2013, pp. 336–339. doi: 10.1109/WCSE.2013.63.

[3] 'Face recognition with learning-based descriptor | IEEE Conference Publication | IEEE Xplore'. Accessed: Oct. 28, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/5539992

[4] Y. Jing, X. Lu, and S. Gao, '3D face recognition: A comprehensive survey in 2022', *Comput. Vis. Media*, vol. 9, no. 4, pp. 657–685, Dec. 2023, doi: 10.1007/s41095-022-0317-1.

[5] S. Dev and T. Patnaik, 'Student Attendance System using Face Recognition', in *2020 International Conference on Smart Electronics and Communication (ICOSEC)*, Sep. 2020, pp. 90–96. doi: 10.1109/ICOSEC49089.2020.9215441.

[6] A. Arjun Raj, M. Shoheb, K. Arvind, and K. S. Chethan, 'Face Recognition Based Smart Attendance System', in *2020 International Conference on Intelligent Engineering and Management (ICIEM)*, Jun. 2020, pp. 354–357. doi: 10.1109/ICIEM48762.2020.9160184.

[7] G. Singh, M. kumari, V. Tripathi, and M. Diwakar, 'Attendance Monitoring System Using Facial and Geo-Location Verification', in *Intelligent Human Computer Interaction*, B. J. Choi, D. Singh, U. S. Tiwary, and W.-Y. Chung, Eds., Cham: Springer Nature Switzerland, 2024, pp. 406–416. doi: 10.1007/978-3-031-53827-8_36.

[8] F. Boutros, V. Struc, J. Fierrez, and N. Damer, 'Synthetic data for face recognition: Current state and future prospects', *Image Vis. Comput.*, vol. 135, p. 104688, Jul. 2023, doi: 10.1016/j.imavis.2023.104688.

[9] Z. Liu *et al.*, 'Swin Transformer: Hierarchical Vision Transformer using Shifted Windows', Aug. 17, 2021, *arXiv*: arXiv:2103.14030. doi: 10.48550/arXiv.2103.14030.

[10] A. Dosovitskiy *et al.*, 'An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale', Jun. 03, 2021, *arXiv*: arXiv:2010.11929. doi: 10.48550/arXiv.2010.11929.