Revolutionizing Tomato Production and quality assessment using AI

Project ID: 24-25J-337

Project Proposal Report

Herath H.M.R.K

B.Sc. (Hons) in Information Technology Specialized in Software Engineering

Department of Information Technology

Sri Lanka Institute of Information Technology Sri Lanka

July 2024

Defect detection and quality grading of tomatoes

Project ID: 24-25J-337

Project Proposal Report

Herath H.M.R.K

B.Sc. (Hons) in Information Technology Specialized in Software Engineering

Department of Information Technology

Sri Lanka Institute of Information Technology Sri Lanka

July 2024

DECLARATION OF THE CANDIDATE & SUPERVISOR

I declare that this is my own work, and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Herath H.M.R.K – IT20665548 Date

08/12/24

………………………. ………………….

The above candidate is carrying out research for the undergraduate Dissertation under my supervision

Signature of the supervisor Date

09/12/24

………………………………….. …………………..

(Mr. Amila Senarathna)

Signature of the Co-supervisor Date

09/12/24

…………………………………. ….…………………

(Dr. Lakmini Abeywardhana)

ABSTRACT

Ensuring the quality and market readiness of tomatoes is a critical challenge in the agricultural sector. Current grading systems primarily rely on manual inspection, which is inherently subjective, inconsistent, and labor-intensive, making it inefficient for large-scale operations. These methods often fail to identify subtle defects or account for environmental factors such as lighting variations and background noise. As the demand for high-quality produce increases, there is a pressing need for automated systems that can accurately and consistently assess tomato quality.

This research aims to address these challenges by developing an integrated system for Tomato Defect Detection and Quality Grading using advanced machine learning models.

consists of two key components:

1. **Defect Detection**: Leveraging **EfficientNet** to identify defects such as cracks, shriveling, and rot with high precision. EfficientNet's scalability and accuracy make it ideal for detecting both prominent and subtle defects under diverse conditions.
2. **Quality Grading**: Utilizing **MobileNetV2**, a lightweight yet powerful model, to classify tomatoes into predefined commercial grades (Grade A, B, C, and Rejected). MobileNetV2's efficiency ensures rapid grading without compromising accuracy, making it suitable for real-time applications.

The system is trained on a comprehensive dataset containing annotated images of tomatoes with varying defect types and quality attributes. Data preprocessing techniques such as normalization and augmentation are applied to enhance robustness and address class imbalance. Both models are fine-tuned to optimize performance, achieving high accuracy and consistency across training and validation datasets.

This solution represents a significant advancement in agricultural AI by providing a scalable and objective framework for quality assessment. The integration of **EfficientNet** and **MobileNetV2** ensures accurate defect detection and precise grading, even in challenging environments. Additionally, the system is designed to support potential deployment in real-world scenarios, including IoT-based field operations and mobile applications for farmers and supply chain managers.

By automating the defect detection and quality grading processes, this research not only reduces the reliance on manual labor but also minimizes errors, enhances productivity, and ensures uniformity in quality standards. Furthermore, the proposed system aligns with the broader goal of leveraging technology to promote sustainable and efficient agricultural practices.

Keywords: Tomato Quality Assessment, Defect Detection, EfficientNet, MobileNetV2, Machine Learning, Agricultural AI, Image Classification, Quality Grading, Automation in Agriculture.

Table of Contents

1. INTRODUCTION

1.1. Background & Literature survey

In the agricultural sector, quality assessment and grading of produce play a crucial role in ensuring market standards and customer satisfaction. Tomatoes, as one of the most widely consumed agricultural products, require meticulous grading processes to meet consumer and commercial expectations. Traditionally, these processes have relied on manual inspection, which is inherently subjective, labor-intensive, and inconsistent, especially for large-scale operations. Research has shown that such methods often fail to detect subtle defects and lack the scalability to cater to modern agricultural demands.

Advancements in machine learning and computer vision have paved the way for automated systems that address these limitations. Studies by researchers such as Zhang et al. [1] demonstrate the efficacy of EfficientNet in detecting visual anomalies in agricultural products. Automated defect detection systems leverage these technologies to identify defects like cracks, shriveling, and rot with greater precision and reliability compared to manual methods.

Quality grading, another critical component, requires systems that classify produce based on predefined standards, such as ripeness, size, and texture. MobileNetV2, as highlighted by Sandler et al. [2], offers a lightweight yet powerful framework suitable for classification tasks, making it an ideal choice for real-time grading applications. The integration of such models ensures scalability and consistency in grading outcomes.

However, despite these advancements, challenges persist. Variability in environmental conditions, such as lighting and background noise, often affects the accuracy of defect detection and grading models. Additionally, the lack of standardized and diverse datasets for training machine learning models further limits their applicability. Recent efforts, such as those by Patel et al. [3], emphasize the importance of robust preprocessing techniques and data augmentation to mitigate these issues.

This research builds on existing methodologies by integrating **EfficientNet** for defect detection and **MobileNetV2** for quality grading. The proposed system aims to enhance accuracy and scalability, addressing real-world challenges while aligning with commercial standards. By automating these processes, the system seeks to revolutionize tomato quality assessment, benefiting both producers and consumers.

1.2. Research Gap

While advancements in agricultural automation have made significant strides in addressing traditional challenges, a notable gap persists in creating comprehensive systems for tomato defect detection and quality grading that are both scalable and adaptable to real-world conditions. Existing approaches largely focus on individual aspects, such as defect detection or quality grading, but rarely integrate both into a unified framework that aligns with commercial standards.

Key Gaps Identified:

Manual Limitations:

Traditional grading relies on manual inspection, which is prone to human error, inconsistency, and scalability challenges in large-scale operations.

Subtle defects such as small cracks or early signs of rot are often missed, leading to inefficiencies in quality assessment.

Lack of Integration:

Current systems often handle either defect detection or grading independently, failing to provide an end-to-end solution for tomato quality assessment.

There is limited research on combining defect detection with commercial-grade classification in a cohesive system.

Dataset Challenges:

Publicly available datasets for defect detection and grading are scarce, imbalanced, or lack variability, which hampers model training and real-world applicability.

Existing datasets often fail to account for environmental variations such as lighting conditions, occlusions, and complex backgrounds.

Technological Constraints:

Many models are computationally intensive, making them unsuitable for deployment in real-time or resource-constrained environments like farms or mobile devices.

Current systems struggle with generalizing across diverse tomato varieties and environmental conditions.

Proposed Solution:

The proposed system bridges these gaps by:

Utilizing EfficientNet for robust defect detection, capable of identifying subtle defects under diverse conditions.

Employing MobileNetV2 for lightweight, accurate quality grading aligned with commercial standards.

Integrating both components into a unified framework to deliver consistent, scalable, and efficient quality assessment.

Developing a comprehensive dataset with diverse defect types and quality grades, incorporating augmentation techniques to enhance robustness.

Ensuring real-time operability by optimizing models for deployment on IoT platforms and mobile devices.

Features	Proposed System	Existing Systems
Tomato Defect Detection	✅	✅
Quality Grading	✅	❌
Texture Analysis	✅	❌
Commercial Standard Classification	✅	❌
AI Integration for Real-Time Detection	✅	❌
IoT Connectivity	✅	❌
Cloud-based Data Storage	✅	✅
User-friendly Dashboard	✅	❌

Figure 1- Novelty of the proposed system

The research gap lies in the absence of a unified framework that integrates tomato defect detection and quality grading into a single, efficient, and scalable system. Existing methods either focus on defect detection or grading in isolation, lacking the ability to provide a comprehensive solution. Additionally, current systems struggle with real-world variability, such as environmental conditions and diverse tomato varieties, limiting their effectiveness in large-scale agricultural operations.

The proposed system seeks to address this gap by developing a robust AI-driven framework that combines **EfficientNet** for defect detection and **MobileNetV2** for quality grading. This integrated approach ensures precise identification of defects and accurate classification into commercial quality grades, offering a scalable, real-time solution that aligns with market standards and operational needs.

1.3. Research Problem

The The agriculture industry faces persistent challenges in ensuring the consistent quality and grading of tomatoes, a crop that significantly contributes to global food supply chains. Despite advancements in technology, the current methods for defect detection and quality grading remain inadequate, particularly when scaled to meet the demands of commercial production.

Limitations of Existing Methods

Manual Inspection:

Tomato quality assessment primarily relies on human inspectors, which is labor-intensive, prone to fatigue, and subject to inconsistencies.
Subtle defects such as fine cracks, early rot, or uneven ripening are often overlooked, resulting in errors and reduced efficiency.
Manual grading struggles to meet the demands of large-scale operations, particularly in fast-paced supply chains.

Automated Systems:

Existing automated systems often focus on singular tasks, such as defect detection or quality grading, without providing an integrated solution.
Many models lack the robustness to handle real-world variations, such as changes in lighting, background interference, and differences in tomato varieties.
Limited availability of diverse and annotated datasets hampers the training of models, reducing their generalizability and accuracy.
Challenges in Technology Adoption

Computational Constraints:

Advanced machine learning models often require significant computational power, making them less feasible for real-time or on-site deployment.

Scalability:

Current solutions do not sufficiently address the need for scalability, particularly for small-scale farmers or decentralized operations.

Alignment with Standards:

Grading systems often fail to align with commercial quality standards, limiting their adoption by supply chain stakeholders.

The Need for a Comprehensive Solution

Given the importance of consistent quality in tomatoes for both producers and consumers, there is a pressing need for a reliable, scalable, and accurate system. This system should:

Detect defects such as cracks, shriveling, and rot with high precision using advanced machine learning techniques.

Grade tomatoes into commercially recognized categories such as Grade A, B, C, and Rejected.

Operate effectively under diverse real-world conditions, ensuring robust performance in different environmental scenarios.

1.3.1. Proposed Project

This project aims to develop a comprehensive AI-based system for Tomato Defect Detection and Quality Grading, addressing critical challenges in the agricultural sector. The proposed system will integrate cutting-edge machine learning technologies to enhance efficiency, accuracy, and scalability in tomato quality assessment. The key components include:

Defect Detection Using EfficientNet

Develop a defect detection module using EfficientNet model optimized for image classification tasks.

The system will identify defects such as cracks, shriveling, rot, and other visual anomalies with high precision and robustness.

Real-time predictions will ensure faster and more reliable defect identification in large-scale operations.

Quality Grading Using MobileNetV2

Implement a lightweight yet powerful grading module using MobileNetV2 to classify tomatoes into predefined grades: Grade A, Grade B, Grade C, and Rejected.

Ensure alignment with commercial quality standards for consistent and objective grading.

The model will assess attributes like size, texture, and ripeness while maintaining scalability for real-world applications.

Data Collection and Preprocessing

Curate a comprehensive dataset comprising diverse tomato images representing various defect types and quality grades.

Apply preprocessing techniques such as data augmentation, normalization, and resizing to improve model performance under different environmental conditions.

System Integration

Combine the defect detection and quality grading modules into a unified pipeline to provide seamless quality assessment from raw image input to classification output.

The integrated system will handle input variability, ensuring robust performance in different lighting conditions and environments.

Real-Time Deployment and Accessibility

Design the system for potential integration with IoT-enabled devices and mobile applications to facilitate real-time deployment.

Develop a user-friendly interface for farmers, supply chain stakeholders, and quality control personnel to access grading results efficiently.

2. OBJECTIVES

2.1. Main Objective

To develop an AI-based system for **Tomato Defect Detection and Quality Grading** that leverages advanced machine learning models, such as EfficientNet and MobileNetV2, to automate the classification of tomato defects and grades. The system aims to enhance accuracy, consistency, and scalability in quality assessment, addressing the limitations of traditional manual methods and supporting agricultural supply chain stakeholders with a robust, real-time solution.

2.2. Specific Objectives

Develop a Defect Detection Module:

Implement an EfficientNet-based model to identify common tomato defects such as cracks, rot, and shriveling.

Ensure high precision in defect classification under diverse environmental conditions, including varying lighting and backgrounds.

Design a Quality Grading System:

Utilize MobileNetV2 to classify tomatoes into predefined commercial grades (Grade A, Grade B, Grade C, and Rejected).

Align the grading system with standardized quality benchmarks to ensure consistent outputs.

Curate and Preprocess Datasets:

Collect and annotate a comprehensive dataset representing various defects and quality grades.

Apply preprocessing techniques like augmentation and normalization to address class imbalance and enhance model robustness.

Integrate Defect Detection and Grading Modules:

Combine the two models into a unified framework capable of delivering both defect and grade classification in a single workflow.

Ensure seamless data flow and processing from raw image input to final quality assessment output.

Optimize for Real-Time Deployment:

Design the system for integration with IoT-enabled devices or mobile platforms to facilitate on-field usage by farmers and supply chain operators.

Optimize computational efficiency to ensure fast processing and scalability.

Evaluate Model Performance:

Conduct comprehensive testing using validation datasets and real-world tomato samples.

Measure performance through metrics such as accuracy, precision, recall, and F1 score to validate system reliability.

Promote Usability and Accessibility:

Develop a user-friendly interface for stakeholders to access defect and grading results efficiently.

Provide deployment options suitable for small and large-scale agricultural operations, ensuring inclusivity.

Contribute to Sustainable Agriculture:

Enhance productivity and reduce waste by ensuring consistent quality assessment of tomatoes.

Support data-driven decision-making for farmers and supply chain stakeholders.

3. METHODOLOGY

The development of the **Tomato Defect Detection and Quality Grading System** involves several key stages, including data collection and preprocessing, model training, and system integration. The methodology outlined below provides a step-by-step process for creating an AI-based system that ensures accurate defect detection and robust quality grading, addressing the challenges of traditional methods.

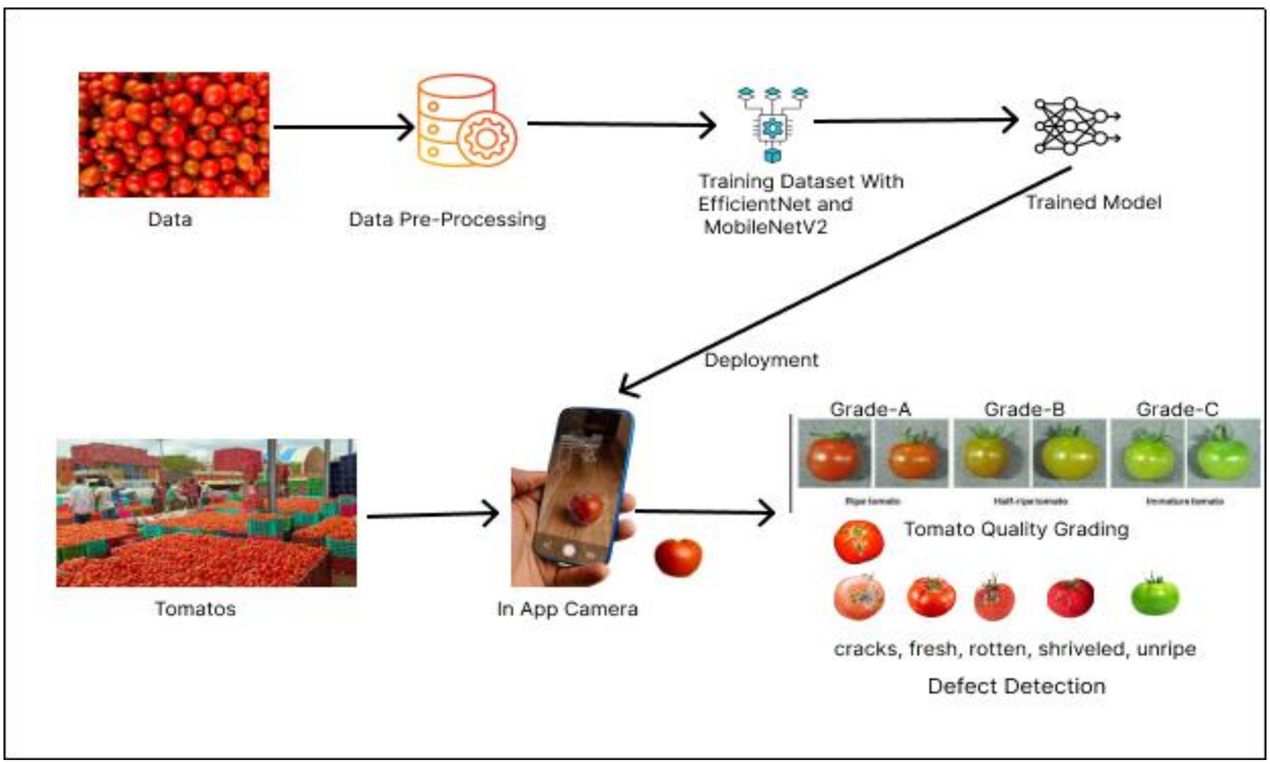

Figure 2 System overview

3.1. Conceptual Design

The conceptual phase focuses on defining the system's structure, identifying key objectives, and outlining its features:

Defining Objectives: Develop a system to detect defects (e.g., cracks, rot) and classify tomatoes into predefined quality grades (Grade A, B, C, Rejected).
System Flow: Conceptualize a two-stage process—EfficientNet for defect detection and MobileNetV2 for grading.
Interface Design: Plan a user-friendly interface for real-time results accessible via IoT or mobile platforms..

3.2. Data Collection and Preprocessing

Dataset Acquisition:

Collect diverse datasets representing tomato defects and quality grades from farms, marketplaces, and online repositories.

Labeling:

Manually annotate images with defect categories and grade classifications. Use tools like LabelImg to ensure consistency.

Preprocessing:

Normalization: Scale pixel values to a range of 0–1.
Augmentation: Introduce variability through rotation, flipping, brightness adjustment, and cropping.
Resizing: Resize images to 224x224 pixels for compatibility with machine learning models.

3.3. Model Development and Training

3.3.1 Defect Detection with EfficientNet

Train an EfficientNet model for multi-class classification of defects.
Fine-tune the model to optimize metrics like accuracy, precision, and recall.
Validate using an 80-20 training-validation split to evaluate model performance.

3.3.2 Quality Grading with MobileNetV2

Train MobileNetV2 to classify tomatoes into grades (Grade A, B, C, Rejected).
Use techniques such as dropout to prevent overfitting and ensure lightweight deployment for real-time applications.

3.4. System Integration

Pipeline Integration:

Combine defect detection and quality grading into a seamless workflow. Images are processed first for defects, followed by grading based on predefined standards.

Interface Development:

Create a web or mobile-based interface for farmers and quality controllers to upload images and receive predictions.

Real-Time Feedback:

Optimize inference speed to provide instant results during real-world operations.

3.5. Evaluation and Testing

Performance Metrics:

Evaluate the system using precision, recall, F1 score, and confusion matrices. Test on both validation datasets and real-world samples.

User Feedback:
Collect feedback from stakeholders, including farmers and supply chain operators, to refine system usability and functionality.

Environmental Testing:

Validate robustness against variations like lighting, occlusions, and backgrounds.

3.6. Project Requirements

3.6.1. Functional Requirements

- Defect Identification: Accurately detect defects like cracks, rot, and shriveling.
- Quality Classification: Grade tomatoes into categories aligned with commercial standards.
- Real-Time Processing: Provide instant predictions suitable for large-scale operations.
- User Accessibility: Develop an interface for easy data upload and results visualization.

3.6.2. Non-Functional Requirements

- Accuracy: Ensure models achieve high precision and recall across all categories.
- Scalability: Adapt the system for small-scale and industrial-level use cases.
- Device Compatibility: Support integration with IoT devices and mobile platforms.
- Robustness: Handle environmental variability without performance degradation.

3.6.3. Software Requirements

- TensorFlow and Keras: For building and training EfficientNet and MobileNetV2 models.
- OpenCV: For preprocessing and image analysis tasks.
- Flask/Django: For backend development of the interface.
- Database: Use PostgreSQL or MongoDB for storing annotated data and user inputs.

3.7. Commercialization

3.7.1 Target Market

Focus on farmers, distributors, and supply chain managers in agriculture. Emphasize how the system reduces manual labor and increases quality assurance.

3.7.2 Pricing Strategy

Offer a subscription-based model:

- **Basic Plan**: Single-user access for small-scale farmers.
- **Enterprise Plan**: Unlimited access for large agricultural corporations.

3.7.3 Special Offers

Provide initial trials for early adopters and discounts for bulk usage to encourage adoption.

3.8 Conclusion

This methodology ensures a robust framework for developing and deploying the Tomato Defect Detection and Quality Grading System. By integrating EfficientNet and MobileNetV2, the project delivers a scalable and efficient solution tailored to the needs of modern agriculture. The system promises to revolutionize quality assessment, ensuring consistency, scalability, and usability across the supply chain.

1. BUDGET AND JUSTIFICATION

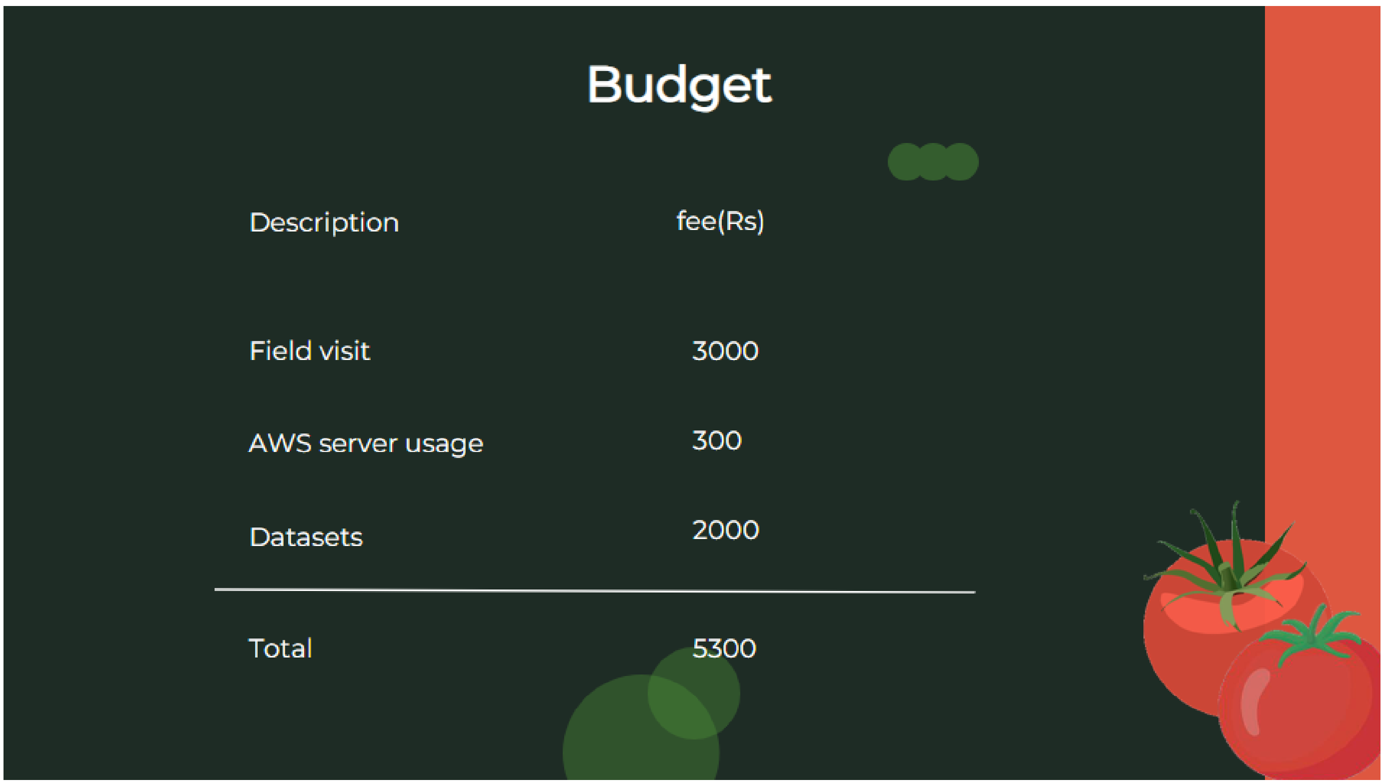

Figure 3 Budget chart

REFERENCES

1. A microcontroller based machine vision approach for tomato grading and sorting using SVM classifier [https://www.sciencedirect.com/science/article/abs/pii/S0141933119307057]

2. Fruit-Based Tomato Grading System Using Features Fusion and Support Vector Machine [https://link.springer.com/chapter/10.1007/978-3-319-11310-4_35]

3. Computer Vision Based Fruit Grading System for Quality Evaluation of Tomato in Agriculture industry [https://www.sciencedirect.com/science/article/pii/S1877050916001861]

4. A cost effective tomato maturity grading system using image processing for farmers [https://ieeexplore.ieee.org/abstract/document/7019591]

APPENDICES

Gantt chart

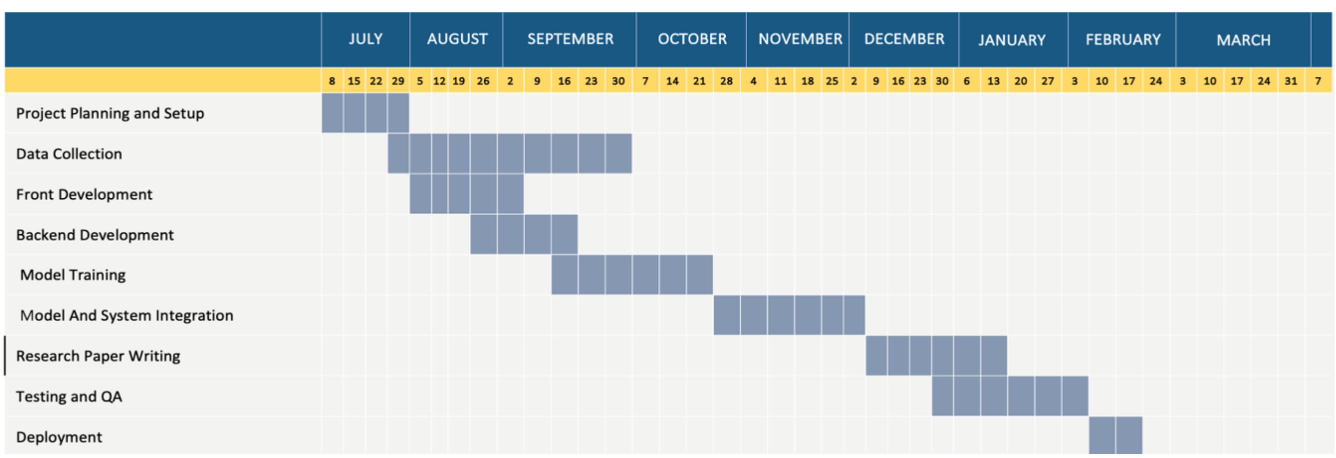

Figure 4 Gantt char

Work Breakdown Structure

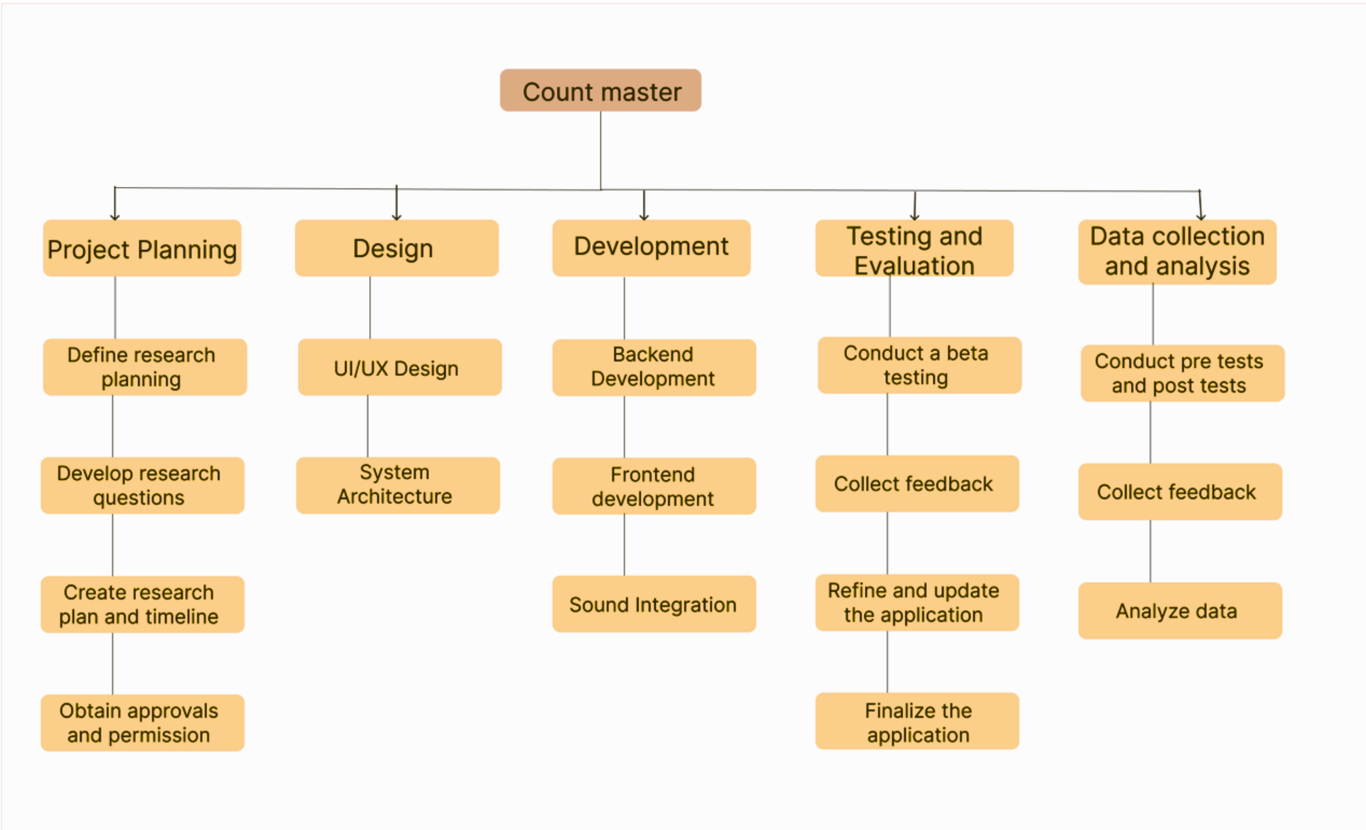

Figure 5 Work Breakdown Structure

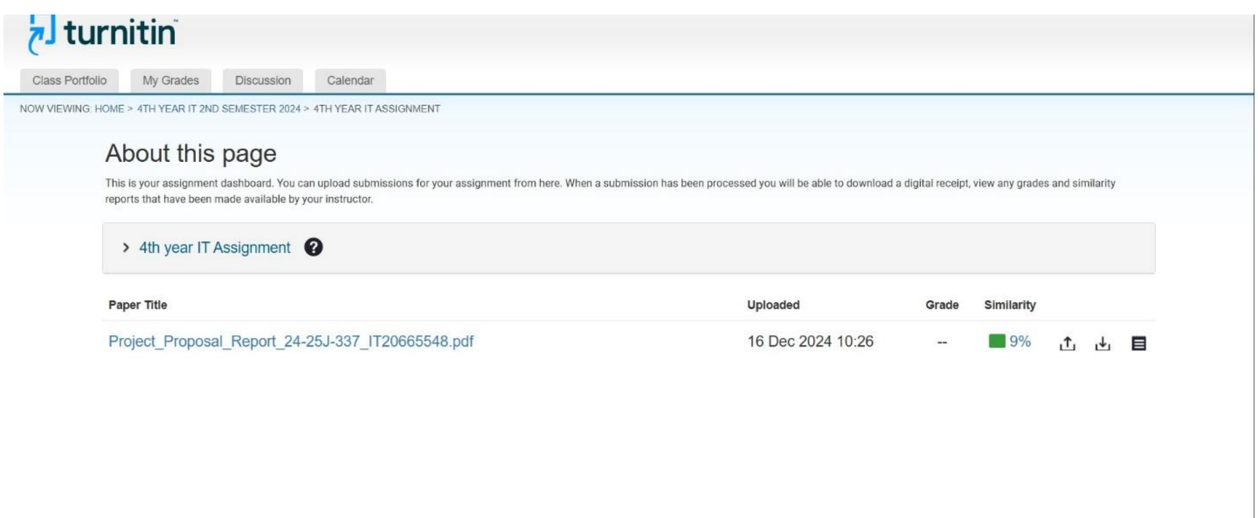

Figure 6 Plagiarism Report