Interim Report

Level 04

AI-Driven Educational Solutions for Visually Impaired Students in Sri Lanka

Group name: Storm Riders

Faculty of Information Technology

University of Moratuwa

2025

Interim Report

Level 04

AI-Driven Educational Solutions for Visually Impaired Students in Sri Lanka

Supervisor: Mr. B.H. Sudantha

Group members

Index Number	Name
204136G	Nedungamuwa P.D.M.C.C.B
204106P	Kumara P.M.M.D.S
204115R	Liyanage L.D.S
204049R	Dissanayake J.W

Storm Riders

Faculty of Information Technology

University of Moratuwa

2025

Table of Content

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Introduction

The integration of technology in education has revolutionized learning, offering new opportunities for students worldwide. However, visually impaired learners, particularly those pursuing the Mathematical stream, continue to face significant challenges. These students struggle to access graphical and spatial information, such as geometric diagrams, mathematical graphs, and data visualizations, which are essential for understanding complex mathematical concepts. While traditional tools like Braille and raised-line drawings provide some assistance, they are static and lack dynamic interaction or real-time feedback, limiting their effectiveness in teaching abstract and spatial concepts [3]. As a result, visually impaired students often lag behind their sighted peers in academic performance, not due to a lack of ability, but because existing resources and methodologies fail to meet their unique needs. This disparity is particularly evident in mathematics, where understanding core concepts, completing assignments, and excelling in exams remain significant hurdles. The shortage of specialized educators further exacerbates the problem.

1.2 Background and Motivation

The challenges faced by visually impaired students in mathematics education are well-documented. Studies reveal that 75% of visually impaired students are at least one grade level behind in math compared to their sighted peers, highlighting a persistent achievement gap [10]. This gap stems from the inherent difficulties in accessing dynamic mathematical concepts and educational materials, which are predominantly designed for sighted learners. Traditional resources, such as tactile materials and Braille books, while valuable, are inherently static and fail to convey the dynamic nature of mathematical concepts like graphs, geometric transformations, and algebraic equations. The limitations of these tools became even more pronounced during the COVID-19 pandemic, where visually impaired students faced severe barriers in accessing online math content. Unlike their sighted counterparts, who had access to a variety of digital resources, visually impaired students were largely restricted to screen readers and refreshable Braille displays, which proved insufficient for understanding complex mathematical materials [1].

The importance of addressing these challenges cannot be overstated. Mathematics is a foundational subject that plays a critical role in academic and professional success. For visually impaired students, the inability to access and comprehend mathematical content not only hinders their academic performance but also limits their future opportunities. Advances in artificial intelligence (AI) and assistive technologies offer a transformative opportunity to bridge this gap. AI-driven learning systems can provide personalized, adaptive, and interactive experiences, enabling visually impaired students to engage with mathematical concepts in innovative ways. Moreover, such systems can empower educators, many of whom feel unprepared to address the unique needs of visually impaired students, by providing them with tools to deliver more effective and accessible instruction [2].

1.3 Problem in Brief

Visually impaired students face significant barriers in accessing and understanding mathematical and graphical content, which hinders their ability to perform on par with their sighted peers. Current tactile tools, while helpful, lack dynamic guidance, making it difficult for students to accurately trace shapes and comprehend geometric properties. Traditional graphs are entirely inaccessible, requiring alternative approaches like sonification to convey patterns and trends effectively. Moreover, The lack of adaptive learning systems fails to meet diverse needs, while generic Q&A models lack the contextual understanding necessary for addressing specific mathematical challenges. These limitations underscore the need for innovative solutions to enhance accessibility and inclusivity in mathematics education [3].

1.4 Objectives and Aims

Aim

This research aims to develop an AI-Enhanced Educational Framework specifically designed for 9th-grade mathematics education for visually impaired students in Sri Lanka. This innovative framework will utilize cutting-edge AI technologies to create a personalized, multi-sensory learning environment. By focusing on key mathematical concepts such as algebra, geometry, and graphs, the framework aims to enhance accessibility, improve comprehension of abstract and visual content, and foster an inclusive approach to mathematics education [10].

Objectives

1. Study of the concepts, technologies relating to AI-Driven Educational Solutions for Visually Impaired Students in Sri Lanka.
2. Develop a personalized learning tool for visually impaired students in Sinhala.
3. Implement an interactive Q&A system with real-time feedback.
4. Create a sonification system for understanding graphs through audio cues.
5. Develop a video analysis system for shape recognition with tactile feedback.
6. Ensure compatibility with assistive technologies like screen readers and Braille displays.
7. Build a repository of accessible educational materials aligned with Sri Lanka's 9th-grade curriculum.
8. Develop training materials to help educators effectively use the AI framework in mainstream classrooms.
9. Significantly improve learning outcomes for visually impaired students.
10. Increase engagement and motivation through tailored learning experiences.
11. Foster greater independence in the learning process for visually impaired students.
12. Evaluation of the proposed solution.

By achieving these objectives, the project aims to contribute significantly to the field of stress management, offering a scalable and practical solution that enhances emotional well-being in various settings.

1.5 Proposed Solution

Our proposed AI-Enhanced Educational Framework for Visually Impaired Students focuses on addressing the unique challenges faced by 9th-grade students in Sri Lanka, specifically in mathematics. The framework consists of four integrated modules that utilize advanced AI technologies to create an accessible, personalized, and effective learning environment tailored to mathematical concepts [8].

1. Personalized Learning Model for Visually Impaired Students

The proposed personalized learning module for visually impaired 9th-grade students in Sri Lanka aims to tailor educational content based on two primary factors: the student's visual impairment state and their past academic performance. This innovative approach addresses the diverse needs of visually impaired students by creating a more accessible and appropriately challenging learning environment.

Key Components:

- Student Categorization: Visual Impairment State: Classify students as blind since birth, partially sighted, or fully blind due to later-life causes. Academic Performance: Categorize students as average, below average, or high performing based on past scores.
- AI-Driven Content Generation: Utilize artificial intelligence and machine learning algorithms to create and adapt educational materials. Generate content that is easily understandable and accessible for each student category.
- Multi-Modal Content Delivery: For fully blind students: Focus on audio-based learning materials representations. For partially sighted students: Implement high-contrast visual aids combined with audio support. Incorporate assistive technologies such as screen readers and Braille displays for seamless integration [9].
- Adaptive Learning Paths: Implement dynamic learning paths that adjust based on student performance and engagement. Continuously optimize the learning experience through real-time feedback and progress tracking.
- Natural Language Processing (NLP) Integration: Develop and train NLP models specifically for the Sinhala language to ensure cultural relevance and

accessibility. Implement sentiment analysis to gauge student engagement and emotional state during learning sessions.

This module addresses the critical need for personalized education in the context of visual impairment, potentially revolutionizing how these students access and interact with educational content.

2. Specialized Question and Answer Module for Visually Impaired Students.

This section discusses research on a specialized question and answer (Q&A) module for visually impaired students, focusing on mathematics education. Unlike general Q&A models, this module addresses the unique challenges faced by visually impaired learners, providing tailored content that meets their specific needs.

Current Q&A models present several challenges for visually impaired students. These models often lack accessibility, relying on visual inputs that don't accommodate the needs of visually impaired users, resulting in inaccurate responses. They also struggle to understand context, leading to poor performance on specialized queries, and typically lack interactive features like audio descriptions and tactile feedback, which are crucial for visually impaired learners [1].

The proposed solution addresses limitations by providing tailored responses for various visual impairments, ensuring clarity and relevance. It incorporates multi-sensory features like auditory descriptions and tactile materials for inclusiveness. AI-driven contextual awareness ensures accurate responses, while voice interaction, real-time feedback, and adaptive learning paths enhance engagement and accessibility in mathematics education. This approach delivers a personalized and effective learning experience for visually impaired students [11].

Key Features:

- Language Support: The module will be developed in Sinhala to ensure accessibility for Sri Lankan students.
- Adaptive Responses: Questions and answers will be generated according to the type of visual impairment (fully blind from birth, fully blind due to incidents, partially blind), ensuring that content is relevant to each student's experience.
- Natural Language Processing (NLP): Implementing NLP for intuitive interactions and natural question phrasing.

- Real-time Feedback Mechanisms: Immediate responses to improve mathematical comprehension.

This specialized Q&A module aims to enhance educational outcomes by addressing critical issues and providing equitable access to mathematics education for visually impaired students [2].

3. AI-Driven Sonification and Natural Language Generation for Accessible Mathematical Graph Visualization for blind students

The research focuses on leveraging Artificial Intelligence (AI) to develop advanced sonification techniques and natural language generation for graph visualization, with a specific emphasis on enhancing accessibility for blind and visually impaired students [5]. AI-driven sonification systems to make graphs and data visualizations accessible for blind and visually impaired students, promoting equal access to data-driven education. Visually Impaired students typically need more time to explore and understand graphical information compared to their sighted peers. This can lead to challenges in keeping pace with the curriculum and may require additional one-on-one instruction. Visually impaired students may struggle with developing spatial concepts necessary for understanding graphs, such as direction, distance, and relative positioning [7]. This fundamental challenge can impact their overall comprehension of graphical information. Sonification transforms visual data into sound, making it accessible to students who cannot perceive visual graphs. This allows blind students to engage with data in a meaningful way, similar to their sighted peers who use visual graphs. Unlike visual graphs, which may require careful analysis, sonified data can be interpreted quickly through sound. Students can hear changes in pitch or volume that correspond to data variations, enabling them to grasp trends and patterns immediately [4].

- Automated Mapping of Data Attributes to Sound Parameters-automate the translating numerical data values (e.g., magnitude, frequency) into auditory cues such as pitch, volume, and tempo.
- Pattern Recognition and Identification Through Sonification-identify patterns, trends, and key features within the graph data and translate these into recognizable auditory cues.
- Dynamic Graph Descriptions-NLG can generate real-time textual or spoken narratives that describe the features and patterns of a graph [5].

4. Smart Learning for the Visually Impaired: "AI-Driven Real-Time 3D Shape Detection and Position based description for Blind Students Using Finger Tracking and Video Processing"

The AI-Driven Real-Time Shape Recognition with Finger Tracking for Blind Students uses video analysis to enhance the tactile exploration of geometric shapes by providing real-time auditory guidance. Combining AI, computer vision, and finger-tracking technology, it detects finger movements on tactile shapes and delivers precise audio feedback, enabling students to understand geometric properties through touch and sound [12]. When a student traces a shape, the system tracks their finger using algorithms like MediaPipe and offers real-time auditory cues. For instance, if the student approaches a corner, they might hear, "You are approaching a 90-degree curve. Move slowly to trace the bend." The system continues, saying things like, "You have reached the next face. This edge is longer and straight." This dynamic feedback helps students visualize spatial relationships and structural details without sight. The tool fosters an understanding of geometric concepts, such as the transitions between faces of a cube, by describing properties like parallel edges, perpendicular intersections, and varying angles. The guidance adjusts to the student’s progress, ensuring accurate feedback at each point. This interactive, multi-sensory approach enhances spatial reasoning and builds confidence in exploring shapes. It bridges the gap between tactile exploration and cognitive understanding, making geometry and spatial subjects more accessible and engaging for blind students [13] [14] [15].

Key features:

1) Shape Detection & Initialization: When a shape is detected, the system announces its type (e.g., “You are near a triangle”).
2) Finger Placement & Tracking: Once the student places their finger on the shape, the system identifies the starting point and begins tracking their movement across the shape’s edges or axes.
3) Real-Time Guidance: As the student traces:
    - Audio prompts guide them through edges and vertices.
    - Feedback on shape properties (e.g., “You’ve reached an angle of 90 degrees”) is provided.

4) Correction & Encouragement: If the student moves off the shape, the system offers corrective feedback (e.g., “Move back to the left edge”) and encourages exploration.

1.6 Summary

This chapter discussed the introduction of the research topic, the background and motivation behind implementing the solution, the problem in brief and the aim and objectives have elaborated. In chapter two a literature review of existing works will be presented. Chapter three addresses the technologies that are adopted and the reasons for those selections over the other available options while Chapter four is about the approach adopted in developing the solution. The analysis and design is explained in Chapter five whereas Chapter six has detailed information on the implementation of each module up to now. Evaluation of each module is explained in Chapter seven as the discussion and the conclusion of the report followed by references and appendices.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This literature review explores the existing research on the challenges faced by visually impaired students in mathematics education, the limitations of current tools, and the potential of AI-powered solutions to bridge the gap. By examining these areas, this review aims to provide a foundation for developing innovative educational systems that enhance accessibility and inclusivity for visually impaired learners.

2.2 Research Paper Review

2.2.1 Module 01: Personalized Learning Model for Visually Impaired Students

Introduction

The integration of technology in education has revolutionized learning methodologies, yet visually impaired students continue to face significant barriers, particularly in mathematics education. Research indicates that these students often struggle to access and comprehend graphical and spatial information, creating a persistent achievement gap between visually impaired and sighted learners. Studies reveal that approximately 75% of visually impaired students perform at least one grade level behind in mathematics, highlighting an urgent need for innovative educational solutions. This review examines current research on personalized learning approaches, assistive technologies, and the potential of artificial intelligence in creating more accessible and effective learning environments for visually impaired students [23].

Theoretical Framework

The foundation of personalized learning for visually impaired students rests on two primary theoretical pillars: the Social Model of Disability and Vygotsky's Socio-Cultural Theory. The Social Model of Disability emphasizes that educational barriers arise not from individual impairments but from environmental and systemic limitations that fail to accommodate diverse learning needs. This perspective has driven the development of adaptive technologies and inclusive pedagogical approaches. Vygotsky's theory, particularly relevant in this context, emphasizes the crucial role of social interaction and guided exploration in learning, suggesting that visually impaired students benefit most from interactive, multi-sensory learning environments that facilitate both independent discovery and collaborative learning.

Current State of Assistive Technologies

Traditional assistive tools like Braille and raised-line drawings, while foundational, present inherent limitations in conveying dynamic mathematical concepts. Recent technological advances have introduced more sophisticated solutions, including electronic textbooks and screen readers. However, these tools often fall short in addressing the unique challenges of mathematical education, where spatial reasoning and graphical interpretation are crucial. Research demonstrates that while these technologies improve access to textual content, they struggle to effectively communicate complex mathematical concepts, especially in areas requiring visual-spatial understanding [24].

Challenges and Barriers

The implementation of personalized learning models faces several significant challenges. Studies in various educational contexts, particularly in developing regions, reveal systemic gaps in infrastructure and training. For instance, research indicates that 83% of schools lack adequate braille printing facilities, and 67% of teachers report insufficient training in assistive technology implementation. These challenges are compounded by the high costs of specialized equipment and the need for continuous technical support and maintenance [26].

Emerging Solutions and Innovations

Recent innovations in AI-driven educational technology show promising potential for addressing these challenges. Adaptive learning systems utilizing machine learning algorithms can now customize content delivery based on individual student needs and learning patterns. Research demonstrates successful implementations of sonification techniques for graph comprehension and video analysis systems for shape recognition. These technologies, when combined with traditional teaching methods, create more engaging and effective learning environments for visually impaired students.

Pedagogical Strategies and Adaptations

Successful implementation of personalized learning models requires careful consideration of pedagogical approaches. Studies indicate that flipped classroom models, where students review audio materials before class and engage in tactile activities during sessions, show particular promise. Assessment modifications, including voice-controlled navigation and extended time allowances for tactile exploration, have proven essential for equitable evaluation. The integration of AI-driven personalization with teacher-led instruction creates a balanced approach that maintains the human element while leveraging technological advantages [24].

Research Gaps and Future Directions

Several critical areas remain understudied in the current literature. Longitudinal studies examining the sustained impact of personalized learning models on academic achievement and career readiness are notably absent. Additionally, research on cost-effective solutions for resource-limited settings and the potential of emerging technologies like VR/AR in simulating visual concepts requires further exploration. The development of comprehensive frameworks for teacher training in assistive technology implementation also represents a significant gap in current research [26].

Conclusion

The literature reveals significant progress in developing personalized learning solutions for visually impaired students, particularly through the integration of AI-driven technologies and adaptive pedagogical approaches. However, substantial challenges remain in areas of implementation, scalability, and long-term effectiveness. Future research should prioritize developing more accessible, cost-effective solutions while maintaining focus on pedagogical effectiveness and student engagement. This review underscores the need for continued innovation in creating inclusive, technology-enhanced learning environments that effectively serve visually impaired students' educational needs.

2.2.2 Module 02: Specialized Question and Answer Module for Visually Impaired Students

The development of question-answering systems for visually impaired students represents a crucial advancement in accessible education technology. This review examines the specialized Question and Answer (Q&A) module implemented in the Mahoshadha system, focusing on its architectural design, processing methodology, and technological innovations that address the unique challenges faced by visually impaired learners.

Traditional Q&A systems have historically presented significant limitations for visually impaired users. These systems often rely heavily on visual inputs and lack the necessary accessibility features, resulting in compromised user experience and reduced accuracy in response generation. The conventional models struggle with context comprehension and typically lack essential interactive features such as audio descriptions and tactile feedback, which are crucial for visually impaired learners to engage effectively with educational content [18].

The Mahoshadha system introduces an innovative approach to Q&A processing through its dual-component architecture. The first component focuses on question analysis, incorporating Part of Speech (POS) tagging and language modeling. This foundational layer ensures accurate syntactic categorization of words within their contextual framework, utilizing the tagged Sinhala corpus from the University of Colombo's school of computing. The system's ability to process questions in Sinhala demonstrates a significant advancement in localized accessibility solutions.

A particularly noteworthy aspect of the system is its answer type identification mechanism. The implementation utilizes a rule-based pattern matching approach, which proves effective in categorizing questions into specific types such as person-type or location-type queries. This categorization is crucial for generating precise and relevant responses. The system employs Support Vector Machines (SVM) for question classification, enhancing the accuracy of answer type identification through automated pattern recognition.

The processing methodology incorporates sophisticated natural language processing techniques. The system utilizes a question-gram approach, which creates tokens from the query and matches them against the document corpus. This technique enables more accurate passage retrieval and answer generation. The implementation of PPMI (Positive Pointwise Mutual Information) values for keyword identification represents a significant advancement in understanding question context and identifying relevant answer candidates [17].

Performance evaluation of the system reveals promising results, with the authors reporting high accuracy rates in question processing and answer generation. The integration of SVM for pattern recognition has proven particularly effective in handling unseen text corpus, demonstrating the system's adaptability and robustness. This success can be attributed to the careful consideration of visually impaired users' needs in the system's design phase [17].

However, the literature also identifies certain limitations and areas for future development. The current implementation is primarily focused on theoretical questions and direct queries, with limitations in handling mathematical calculations and complex reasoning tasks. This presents opportunities for future research in expanding the system's capabilities to address these gaps.

In conclusion, the specialized Q&A module represents a significant advancement in accessible education technology. Its integration of modern machine learning techniques with accessibility-focused design principles offers a promising framework for developing more inclusive educational tools. Future research directions could focus on expanding the system's capabilities to handle more complex query types and incorporating additional interactive features to enhance the learning experience for visually impaired students [18].

This literature review elucidates the importance of specialized Q&A systems in educational accessibility while highlighting the innovative approaches used in their implementation. The findings suggest that continued development in this field could significantly impact the quality of education available to visually impaired students [17].

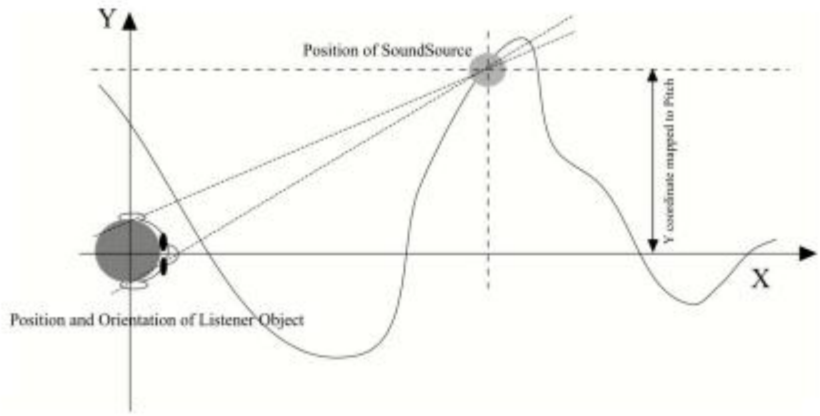
2.2.3 Module 03: AI-Driven Sonification and Natural Language Generation for Accessible Mathematical Graph Visualization for blind students

Constructing Sonified Haptic Line Graphs for the Blind Student: First Steps

This research paper, "Constructing Sonified Haptic Line Graphs for the Blind Student: First Steps," explores the potential of auditory and haptic technologies to make line graphs accessible to blind students, addressing the challenge that these graphs are traditionally visual tools. The authors argue against the prevailing view that blind individuals cannot effectively use line graphs for analysis, emphasizing that the usefulness of such graphs does not depend on their visual representation. Rather, access can be facilitated through non-visual media such as sound and touch. The study's focus is on making line graph functionality accessible, with a particular interest in understanding how auditory and haptic technologies can be employed as analytic tools for blind students, a topic that lacks substantial evidence in existing research. While much has been done in converting visual information for graphic user interfaces, this research extends these efforts by applying them specifically to line graphs, an area where fewer tools exist.

A significant portion of the research is dedicated to the design and implementation of sonified haptic line graphs, which combine spatial sound and force feedback. The paper discusses the design considerations necessary for mapping data to auditory parameters and creating intuitive haptic surfaces that allow users to experience the graph in an accessible and meaningful way. A key feature of the design involves fixing the listener or headphone object at the origin of the graph, thus aligning the user’s avatar with the direction of the X-axis. This design choice aids in quickly detecting the intersection points of the curve with the X-axis, making it easier for the user to follow the data. As technology advances, particularly with the increasing availability of force feedback devices, there is a growing interest in creating refreshable or dynamic audio-haptic displays, further enhancing the accessibility of line graphs. Additionally, the research employs a participatory design approach, collaborating closely with blind students and educators to ensure that the solutions developed are practical and truly meet the needs of the target audience. The goal of the research is to create tools that empower blind individuals to independently analyze and interpret graphical data, fostering greater independence and inclusion in educational settings.

In conclusion, the research offers a promising step forward in making graphical data accessible to blind students. By integrating auditory and haptic technologies, the authors demonstrate that line graphs can be transformed from a visual tool into a multi-sensory experience. Their approach, which includes close collaboration with the blind community, highlights the importance of user-centered design in developing accessible technologies. As these tools evolve, they hold the potential to provide blind individuals with powerful new methods for data analysis, contributing to greater academic and professional opportunities.



Fgure 1: Sound mapping for the tactile graph

Figure 2: Exploring a line graph in the auditory and haptic media

Interactive web-based image and graph analysis using Sonification for the Blind

The research paper "Interactive web-based image and graph analysis using Sonification for the Blind emphasizes that the increasing reliance on digital interfaces has created significant accessibility challenges for the 285 million visually impaired individuals worldwide. Traditional assistive technologies, such as tactile graphics and braille displays, are expensive, lack scalability, and are often incompatible with dynamic web content. Sonification, the process of converting visual data into non-speech audio, has emerged as a promising alternative. However, existing implementations face critical limitations in real-time interactivity, adaptability across different data types, and seamless integration with modern web architectures.

Current sonification systems suffer from several key limitations. Early tools, such as PLUMB and OpenCV-based systems, provided only static audio mappings and lacked real-time responsiveness. Mobile applications improved portability but constrained users to predefined object categories without adaptive sound generation for new visual elements. Computational overhead and latency issues further limit usability, with depth-sensor and CNN-based approaches suffering from processing delays exceeding 500ms. Traditional computer vision methods achieve lower latency but recognize only a limited number of object classes. Additionally, platform dependency remains a major challenge, as desktop applications require specialized hardware and OS-specific installations, while browser-based implementations often lack deep learning integration. Furthermore, most existing sonification approaches rely on either synthetic tones or speech synthesis alone, leading to high error rates in complex environments due to inadequate multimodal feedback.

To address these gaps, this study proposes a unified web-based sonification framework that integrates real-time object detection, adaptive sound mapping, and multimodal feedback. A pruned YOLOv5s model is employed to reduce inference latency while maintaining high object recognition accuracy across diverse categories. An adaptive sound mapping engine dynamically assigns natural sounds to biological entities and spatialized speech synthesis to inanimate objects, with depth estimation-based Doppler effect simulations for spatial awareness. The Web Audio API enables efficient client-side processing, while contextual haptic feedback, delivered through touchscreen vibration cues, enhances user interaction. This system ensures low-latency, real-time sonification and improves accessibility without the need for specialized hardware.

By combining modern web standards, deep learning, and multimodal interaction techniques, this research establishes a new paradigm for accessible digital content navigation for visually impaired users.

2.2.4 Module 04: AI-Driven Real-Time 3D Shape Detection and Position based description Using Finger Tracking and Video Processing

A 3D Shape Recognition Method Using Hybrid Deep Learning Network CNN–SVM

In this research article,3D shape recognition has traditionally lagged behind 2D object recognition due to the inherent challenges in representing 3D models as structured data suitable for deep learning techniques. Unlike 2D images, which are represented as pixel matrices, 3D shapes exist in formats such as point clouds, meshes, and volumetric data, each with unique complexities. The absence of a standard approach for storing 3D geometric shape information further complicates the task, making it difficult for conventional CNN architectures to generalize effectively from 3D data. Many existing methods rely on multiple 2D projections of 3D shapes, increasing model complexity and computational costs. Additionally, direct voxel-based approaches often suffer from high memory consumption and loss of fine structural details.

To address these challenges, a hybrid deep learning framework combining Convolutional Neural Networks (CNN) with a Support Vector Machine (SVM) is proposed for 3D shape recognition. The method begins by converting the vertices of a 3D mesh into point clouds, providing a structured yet simplified representation of the 3D shape. These point clouds are then subjected to rotational transformations for data augmentation, enhancing the model's ability to recognize shapes from different viewpoints. A key step in the approach is the transformation of 3D point cloud data into a 2D format, stored in a **32 × 32 × 12 matrix**. This compact representation allows CNN to process 3D information efficiently without requiring parallel CNN structures for multi-view analysis. An eight-layer CNN is then employed for feature extraction, leveraging its capability to learn spatial hierarchies from the 2D-projected data. Unlike conventional CNN-based classifiers, which rely on fully connected layers, this method utilizes an SVM with a polynomial kernel for classification. The SVM enhances generalization by effectively handling high-dimensional feature spaces, leading to improved accuracy in recognizing 3D shapes.

The CNN-SVM hybrid model offers several advantages over traditional approaches. By utilizing a single 2D projection of interpolated point clouds, the method captures spatial information effectively while reducing the number of model parameters. This not only enhances computational efficiency but also eliminates the need for complex network architectures. The use of SVM further refines the classification process, making it more robust against variations in shape orientation and noise. Experimental validation on the Model Net dataset demonstrates that the proposed method outperforms existing CNN-based and voxel-based models in terms of accuracy and efficiency.

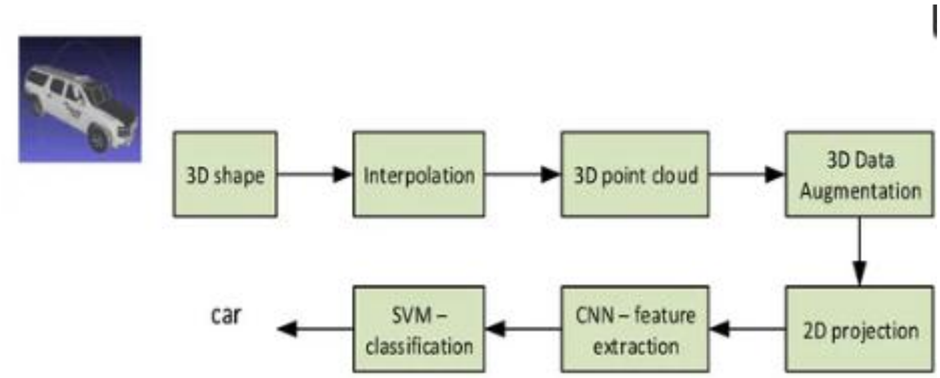
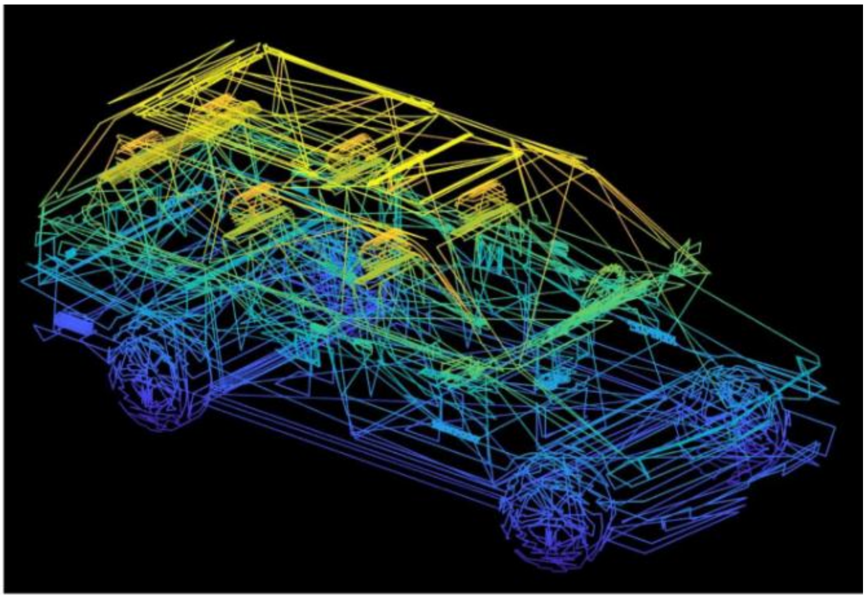


Figure 3: The proposed method of hybrid model



.

Figure 4: The interpolation Point Clouds

2.3 Summary

This chapter provides a detailed analysis of the approaches that have been used to solve this problem in the past. The next chapter discusses the technologies that are expected to be used in the implementation of the system and why those technologies are chosen.

CHAPTER 3

TECHNOLOGY ADAPTED

3.1 Introduction

This chapter elaborates on a description of technologies used during the implementation of the system. There are different reasons to use the following technologies for our implementation.

3.2 Technologies Adopted for Implementation

This section will be mainly focused on the technologies which will be used throughout the development of the proposed system, AI-Driven Educational Solutions for Visually Impaired Students in Sri Lanka. Through this research as a programming language, we will be using Python and Python Libraries such as OpenCV, NumPy, etc. Apart from that, we will be using Natural Language Processing and Convolutional Neural Networks such as YOLOv7 for our research. The following are some brief descriptions about some of the basic technologies which will be used in this research.

3.2.1 Programming Language

Python

Python is a popular, high-level programming language that is widely used in various fields such as web development, data analysis, machine learning, and scientific computing. It is known for its simplicity, readability, and flexibility, which makes it a good choice for beginners and experienced programmers alike. Some key features of Python include:

- Dynamically typed: Python is a dynamically typed language, which means that the type of a variable is determined at runtime rather than at compile-time. This allows for greater flexibility in coding, but it can also make it more difficult to catch certain types of errors.
- Object-oriented: Python is an object-oriented language, which means that it is based on the concept of "objects" that contain both data and methods (functions) that operate on that data. This makes it easier to modularize and reuse code.

- Rich standard library: Python has a large and comprehensive standard library that includes a wide range of functionality, from basic data types and control structures to advanced features such as regular expressions and network programming.
- Third-party libraries: In addition to the standard library, Python has a large and active community of developers who have created a vast array of third-party libraries that extend the functionality of Python in many different areas. These libraries can be easily installed and used in Python programs.
- Interoperability: Python can be easily integrated with other languages and tools, making it a good choice for building hybrid applications that combine the strengths of different technologies.

Overall, Python is a powerful and versatile programming language that is widely used in many different fields.

3.2.2 Development Tools

- **PyCharm**

PyCharm, a Python-focused Integrated Development Environment (IDE) crafted by JetBrains, offers a suite of powerful tools tailored for efficient Python programming. It features an intelligent code editor equipped with syntax highlighting, auto-completion, and code analysis for streamlined coding. PyCharm excels in testing and debugging, facilitating test creation, code debugging, and seamless integration with version control systems like Git. Additionally, it supports popular web frameworks like Django and Flask, simplifying web development tasks. With its user-friendly interface and customizable features, PyCharm stands out as a versatile and efficient IDE for Python developers.

- **Google Colab**

Google Colab offers a user-friendly, cloud-based platform for Python coding. It simplifies code execution, supports collaboration among multiple users, leverages cloud computing resources, and seamlessly integrates with Google Drive. It's a convenient and accessible environment for Python development, ideal for collaborative projects and hassle-free coding.

3.2.3 Libraries

- **OpenCV Library**

OpenCV (Open-Source Computer Vision) is a free and open-source library of computer vision and machine learning algorithms that are widely used for image and video processing tasks. OpenCV provides a range of functions for image pre-processing, including filtering, color space conversions, and feature detection.

- **Pandas**

Pandas is an open-source Python package for data analysis and manipulation. Data scientists and analysts choose it because it offers strong data structures and tools for data analysis. Large datasets may be handled well with the Data Frame object from Pandas, a two-dimensional table-like data structure. It enables users to carry out several tasks, including data aggregation, sorting, grouping, and filtering. Tasks like data cleansing, combining, and reshaping are also supported by Pandas. Pandas facilitate data manipulation and exploration with its vast capability and easy syntax, making it a great tool for Python data professionals dealing with structured data.

- **NumPy**

NumPy is an open-source, cross-platform toolkit that enables processing multi-dimensional arrays and provides mathematical operations that may be used on the data contained in arrays, such as linear algebra, the Fourier transform, and matrices. While working with lists is an option in Python, arrays are used more often because of their faster processing speed compared to lists. In order to store the pixel values of the photos utilizing data structures and process them suitably and effectively, NumPy is heavily used in this research study.

- **TensorFlow**

TensorFlow is an open-source library developed by Google Brain for machine learning and deep learning tasks. It provides a computational framework based on data flow graphs, supporting various model architectures. TensorFlow is known for its scalability and distributed training capabilities. It offers a high-level API called Keras for simplified model development and includes pre-trained models and tools for visualization. TensorFlow is widely used in applications like image and speech recognition, natural language processing, and recommender systems. Overall, it is a popular and powerful library for artificial intelligence tasks.

- **Matplotlib.pyplot**

Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. It is widely used for generating high-quality 2D and 3D plots, including line plots, bar charts, scatter plots, histograms, and more. Within Matplotlib, the pyplot module provides a MATLAB-like interface for easily plotting graphs and charts.One of the key advantages of Matplotlib is its extensive customization options, allowing users to modify colors, labels, titles, and annotations to create visually appealing and informative graphics. Additionally, it integrates seamlessly with other popular Python libraries like pandas and numpy, making it an essential tool for data analysis and visualization.

- **MediaPipe**

MediaPipe is an open-source framework developed by Google for building real-time machine learning pipelines, particularly for computer vision tasks such as hand tracking, face detection, pose estimation, and object recognition. It is optimized for efficiency, supporting multiple platforms including Android, iOS, Windows, and Linux, while running seamlessly on CPUs, GPUs, and edge devices. MediaPipe provides pre-built models and low-latency processing, making it ideal for applications in augmented reality, virtual reality, fitness tracking, and robotics. Its lightweight architecture and cross-platform compatibility allow developers to integrate advanced AI-powered vision capabilities with minimal effort and high performance.

- **PyAudio**

PyAudio is a Python library that provides bindings for PortAudio, enabling seamless audio input and output across multiple platforms. It allows developers to work with real-time audio streams, making it ideal for applications such as speech recognition, sound processing, music analysis, and voice communication systems. PyAudio simplifies handling audio data by providing an easy-to-use interface for recording, playing, and manipulating audio in various formats. Its support for both blocking and non-blocking (callback-based) modes makes it flexible for real-time audio applications, including AI-driven speech analysis, audio sonification, and interactive voice assistants. Due to its efficiency and cross-platform support, PyAudio is widely used in research, multimedia applications, and machine learning projects involving sound processing.

- **Open3D**

Open3D is an open-source library designed for 3D data processing, offering efficient tools for working with point clouds, meshes, and RGB-D images. It provides a flexible and high-performance framework for tasks such as 3D reconstruction, shape analysis, object recognition, and visualization. Open3D is widely used in computer vision, robotics, augmented reality, and AI-based 3D modeling due to its optimized algorithms for point cloud registration, surface reconstruction, and deep learning integration. It supports multiple platforms and can interface with libraries like TensorFlow and PyTorch for advanced machine learning applications. With its user-friendly API and GPU acceleration, Open3D is an essential tool for researchers and developers working with 3D spatial data.

- **SciPy**

SciPy is an open-source Python library built on NumPy, providing advanced mathematical, scientific, and engineering functions. It is widely used for numerical computing, signal processing, optimization, statistical analysis, and machine learning. SciPy offers efficient algorithms for linear algebra, interpolation, integration, and differential equations, making it essential for scientific computing and data analysis. Its performance-optimized functions allow researchers and engineers to handle complex mathematical operations with ease. SciPy is particularly useful in fields like physics, bioinformatics, finance, and AI, where high-precision computations and modeling are required. With its extensive functionality and seamless integration with other Python libraries, SciPy is a powerful tool for solving scientific and engineering problems efficiently.

3.3 Summary

This chapter gives an overview of the technologies that will be used in the implementation of the system and why those technologies are chosen. Next chapter illustrates the approach of implementing the system.

CHAPTER 4

OUR APPROACH

4.1 Introduction

In the previous chapter, we discussed the technologies adopted to implement the proposed solution. The solution can be interpreted in terms of users, inputs, processes, and outputs and the system implementation approaches for each module will be discussed.

4.2 Proposed system

The AI-Enhanced Educational Framework for Visually Impaired Students is a comprehensive system designed to address the unique challenges faced by 9th-grade visually impaired students in Sri Lanka, particularly in mathematics. The framework integrates four specialized modules, each leveraging advanced AI technologies to create an accessible, personalized, and effective learning environment. These modules work together to ensure that students with varying degrees of visual impairment can engage with mathematical concepts in a way that is tailored to their individual needs and abilities.

1. **Personalized Learning Model**: This module tailors educational content based on the student's visual impairment state (e.g., blind since birth, partially sighted, or fully blind due to later-life causes) and their past academic performance (e.g., average, below average, or high-performing). It uses AI-driven content generation, multi-modal content delivery (audio, high-contrast visuals, Braille), adaptive learning paths, and NLP for Sinhala language support to create a customized learning experience.

2. **Specialized Question and Answer (Q&A) Module**: This module provides tailored responses to mathematical questions based on the student's type of visual impairment. It incorporates multi-sensory features like auditory descriptions and tactile materials, along with NLP for intuitive interactions and real-time feedback to enhance comprehension and engagement.

3. **AI-Driven Sonification and Natural Language Generation for Graph Visualization**: This module transforms visual data (e.g., graphs) into auditory

cues using sonification and natural language generation. It helps blind students understand spatial and graphical concepts by translating numerical data into sound parameters and providing real-time descriptions of graph features.

4. **AI-Driven Real-Time 3D Shape Detection and Position based description Using Finger Tracking and Video Processing**: This module uses video analysis and AI to provide real-time auditory guidance as students explore geometric shapes through touch. It tracks finger movements, offers feedback on shape properties (e.g., angles, edges), and corrects errors, enabling students to understand geometric concepts without visual input.

4.3 Users

The system is designed for visually impaired 9th-grade students in Sri Lanka as primary users, with teachers, parents, and educational institutions as secondary users. Developers and content creators support the system's technical and educational aspects.

- Primary Users:
  Visually Impaired Students: 9th-grade students in Sri Lanka with varying degrees of visual impairment (blind since birth, partially sighted, or fully blind due to later-life causes). These students interact with the system to learn mathematical concepts through personalized, multi-sensory content.
- Secondary Users:
  - Teachers and Educators: Teachers use the system to monitor student progress, customize learning paths, and provide additional support. They can also use the system to generate accessible teaching materials.
  - Parents and Guardians: Parents can use the system to support their children's learning at home, track progress, and understand their child's educational needs.
  - Educational Institutions: Schools and educational organizations use the system to implement inclusive education practices and ensure compliance with accessibility standards.
- Technical Users:
  - Developers and AI Engineers: Responsible for maintaining, updating, and improving the system's AI algorithms, NLP models, and assistive technologies.

- Content Creators: Professionals who develop and adapt educational materials for visually impaired students, ensuring they are compatible with the system's multi-modal delivery methods.

4.4 Inputs and Outputs

The system is designed for visually impaired 9th-grade students in Sri Lanka as primary users, with teachers, parents, and educational institutions as secondary users. Developers and content creators support the system's technical and educational aspects.

Table 1: Inputs and outputs of each module

Module	Inputs	Outputs
1. Personalized Learning Model	• Students' visual impairment state (e.g., blind since birth, partially sighted, fully blind due to later-life causes). • Past academic performance data (e.g., test scores, grades). • Learning preferences (e.g., preference for audio, Braille, or high-contrast visuals). • Real-time engagement and performance data during learning sessions.	• Customized educational content (audio, Braille, or high-contrast visuals). • Adaptive learning paths based on performance and engagement. • Sentiment analysis and emotional state feedback.
2. Specialized Q&A Module	• Student's type of visual impairment (e.g., fully blind, partially blind). • Mathematical questions posed by the student (in Sinhala or English).	• Tailored responses with auditory descriptions or tactile materials. • Real-time feedback and explanations to enhance understanding.

	• Contextual information about the student's current learning topic.	• Adaptive learning paths based on question difficulty and student performance.
3. AI-Driven Sonification and Natural Language Generation for Graph Visualization	• Numerical data or graph data (e.g., coordinates, values). • Student's preferences for auditory cues (e.g., pitch, volume, tempo). • Contextual information about the graph (e.g., type of graph, key features).	• Sonified data (e.g., pitch changes for data trends). • Real-time textual or spoken descriptions of graph features. • Pattern recognition and key insights delivered through sound.
4. AI-Driven Real-Time 3D Shape Detection and Position based description Using Finger Tracking and Video Processing	• Video input of the student's hand and finger movements on a tactile shape. • Geometric shape data (e.g., type of shape, edge lengths, angles).	• Real-time auditory feedback (e.g., "You are approaching a 90-degree curve"). • Corrective guidance if the student moves off the shape. • Descriptions of shape properties (e.g., parallel edges, perpendicular intersections).

4.5 Process

This section explains how each module processes the inputs to deliver the desired outputs. It involves AI algorithms, multi-sensory integration, and real-time adaptations to create personalized, accessible, and effective learning experiences for visually impaired students. Each module uses advanced technologies like machine learning, NLP, sonification, and computer vision to transform raw input into meaningful educational interactions.

Table 2: Processes of each module

Module	Process
1. Personalized Learning Model	• Language Support: Process questions in Sinhala using NLP for intuitive interactions. • Adaptive Responses: Generate answers tailored to the student's visual impairment and learning context. • Real-Time Feedback: Provide immediate responses and explanations to enhance comprehension.
2. Specialized Q&A Module	• Language Support: Process questions in Sinhala using NLP for intuitive interactions. • Adaptive Responses: Generate answers tailored to the student's visual impairment and learning context. • Real-Time Feedback: Provide immediate responses and explanations to enhance comprehension.
3. AI-Driven Sonification and Natural Language Generation for Graph Visualization	• Automated Mapping: Translate numerical data into auditory cues (e.g., pitch, volume, tempo). • Pattern Recognition: Identify trends and patterns in the data and convert them into recognizable auditory cues. • Natural Language Generation (NLG): Generate real-time textual or spoken descriptions of graph features.
4. AI-Driven Real-Time 3D Shape Detection and	• Shape Detection: Identify the type of shape (e.g., triangle, cube) and initialize tracking.

Position based description Using Finger Tracking and Video Processing	• Finger Tracking: Use computer vision (e.g., MediaPipe) to track finger movements on the shape. • Real-Time Guidance: Provide auditory feedback on shape properties (e.g., angles, edges) as the student traces the shape. • Correction and Encouragement: Offer corrective feedback if the student moves off the shape and encourages exploration.

4.6 Summary

This chapter explains our approach to the proposed solution. We describe the input, output and process of the system that we are going to implement. In the next chapter, we describe on analysis and design part of our proposed solution.

CHAPTER 5

ANALYSIS AND DESIGN

5.1 Introduction

In the previous chapter, our approach to the project in terms of users, inputs, outputs, features and process were discussed. This chapter presents the top-level architecture diagram of the system and describes each module of the diagram indicating how each module interacts with other modules.

5.2 High-level Architecture of the System

The architecture of the AI-Enhanced Educational Framework for Visually Impaired Students is designed as an integrated, modular system that leverages advanced AI technologies to deliver personalized and accessible learning experiences. The system consists of four core modules, Personalized Learning, Specialized Q&A, Sonification for Graph Visualization, and Real-Time Shape Recognition. Each module interacts seamlessly to address the unique needs of visually impaired students.

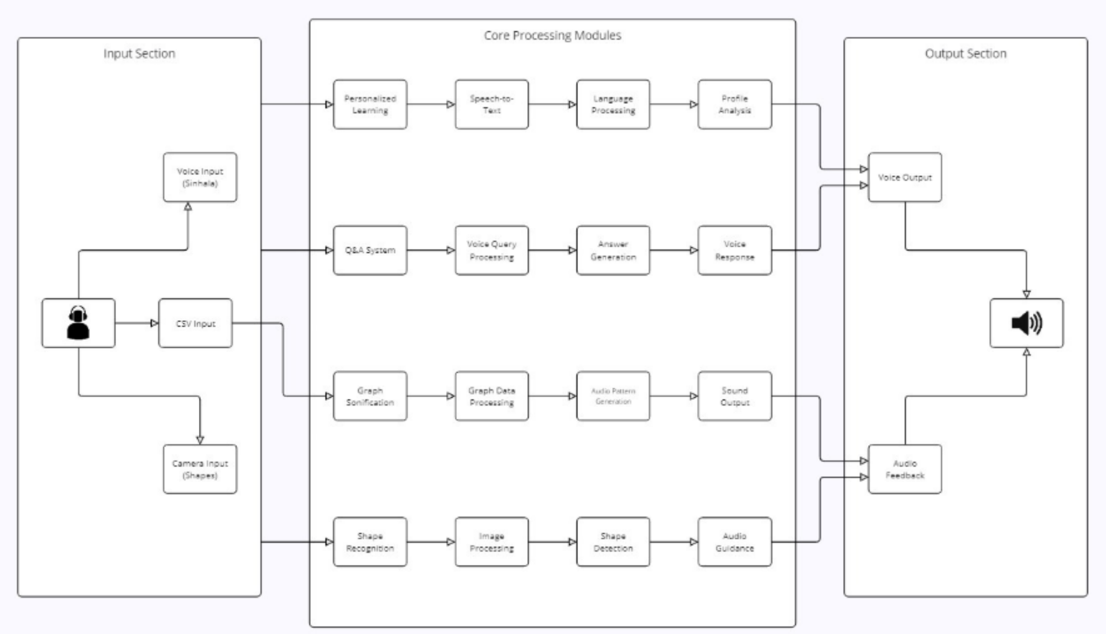


Figure 5: High-level Architecture of the System

5.3 High-level Architecture of the Modules

5.3.1 Module 01: Personalized Learning Model for Visually Impaired Students.

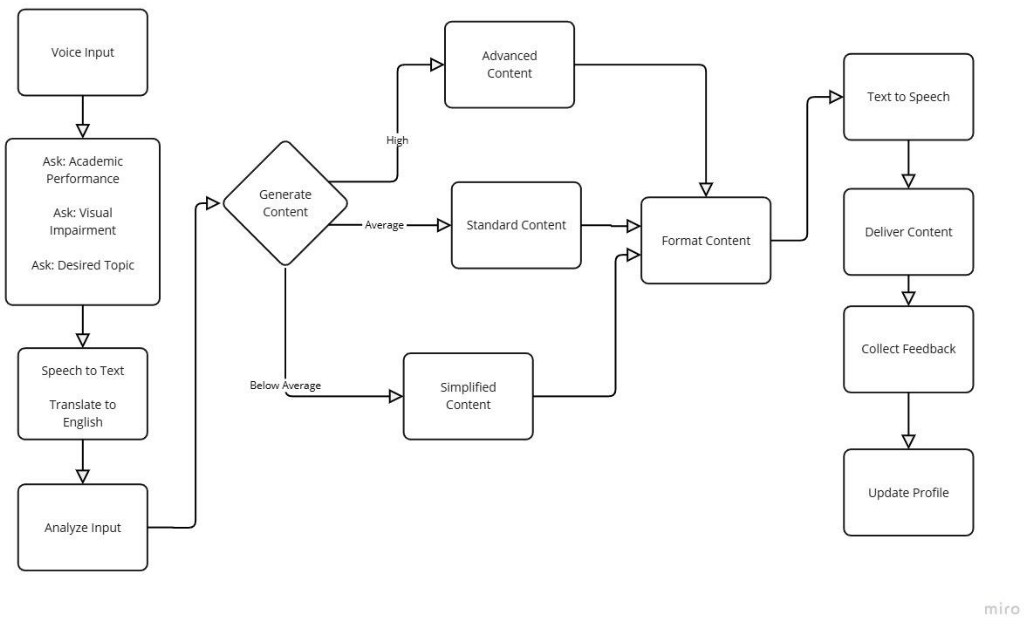


Figure 6: High-level Architecture of Module 1

The process of *Personalized Learning Model for Visually Impaired Students* begins with the student providing their name, allowing the system to retrieve their profile and identify their visual impairment state (e.g., blind since birth, partially sighted). The system then asks the student to specify the topic they wish to learn. Once the topic is provided, the system uses a Content Generation Model to analyze the knowledge base and retrieve relevant educational materials. This content is first translated into Sinhala using an English-to-Sinhala Translator API to ensure cultural relevance and accessibility. Finally, the translated text is converted into audio through a Text-to-Speech Model, delivering personalized, auditory learning materials tailored to the student's needs. This seamless flow ensures an inclusive and effective learning experience for visually impaired students.

5.3.2 Module 02: Specialized Question and Answer Module for Visually Impaired Students.

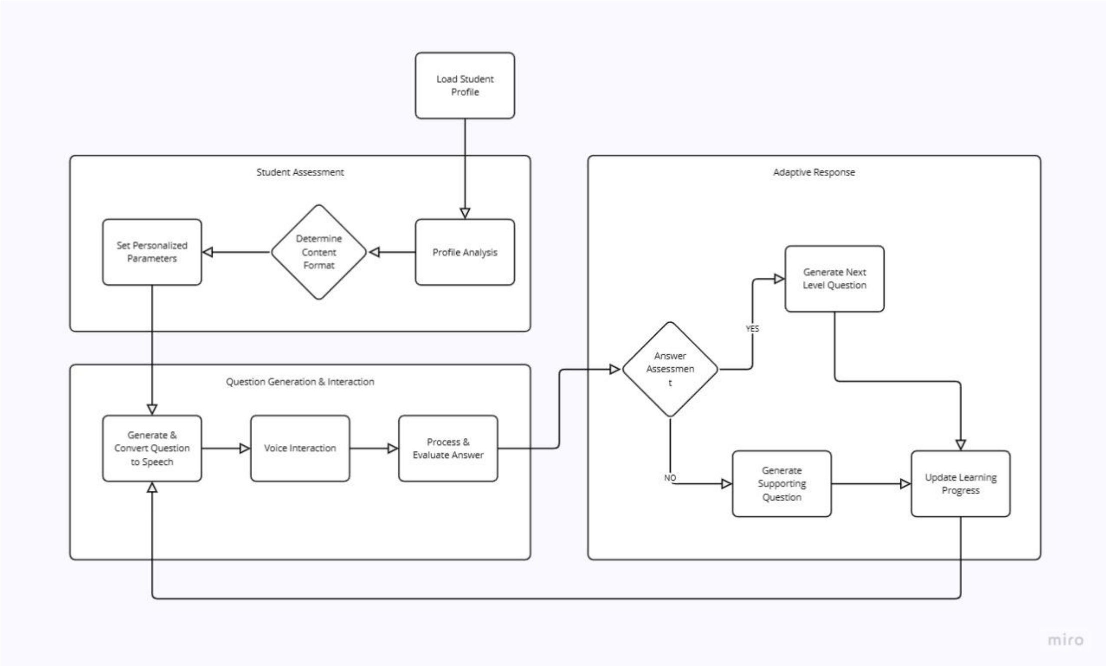


Figure 7: High-level Architecture of Module 2

The Specialized Question and Answer (Q&A) Module begins by generating a tailored question based on the student's visual impairment state (e.g., blind since birth, partially sighted) identified in Module 1. After the student answers the question, the system checks the correctness of the response using an AI-based validation model and generates a detailed explanation. This explanation is then translated into Sinhala via an English-to-Sinhala Translator API to ensure cultural relevance and accessibility. Finally, the translated explanation is converted into audio using a Text-to-Speech Model, delivering a clear, auditory explanation to the student. This seamless flow ensures an interactive and accessible learning experience, allowing visually impaired students to engage with mathematical concepts effectively.

5.3.3 Module 03: AI-Driven Sonification and Natural Language Generation for Accessible Mathematical Graph Visualization for blind students

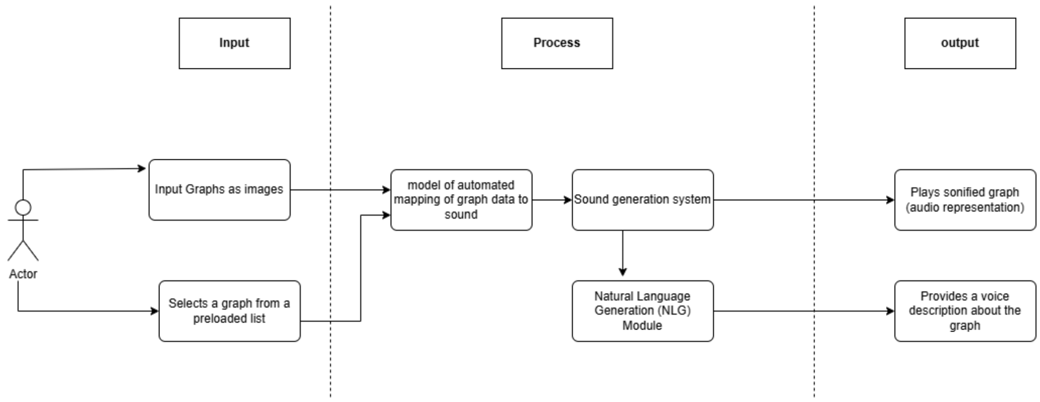


Figure 8: High-level Architecture of Module 3

In the third module of the proposed system, Users can input a function via voice, upload a graph image, or select from preloaded graphs. The Graph Processing Module extracts and analyzes the graph's key features, such as slope and intercepts. These features are then translated into sound through the Sonification Module, mapping the X-axis to time and the Y-axis to pitch, allowing users to "hear" the graph's shape. Simultaneously, the NLG Module generates a spoken description of the graph's characteristics, which is delivered through Text-to-Speech (TTS). This multi-modal approach makes mathematical graphs more accessible by offering both auditory and verbal representations.

5.3.4 Module 04: AI-Driven Real-Time 3D Shape Detection and Position based description Using Finger Tracking and Video Processing

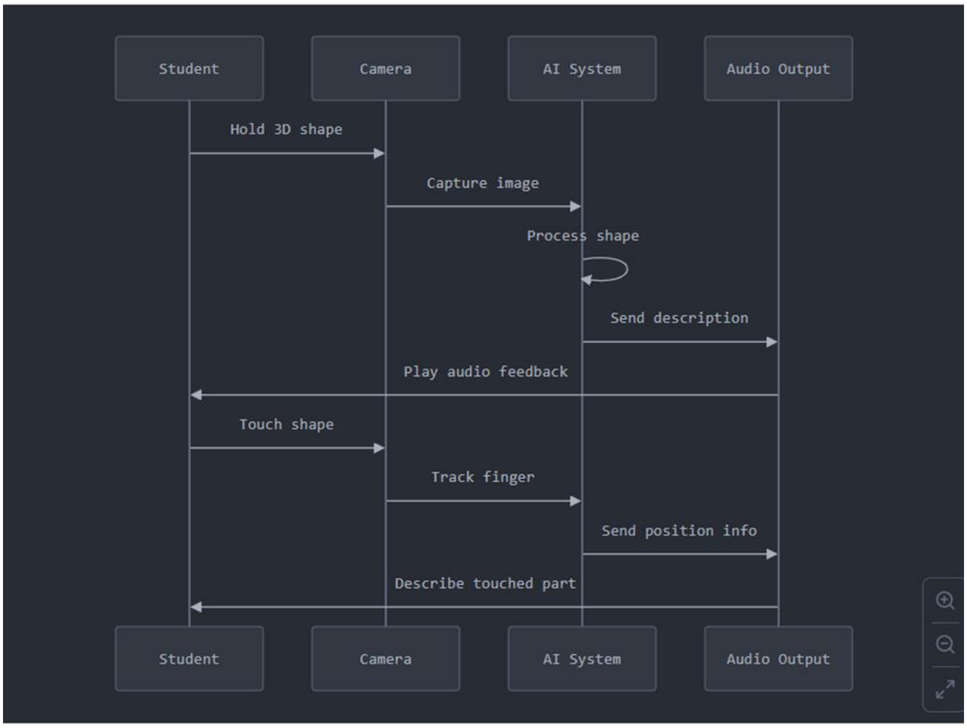


Figure 9: High-level Architecture of Module 4

The process begins with the student holding a physical 3D shape, which is then captured by the camera and sent to the AI system for processing. The AI analyzes the image to identify the shape, its structure, and features, then generates a textual description that is relayed to the audio output system. The student listens to the description and interacts with the shape by touching different parts to explore its details. As the student moves their finger across the surface, the camera tracks the movement and sends the finger position data to the AI system. The AI determines the specific part of the shape being touched and generates a real-time verbal description, which is then played through the audio output. This interactive feedback loop enables tactile and auditory learning, helping students, particularly those with visual impairments, to better understand geometric shapes by combining touch-based exploration with AI-driven audio guidance.

5.4 Summary

In this chapter we discussed how each module in the proposed system is designed to address specific challenges faced by visually impaired students, leveraging advanced AI technologies like NLP, sonification, and computer vision. Together, these modules create a comprehensive, multi-sensory learning environment that ensures accessibility, personalization, and engagement for visually impaired students in mathematics.

CHAPTER 6

IMPLEMENTATION

6.1 Introduction

The analysis and design of the solution were discussed in the previous chapter. This chapter discusses the implementation details of approaches used, techniques, tools, and models based on each module.

6.2 Implementation of each module

6.2.1 Module 01: Personalized Learning Model for Visually Impaired Students

The Personalized Learning Model module is currently in its initial training phase, centered on fine-tuning the LLAMA 3 model using the Unsloth optimization framework. We are currently assembled a comprehensive dataset combining mathematical content from Kaggle and Hugging Face, specifically focused on fundamental mathematics concepts aligned with the Sri Lankan 9th-grade curriculum.

Our development efforts are focused on training the model to process basic mathematical concepts and generate responses tailored to different visual impairment categories. The implementation utilizes LLAMA 3 as the base model, with ongoing work to integrate Sinhala language support. Current challenges include optimizing real-time content generation performance and developing appropriate evaluation metrics for visually impaired user scenarios.

The module is progressing according to schedule and represents a crucial component of our larger framework aimed at providing accessible, personalized mathematical education for visually impaired students in Sri Lanka. Next steps include completing the initial training phase, implementing evaluation metrics, and beginning integration with the multi-modal content delivery system.

- **Dataset loading into unsloth library.**

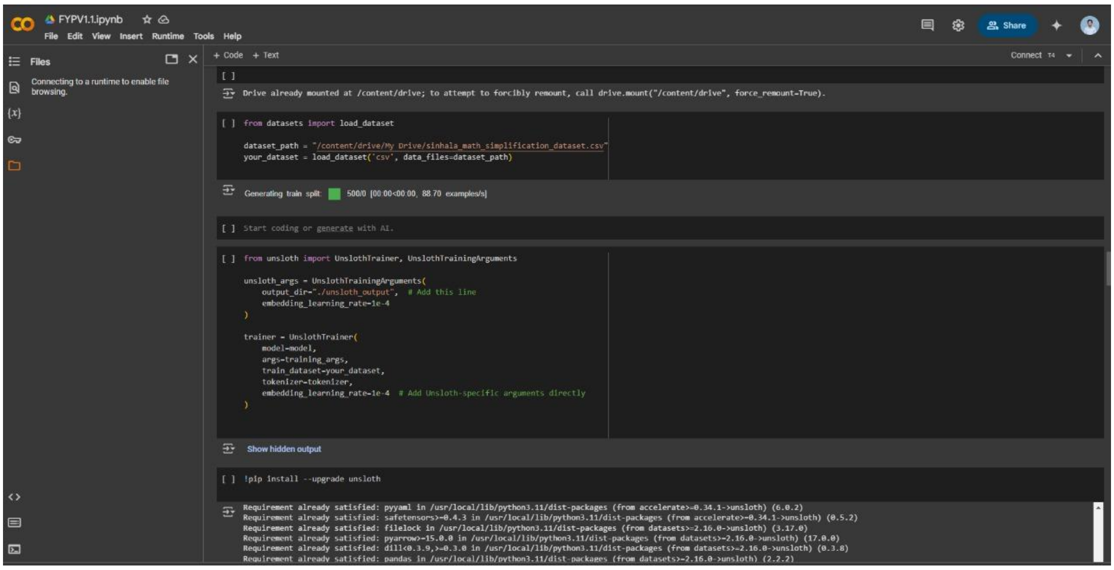


Figure 10: Screenshot of dataset loading into unSloth library for module 01

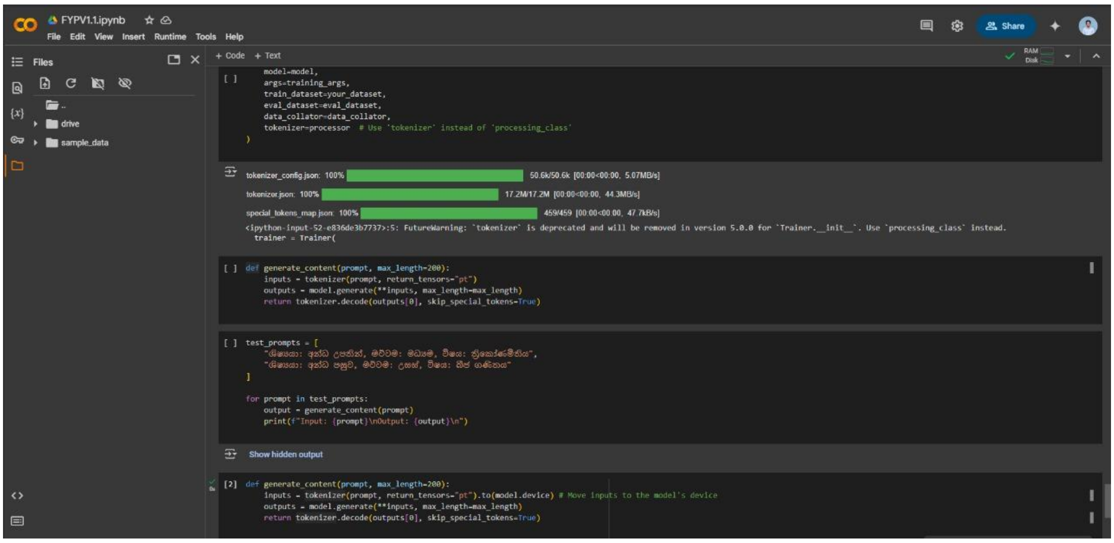


Figure 11: Screenshot of Model testing in module 01

6.2.2 Module 02: Specialized Question and Answer Module for Visually Impaired Students

The development of the Specialized Question and Answer Module for visually impaired students has made significant progress in its initial phase. The foundation of this module has been established through the strategic collection and curation of mathematical content from both Kaggle and Hugging Face repositories, specifically focusing on fundamental mathematical concepts. This educational content forms the cornerstone of the module's knowledge base, ensuring alignment with the core mathematical principles that ninth-grade students need to master.

A notable advancement in the module's development is the implementation of the Llama 3 model, which has been pre-trained using the Unsloth framework to optimize performance. This pre-training process has been specifically tailored to generate contextually relevant questions and answers in mathematics education. The training approach focuses on creating responses that are not only academically accurate but also accessible and comprehensible for visually impaired students, taking into account their unique learning needs and challenges in mathematical concept visualization.

The module's current architecture incorporates specialized features designed to address the diverse needs of visually impaired learners. By leveraging the pre-trained Llama 3 model, the system can generate questions and answers that adapt to different levels of visual impairment, ensuring that mathematical concepts are conveyed effectively regardless of the student's specific visual limitations. This adaptive approach represents a significant step forward in creating an inclusive educational tool that can cater to various learning styles and accessibility requirements.

As the development continues, the focus remains on refining the question-answer generation capabilities and ensuring seamless integration with other modules of the AI-Enhanced Mathematics Learning Framework.

- **Dataset loading into unsloth library.**

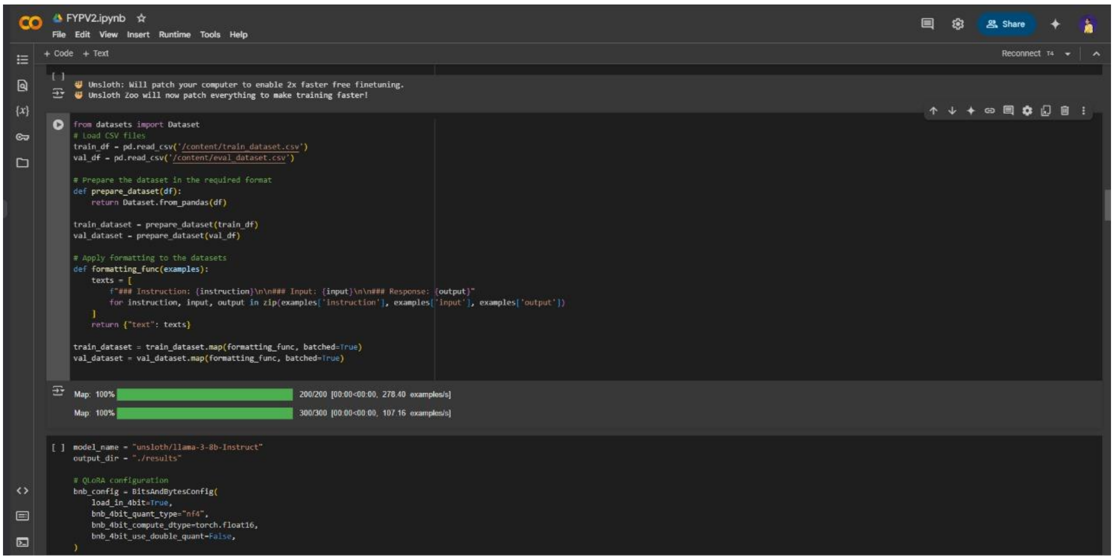


Figure 12: Screenshot of dataset loading into unSloth library for module 02

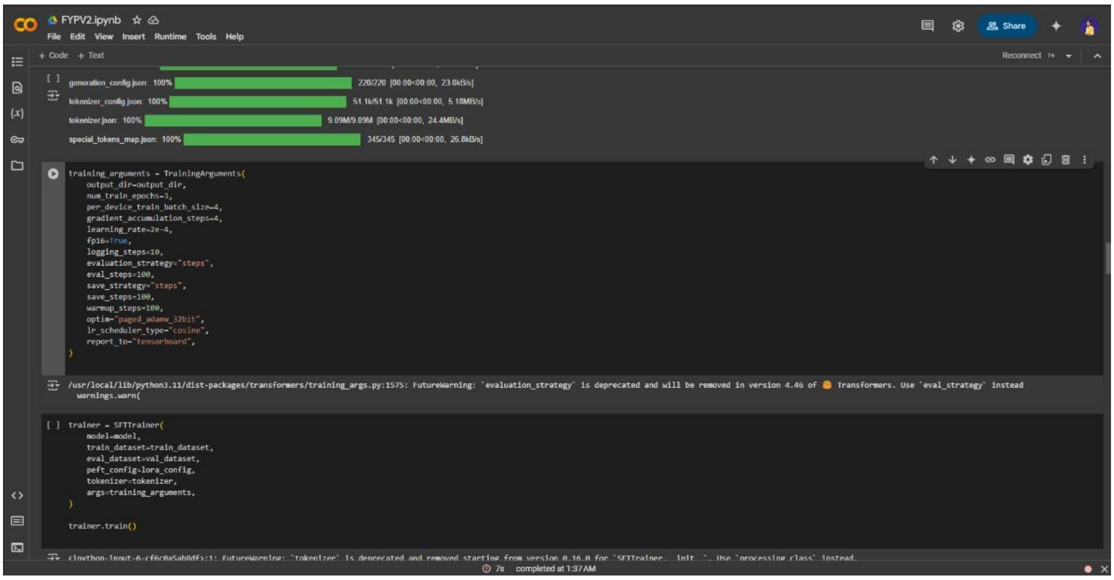
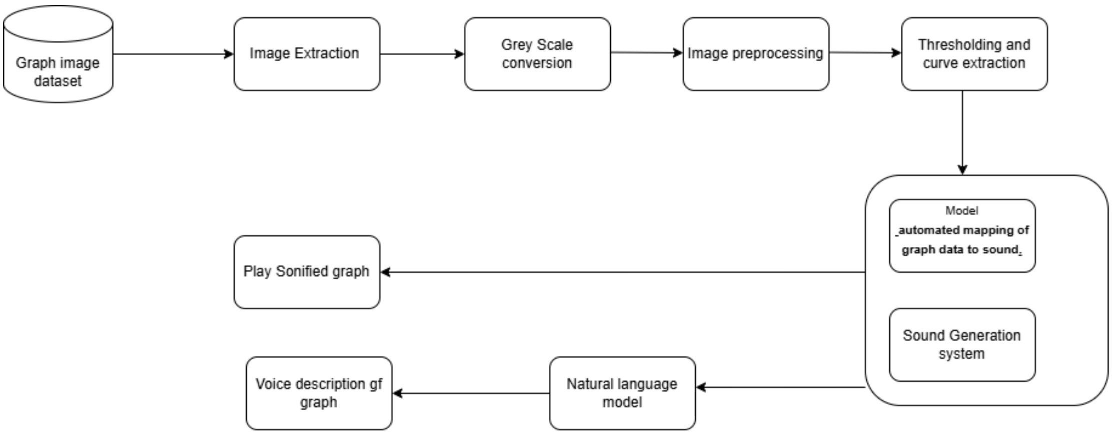


Figure 13: Screenshot of Model testing in module 02

6.2.3 Module 03: AI-Driven Sonification and Natural Language Generation for Accessible Mathematical Graph Visualization for blind students



- **Import Dataset to googleColab**

```
!pip install kagglehub
```

```
Requirement already satisfied: kagglehub in /usr/local/lib/python3.11/dist-packages (0.3.9)
Requirement already satisfied: packaging in /usr/local/lib/python3.11/dist-packages (from kagglehub) (24.2)
Requirement already satisfied: pyyaml in /usr/local/lib/python3.11/dist-packages (from kagglehub) (6.0.2)
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from kagglehub) (2.32.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-packages (from kagglehub) (4.67.1)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->kagglehub) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests->kagglehub) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests->kagglehub) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests->kagglehub) (2025.1.31)
```

```
[7] import kagglehub
    print("KaggleHub installed successfully!")
```

```
KaggleHub installed successfully!
```

```
[8] from google.colab import drive
    drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```
[9] import kagglehub

    # Download latest dataset version
    dataset_path = kagglehub.dataset_download("kopfgeldjaeger/function-graphs-polynomial")

    print("Path to dataset files:", dataset_path)
```

```
Downloading from https://www.kaggle.com/api/v1/datasets/download/kopfgeldjaeger/function-graphs-polynomial?dataset_version_number=1...
100%|██████████| 81.0M/81.0M [00:00<00:00, 89.8MB/s]Extracting files...

Path to dataset files: /root/.cache/kagglehub/datasets/kopfgeldjaeger/function-graphs-polynomial/versions/1
```

```
[10] import os

     # Check downloaded dataset files
     dataset_files = os.listdir(dataset_path)
     print("Files in dataset:", dataset_files)
```

```
Files in dataset: ['function_graphs', 'function_graphs.csv']
```

```
import os

# Path to the dataset root
dataset_path = "/path/to/dataset"  # Replace with your actual path
function_graphs_dir = os.path.join(dataset_path, "function_graphs")

# List immediate contents of the function_graphs directory
if os.path.exists(function_graphs_dir):
    print("Contents of 'function_graphs':")
    for item in os.listdir(function_graphs_dir):
        item_path = os.path.join(function_graphs_dir, item)
        if os.path.isdir(item_path):
            print(f"[Directory] {item}")
        else:
            print(f"[File] {item}")
else:
    print("Directory 'function_graphs' not found!")
```

```
Directory 'function_graphs' not found!
```

- **Extracting Curve Points create a binary image**

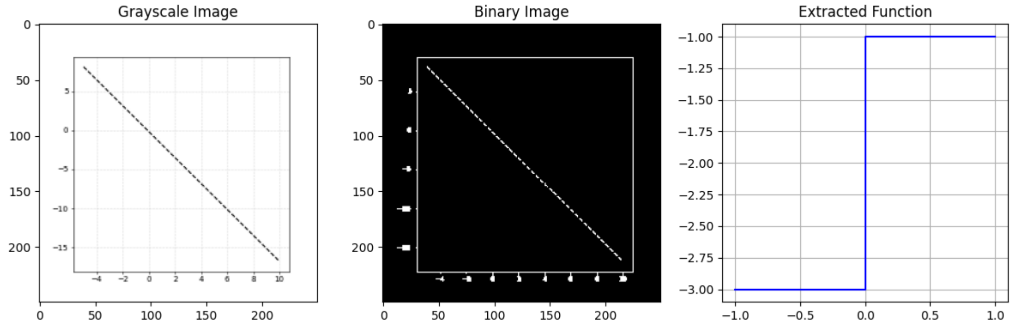
```python
def _extract_curve_points(self, binary_img):
    """Extract (x,y) points from binary image"""
    y_coords, x_coords = np.nonzero(binary_img)
    points = np.column_stack((x_coords, y_coords))

    points = points[points[:, 0].argsort()]

    # Get unique x values and corresponding y values
    unique_x = np.unique(points[:, 0])
    curve_points = [[x, np.min(points[points[:, 0] == x, 1])] for x in unique_x]

    return np.array(curve_points)

def _normalize_points(self, points):
    """Manually normalize points to [-1, 1]"""
    points[:, 0] = (points[:, 0] - points[:, 0].min()) / (points[:, 0].max() - points[:, 0].min()) * 2 - 1
    points[:, 1] = (points[:, 1].min() - points[:, 1]) / (points[:, 1].max() - points[:, 1].min()) * 2 - 1
    return points

def _visualize_steps(self, gray, binary, points):
    """Visualize processing steps"""
    fig, axes = plt.subplots(1, 3, figsize=(12, 4))

    axes[0].imshow(gray, cmap='gray')
    axes[0].set_title('Grayscale Image')

    axes[1].imshow(binary, cmap='gray')
    axes[1].set_title('Binary Image')

    axes[2].plot(points[:, 0], points[:, 1], 'b-')
    axes[2].grid(True)
    axes[2].set_title('Extracted Function')

    plt.tight_layout()
    plt.show()
```

```python
def process_dataset(self, input_dir, csv_path, output_dir):
    """Process entire dataset and match with labels"""
    os.makedirs(output_dir, exist_ok=True)

    # Load labels from CSV
    df = pd.read_csv(csv_path)

    for index, row in df.iterrows():
        img_file = row['filename']
        label = row['label']

        img_path = os.path.join(input_dir, img_file)
        if not os.path.exists(img_path):
            print(f"File not found: {img_file}")
            continue

        try:
            points = self.process_single_image(img_path)

            # Save extracted points
            np.save(os.path.join(output_dir, f"{img_file[:-4]}_points.npy"), points)

            # Save label
            with open(os.path.join(output_dir, f"{img_file[:-4]}_label.txt"), "w") as f:
                f.write(str(label))

        except Exception as e:
            print(f"Error processing {img_file}: {str(e)}")

    print("\nProcessing complete!")
```

- **output**

```
# Display sample outputs for verification
input_dir = "/content/filtered_function_graphs"
processor.display_sample_outputs(input_dir, num_samples=5)
```

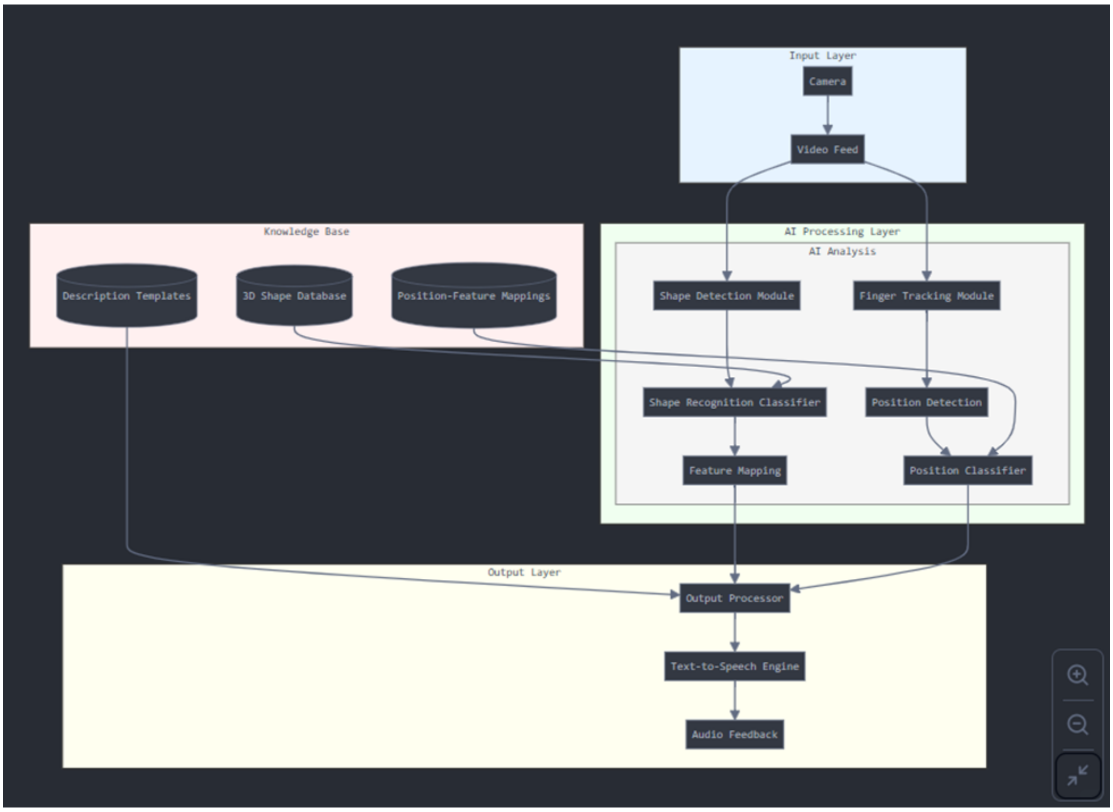
```
Displaying 5 samples from lfunc:

Processing 940_1.png...
```




```
Extracted Points (First 10):
```


```
Extracted Points (First 10):
[[-1 -3]
 [ 0 -3]
 [ 0 -3]
 [ 0 -3]
 [ 0 -3]
 [ 0 -2]
 [ 0 -2]
 [ 0 -2]
 [ 0 -2]
 [ 0 -1]]
Number of points: 210
X range: [-1.00, 1.00]
Y range: [-3.00, -1.00]
```

6.2.4 Module 04: AI-Driven Real-Time 3D Shape Detection and Position based description for Blind students using Finger Tracking and Video Processing



- **Data Collection and Initial Filtering**

```python
class ShapeValidator:
    def extract_features(self, image):
        features = {
            'geometric': self._get_geometric_features(image),
            'statistical': self._get_statistical_features(image)
        }
        return features

    def _get_geometric_features(self, image):
        gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
        edges = cv2.Canny(gray, 50, 150)
        contours, _ = cv2.findContours(edges, cv2.RETR_EXTERNAL, cv2.CHAIN_APPROX_SIMPLE)

        if not contours:
            return None

        largest_contour = max(contours, key=cv2.contourArea)
        approx = cv2.approxPolyDP(largest_contour, 0.02 * cv2.arcLength(largest_contour, True), True)

        return {
            'vertices': len(approx),
            'area': cv2.contourArea(largest_contour),
            'perimeter': cv2.arcLength(largest_contour, True)
        }
```

- **Feature Extraction and Validation**

```python
class ShapeValidator:
    def extract_features(self, image):
        features = {
            'geometric': self._get_geometric_features(image),
            'statistical': self._get_statistical_features(image)
        }
        return features

    def _get_geometric_features(self, image):
        gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
        edges = cv2.Canny(gray, 50, 150)
        contours, _ = cv2.findContours(edges, cv2.RETR_EXTERNAL, cv2.CHAIN_APPROX_SIMPLE)

        if not contours:
            return None

        largest_contour = max(contours, key=cv2.contourArea)
        approx = cv2.approxPolyDP(largest_contour, 0.02 * cv2.arcLength(largest_contour, True), True)

        return {
            'vertices': len(approx),
            'area': cv2.contourArea(largest_contour),
            'perimeter': cv2.arcLength(largest_contour, True)
        }
```

- Advanced Filtering Implementation

```python
def advanced_filtering(self, images, features):
    filtered_indices = []

    # Convert features to array for processing
    feature_array = np.array([[
        f['geometric']['vertices'],
        f['geometric']['area'],
        f['geometric']['perimeter']
    ] for f in features])

    # Normalize features
```

6.3 Summary

This chapter outlines the progress made in developing the AI-Enhanced Educational Framework for Visually Impaired Students up to the interim evaluation. It focuses on the implementation details of the modules that have been partially completed, ensuring consistency between the design and the current state of implementation. This chapter describes the software and hardware utilized, along with flowcharts and code segments that demonstrate the implementation progress.

CHAPTER 7

CONCLUSION

7.1 Introduction

In this chapter, we are going to discuss the current progress, assumptions made, and limitations of the system, and future work in our system.

7.2 Current Progress

To gain a comprehensive understanding of existing methodologies in web-based sonification, we conducted a detailed literature review by referring to various research papers, documentation, and online resources. This review helped us analyze the strengths and limitations of current tools and approaches while identifying key research gaps in interactivity, computational efficiency, and multimodal feedback. In addition to reviewing existing studies, we explored relevant datasets from Kaggle and collected additional data from blind students in specialized schools to ensure the system's real-world applicability.

To build a strong foundation in machine learning technologies before implementing our modules, we performed small practical exercises, experimenting with different models and techniques. This hands-on approach allowed us to refine our understanding of key concepts and select the most suitable methods for our project. Our implementation process was carried out using Google Colab, which provided an efficient and accessible environment for model development. So far, we have successfully completed data preprocessing and data extraction from input sources, laying the groundwork for future implementations.

7.3 Assumptions and Limitations of the system

Assumptions

- The effectiveness of each module relies on the availability of sufficient and diverse training data.
- The system assumes that users will provide consistent and genuine input regarding their visual impairment state.

Limitations

- Collecting suitable datasets specifically for blind students can be challenging due to the limited availability of data
- Accuracy Issues Due to Limited Datasets

7.4 Future Works

The next steps involve completing the remaining modules as outlined in the project's scope, finalizing the system's functionality. Once integrated, feedback will be gathered from blind schools to evaluate the system's effectiveness for visually impaired students. This feedback will help assess how well the system addresses their needs and highlight areas for improvement. Additionally, efforts are being made to enhance the accuracy of each module to ensure better performance and reliability.

7.5 Summary

In this chapter, we provide an update on the progress of our research. We are currently finalizing the main modules of our framework and conducting detailed experiments with datasets, using knowledge from existing projects. This phase involves thorough testing and improvement of the modules to ensure they work well and meet our research goals. These efforts in module development and dataset testing are crucial steps that prepare the groundwork for further analysis and advancements.

References

[1] R. D. V. A. D. e. a. Jayathilaka, "Personal well-being index as a measure of quality of life of diverse groups of people with visual impairment and blindness," p. 1665–1684, 2023.

[2] M. Hamash, H. Ghreir and P. Tiernan, "Breaking through Barriers: A Systematic Review of Extended Reality in Education for the Visually Impaired," Education Sciences, vol. 14, no. 4, p. 365, 2024.

[3] S. N. S. N. &. N. Z. Rosyada, "Mathematical reasoning and self-regulated learning differences by using mathematical literacy-based e-module," Jurnal Elemen, vol. 10, no. 2, pp. 222 - 238, 2024.

[4] Constantinescu et al., "Bring the Environment to Life: A Sonification Module for People with Visual Impairments to Improve Situation Awareness," pp. 50-59, 2020.

[5] L. M. Brown and S. A. Brewster, "Drawing by ear: Interpreting sonified line graphs. ACM Transactions on Applied Perception," vol. 18, no. 3, pp. 1-22, 2021.

[6] "The Importance of Accessible Education for Visually Impaired Students," 05 November 2024. [Online]. Available: https://sebetaschoolforblind.com/the-importance-of-accessible-education-for-visually-impaired-students/.

[7] W. Hu, K. Wang, K. Yang, R. Cheng, Y. Ye, L. Sun and Z. Xu, "A Comparative Study in Real-Time Scene Sonification for Visually Impaired People," vol. 20, no. 11, p. 3222, 2020.

[8] Z. Temesgen, "School Challenges of Students with Visual Disabilities," INTERNATIONAL JOURNAL OF SPECIAL EDUCATION, vol. 33, no. 3, 2018.

[9] Tsouktakou and A. Hamouroudis, "The use of artificial intelligence in the education of people with visual impairment," World Journal of Advanced Engineering Technology and Sciences,, vol. 13, no. 01, pp. 734-744, 2024.

[10] C. F. M. G. B. M. R. S. G. M. E. R. F. &. C. K. M. Alves, "Assistive technology applied to education of students with visual impairment.," Revista Panamericana de Salud Pública, vol. 26, pp. 148-152, 2009.

[11] Academics, "The effective methods of teaching mathematics to visually impaired students," [Online]. Available: https://www.myprivatetutor.co.za/blog/the-effective-methods-of-teaching-mathematics-to-visually-impaired-students.

[12] "Geometry Resources for Students with Visual Impairment," [Online]. Available: https://www.pathstoliteracy.org/resource/geometry-resources-students-visual-impairment/.

[13] S. A. A. D. P. D. S. K. P. P. R. Sohom Mukherjee, "Fingertip detection and tracking for recognition of air-writing in videos," Expert Systems with Applications, vol. 136, pp. 217-229, 2019.

[14] H. Rintaro et al., "Artificial intelligence using convolutional neural networks for real-time detection of early esophageal neoplasia in Barrett’s esophagus (with video)," Gastrointestinal Endoscopy, vol. 91, no. 6, pp. 1264-1271, 2020.

[15] M. a. H. R. Aradhya, "Object Detection and Tracking using Deep Learning and Artificial Intelligence for Video Surveillance Applications," International Journal of Advanced Computer Science and Applications(IJACSA), vol. 10, no. 12, 2019.

[16] Roy, T., & Boppana, L. (2022). Interactive web-based image and graph analysis using Sonification for the Blind. *2017 IEEE Region 10 Symposium (TENSYMP)*, 1–6. https://doi.org/10.1109/tensymp54529.2022.9864411

[17] Jayakody, J. a. T. K., Gamlath, T. S. K., Lasantha, W. a. N., Premachandra, K. M. K. P., Nugaliyadde, A., & Mallawarachchi, Y. (2016). "Mahoshadha", the Sinhala tagged corpus based question answering system. In Smart innovation, systems and technologies (pp. 313–322). https://doi.org/10.1007/978-3-319-30933-0_32

[18] Budler, L. C., Gosak, L., & Stiglic, G. (2023). Review of artificial intelligence-based question-answering systems in healthcare. Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery, 13(2). https://doi.org/10.1002/widm.1487

[19] Ramloll, R., Yu, W., Brewster, S., Riedel, B., Burton, M., & Dimigen, G. (2000). Constructing sonified haptic line graphs for the blind student. *Constructing Sonified Haptic Line Graphs for the Blind Student*. https://doi.org/10.1145/354324.354330

[20] Vines, K., Hughes, C., Alexander, L., Calvert, C., Colwell, C., Holmes, H., Kotecki, C., Parks, K., & Pearson, V. (2019). Sonification of numerical data for education. *Open Learning the Journal of Open Distance and e-Learning*, *34*(1), 19–39. https://doi.org/10.1080/02680513.2018.1553707

[21] DIAGRAM Center. (2020, December 16). *DIAGRAM Report: Sonification 2019 - DIAGRAM Center*. http://diagramcenter.org/diagram-reports/diagram-report-2019/sonification.html

[22] Ohshiro, K., Hurst, A., & DuBois, L. (2021). Making Math Graphs More Accessible in Remote Learning: Using Sonification to Introduce Discontinuity in Calculus. *Making Math Graphs More Accessible in Remote Learning: Using Sonification to Introduce Discontinuity in Calculus*, 1–4. https://doi.org/10.1145/3441852.3476533

[23] Journal of Dynamics and Control. (2024, December 30). AN INNOVATIVE AND INCLUSIVE E-LEARNING APPLICATION FOR VISUALLY IMPAIRED STUDENTS - Journal of Dynamics and Control. https://jodac.org/an-innovative-and-inclusive-e-learning-application-for-visually-impaired-students/

[24] Butucea, D. & University of Economic Studies, Bucharest, Romania. (2013). Personalized e-learning software systems. Extending the solution to assist visually impaired users. In Database Systems Journal: Vol. IV (Issue 3, p. 41). http://dbjournal.ro/archive/13/13_5.pdf

[25] Nees, M. A. (2018). Auditory graphs are not the "Killer app" of sonification, but they work. *Ergonomics in Design the Quarterly of Human Factors Applications*, *26*(4), 25–28. https://doi.org/10.1177/1064804618773563

[26] Rong, A., & Li, G. (2023). Visually Impaired Children with Special Educational Needs: Identifying Suitable Tactile Graphics Learning Materials. SHS Web of Conferences, 174, 01003. https://doi.org/10.1051/shsconf/202317401003

[27] Hoang, L., Lee, S., & Kwon, K. (2020). A 3D shape recognition method using hybrid deep learning network CNN–SVM. *Electronics*, *9*(4), 649. https://doi.org/10.3390/electronics9040649

[28] Qi, S., Ning, X., Yang, G., Zhang, L., Long, P., Cai, W., & Li, W. (2021). Review of multi-view 3D object recognition methods based on deep learning. *Displays*, *69*, 102053. https://doi.org/10.1016/j.displa.2021.102053

Appendix A

Individuals Contribution to the Project

204106P - Kumara P.M.M.D.S

Module 1- Personalized Learning Model for Visually Impaired Students.

In this research project, I am developing a Personalized Learning Model designed to cater to the needs of visually impaired students. Currently, the module is in its initial training phase, focusing on fine-tuning the LLAMA 3 model using the Unsloth optimization framework. To build the foundation for the model, I have compiled a comprehensive dataset that includes mathematical content sourced from Kaggle and Hugging Face, specifically tailored to fundamental mathematics concepts from the Sri Lankan 9th-grade curriculum.

The primary goal of this module is to train the LLAMA 3 model to process and generate responses that are personalized for different visual impairment categories. A key focus is integrating Sinhala language support to ensure accessibility for students in Sri Lanka. As part of the development process, I am addressing the challenge of optimizing real-time content generation performance and developing effective evaluation metrics tailored to visually impaired users.

The module is progressing as planned and represents an essential part of a larger framework aimed at providing accessible and personalized mathematical education for visually impaired students in Sri Lanka. Moving forward, the next steps include completing the initial training phase, implementing evaluation metrics, and integrating the model with a multi-modal content delivery system to further enhance the learning experience.

204115R - Liyanage L.D.S

Module 2 - Specialized Question and Answer Module for Visually Impaired Students

The development of the Specialized Question and Answer Module began with extensive research into existing literature and current technological solutions. I conducted a comprehensive literature review, analyzing over 8 research papers focusing on AI-based educational systems for visually impaired students. This review concentrated particularly on papers examining both theoretical frameworks and practical implementations of similar systems. Special attention was given to studies involving natural language processing applications in educational contexts and adaptive learning systems for visually impaired students.

Following the literature review, I developed a detailed technical specification for the module's architecture. This involved creating comprehensive documentation outlining the system's requirements, focusing specifically on accessibility needs for visually impaired students. I identified key integration points with other system modules and established the necessary technical parameters for successful implementation, ensuring alignment with the overall project objectives.

In the implementation phase, I took responsibility for data collection and model training. This involved curating a specialized dataset combining mathematical content from multiple sources, including Kaggle datasets and Hugging Face repositories. The content was carefully selected to cover fundamental mathematical concepts aligned with the Sri Lankan 9th-grade curriculum. I processed and standardized this data to ensure consistency and quality, creating a robust foundation for the model training process.

The technical implementation phase focused on adapting and fine-tuning the Llama 3 model using the Unsloth framework. This involved writing custom preprocessing scripts to format the mathematical content appropriately for training, developing optimization routines to enhance model performance, and implementing specialized prompting strategies to generate mathematics-focused questions and answers. I also developed evaluation metrics to assess the quality and relevance of the generated content, ensuring it met the educational standards required for the target audience.

My contribution also extended to integration testing with other modules of the system. I worked on developing APIs and interface specifications that would allow the Question-and-Answer Module to communicate effectively with the other components of the system, particularly the Personalized Learning Mode. This involved creating detailed technical documentation for the integration points and developing test cases to ensure smooth interaction between different system components.

204049R –Dissanayake J.W

Module 3 - AI-Driven Sonification and Natural Language Generation for Accessible Mathematical Graph Visualization for blind students

In this research project, I am working on processes mathematical function graph images to extract meaningful data points for analysis and generate the sonification sound with NLG voice. To begin, I referred to various research papers, documentation, and online resources to gain a clear understanding of existing methodologies for function graph analysis and extraction. This helped me explore different approaches used in mathematical graph interpretation, including image-processing techniques, curve extraction, and normalization methods. Additionally, I performed small practical exercises to test various image-processing techniques before implementing the module.

After gaining sufficient background knowledge, I moved on to the implementation of Module 1. The first step was to prepare a dataset by collecting relevant mathematical function graph images. Once the dataset was ready, I performed image pre-processing, which included converting images to grayscale to simplify processing. Then, I applied morphological operations to clean the images and remove noise, ensuring that unnecessary artifacts such as grid lines and labels do not interfere with curve extraction. Following this, I implemented Otsu's Thresholding, which binarizes the images, making it easier to distinguish the graph curve from the background.

Once the pre-processing was completed, I proceeded with curve extraction, where I identified and extracted the key data points representing the function. I sorted these extracted points based on their x-coordinates and performed normalization, scaling the values into the range of [-1, 1] to ensure consistency across different graphs. The extracted data points were then stored for further analysis, visualization, and potential applications in machine learning or accessibility tools.

Moving forward, I plan to integrate these processed data points into a sonification module, where the extracted function curves will be mapped to auditory cues. Additionally, I will work on a natural language generation (NLG) module to automatically describe the extracted graphs in textual and spoken formats, making mathematical graphs more accessible for visually impaired users. These days, I am studying different techniques for effective sonification and exploring advanced models for automated text generation to enhance the system's usability.

204136G – Nedungamuwa P.D.M.C.C.B

Module 4 - AI-Driven Real-Time 3D Shape Detection and Position based description Using Finger Tracking and Video Processing

In this research project, I focused on learning and implementing key technologies for image processing, object detection, and video processing to develop a system for visually impaired students. As all the technologies were new to me, I began by enrolling in courses to learn the fundamentals of image processing and object detection. I referred to various research papers and resources to understand the concepts behind these technologies and identify the most appropriate methods to use for my work.

After gaining foundational knowledge, I proceeded to search for relevant datasets, specifically image datasets, to begin the project. Once the datasets were acquired, I applied image processing techniques for pre-processing the data, ensuring it was ready for further analysis. This included converting the images to grayscale, applying morphological operations, and cleaning the images to remove noise and unnecessary artifacts.

Moving forward, I plan to develop the next steps for the system, which will involve real-time tracking and guidance for the visually impaired student as they trace shapes. The system will offer audio prompts to guide them through the edges and vertices of the shape, while also providing feedback on shape properties (e.g., angles). If the student moves off course, corrective feedback will be provided, along with encouragement to continue exploring the shape. This approach aims to make mathematical concepts more accessible through real-time interaction and feedback.