

AR Educational Suite for Primary Students'

Geometry and Literacy

R26-IM-005

Project Proposal Report

Abhilash K.A.T.W- IT22081698

B.Sc. (Hons) Degree in Information Technology

Specialization in Interactive Media (IM)

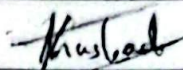
Department of Information Technology

Sri Lanka Institute of Information Technology

March 2026

DECLARATION

I declare that this is my own work, and this proposal does not knowingly incorporate any material previously submitted for a degree or diploma at any other university or higher education institution, nor does it contain any material that has been previously published or written by another person apart from where proper acknowledgment is given in the text.

Name	Student ID	Signature
Abhilash K.A.T.W.	IT22081698	

The above candidate is conducting research for their undergraduate Dissertation under my supervision.

Name	Role	Date	Signature
Mr. Aruna Ishara Gamage	Supervisor	19/03/2016	

ABSTRACT

This research addresses the issue of incorrect letter stroke habits in primary school students (ages 6–12) due to the lack of real-time, individualized guidance in classroom settings. Traditional literacy applications focus only on the final shape of characters, ignoring the dynamic process of stroke ordering and direction. The proposed solution is a mobile-based AI Air-Writing Trajectory Tracking system that uses a CNN-based model with MediaPipe hand landmark detection to track and correct the specific path and stroke order of English letters in 3D space. By capturing real-time 3D fingertip coordinates (x, y, z), the system ensures that students develop proper handwriting habits, providing stroke-direction awareness and real-time corrective feedback. The solution is designed for mobile optimization and is validated through user testing, with expected outcomes including improved writing accuracy and sustained engagement through gamified AR feedback.

Keywords: AI, Augmented Reality, Air-Writing, CNN, Trajectory Recognition, Stroke Order, Handwriting, Mobile Computer Vision, Primary Education, Stroke Direction Awareness, Real-Time Feedback, Gesture Recognition, 3D Coordinates, Educational Technology.

ACKNOWLEDGEMENT

I wish to express my sincere gratitude to my supervisor, Mr. Aruna Ishara Gamage, and my co-supervisor, Mr. Nushkan Nismi, for their expert guidance, consistent feedback, and encouragement throughout this research project.

I extend my appreciation to the Faculty of Computing at SLIIT for providing access to research resources and the Mixed Reality research cluster, which formed the academic foundation of this work.

Special thanks are owed to my project group members Himsara M.K.D.L., De Costa M.M.P., and Wickramasingha N.N. for their collaboration and the stimulating discussions that continually refined the design of this system.

I am also grateful to the primary school teachers and students who participated in preliminary consultations and testing, and to my family for their unwavering support.

TABLE OF CONTENTS

DECLARATION	ii
ABSTRACT	iii
ACKNOWLEDGEMENT	iv
TABLE OF CONTENTS	v
LIST OF FIGURES.....	vii
LIST OF TABLES	vii
LIST OF ABBREVIATIONS	viii
CHAPTER 1: INTRODUCTION	1
1.1 Background and Context.....	1
1.2 Literature Review.....	2
1.2.1 Augmented Reality in Education	2
1.2.2 Air-Writing Recognition Systems	2
1.2.3 Hand Gesture and Pose Estimation	3
1.2.4 Sketch-to-3D and Recognition Validation	3
1.3 Research Gap	3
CHAPTER 2: OBJECTIVES	6
2.1 Main Objective	6
2.2 Specific Objectives.....	6
CHAPTER 3: METHODOLOGY	7
3.1 Research Approach	7
3.2 Technical Methods and Algorithms	7
3.2.1 Hand Landmark Detection	7
3.2.2 Trajectory Buffer and Feature Extraction	7
3.2.3 CNN Trajectory Recognition Model.....	8
3.2.4 Stroke-Order Validation Engine.....	10
3.2.5 AR Corrective Overlay.....	10
3.3 Tools and Platforms	11

3.4 Data Collection Strategy	11
3.4.1 Primary Technical Data.....	11
3.4.2 Secondary Data	12
3.4.3 Ethical Considerations.....	12
3.5 Validation Strategy and Performance Metrics	12
CHAPTER 4: HIGH LEVEL SYSTEM ARCHITECTURE.....	13
4.1 Overview of the Proposed Solution	14
4.2 Component Architecture Diagram	15
4.3 Process Flow	17
.....	17
4.4 Data Flow	18
4.5 Component Interactions	19
4.6 Scalability and Security Considerations.....	19
CHAPTER 5: USER REQUIREMENTS	20
5.1 Stakeholder Identification	20
5.2 Functional Requirements.....	20
5.3 Non-Functional Requirements	21
CHAPTER 6: COMMERCIALISATION PLAN	22
6.1 Target Market and Customer Persona.....	22
6.2 Value Proposition.....	22
6.3 Revenue Model	22
6.4 Competitive Advantage.....	22
6.5 Intellectual Property Considerations	23
CHAPTER 7: BUDGET AND JUSTIFICATION.....	24
CHAPTER 8: WORK BREAKDOWN STRUCTURE	25
8.1 Task Breakdown.....	25
8.2 Gantt Chart.....	26
REFERENCES.....	27
APPENDICES.....	28

Appendix A – Topic Assessment Form Summary	28
Appendix B – Canonical Stroke Database Format (Sample)	28
Appendix C – Ethical Clearance Checklist	28

LIST OF FIGURES

Figure 1: CNN-Based Trajectory Recognition Model Architecture
Figure 2: High Level System Architecture
Figure 3: AI Air-Writing Trajectory Tracking – System Architecture
Figure 4: Air-Writing Recognition Process Flow
Figure 5: Data Flow – Air-Writing Trajectory Tracking Component

LIST OF TABLES

Table 1: Comparison of Existing Air-Writing and Literacy Systems
Table 2: Functional Requirements
Table 3: Non-Functional Requirements
Table 4: Budget and Justification
Table 5: Work Breakdown Structure

LIST OF ABBREVIATIONS

Abbreviation	Full Form
AR	Augmented Reality
CNN	Convolutional Neural Network
MR	Mixed Reality
ML	Machine Learning
API	Application Programming Interface
UAT	User Acceptance Testing
SDG	Sustainable Development Goal
FPS	Frames Per Second
DNN	Deep Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
CV	Computer Vision
HCI	Human Computer Interaction

LIST OF APPENDICES

Appendix A	Topic Assessment Form Summary
Appendix B	Canonical Stroke Database Format
Appendix C	Ethical Clearance Checklist

CHAPTER 1: INTRODUCTION

1.1 Background and Context

Foundational handwriting skills remain one of the most critical yet underserved areas in primary education. The ability to form English letters with correct stroke order and directionality is not merely an aesthetic concern it is a neurological and habitual process that, once ingrained incorrectly, is exceptionally difficult to reverse. Researchers in educational psychology have consistently demonstrated that improper stroke habits formed in early childhood lead to academic frustration, reduced writing speed in later years, and persistent "engagement issues" with written tasks [7][8].

In a typical Sri Lankan primary classroom, a single teacher manages 30 to 40 students simultaneously. Under these conditions, providing real-time, individualized correction of each child's stroke technique is a practical impossibility. Students are left to self-correct based solely on the final appearance of their written characters an approach that completely ignores whether the correct movement sequence was used. Applications currently used in schools, such as Letter School and Write, limit their feedback to touchscreen tracing and binary "Correct/Incorrect" responses with no capacity to capture three-dimensional motion or stroke directionality.

The proliferation of low cost Android smartphones in Sri Lanka even in rural areas presents an unprecedented opportunity to deploy a computer vision solution that turns an ordinary camera into a real-time literacy coach. By tracking a student's fingertip in three-dimensional space (air-writing) and analyzing the trajectory against a ground-truth stroke model, the system can detect the precise moment a stroke error occurs and provide immediate, actionable visual guidance.

This component is part of the broader AR Educational Suite for Primary Students' Geometry and Literacy, which sits within the Mixed Reality (MR) research cluster at SLIIT. The full suite integrates geometry learning, air-writing, and vocabulary development into a single mobile AR environment. This proposal focuses exclusively on the AI Air-Writing Trajectory Tracking component (Sub-objective of IT22081698).

This work aligns directly with SDG 4: Quality Education, by delivering an inclusive, low-cost tool that is equally accessible to urban and rural learners regardless of the availability of specialist teaching staff.

1.2 Literature Review

1.2.1 Augmented Reality in Education

Ibáñez et al. [1] demonstrated through experimental studies that AR based learning environments significantly increase student engagement and learning effectiveness compared to traditional methods. Their work showed that immersive, interactive feedback loops rather than static instruction are crucial for retaining spatial and procedural knowledge in young learners. Radu [2] conducted a meta-review confirming that AR's greatest educational impact is in domains requiring spatial understanding and repeated procedural practice, which directly justifies its application to handwriting stroke guidance.

Kalawsky and Schmalstieg [3] demonstrated that collaborative AR environments for mathematics education improved geometric comprehension among primary students. While their focus was on geometry, the pedagogical framework using real-time spatial feedback to bridge the gap between abstract instruction and physical understanding translates directly to the challenge of teaching correct motor sequences in handwriting.

1.2.2 Air-Writing Recognition Systems

Amma et al. [7] introduced a landmark air-writing system that uses inertial sensors to track hand movement and recognize words written in free space. Their key insight was that trajectory-based recognition capturing the temporal flow of a stylus's movement is fundamentally superior to static image classification for identifying letters. Their system achieved high accuracy but was constrained to wearable sensors, making it impractical for a standard classroom environment.

Chen et al. [8] extended online handwriting recognition to recurrent neural networks, demonstrating that sequential models could accurately model the temporal dependencies inherent in letter strokes. Critically, their work showed that a network trained on coordinate sequences (rather than pixel images) learns stroke order implicitly providing the theoretical foundation for the CNN trajectory recognition approach proposed in this component.

Yang et al. [9] advanced the field further by demonstrating vision-based air-writing recognition using deep CNNs applied to video frames. Their approach, processing fingertip position over time from a standard camera, showed that high-accuracy recognition is achievable without specialized sensors directly validating the mobile-camera approach proposed here.

1.2.3 Hand Gesture and Pose Estimation

Molchanov et al. [10] demonstrated that 3D convolutional neural networks could classify hand gestures with high accuracy, but their approach required substantial computational resources. Cao et al. [11] introduced Part Affinity Fields for real-time multi-person pose estimation, the methodological precursor to MediaPipe Hands, which makes real-time fingertip tracking feasible on mobile CPUs. MediaPipe, developed by Google, provides a pre-trained landmark detection model that extracts 21 hand landmarks in real time at 30+ FPS on modern Android devices enabling technology for this component.

1.2.4 Sketch-to-3D and Recognition Validation

Wang et al. [4] and Han et al. [6] explored deep learning approaches to transforming free-hand 2D sketches into 3D models. While their work focuses on geometry rather than literacy, their underlying insight that deep networks can learn the spatial and structural grammar of hand-drawn inputs reinforces the viability of training CNN models to understand the grammar of letter strokes. Their validation approaches also informed the error-detection methodology used in this component.

1.3 Research Gap

Although Augmented Reality (AR) and Artificial Intelligence (AI) technologies have increasingly been applied in educational environments, many existing mobile learning systems still fail to effectively support primary school students in developing correct foundational English handwriting habits. Traditional classroom instruction relies heavily on static tracing worksheets and touchscreen-based letter-tracing applications that evaluate only the final visual appearance of a written character, without capturing or analyzing the dynamic process by which that character was produced. As a result, students frequently develop and reinforce incorrect stroke sequences and directional habits during the critical early learning window of ages 6 to 12 that, once automated at the motor level, are exceptionally difficult to correct in later years. At the same time, current systems do not leverage the capabilities of mobile cameras and computer vision to provide real-time, three-dimensional stroke guidance in free-space writing environments, leaving a significant gap in accessible and affordable literacy support for primary school learners. As identified in the project topic assessment document, the overall research problem of this project can therefore be expressed as the following research question:

Research Problem (Main): How can a mobile-based AI and Augmented Reality (AR) system be designed to accurately convert student drawn 2D sketches into validated 3D geometric models, verify geometric net correctness, and simultaneously support English letter stroke

guidance and vocabulary learning in a single low cost, real time interactive educational environment optimized for primary school learners?

Within this broader research problem, several limitations exist in current AR-based geometry learning systems. First, many applications focus mainly on visualizing predefined three-dimensional models rather than allowing students to create and validate their own geometric constructions. Second, most systems provide only binary feedback indicating whether a student's answer is correct or incorrect, which does not help learners understand the cause of their spatial mistakes. Third, existing educational tools rarely incorporate visual scaffolding techniques that guide students in correcting errors through direct visual comparison with the correct geometric structure. Additionally, most systems do not include volumetric slicing visualizations, which could help learners understand the internal structure of geometric solids and concepts such as fractional volumes. Furthermore, there is limited research exploring the integration of ghost overlay visualization techniques within mobile AR learning environments to assist students during geometric construction tasks.

Therefore, the specific research problem addressed in this component can be expressed as the following question:

Research Problem (Sub-Component): How can a CNN-based air-writing trajectory tracking system be designed for mobile AR environments to accurately capture three-dimensional fingertip movement in real time, recognize English letter strokes by their directional sequence and stroke order, and deliver immediate corrective AR visual guidance to primary school students using only a standard device camera without specialized hardware?

To address this gap, this research proposes an AI Air-Writing Trajectory Tracking framework that employs MediaPipe hand-landmark detection to extract real-time 3D fingertip coordinates from a standard mobile camera, a trained CNN trajectory recognition model to classify letter strokes by their movement path and directionality, and a stroke-order validation engine that compares the student's live trajectory against a canonical stroke database to detect deviations at the sub-stroke level. By combining lightweight on-device computer vision with real-time AR corrective overlays, the proposed approach aims to establish correct stroke habits, reduce the formation of long-term handwriting errors, and deliver individualized literacy guidance to primary school learners in environments where specialist teacher supervision is unavailable.

Feature	Air-Writing [7]	Online HWR [8]	Gesture CNN [10]	Proposed System
Real-time stroke monitoring	Yes	Partial	No	Yes
Trajectory-based recognition	Yes	Yes	Partial	Yes
Mobile-only (no sensors)	No	Partial	No	Yes
Stroke-direction awareness	Partial	No	No	Yes
Real-time corrective alerts	No	No	No	Yes
Integrated AR feedback	No	No	No	Yes

Table 1: Comparison of Existing Air-Writing and Literacy Systems

CHAPTER 2: OBJECTIVES

2.1 Main Objective

The main objective of this research is to design and develop an integrated mobile-based Augmented Reality (AR) and Artificial Intelligence (AI) educational system that transforms traditional static two-dimensional learning into an interactive three-dimensional learning experience for primary school students. The system aims to accurately convert student-drawn 2D sketches into validated 3D geometric models, verify the correctness of geometric nets, and simultaneously support English letter stroke guidance and vocabulary learning within a single low-cost, real-time AR learning environment optimized for children aged 6–12.

2.2 Specific Objectives

To achieve the main objective, the research project is divided into four specialized components corresponding to the contributions of the project members.

- **2D-to-3D Geometry Validation** – To develop a computer vision-based recognition system capable of detecting hand-drawn geometric nets and mathematically validating whether they can form correct three-dimensional shapes before generating interactive 3D models.
- **Adaptive Scaffolding and Virtual Slicing** – To implement interactive visual feedback mechanisms that explain spatial errors using ghost overlay visualizations and volumetric slicing techniques within an augmented reality environment.
- **AI Air-Writing Trajectory Tracking** – To design a CNN-based recognition system that tracks and analyzes finger movement trajectories in three-dimensional space in order to monitor and correct English letter stroke order and writing direction.
- **Multisensory Phonics and Vocabulary Integration** – To develop an interactive learning module that connects sketches, air-writing, and pronunciation with animated 3D objects to reinforce vocabulary learning and phonics development.

CHAPTER 3: METHODOLOGY

3.1 Research Approach

This component adopts a Design Science Research (DSR) methodology, which is appropriate for projects that create and evaluate a novel technical artefact (the trajectory tracking system) to solve an identified practical problem (incorrect stroke habits in primary education). The research proceeds through iterative design, build, and evaluate cycles, with each cycle informing the next refinement of the CNN model and feedback system.

3.2 Technical Methods and Algorithms

3.2.1 Hand Landmark Detection

MediaPipe Hands (Google) is selected as the hand-detection backbone. It provides a pre-trained, mobile optimized model that detects 21 hand landmarks in real time using a single RGB camera frame with no depth sensor or wearable required. Landmark index 8 (index fingertip) provides the primary coordinate source. The raw (x, y, z) coordinates are normalized relative to the wrist landmark to ensure scale and distance invariance.

3.2.2 Trajectory Buffer and Feature Extraction

The normalized fingertip coordinates are accumulated into a sliding temporal buffer of fixed length N (determined empirically, expected 60–100 frames at 30 FPS corresponding to 2–3 seconds of writing).

Each buffer snapshot represents one potential letter stroke. Feature vectors include:

- Raw coordinate sequence: $[(x_1, y_1, z_1), \dots, (x_n, y_n, z_n)]$
- Velocity vectors: $\Delta x, \Delta y, \Delta z$ between consecutive frames
- Direction angles: azimuth and elevation derived from velocity
- Start-point normalized position: critical for stroke-direction detection

3.2.3 CNN Trajectory Recognition Model

A 1D Convolutional Neural Network is selected over RNN/LSTM alternatives for three reasons: lower latency on mobile hardware, superior generalization on the size of training dataset expected, and proven performance on sequential gesture data [10]. The architecture comprises:

- Input layer: trajectory feature sequence of shape (N, features)
- Two Conv1D blocks with Max Pooling and Dropout
- Two deeper Conv1D blocks with Global Average Pooling
- Dense layers
- Output SoftMax layer: 26 classes (A–Z)

The model is trained in TensorFlow/Kera's and exported to TensorFlow Lite for on-device inference, ensuring the entire recognition pipeline runs locally without internet dependency.

CNN-Based Trajectory Recognition Model Architecture

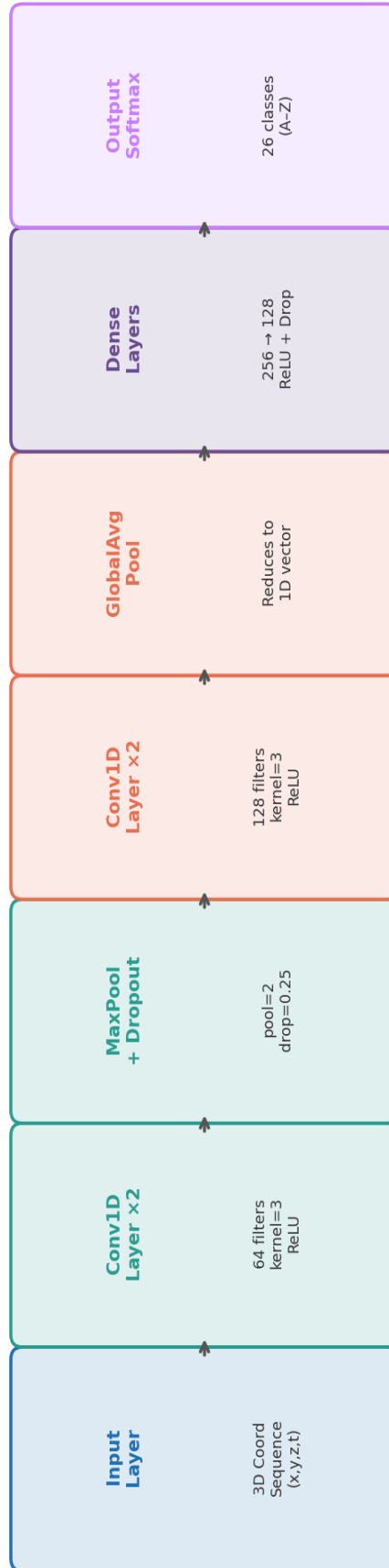


Figure 1: *CNN-Based Trajectory Recognition Model Architecture*

3.2.4 Stroke-Order Validation Engine

Beyond letter classification, a rule-based validation engine compares the student's trajectory against a ground truth stroke database containing the canonical (British National Curriculum) stroke paths for all 26 letters. Validation checks include:

- Start point correctness: the fingertip's position at stroke initiation compared to the expected start region of the canonical path
- Direction agreement: the initial velocity vector of the student's stroke compared to the expected initial direction
- Waypoint adherence: at defined temporal checkpoints, the student's position is compared to the expected waypoint $\pm$ a tolerance radius

If any validation check fails, the system immediately raises a corrective event before the stroke is completed enabling real-time intervention rather than post-hoc assessment.

3.2.5 AR Corrective Overlay

On detection of a stroke error, the system triggers an AR overlay within the Unity 3D framework: a translucent ghost path (the canonical stroke trajectory rendered in the AR space at the student's hand position) and a directional arrow indicating where the fingertip should move next. The overlay persists for 1.5 seconds and is accompanied by a gentle audio cue. On successful completion of the correct stroke sequence, a celebratory 3D animation (integrated with the broader suite's vocabulary reward system) is triggered.

3.3 Tools and Platforms

Tool / Platform	Purpose	Justification
MediaPipe Hands	Real-time hand tracking	Mobile-optimized, open-source, no sensor required
TensorFlow / Keras	CNN model training	Industry standard, extensive mobile deployment support
TensorFlow Lite	On-device inference	Optimized for ARM CPUs, low latency, offline capable
Unity 3D (2022 LTS)	AR application framework	Cross-platform; used by full project team
ARCore	Marker less AR tracking	Robust real-world tracking for mobile AR
Python 3.10	Data preprocessing & training	Ecosystem for ML development
Android (API 26+)	Target deployment platform	Most accessible device category in target schools
Blender	Ghost-path 3D asset creation	Team proficiency; used across all sub-components

3.4 Data Collection Strategy

3.4.1 Primary Technical Data

Air-writing trajectory data will be collected in two phases: (1) an initial dataset of capital letter traces from 10 adult volunteers (team members and peers) to bootstrap model training; (2) a supplementary dataset from 20 primary school children (ages 6–12) to capture child-specific hand-size and movement characteristics. Each participant will write each letter 20 times under varied conditions (distance from camera, ambient lighting). Ethical approval from the SLIIT Ethics Review Committee will be obtained before data collection from minors.

3.4.2 Secondary Data

The British National Curriculum stroke-order specifications and SLIIT Faculty of Computing Research guidelines on Mixed Reality development will be reviewed to ensure the ground-truth stroke database and system architecture conform to both educational standards and institutional research requirements.

3.4.3 Ethical Considerations

All participants under 18 require informed written consent from parents or guardians. Video data will be processed on-device and will not be stored; only anonymized coordinate sequences will be saved to a secure, access-controlled repository. The SLIIT Ethics Review Committee approval will be sought and obtained before any data collection activity involving minors begins.

3.5 Validation Strategy and Performance Metrics

Metric	Measurement Method
Letter recognition accuracy	CNN test-set evaluation (80/20 split)
Stroke-direction detection rate	Confusion matrix on validation trajectories
Real-time processing latency	Device benchmarking on mid-range Android
User satisfaction (SUS score)	System Usability Scale questionnaire post-UAT
Writing accuracy improvement	Pre-test / Post-test stroke correctness assessment
False positive rate	Error-trigger analysis on correct traces

CHAPTER 4: HIGH LEVEL SYSTEM ARCHITECTURE

High-Level System Architecture of the Mobile-Based AI and AR Educational System for Primary Students

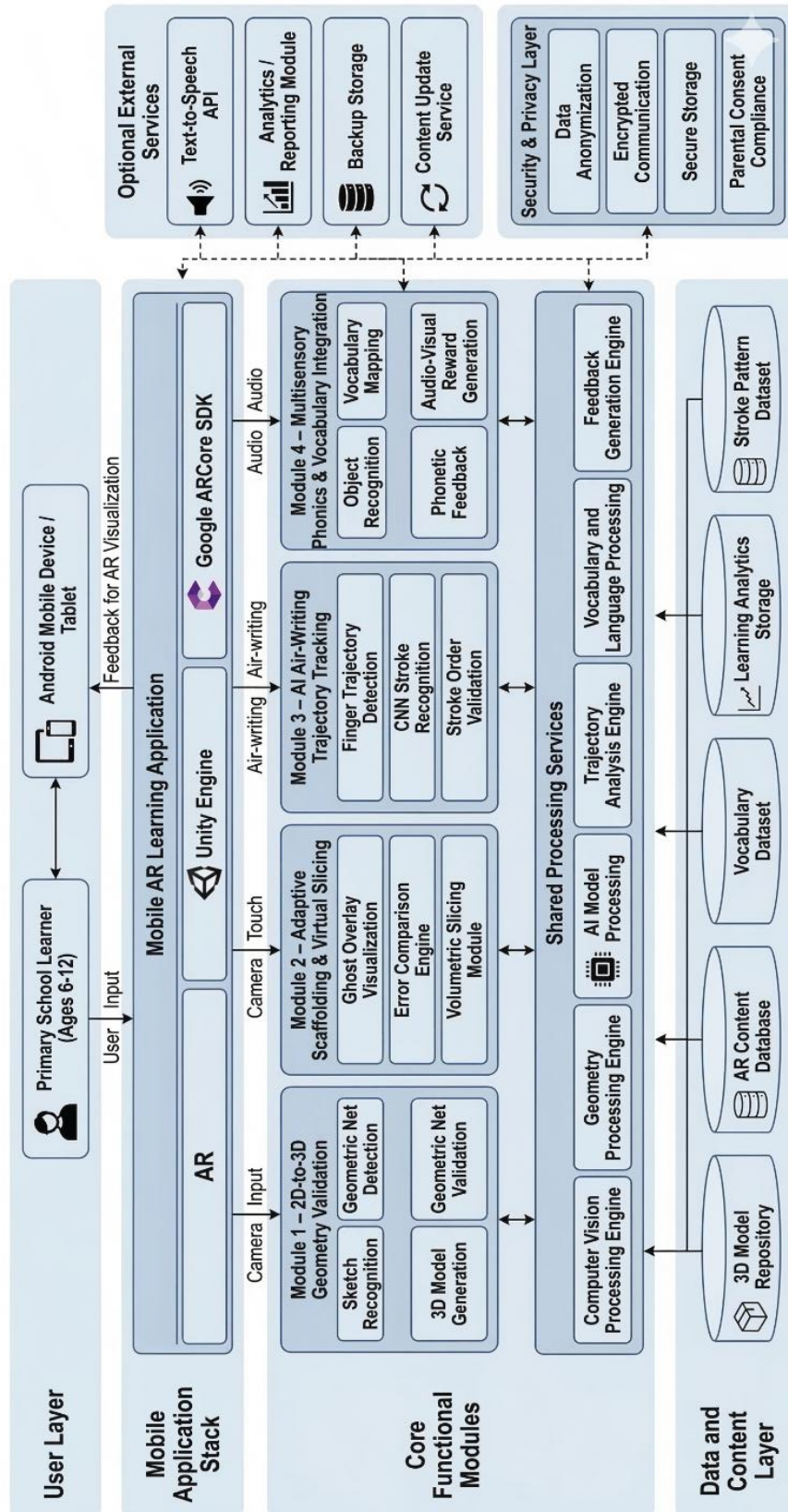


Figure 2: High Level System Architecture

4.1 Overview of the Proposed Solution

The AI Air-Writing Trajectory Tracking component sits within the broader AR Educational Suite as the Literacy Engine. It receives live camera frames from the Unity AR application, performs all processing locally on the device, and returns corrective feedback events to the AR presentation layer. The component is designed as a self-contained module with a clean API, allowing integration with the other three subcomponents (2D-to-3D Geometry Validation, Adaptive Scaffolding, and Multisensory Phonics) without tight coupling.

4.2 Component Architecture Diagram

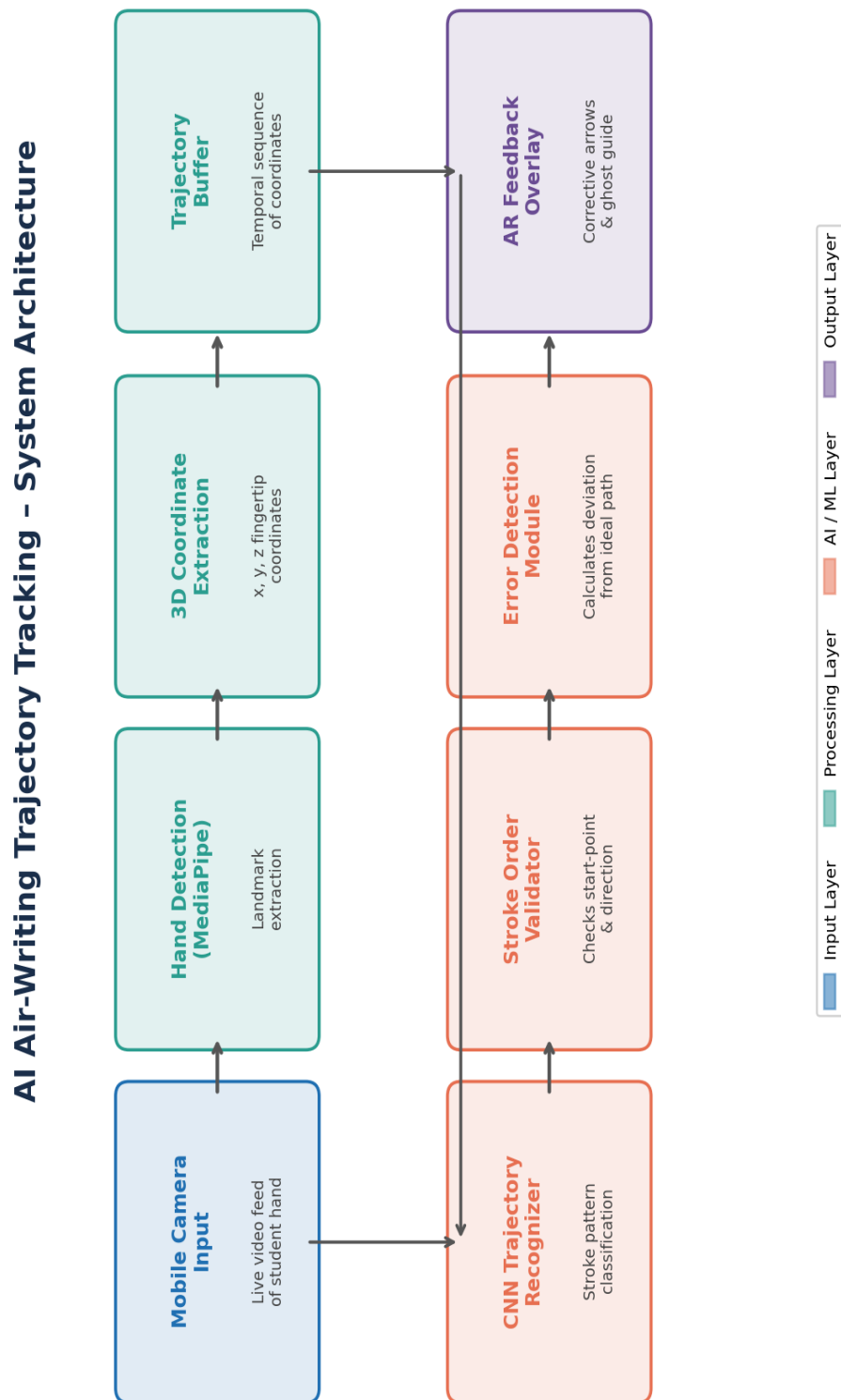


Figure 3: *AI Air-Writing Trajectory Tracking – System Architecture*

The architecture is structured in four logical layers:

Input Layer: The mobile camera delivers a live RGB video stream directly into the pipeline. MediaPipe Hands detects the student's hand within each frame and extracts twenty-one hand landmarks, providing precise positional data for all key joints and fingertips in real time.

Processing Layer: The extracted landmark coordinates are filtered using Kalman smoothing to remove jitter and noise caused by natural hand tremors or minor camera instability. The smoothed coordinates are then normalized relative to the wrist landmark, ensuring that recognition remains consistent regardless of how close or far the student holds their hand from the camera. These normalized coordinates are continuously accumulated into a trajectory buffer that stores the recent movement history of the index fingertip. A gesture-start detector runs alongside this process, identifying when the student has intentionally begun an air-writing action and preventing ambient or accidental hand movements from triggering the recognition pipeline.

AI / ML Layer: Once sufficient trajectory data is accumulated in the buffer, the TensorFlow Lite CNN model analyses the complete movement sequence and classifies the letter being written based on its trajectory shape and directional flow. Simultaneously, the stroke-order validation engine compares the student's live fingertip path against the corresponding entry in the canonical stroke database, checking the start point, direction, and movement waypoints of each stroke as it is being drawn. When a deviation from the correct stroke pattern is detected, the engine immediately generates a corrective event without waiting for the letter to be completed.

Output Layer: Corrective events produced by the validation engine are passed to the Unity AR overlay system, which responds by rendering ghost-path guides and directional arrows anchored to the student's current fingertip position in the AR space, visually showing the correct path forward. When a letter is completed with the correct stroke sequence, a positive completion event is raised and consumed by the Multisensory Phonics sub-component, which triggers the corresponding vocabulary reward animation as part of the integrated cross-domain learning experience.

4.3 Process Flow

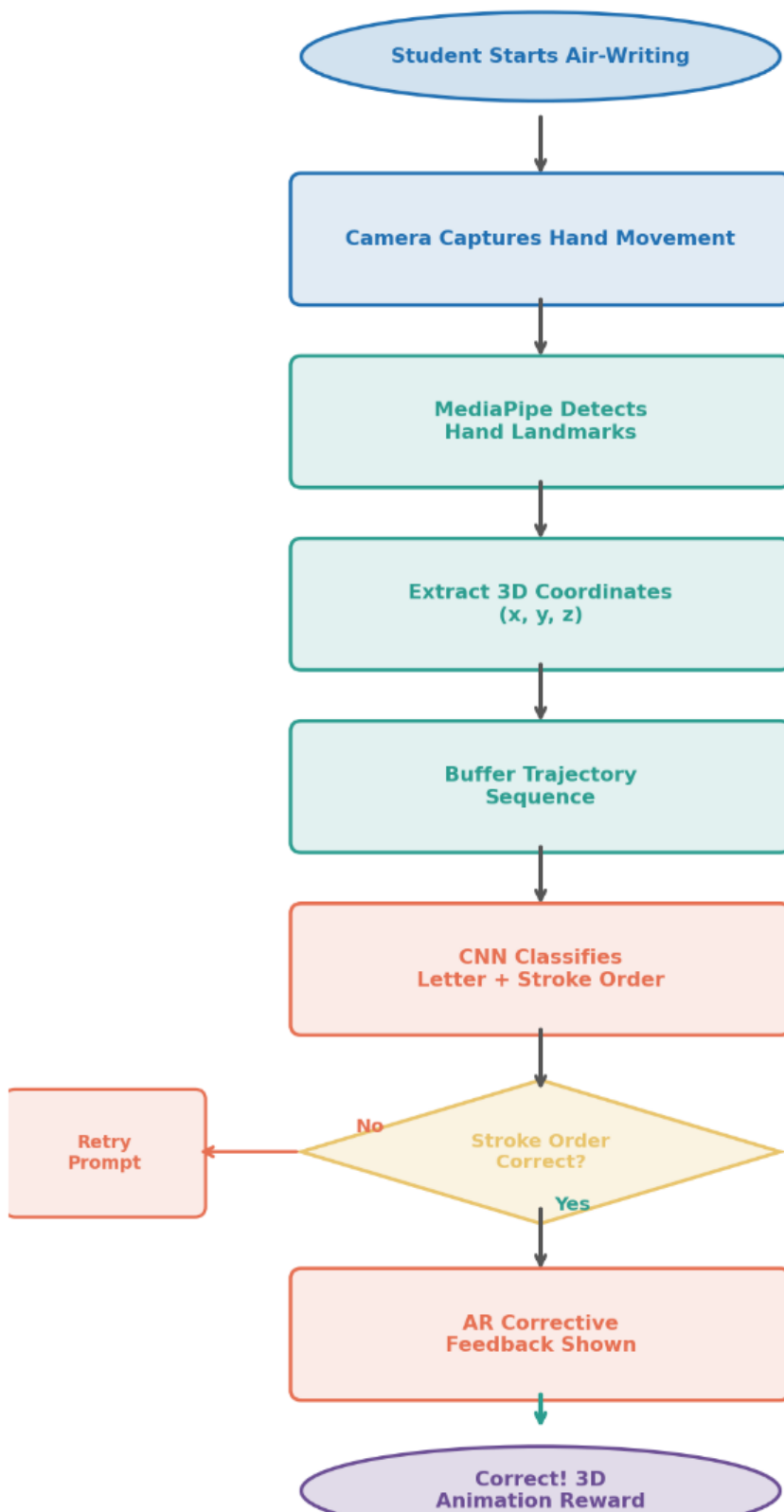


Figure 4: *Air-Writing Recognition Process Flow*

4.4 Data Flow

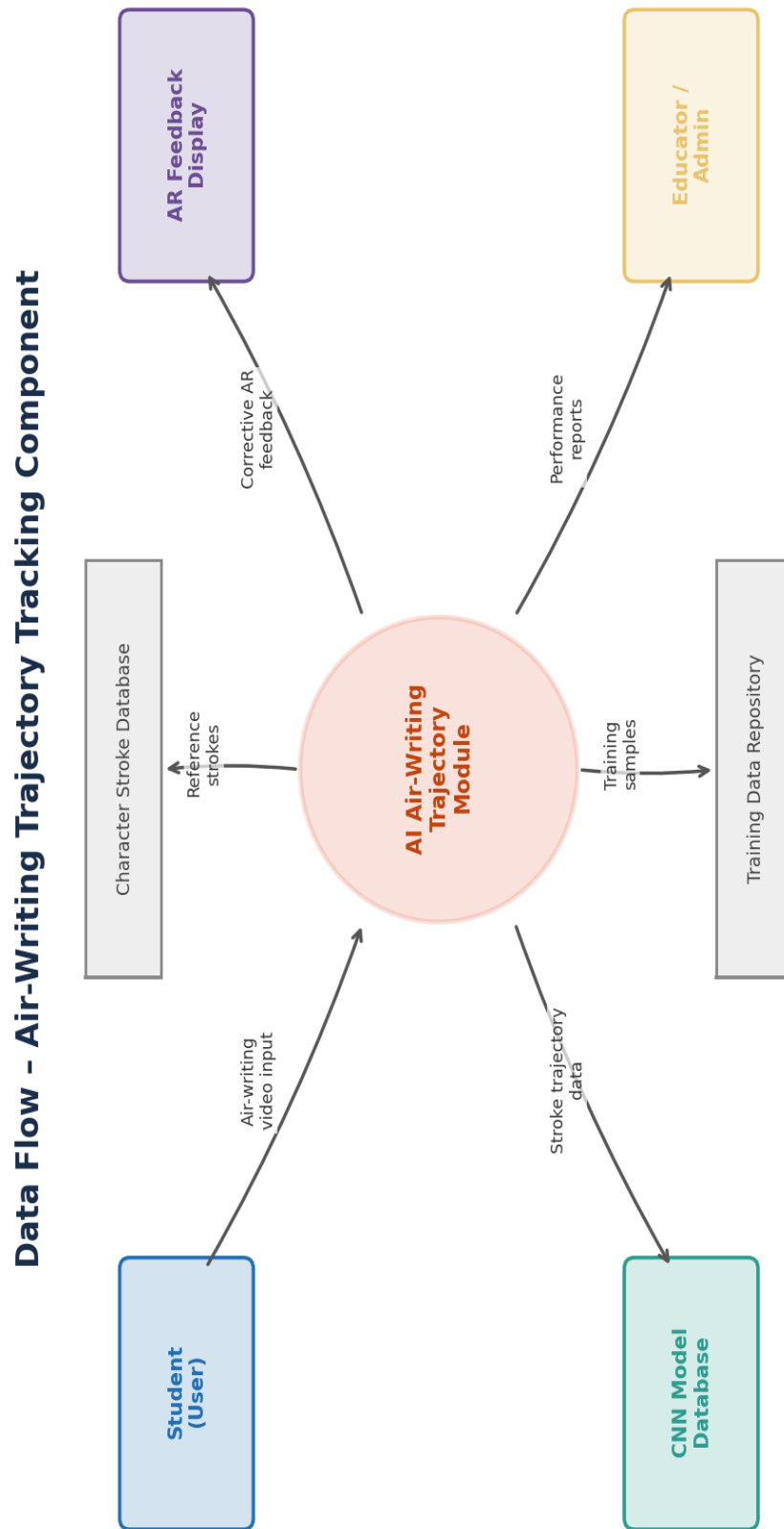


Figure 5: Data Flow – Air-Writing Trajectory Tracking Component\

4.5 Component Interactions

The Air-Writing component interacts with two other subcomponents in the suite:

- **Multisensory Phonics Integration:** On successful completion of a correctly stroked letter sequence, this component raises a "letter-confirmed" event consumed by the Phonics sub-component, which triggers the corresponding 3D vocabulary animation (e.g., correct air-writing of "C-A-T" triggers the animated cat model).
- **Adaptive Scaffolding:** Stroke error events from this component include spatial deviation data (the vector between the student's path and the canonical path), which the Scaffolding component uses to render Ghost Image overlays at the appropriate 3D world position.

4.6 Scalability and Security Considerations

All processing occurs entirely on the student's device, no audio, video, or biometric data is transmitted to external servers. The TensorFlow Lite model is bundled within the application package and updated only through controlled app-store releases. The trajectory coordinated data (anonymized) stored for research analysis is encrypted at rest using AES-256 and stored on a SLIIT-approved secure research server accessible only to the project team under approved protocols.

The system is designed for horizontal scalability in a classroom context: each student's device operates fully independently, with no server-side session management required. Future extension to additional languages (e.g., Sinhala script) is architecturally possible through retraining the CNN model with new stroke databases without changes to the core pipeline.

CHAPTER 5: USER REQUIREMENTS

5.1 Stakeholder Identification

Stakeholder	Role	Primary Need
Primary students (ages 6–12)	End user	Engaging, forgiving guidance that helps without frustrating
Primary school teachers	Indirect user / facilitator	A tool that reduces supervision burden and shows student progress
Parents / Guardians	Consenting party	Assurance of data privacy and educational benefit
SLIIT Research Team	Developer / researcher	Measurable learning outcomes and publishable novelty
School administrators	Deployment decision-maker	Low cost, no infrastructure requirements, curriculum alignment

5.2 Functional Requirements

1. The system should detect and track the student's index fingertip in 3D space using the device camera in real time.
2. The system shall accumulate real-time fingertip coordinate trajectories into a temporal buffer for letter recognition.
3. The system shall classify the air-written letter from the trajectory buffer for all 26 English capital letters.
4. The system shall validate the stroke start-point, direction, and waypoint adherence against the canonical stroke database in real time.
5. The system shall trigger a corrective AR overlay immediately upon detecting a stroke-order violation.
6. The system displays a ghost-path guide showing the correct trajectory at the student's fingertip position when an error is detected.
7. The system shall raise a letter-confirmed event to the Phonics sub-component upon successful completion of a correctly stroked letter.

8. The system shall support gesture-start detection to distinguish intentional writing from background hand movement.
9. The system shall operate fully offline with no internet connection required during classroom use.
10. The system shall record anonymized stroke accuracy data for pre-test and post-test research analysis.

5.3 Non-Functional Requirements

Category	Requirement
Performance	End-to-end latency on a mid-range Android device
Usability	Interface elements sized for child-sized hands; error messages presented as visual cues rather than text to accommodate early readers
Reliability	System shall remain operational for 60-minute continuous classroom sessions without memory degradation or thermal throttling
Portability	Compatible with Android API 26 and above; no root access or special permissions beyond Camera required
Privacy	No raw video or biometric data shall be transmitted off device; stored data shall be anonymized and encrypted
Accuracy	False positive corrective alert rate shall not exceed 10% on correctly formed strokes
Accessibility	System shall function under typical classroom lighting and accommodate left-handed and right-handed writers

CHAPTER 6: COMMERCIALISATION PLAN

6.1 Target Market and Customer Persona

Primary market: Government and private primary schools in Sri Lanka (approximately 4,200 government primary schools serving ~2.9 million students). Secondary market: Educational technology distributors and ed-tech NGOs operating in South and Southeast Asia targeting SDG 4 outcomes. Tertiary market: Individual parents seeking supplementary literacy tools for home use.

Customer persona: A government primary school principal with 300–500 students, limited specialist literacy staff, a desire to demonstrate measurable student outcomes, and a budget constraint that prevents adoption of expensive hardware-dependent solutions. The AR Educational Suite running on teachers' or students' own Android smartphones requires zero capital hardware expenditure.

6.2 Value Proposition

The AI Air-Writing component delivers stroke-level literacy guidance at zero marginal cost per student a capability previously available only through one-on-one specialist tutoring or expensive wearable sensor systems. By converting ubiquitous smartphone cameras into real-time handwriting coaches, the system democratize access to high quality foundational literacy instruction for students in under-resourced schools across Sri Lanka and the wider region.

6.3 Revenue Model

- Institutional SaaS license: Annual per-school fee covering app access, teacher dashboard, and anonymized analytics (estimated LKR 25,000–50,000 per school per year).
- NGO / Government grant partnerships: Position as an SDG 4 aligned tool to attract funding from UNICEF, World Bank EdTech programs, and Sri Lanka's Ministry of Education digital education initiative.
- Freemium consumer app: Free core functionality with premium curriculum packs and multi-child tracking features available through in-app purchase.

6.4 Competitive Advantage

No existing mobile literacy application combines vision-based air-writing with real-time stroke-direction validation and AR corrective overlays on commodity Android hardware. The system's patent-pending novelty (Stroke-Direction Awareness with real-time AR ghost-path correction) creates a defensible differentiation from LetterSchool, Writing Wizard, and similar products that are limited to touchscreen tracing without trajectory analysis.

6.5 Intellectual Property Considerations

The CNN trajectory model architecture, stroke validation engine, and AR corrective overlay pipeline constitute novel technical contributions suitable for a provisional patent application. The research findings will be submitted for publication in an IEEE Educational Technology or ACM CHI venue, establishing academic priority. Open-source release of non-core utilities (data preprocessing scripts) is planned to build community and attract collaborative development.

CHAPTER 7: BUDGET AND JUSTIFICATION

Item	Category	Estimated Cost (LKR)	Justification
AR-Compatible Smartphone	Hardware	60,000	Required to test AR tracking, camera input, and rendering performance
Google Colab Pro subscription (6 months)	Cloud / Compute	25,000	GPU-accelerated CNN training; avoids local GPU hardware cost
Unity Pro license (student)	Software	0	SLIIT institutional Unity education license covers student use
MediaPipe / TensorFlow	Software	0	Open source; no license cost
Data collection stationery & printing	Consumables	15,000	Consent forms, instruction sheets for UAT sessions
External Storage	Hardware	30,000	Required for storing 3D models, datasets, and project files
Transport for school UAT visits	Travel	8,000	Two school visits for pre-test/post-test sessions
3D asset creation (Blender)	Software	0	Open source
Contingency (10%)	Contingency	6,050	Buffer for unforeseen expenses

Total Estimated Budget: LKR 138,000

Note: The budget is deliberately kept minimal by leveraging open-source tools, SLIIT institutional licenses, and cloud computing credits. The primary cost drivers are hardware (required for realistic performance testing) and cloud GPU time for model training. No proprietary dataset purchases are required as training data is collected directly from participants under the approved ethical protocol.

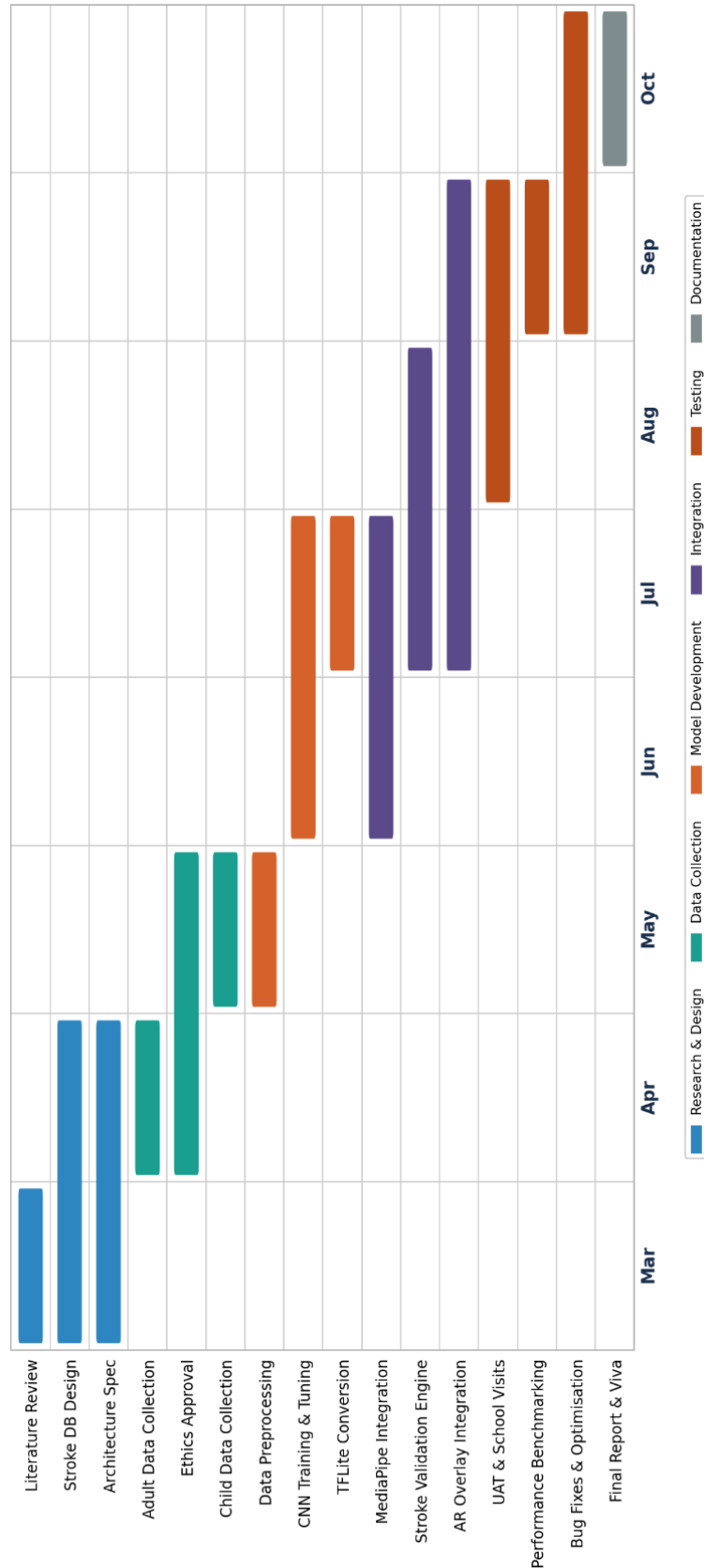
CHAPTER 8: WORK BREAKDOWN STRUCTURE

8.1 Task Breakdown

Phase	Task	Duration	Milestone
1 – Research & Design	Literature review: air-writing & CNN models	Weeks 1–2	
1 – Research & Design	Canonical stroke database design (A–Z paths)	Weeks 2–3	Stroke DB ready
1 – Research & Design	System architecture specification	Weeks 2–4	Architecture doc
2 – Data Collection	Adult volunteer trajectory dataset collection	Weeks 4–5	
2 – Data Collection	Ethics approval submission & approval	Weeks 4–6	Ethics clearance
2 – Data Collection	Child participant dataset collection (with consent)	Weeks 6–7	Dataset complete
3 – Model Development	Data preprocessing pipeline	Weeks 5–6	
3 – Model Development	CNN model design, training & tuning	Weeks 7–9	
3 – Model Development	TensorFlow Lite conversion & benchmarking	Week 9	Model $\geq 90\%$ acc.
4 – Integration	MediaPipe hand tracking Android integration	Weeks 8–10	
4 – Integration	Stroke validation engine implementation	Weeks 9–11	
4 – Integration	AR corrective overlay (Unity integration)	Weeks 10–12	Integration done
5 – Testing & Validation	UAT sessions (school visits, pre/post-test)	Weeks 12–14	
5 – Testing & Validation	System performance benchmarking	Week 14	
5 – Testing & Validation	Bug fixes and accuracy optimization	Weeks 14–15	System validated
6 – Documentation	Final report writing	Weeks 15–16	
6 – Documentation	Viva preparation & demonstration setup	Week 16	Submission

8.2 Gantt Chart

The Gantt chart below illustrates the 16-week project timeline for the AI Air-Writing Trajectory Tracking component, aligned with the overall AR Educational Suite project schedule.



REFERENCES

- [1] A. D. S. D. V. a. C. D. K. M. B. Ibáñez, "Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness," in IEEE Global Engineering Education Conference (EDUCON), 2014.
- [2] I. Radu, "Augmented reality in education: A meta-review and cross-media analysis," in IEEE Int. Symposium on Mixed and Augmented Reality (ISMAR), 2012.
- [3] H. K. a. D. Schmalstieg, "Mathematics and geometry education with collaborative augmented reality," in ACM SIGGRAPH Int. Conf. on Virtual Reality Continuum and Its Applications, 2003.
- [4] X. T. a. L. Z. C. Y. Wang, "Sketch2Model: Sketch-based 3D modeling for novice users," in IEEE Virtual Reality Conference (VR), 2011.
- [5] Q. H. W. X. a. X. Z. P. Xu, "Deep learning for sketch-based 3D shape retrieval," in IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW), 2015.
- [6] Z. L. H. H. E. K. a. Y. Y. X. Han, "DeepSketch2Face: A deep learning based sketching system for 3D face modeling," in Eurographics Symposium on Geometry Processing, 2018.
- [7] M. G. a. T. S. C. Amma, "Airwriting: Hands-free mobile text input by spotting and continuous recognition of 3D-space handwriting with inertial sensors," in ACM Int. Conf. on Multimodal Interaction (ICMI), 2012.
- [8] Y. X. F. Y. a. C. L. X. Chen, "Online handwritten text recognition using recurrent neural networks," in Conf. on Document Analysis and Recognition (ICDAR), 2015.
- [9] P. X. a. Y. W. S. Yang, "Vision-based air-writing recognition using deep convolutional neural networks," in IEEE Int. Conf. on Image Processing (ICIP), 2017.
- [10] S. G. K. K. a. J. K. S. Molchanov, "Hand gesture recognition with 3D convolutional neural networks," in IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW), 2015.
- [11] T. S. S.-E. W. a. Y. S. Z. Cao, "Realtime multi-person 2D pose estimation using part affinity fields," in IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2017.

APPENDICES

Appendix A – Topic Assessment Form Summary

Project ID: R26-IM-005

Topic: AR Educational Suite for Primary Students' Geometry and Literacy

Research Group: MR – Mixed Reality

Specialization: Interactive Media (IM)

Supervisor: Mr. Aruna Ishara Gamage | Co-Supervisor: Nushkan Nismi

Appendix B – Canonical Stroke Database Format (Sample)

Each entry in the stroke database is a JSON object containing letter (string), strokes (array of stroke objects). Each stroke object contains start region (bounding box relative to a 1×1 normalized space), waypoints (array of $[x, y]$ coordinates at 10% intervals), direction (string: e.g., "top-to-bottom", "left-to-right"), order (integer: stroke index within the letter).

Example entry for letter "A": two diagonal strokes meeting at a peak, followed by a horizontal crossbar three strokes total, each with defined start regions, waypoints, and directionality constraints.

Appendix C – Ethical Clearance Checklist

- Research involves participants under 18: YES
- Informed consent from parents/guardians required: YES
- Data anonymization protocol in place: YES (coordinate sequences only, no video stored)
- On-device processing only no external data transmission: YES
- SLIIT Ethics Review Committee submission: Required before data collection
- Right to withdraw at any time without consequence: Communicated to all participant.

