ELLIE'S ADVENTURES

A gamified application to improve **communication**, **emotional**, **verbal** and **motor skills** in children with ASD.

SLIIT FACULTY OF COMPUTING

RP 24-25J-114 *2025-03-20*

PROJECT SUPERVISION

DR. KALPANI MANATHUNGA

Head, Department of Computer Science and Software Engineering

PROF. SAMANTHA THELIJJAGODA

Pro Vice-Chancellor (SLIIT Research & International)

DR. ASIRI HEWAMALAGE

National Program Manager, Child Development and Special Needs, Family Health Bureau, Ministry of Health

MR. SAMITHA VIDANAARACHCHI

Former Lecturer - SLIIT / PhD Researcher at Murdoch University, Australia

PROJECT TEAM

UMAIRA MARIKKAR

Year 4 Undergraduate
SLIIT

THEJANI MALLAWAARACHCHI

Year 4 Undergraduate
SLIIT

JATHURSHAN MANISTAR

Year 4 Undergraduate
SLIIT

PAWAN SENPURA

Year 4 Undergraduate
SLIIT

03 INTRODUCTION

WHAT IS ELLIE'S ADVENTURES?

A Gamified Mobile App

Comprises serious-game exercises addressing several focus areas

Aims to engage the child to improve in these areas

RESEARCH CONDITIONS

Autism is a spectrum of disorders, this research will be conducted for a selected community of the spectrum

1. Children between the age of 3 to 12.

2. Activities to be done under parental supervision.

3. Children who can speak but not comprehensively.

80 % [3]
speech development

87 % [4]
Motor Skills

42 - 50 % [3]
Mental retardation

40.9 % [1]
language comprehension

Every 1 in 100 [1]

7.4% OF CHILDREN [2]

EVERY 1 IN 93 [2]

Affected from ASD

08 ETHNOGRAPHIC STUDIES

Christine Activity Center
Malabe

ETHNOGRAPHIC STUDIES

- Family Health Bureau
- Namaste Program for Autism Awareness in Sri Lanka

03 MAIN OBJECTIVES

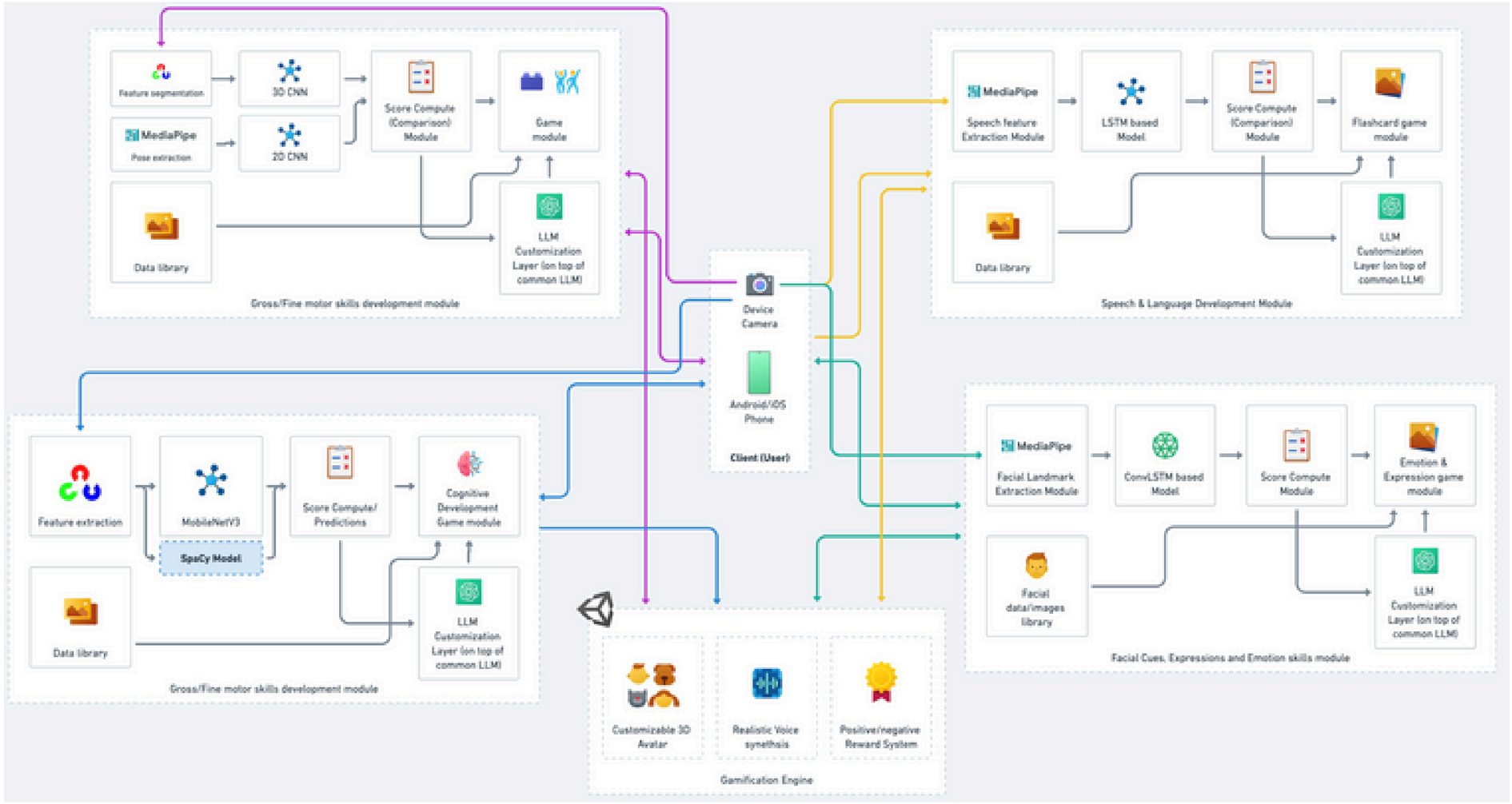
SPEECH AND LANGUAGE DEVELOPMENT

COGNITIVE AND VERBAL SKILL DEVELOPMENT

FINE AND GROSS MOTOR SKILL DEVELOPMENT

FACIAL EXP. AND EMOTION SKILLS DEVELOPMENT

SYSTEM ARCHITECTURE



IMPROVING GROSS & FINE MOTOR SKILLS

Mallawaarachchci T. D. R.
IT21282836

[Software Engineering]

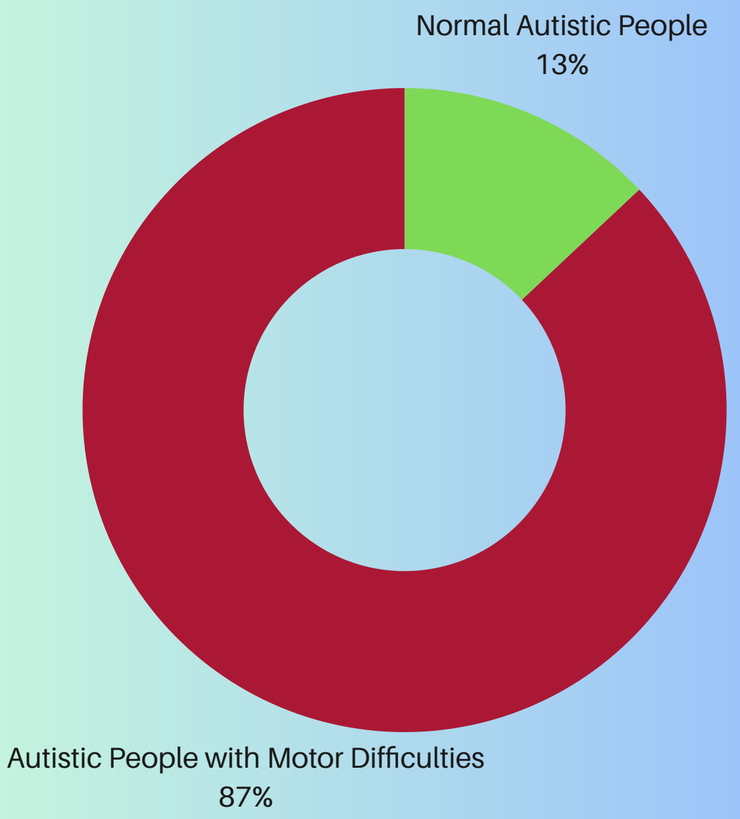
Background

Children with ASD often experience **delays in motor skills development**, affecting their daily activities and overall growth.

Research indicates ASD children has significant **delays in gross motor skills (6.7%) and fine motor skills (38.5%) compared to typically developing children** [2]

Motor skill activities can stimulate **cognitive development** and **improve attention, planning, and problem-solving abilities.**

Addressing both **gross and fine motor skills** is essential for comprehensive motor skill development.



According to the latest estimate **87% of the Autistic people** face some sort of **motor difficulty**. [1]

Research Questions

How to improve gross & fine motor skills of a child with autism ?

1. How to identify and measure a child's **current motor skill abilities** at home?

2. What **milestones and criteria should be used to assess** the child's motor skill progress?

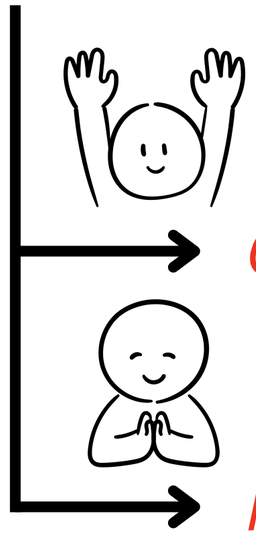
3. What **activities** or games can effectively target **both gross and fine motor skill development** at home?

4. What **methodologies** can be used to enhance motor skills development?

5. How to provide the **necessary guidance** to support the child's motor skill improvements?

Main and Specific Objectives

Improving Motor Skills of Children with ASD



Activities involve larger set of muscles

Gross Motor Skills

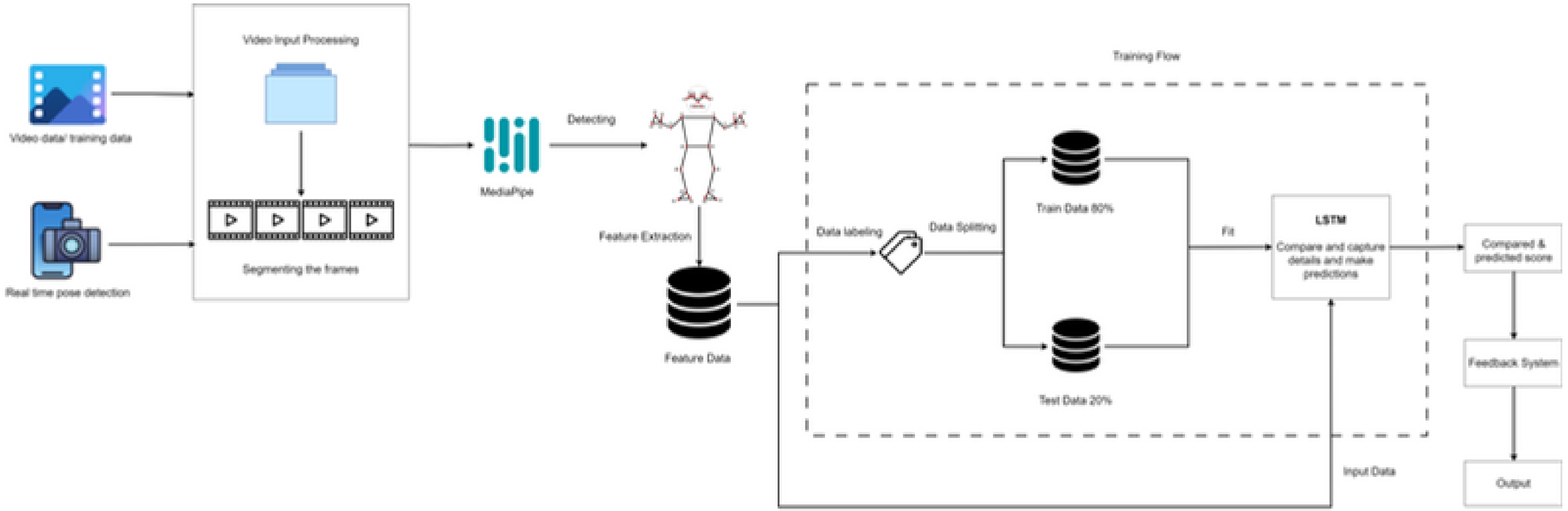
Activities involve small set of muscles

Fine Motor Skills

- Train the model to identify the actions performed by the child.
- Identify and evaluate the child's ability to mimic the given exercises correctly.
- Provide feedback and encouragement based on the performance.
- Train the YOLO model to identify objects correctly.
- Evaluate ability of the child to create the color pattern and give the feedback.

GROSS MOTOR SKILLS DEVELOPMENT

System Diagram



OBJECTIVE 01

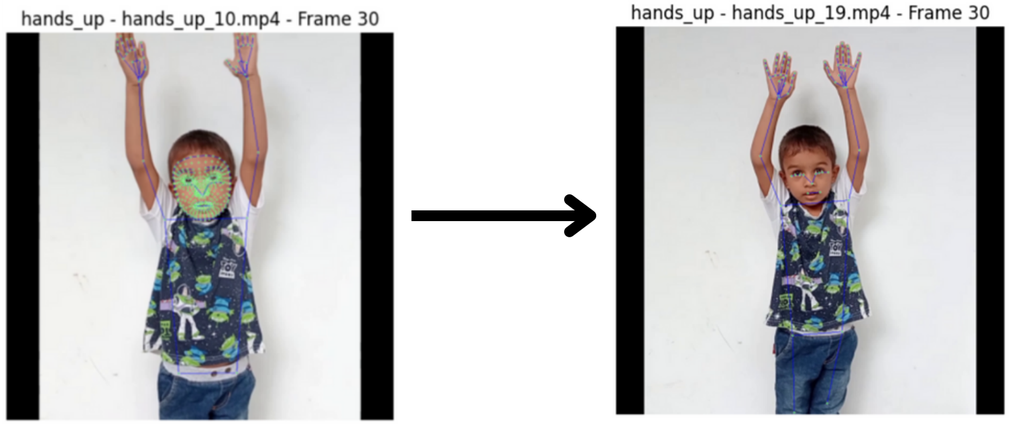
Train the model to identify the actions performed by the child

OVERVIEW

***Objective 01 - 100% completed by PP1**

Objective 01 - Train the model to identify the actions performed by the child.

Evidence : Preprocessed data for keypoint extraction

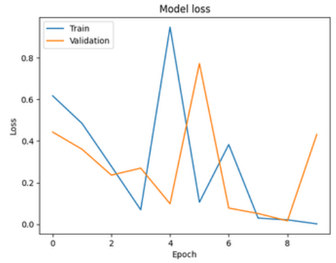
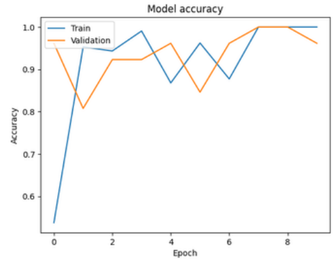
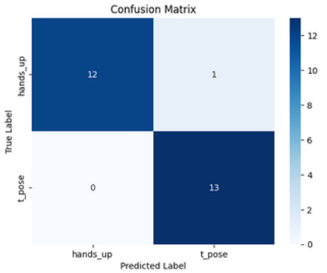
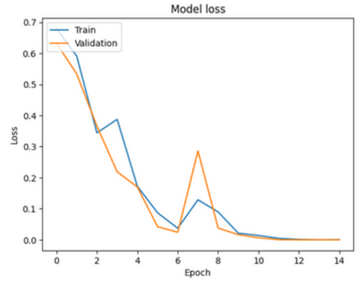
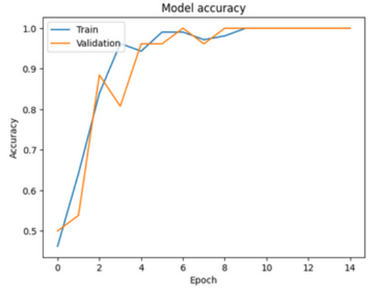
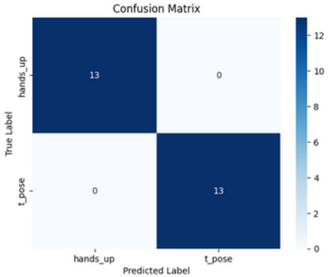
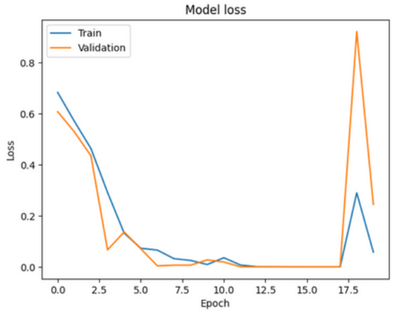
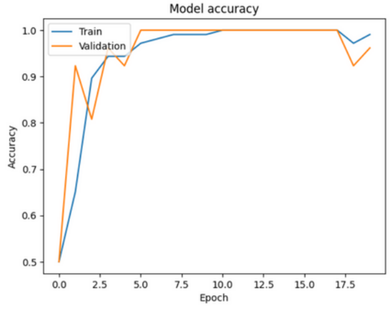
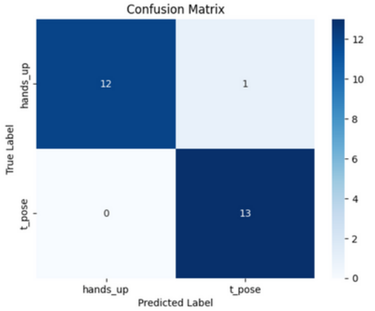


Utilizes MediaPipe Holistic pipeline for keypoint extraction [3]

```
model = Sequential() # sequential API
model.add(LSTM(64, return_sequences=True, activation='relu', input_shape=(sequence_length, 1662)))
model.add(LSTM(128, return_sequences=True, activation='relu'))
model.add(LSTM(64, return_sequences=False, activation='relu')) # not returning sequences into dense
model.add(Dense(64, activation='relu'))
model.add(Dense(32, activation='relu'))
model.add(Dense(actions.shape[0], activation='softmax'))
```

```
model = Sequential() # sequential API
model.add(LSTM(64, return_sequences=True, activation='relu', input_shape=(sequence_length, 258)))
model.add(LSTM(128, return_sequences=True, activation='relu'))
model.add(LSTM(64, return_sequences=False, activation='relu')) # not returning sequences into dens
model.add(Dense(64, activation='relu'))
model.add(Dense(32, activation='relu'))
model.add(Dense(actions.shape[0], activation='softmax'))
```

Objective 01 Contd.

Epochs	No of Videos	Data Split	Loss Graph	Accuracy Graph	Confusion Matrix
10	66 videos per action	Train - 106 Test - 26			
15	66 videos per action	Train - 106 Test - 26			
20	66 videos per action	Train - 106 Test - 26			

OBJECTIVE 02 & 03

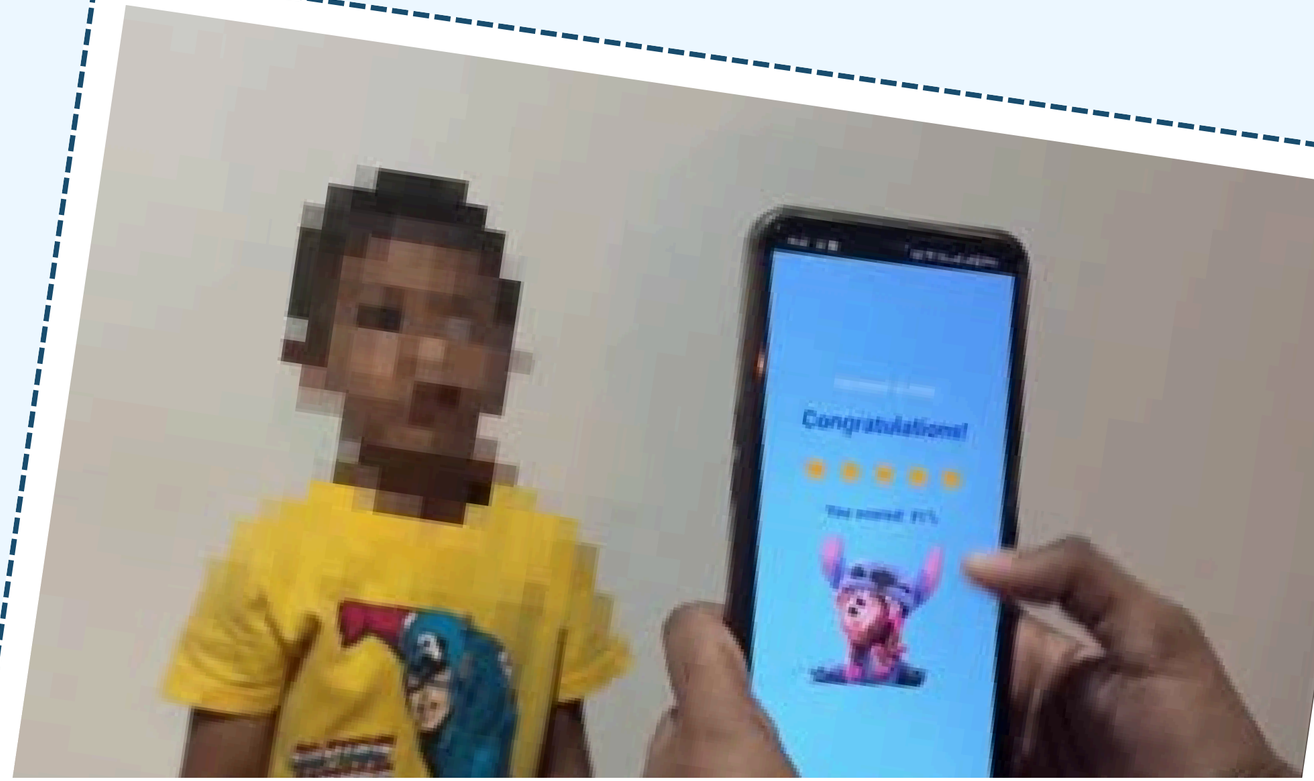
Evaluate the child's ability to mimic exercises correctly Provide feedback with encouragement based on performance.

OVERVIEW

*Objective 02 - 70% completed by PP1

30% completed by PP2

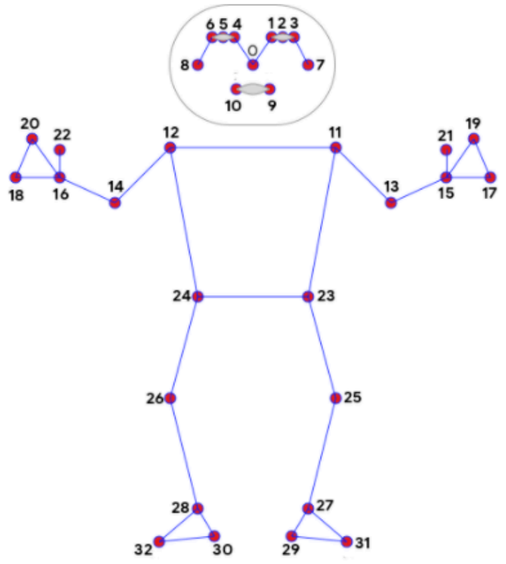
*Objective 03 - 100% completed by PP2



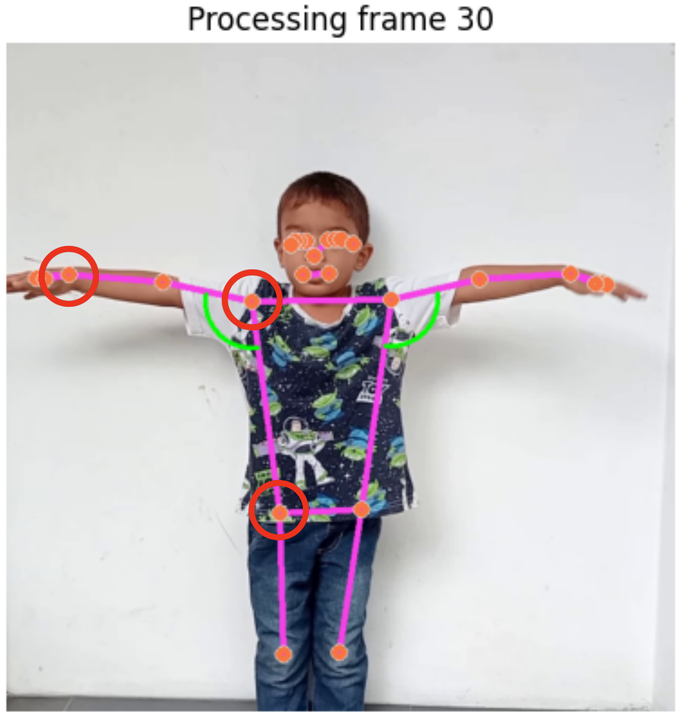
Objective 02 & 03 - Provide feedback and encouragement based on the performance of the child.

Evidence : Implementation of Angle Heuristics and threshold-based validation

- 3 main key landmarks were used to calculate the angle. (wrist, shoulder, hip)
- Python Numpy arctan2() function was used to estimate the angle. [4]
- Normalization to ensure angle is between [0, 360]



0. nose
1. left_eye_inner
2. left_eye
3. left_eye_outer
4. right_eye_inner
5. right_eye
6. right_eye_outer
7. left_ear
8. right_ear
9. mouth_left
10. mouth_right
11. left_shoulder
12. right_shoulder
13. left_elbow
14. right_elbow
15. left_wrist
16. right_wrist
17. left_pinky
18. right_pinky
19. left_index
20. right_index
21. left_thumb
22. right_thumb
23. left_hip
24. right_hip
25. left_knee
26. right_knee
27. left_ankle
28. right_ankle
29. left_heel
30. right_heel
31. left_foot_index
32. right_foot_index



Objective 02 & 03 Contd.

Evidence : Angle calculation and reward based feedback mechanism

- **Deviation from ideal angles** were considered for precise scoring.
- **Midpoints provide balanced threshold** to accommodate natural variations
- Peronalized feedbacks based on the performed angle deviations.
- Studies on upper limb kinematics suggests

Pose	Minimum considered angle	Maximum considered angle	Mid Point (Ideal Angle)
hands_up	150	190	170
t_pose	80	110	95

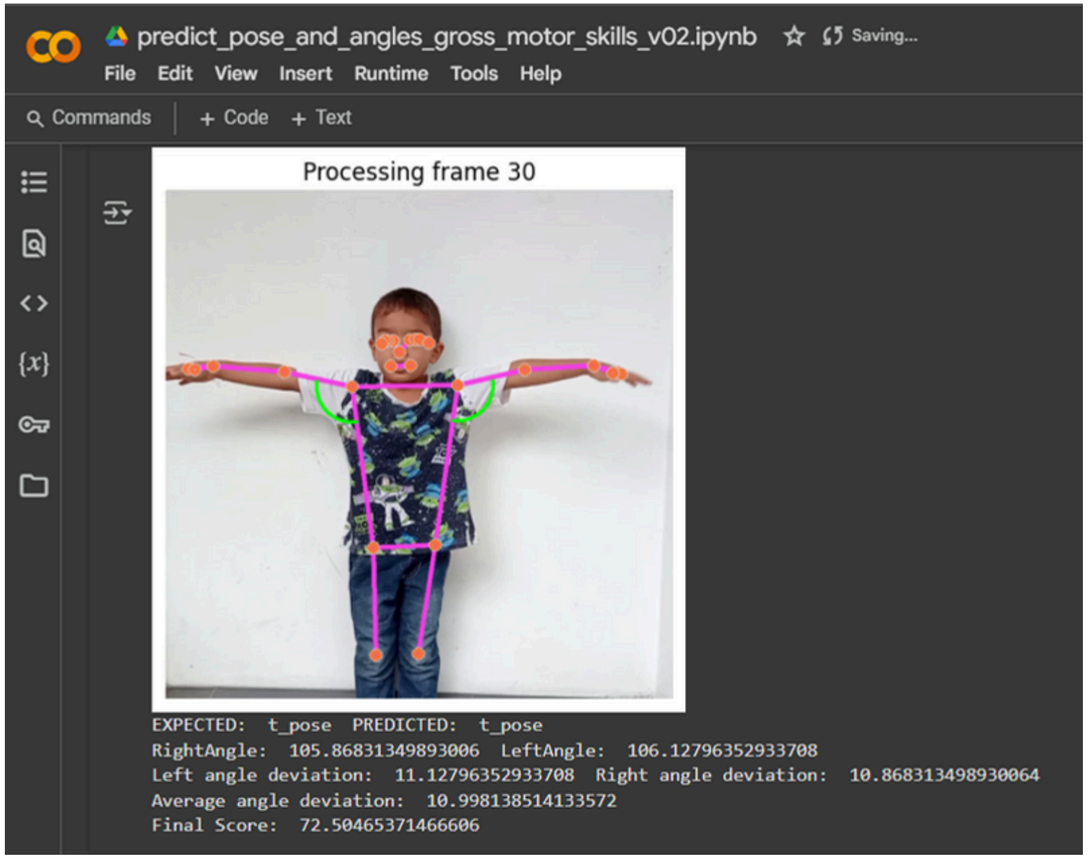
$$\text{Score} = \max\left(100 - \frac{\text{Avg Deviation}}{40} \times 100, 0\right)$$

Avg Deviation = (Left Deviation + Right Deviation) / 2

Max Deviation Tolerance = 40°

Objective 02 & 03 Contd.

Evidence : Angle calculation and reward based feedback mechanism



FINE MOTOR SKILLS DEVELOPMENT

Improving Fine Motor Skills of Children with ASD

[6]

Reasons to change the proposed approach:

- **To follow a more standardized way.**
- **To consider the safety of the child.**

First Box Of Color Tablets

Third Box Of Color Tablets

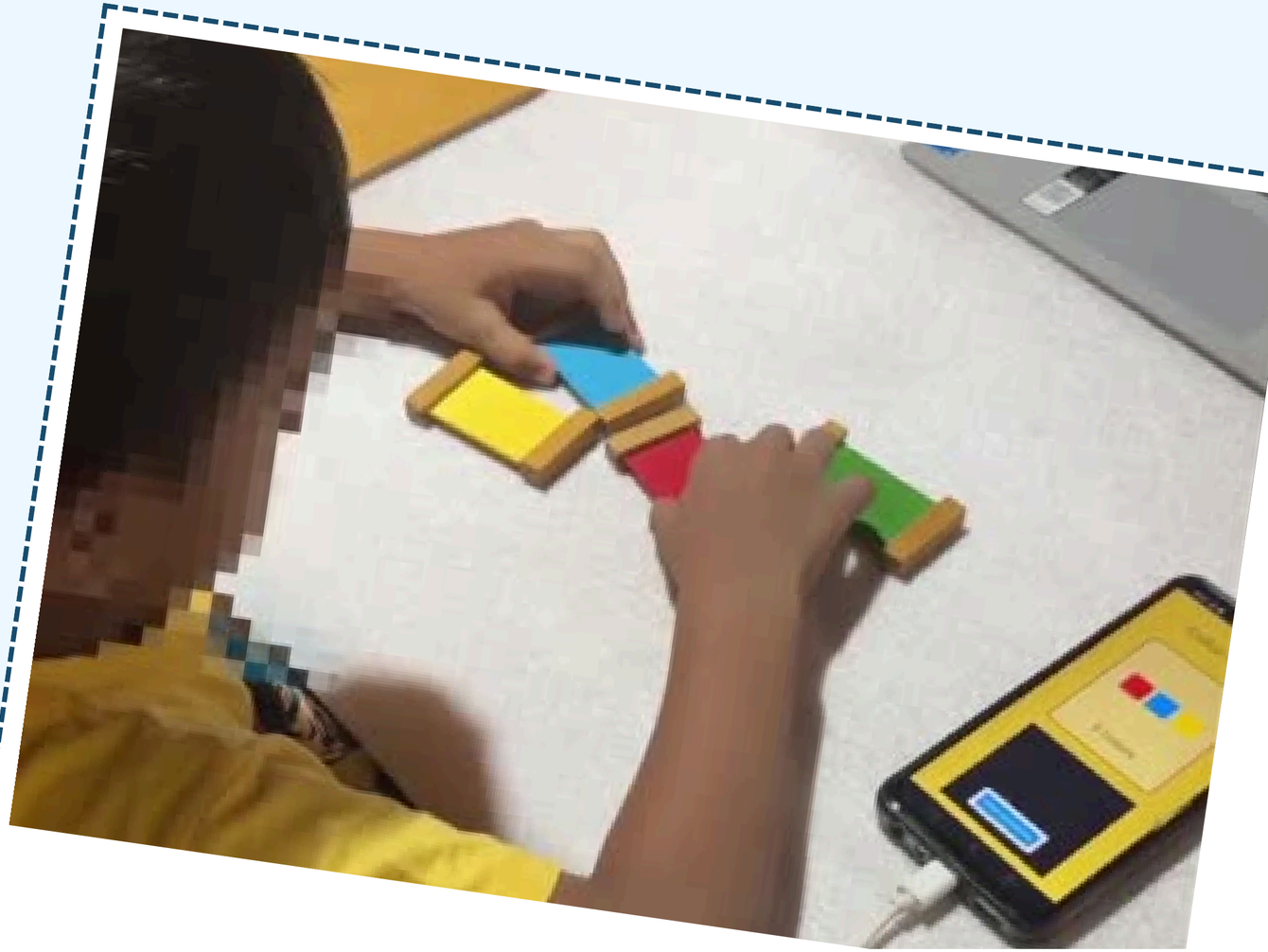
Second Box Of Color Tablets

[5]

OBJECTIVE 04

Train the YOLO model to identify objects correctly.

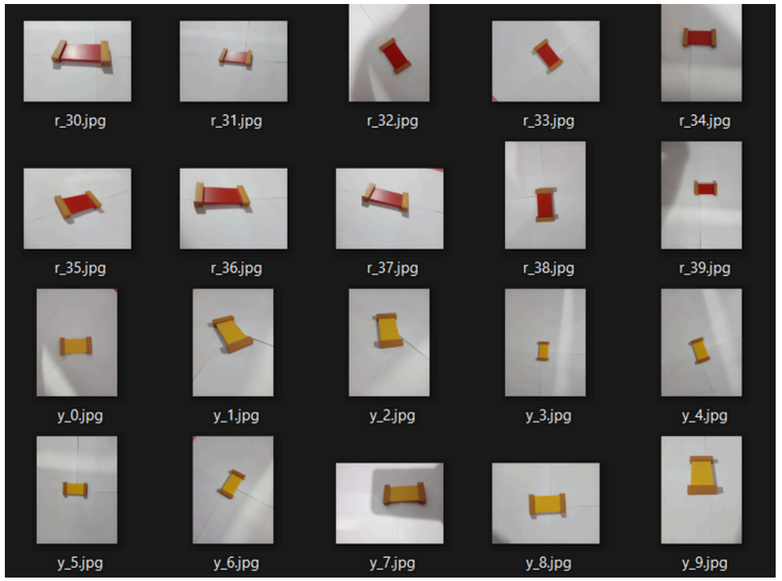
OVERVIEW

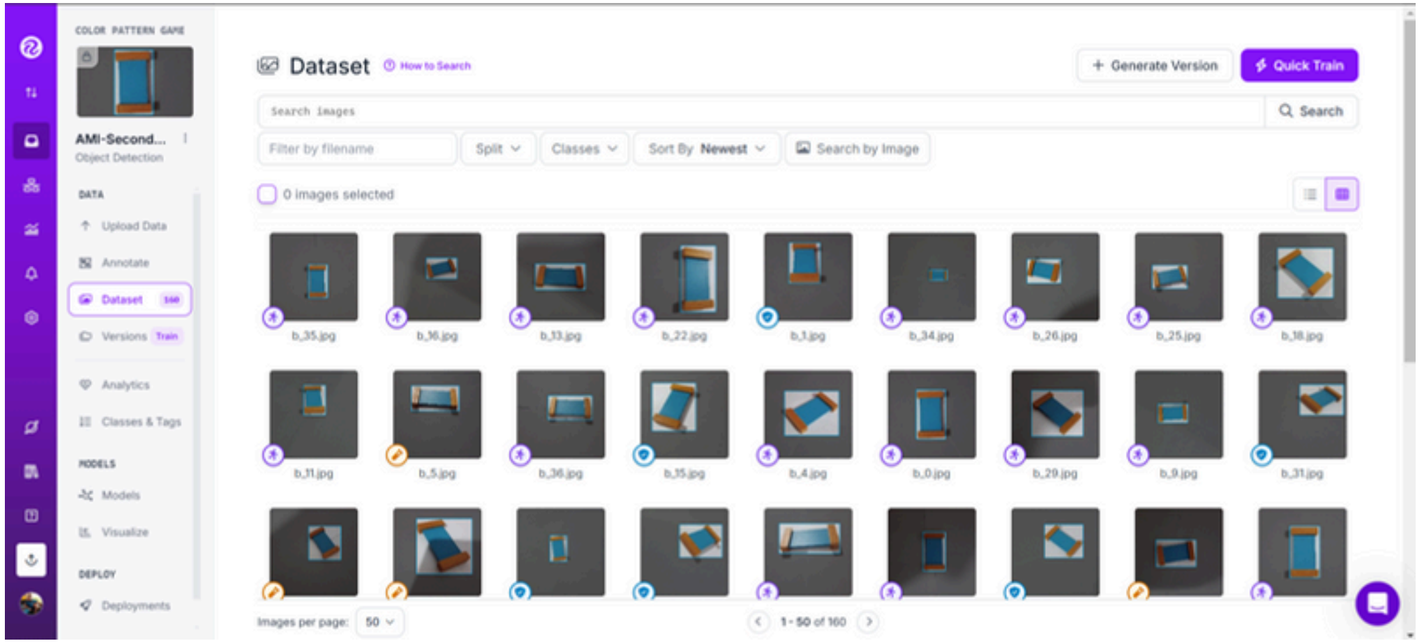


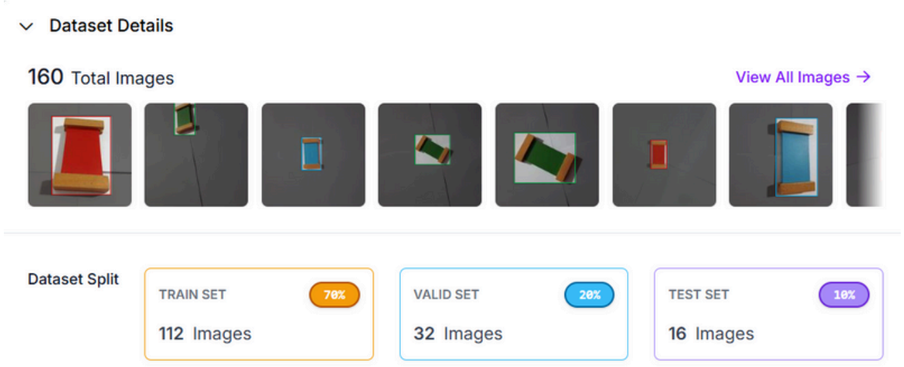
*Objective 04 - 100% completed by PP1

Objective 04 - Train the YOLO model to identify objects.

Evidence: Annotating collected AMI Second Box of Color Pallets dataset.





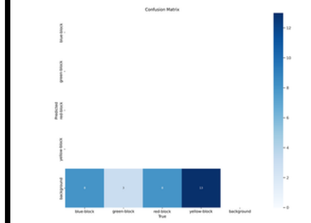
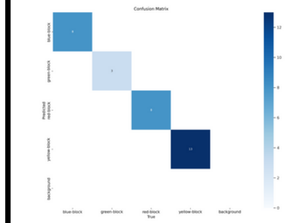
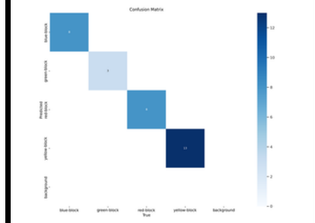


[7]

For object detection on the COCO dataset, YOLO11n model achieves 39.5 mAPval50-95, showing improvements over YOLOv8n [8]

Objective 04 - Train the YOLO model to identify objects.

Comparison between Version 02 (20 Epochs) and Version 03 (50 Epochs) model training.

Metric	Version 01 5 Epochs	Version 02 20 Epochs	Version 03 50 Epochs	Comments
Precision (P)	0.00599	0.982	0.991	Version 3 shows a slightly higher overall precision.
Recall (R)	1.0	1.0	1.0	Both models achieve perfect recall (all true positives detected).
mAP@50	0.399	0.995	0.995	Model 01 has significantly lower mAP@50, which indicates it struggled to detect objects accurately at an IoU threshold of 0.5. Models 02 and 03 perform much better.
mAP@50-95	0.367	0.971	0.98	Version 3 performs better, indicating a stronger ability to generalize across IoU thresholds.
Confusion Matrix				

Objective 04 - Train the YOLO model to identify objects.

Evidence: Testing an external image with multiple color pallets

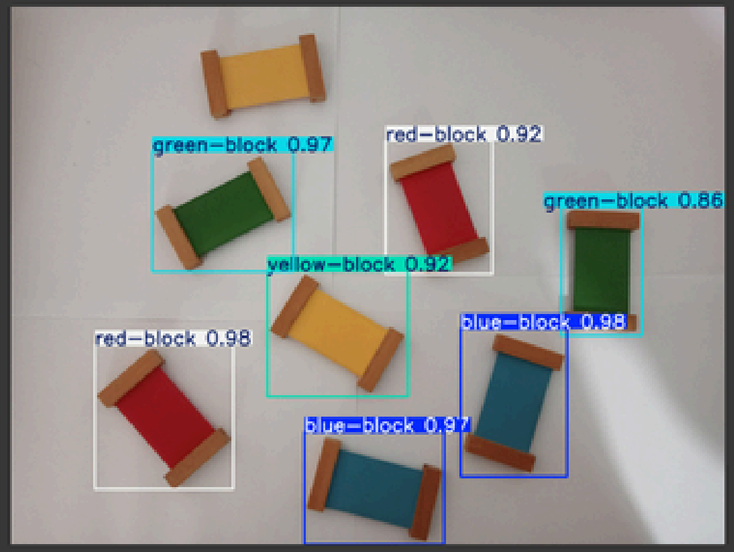
```
[ ]  1 !cp -r runs /content/drive/MyDrive/Thejani_Research_Y4/ColorPatternModel/v02-160_img_with_20_epochs
```

```
[ ]  1 # predict an external image
     2 !yolo task=detect mode=predict model={HOME}/runs/detect/train3/weights/best.pt source=/content/drive/MyDrive/Thejani_Research_Y4/ColorPatternModel/test_img02.jpg conf=0.8 save=true
```

```
Ultralytics 8.3.39 🚀 Python-3.10.12 torch-2.5.1+cu121 CUDA:0 (Tesla T4, 15102MiB)
YOLO11n summary (fused): 238 layers, 2,582,932 parameters, 0 gradients, 6.3 GFLOPs

image 1/1 /content/drive/MyDrive/Thejani_Research_Y4/ColorPatternModel/test_img02.jpg: 512x640 2 blue-blocks, 2 green-blocks, 2 red-blocks, 1 yellow-block, 42.8ms
Speed: 4.5ms preprocess, 42.8ms inference, 565.9ms postprocess per image at shape (1, 3, 512, 640)
Results saved to runs/detect/predict6
💡 Learn more at https://docs.ultralytics.com/modes/predict
```

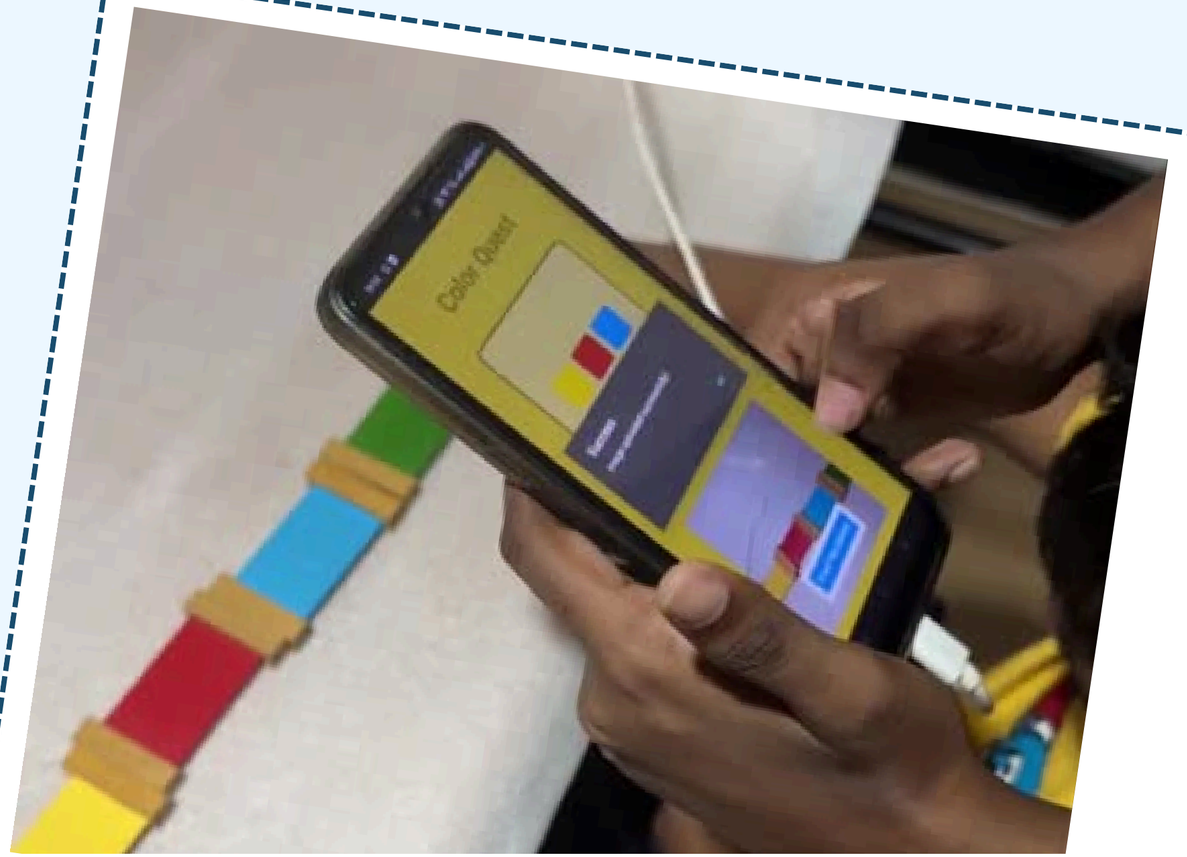
```
[ ]  1 IPyImage(filename=f'{HOME}/runs/detect/predict6/test_img02.jpg', width=600)
```



OBJECTIVE 05

Evaluate ability of the child to create the color pattern and give feedback

OVERVIEW



*Objective 05 - 100% completed by PP2

Objective 05 - Evaluate ability of the child to create the color pattern and give the feedback.

Evidence : Horizontal alignment evaluation

main.py | fine.py | Dockerfile | temp_test02.txt U

runs > detect > predict > labels > temp_test02.txt

```
1 0.223554 0.688182 0.222747 0.173171
2 0.647406 0.695298 0.248687 0.175732
0 0.820207 0.695819 0.257781 0.176134
```

YOLO output:

class_id, x_center, y_center, width, height

But why to pixel values?

Output bounding box coordinates are normalized between 0 and 1.

Small variations in normalized values can result in large pixel differences in high-resolution images

```python
# scales normalized bounding box coordinates to pixel values based on image dimensions
rectangles = [
    (int(x * image_width), int(y * image_height), int(w * image_width), int(h * image_height), class_to_color[cls])
    for cls, x, y, w, h in sorted_yolo_output
]
```

Objective 05 - Evaluate ability of the child to create the color pattern and give the feedback.

Evidence : Horizontal alignment evaluation

```python
# checks all boxes
def is_horizontally_aligned(rectangles, img_height, tolerance_percentage=20):
    # average bounding box height
    avg_box_height = sum(h for _, y, _, h, _ in rectangles) / len(rectangles)

    tolerance = avg_box_height * (tolerance_percentage / 100)
    print(f"Tolerance Percentage: {tolerance_percentage}", f"Tolerance: {tolerance}")
```

Tolerance value based on average bounding box height

```python
# check if difference between each pair of y-positions is within the tolerance
for i in range(len(y_positions)):
    for j in range(i + 1, len(y_positions)):
        difference = abs(y_positions[i] - y_positions[j])
        print("i: ", y_positions[i], " j: ", y_positions[j], " difference: ", difference)
        if difference > tolerance:
            print(f"Boxes at indices {i} and {j} are not aligned.")
            misalignment_score += difference
            isAligned = False

max_possible_misalignment = len(y_positions) * tolerance
alignment_score = max(0, (1 - misalignment_score / max_possible_misalignment)) * 100  # score
```

Alignment score considering the horizontal alignment

Objective 05 - Evaluate ability of the child to create the color pattern and give the feedback.

Evidence : Color pattern evaluation

```
correct_colors_in_position = 0
correct_colors_out_of_position = 0
incorrect_colors = 0

# check how many colors are correct in position
for i in range(len(expected_colors)):
    if i < len(detected_colors):
        if detected_colors[i] == expected_colors[i]:
            correct_colors_in_position += 1
        if detected_colors[i] in expected_colors:
            correct_colors_out_of_position += 1
        else:
            incorrect_colors += 1

total_colors = len(expected_colors)

correct_position_score = (correct_colors_in_position / total_colors) * 70
out_of_position_score = (correct_colors_out_of_position / total_colors) * 30
incorrect_penalty = (incorrect_colors / total_colors) * 30

pattern_score = max(0, correct_position_score + out_of_position_score - incorrect_penalty)
```

*correct_colors_in_position + 70%

*correct_colors_out_of_position + 30%

*incorrect_colors - 30%

Objective 05 - Evaluate ability of the child to create the color pattern and give the feedback.

Evidence : Final score calculation

Pattern Score Calculation:

$$\text{pattern\_score} = \max(0, \text{correct\_position\_score} + \text{out\_of\_position\_score} - \text{incorrect\_penalty})$$

Alignment Score Calculation:

$$\text{alignment\_score} = \max\left(0, \left(1 - \frac{\text{misalignment\_score}}{\text{max\_possible\_misalignment}}\right)\right) \times 100$$

Total Score Calculation:

$$\text{total\_score} = (\text{alignment\_score} \times 0.5) + (\text{pattern\_score} \times 0.5)$$

Objective Completion Status

	Status
Train the model to identify the actions performed by the child.	100%
Identify and evaluate the child's ability to mimic the given exercises correctly.	100%
Provide feedback and encouragement based on the performance.	100%
Train the YOLO model to identify objects correctly.	100%
Evaluate ability of the child to create the color pattern and give the feedback.	100%

100%

References

[1] C. J. Zampella, L. A. L. Wang, M. Haley, A. G. Hutchinson and A. D. Marchena, “Motor Skill Differences in Autism Spectrum Disorder: a Clinically Focused Review,” Current psychiatry reports, 2021.

[2] A. M. Nordin, J. Ismail and N. K. Nor, “Motor Development in Children With Autism Spectrum Disorder,” Frontiers in Pediatrics, 2021.

[3] MediaPipe Holistic

[4] Python Numpy arctan2() - Compute Arc Tangent

[5] AMI Montessori Materials

[6] Second Box Of Color Tablets

[7] roboflow

[8] YOLOv8 vs YOLO11: A Technical Comparison

IMPROVING FACIAL EXPRESSIONS AND EMOTIONAL SKILLS

Helapalla K. O. P. S.
IT21264016

[Software Engineering]

BACKGROUND

- Children with ASD have trouble expressing emotions as facial expressions. [1]
- Most of them have trouble recognizing the emotions of someone in front of them. [1]

LEADS TO →

- Difficulty in communication
- Social isolation
- Frustration of parents [4]

OBJECTIVES

01 Predict the child's facial expression accuracy

02 Skill assessment to understand child's current skills

03 Realtime attentiveness measuring module

04 Gamification with reward based customizations

05 Interactive 3D avatar that supplements the game flow

OBJECTIVE 01

A deep learning model that detects child's FCEs and predicts the accuracy with Computer Vision

Action Unit	Description	Facial Muscle	Example
1	Inner Brow Raiser	*Frontalis, pars medialis*	
2	Outer Brow Raiser (unilateral, right side)	*Frontalis, pars lateralis*	
4	Brow Lowerer	*Depressor Glabellae, Depressor Supercilli, Currugator*	
5	Upper Lid Raiser	*Levator palpebrae superioris*	

OBJECTIVE 01

SELECTING THE BEST MODEL

with Unfrozen Layers

IMAGE NET V2

Light weight
Powerful

```
Evaluating on test set:
789/789 ━━━━━━━━━━━━━━━━━━━━ 12
Test MSE: 0.7749
Test MAE: 0.5138
Test AU Accuracy (±1): 0.2219

Model saved successfully!
```

IMAGE NET V3

More powerful than V2
More suitable for realtime device support

```
Evaluating on test set:
1577/1577 ━━━━━━━━━━━━━━━━━━━━
Test MSE: 0.1614
Test MAE: 0.3426
Test AU Accuracy (±1): 0.4099

Model saved successfully!
```

REST NET 50

```
Evaluating on test set:
1577/1577 ━━━━━━━━━━━━━━━━━━━━ 152
Test MSE: 0.3005
Test MAE: 0.5437
Test AU Accuracy (±1): 0.2037

Model saved successfully!
```

with Unfrozen Layers

REDUCE LEARNING RATE ON PLATEAU

OBJECTIVE 01

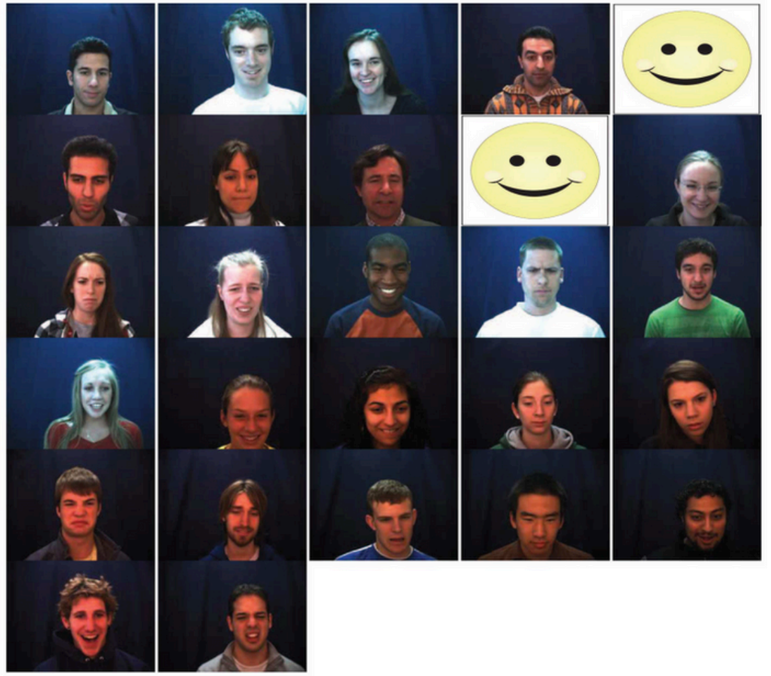
THE DATASET

DISFA+

More than 65000
Labelled Images

Labelled by
FACS expert

Emotion	Action units
Happiness	6+12
Sadness	1+4+15
Surprise	1+2+5B+26
Fear	1+2+4+5+7+20+26
Anger	4+5+7+23
Disgust	9+15+17
Contempt	R12A+R14A



OBJECTIVE 02

Skill assessment module to identify the child’s current abilites

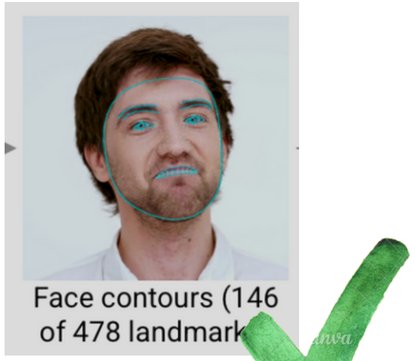
OBJECTIVE 03 & 04

Realtime attention tracking and constructive feedback + gamification

Facial landmark extraction

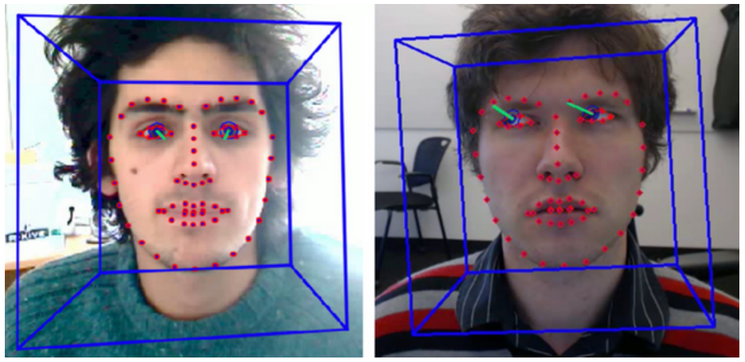
MediaPipe

- Optimized for mobile devices
- Inference time 6ms
- Provides 478 3D face landmarks [3]



OpenFace and OpenPose

- Comparatively slower
- Comparatively higher inference time
- Provides only 68 2D facial landmarks [3]



OBJECTIVE 05

3D avatar that mirrors exact facial cues and blendshapes real-time from camera

REFERENCES

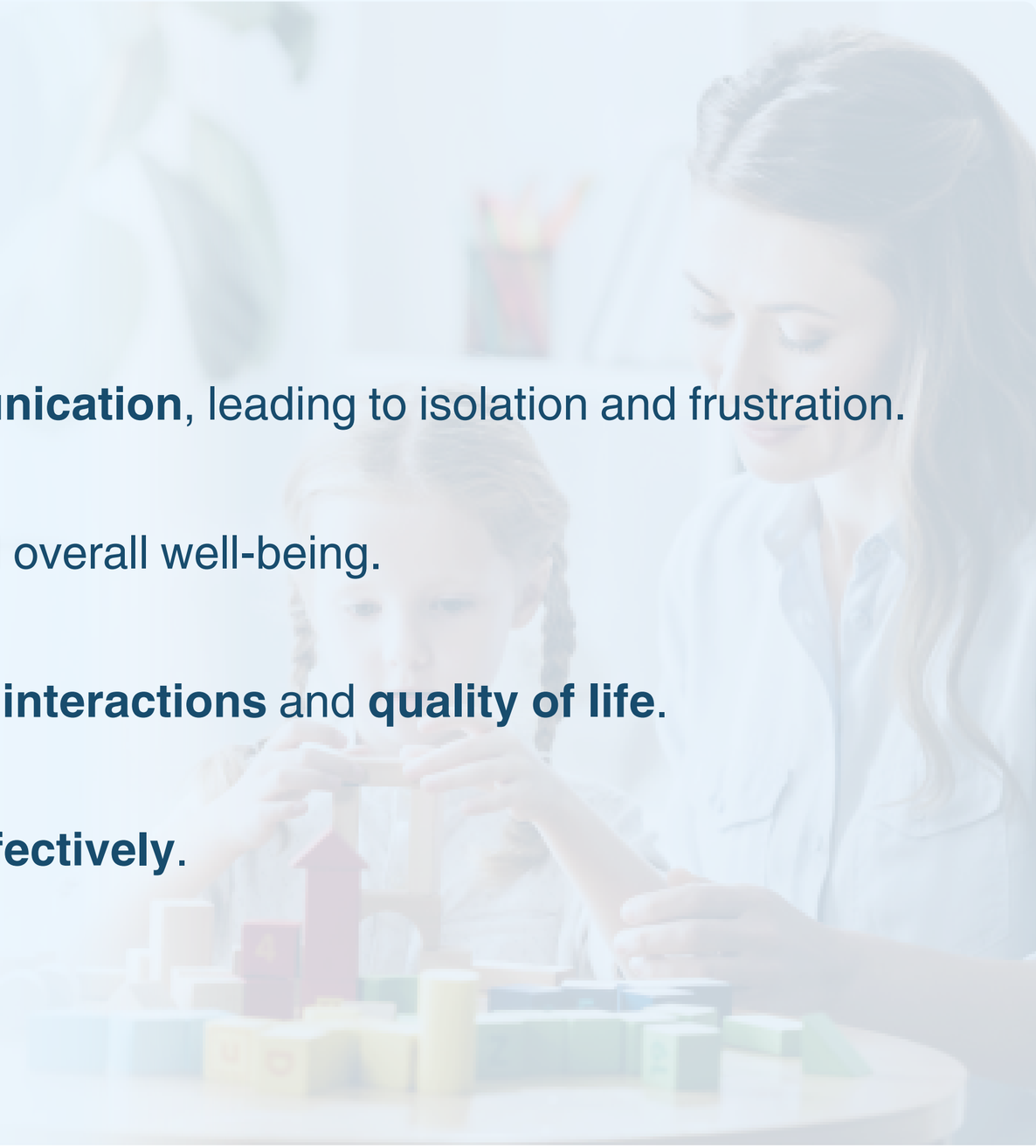
1. C. Tsangouri, W. Li, Z. Zhu, F. Abtahi and T. Ro, "An interactive facial-expression training platform for individuals with autism spectrum disorder," 2016 IEEE MIT Undergraduate Research Technology Conference (URTC), Cambridge, MA, USA, 2016, pp. 1-3, doi: 10.1109/URTC.2016.8284067
2. S. Jain, B. Tamersoy, Y. Zhang, J. K. Aggarwal and V. Orvalho, "An interactive game for teaching facial expressions to children with Autism Spectrum Disorders," 2012 5th International Symposium on Communications, Control and Signal Processing, Rome, Italy, 2012, pp. 1-4, doi: 10.1109/ISCCSP.2012.6217849
3. A. S. Shminan, L. J. Choi and S. Sharif, "AutiTEACCH: Mobile-based Learning in a Structured Teaching Approach for Autistic Children Caregivers," 2020 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), Jakarta, Indonesia, 2020, pp. 259-264, doi: 10.1109/ICIMCIS51567.2020.9354288
4. P. Leijdekkers, V. Gay and F. Wong, "CaptureMyEmotion: A mobile app to improve emotion learning for autistic children using sensors," Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems, Porto, Portugal, 2013, pp. 381-384, doi: 10.1109/CBMS.2013.6627821.

IMPROVING COGNITIVE & VERBAL COMMUNICATION SKILLS

Jathurshan Manistar
IT21246296

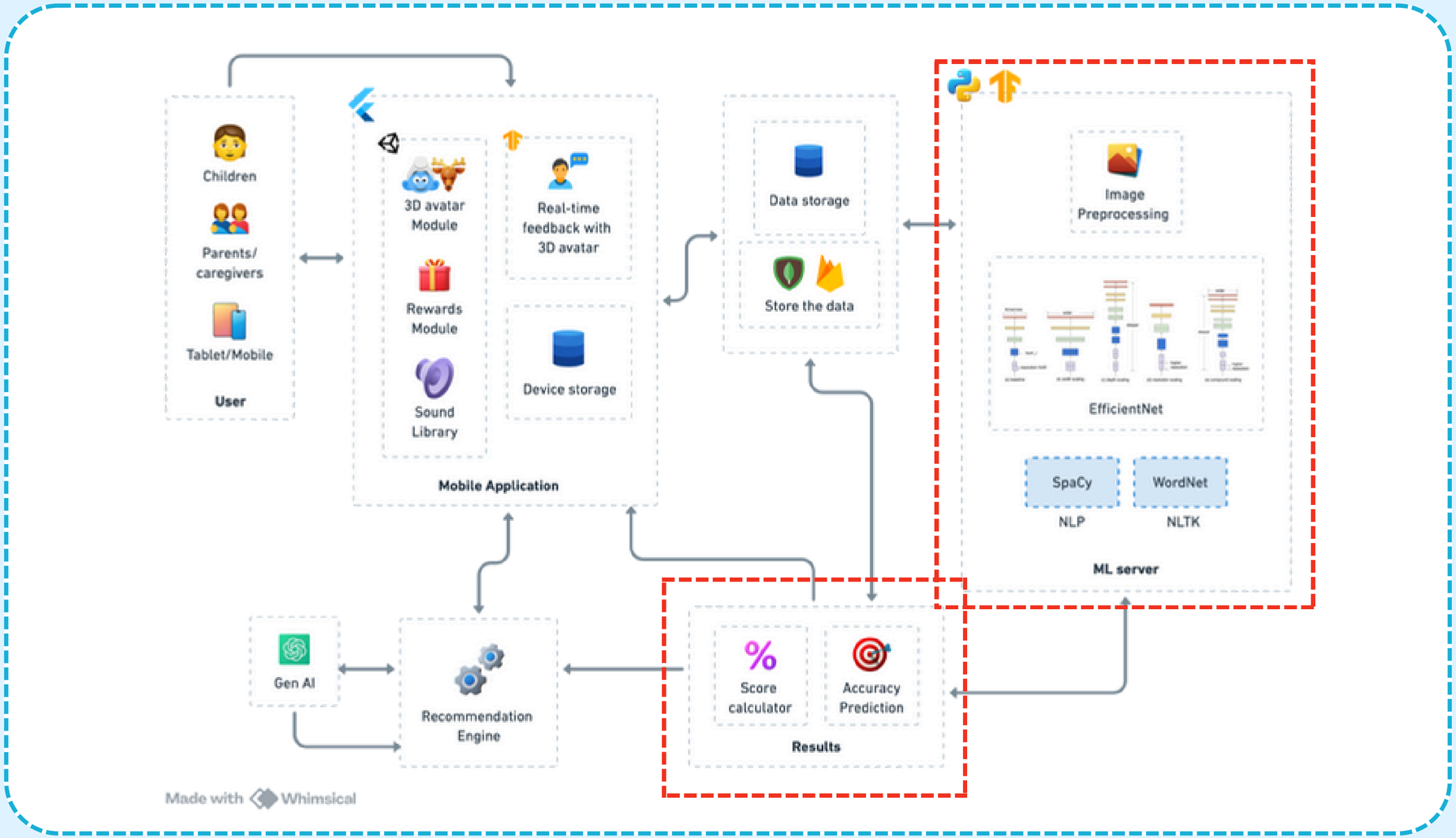
Information Technology

BACKGROUND



- Children with ASD have **difficulties with communication**, leading to isolation and frustration.
- These challenges **impact daily interactions** and overall well-being.
- Enhancing these skills is crucial for **better social interactions** and **quality of life**.
- Enabling the child to **navigate daily life more effectively**.

SYSTEM DIAGRAM



OBJECTIVES AND EVIDENCE OF COMPLETION

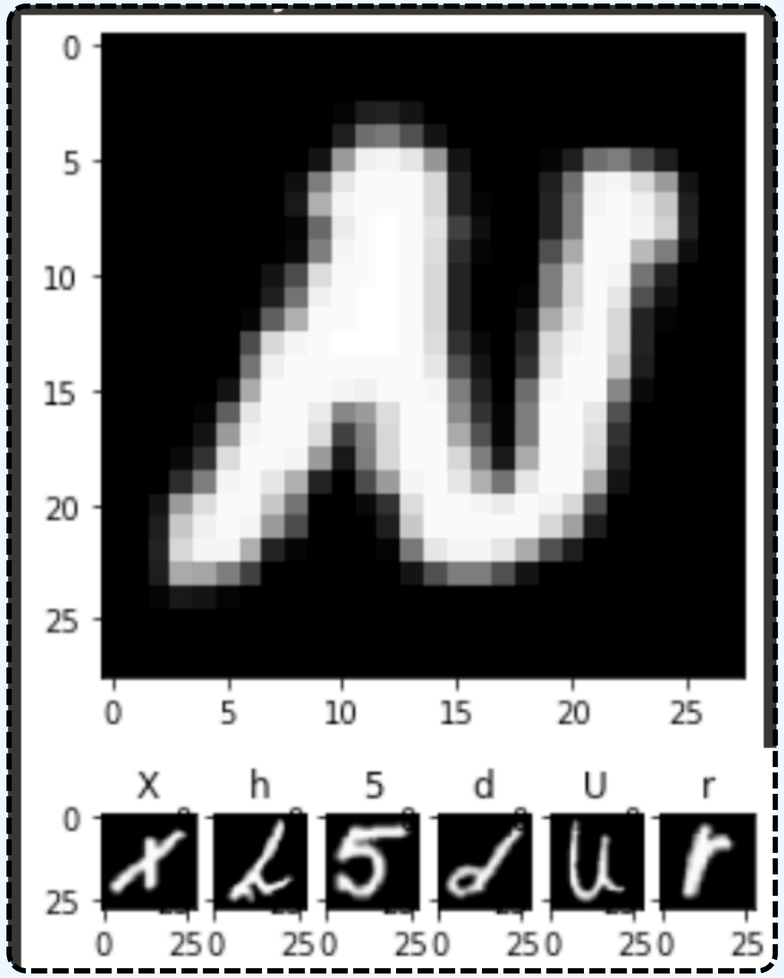
- Identify the child's current cognitive and verbal communication skills.
- **Analyze** written letters from children to identify the current context.
- Predict the **accuracy** of hand-drawn letters for verbal assessment.
- Developing a game to **building shapes from blocks** to enhance cognitive skills.
- Employ DL models and **rule-based systems** for word prediction, accuracy of the letter and **provide feedback**.

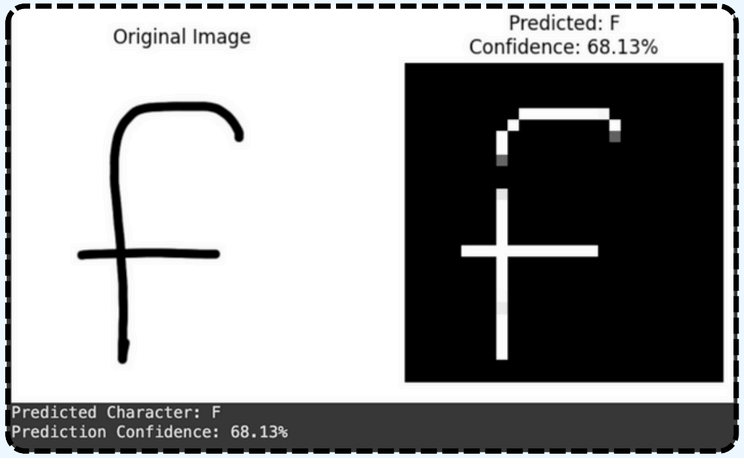
Verbal

Cognitive

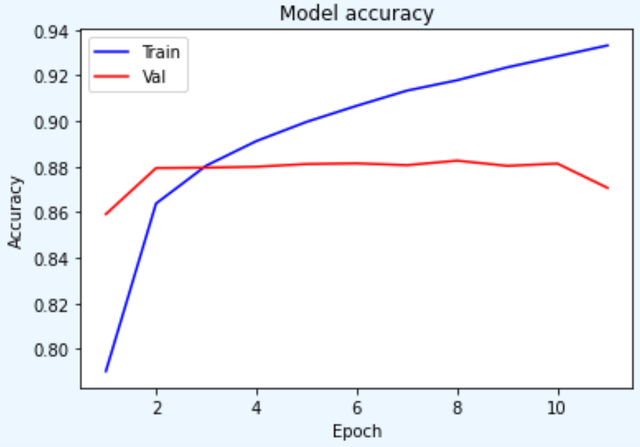
OBJECTIVE 01 & 02:

Analyze written letters from children to identify the current context.

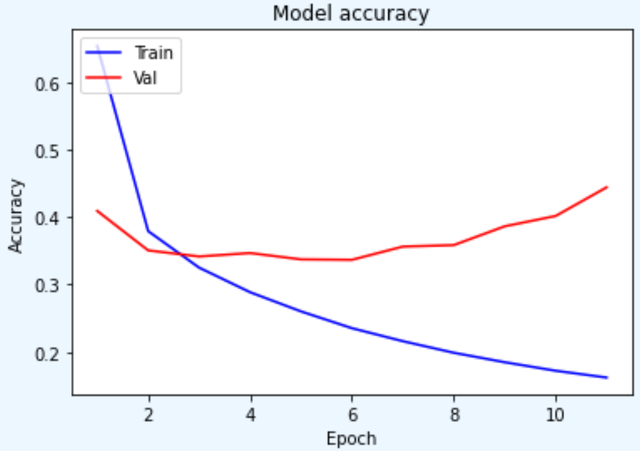




Accuracy curve

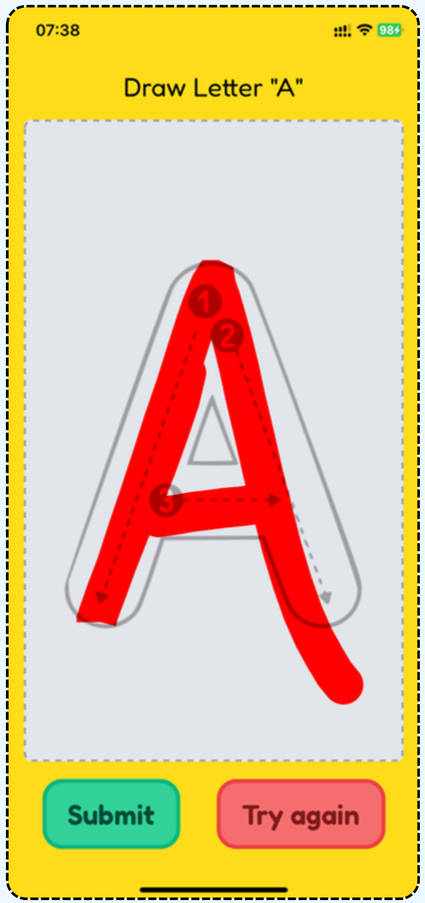


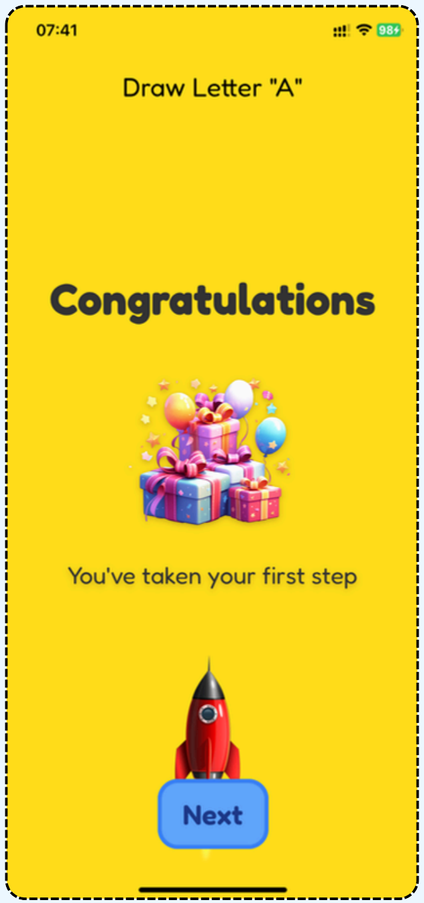
Loss curve

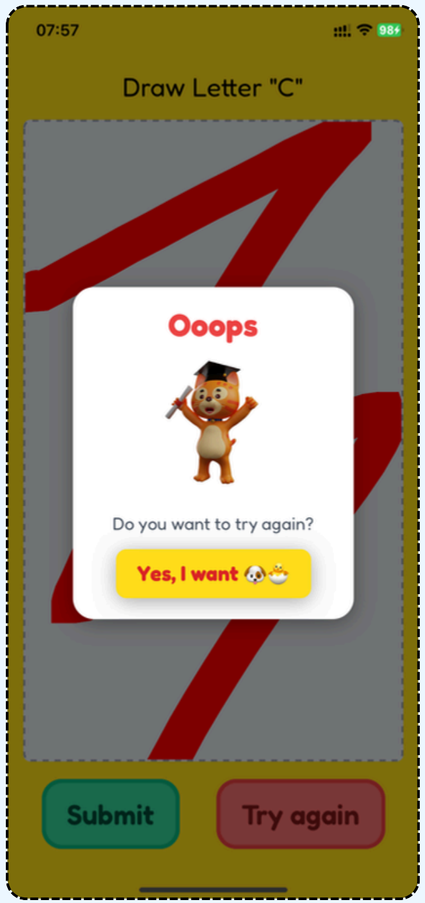


OBJECTIVE 03:

Predict the accuracy of hand-drawn letters for verbal assessment.

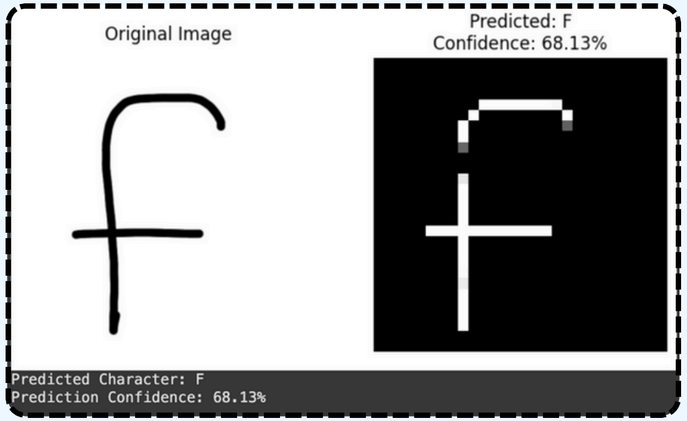






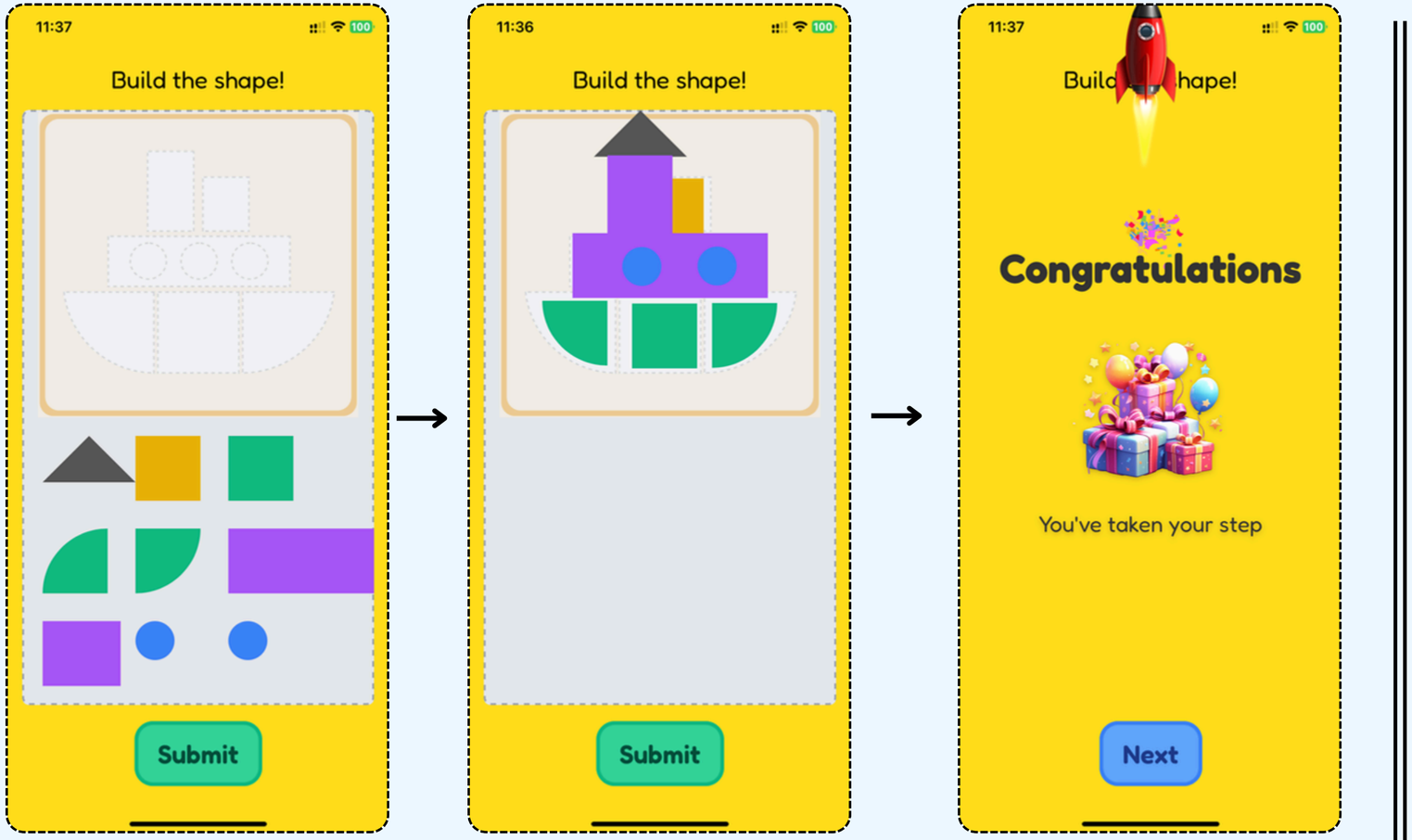

```
confidence: 54.500246
predicted_character: "A"
predicted_character_digit: 10
```

```
confidence: 46.672323
predicted_character: "B"
predicted_character_digit: 11
```



OBJECTIVE 04:

Developing a game to building shapes from blocks to enhance cognitive skills.




```
message: "Images compared successfully."
similarity_score: "70.14%"
```

REFERENCES

[1] L. F. Guerrero-Vásquez et al., "Assessing Children's Perceptions of Live Interactions With Avatars: Preparations for Use in ASD Therapy in a Multi-Ethnic Context," in IEEE Access, vol. 8, pp. 168456-168469, 2020, doi: 10.1109/ACCESS.2020.3023737.

[2] Marvin, Alison & Marvin, Daniel & Lipkin, Paul & Law, Jessica. (2017). Analysis of Social Communication Questionnaire (SCQ) Screening for Children Less Than Age 4. Current Developmental Disorders Reports. 4. 10.1007/s40474-017-0122-1.

IMPROVING SPEECH AND LANGUAGE DEVELOPMENT

Umaira M M
IT21258312
[Software Engineering]

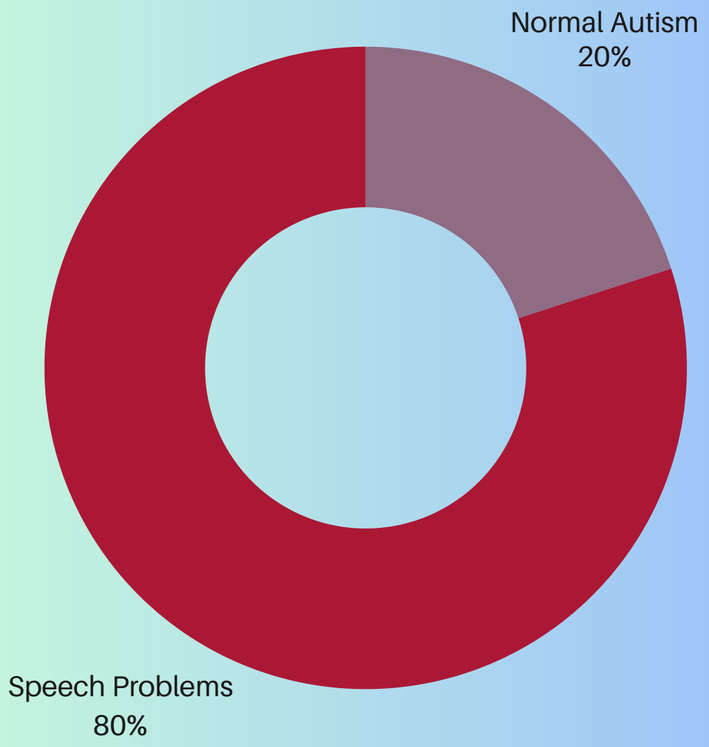
BACKGROUND

Parents suffer from high levels of anxiety due to fear of child being isolated and unable to communicate.

Almost **all children with autism** retain **difficulty in speaking and language** [2].

some may have limited or no speech, while others might struggle with understanding and using language appropriately.

Main focus of the component is to make sure to improve the speech and language development of children in order make them adapt to normal living standards



In Sri Lanka **over 80 % of children with autism** presents with speech related problems [1].

RESEARCH QUESTIONS

How to improve speech and language skills of a child with autism ?

1 .How can we identify the current context of speech and language of a child with autism?

2. What Aspects are considered to identify and evaluate the current level and improvements?

3. What is the methodology use to enhance speech and language development?

4. How to identify if the pronunciation is bad using the proposed mechanism?

5. If an area of words are identified as weak pronunciation, how can it be improved?

MAIN AND SPECIFIC OBJECTIVES

Improvement of speech and language development

Lip pattern Identification of the child when pronouncing a word.

Identifying and evaluating the child's current level of speech and language.

Developing an interactive flash card game with flash cards with words in different categories.

Employing a pipeline to identify areas where the child is finding difficult to pronounce.

Training the LLM to suggest new words using gen AI for the identified difficulty areas.

OBJECTIVE 01

Lip pattern Identification of the child when pronouncing a word.

OVERVIEW

*Objective 01 - 100% completed by PP1

* Improvements added for PP2 addressing panel comments in PP1

OBJECTIVE 01

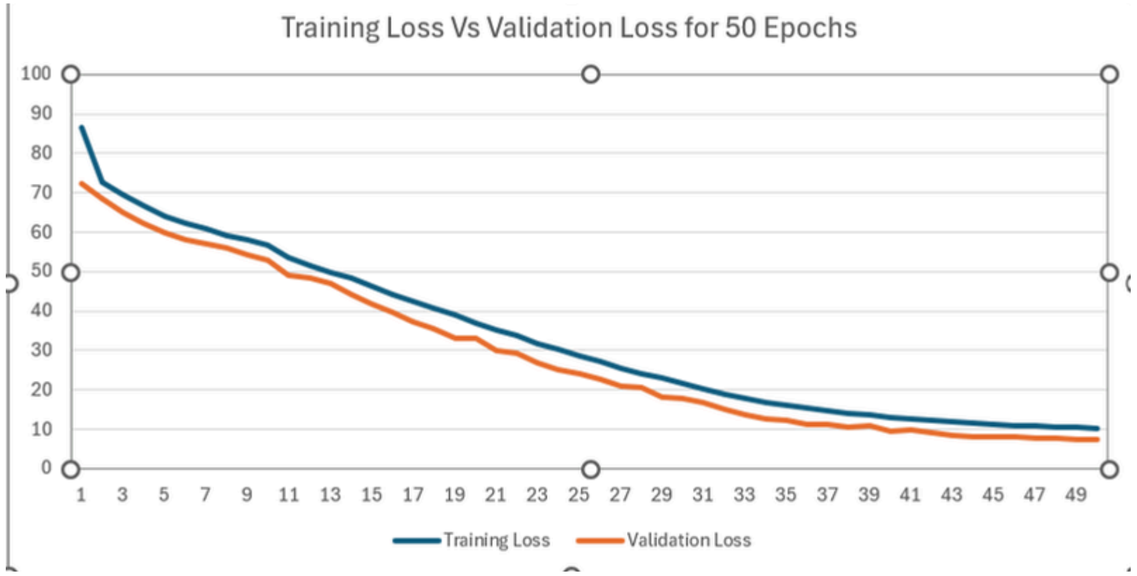
SELECTING THE BEST MODEL - IMPROVEMENTS FROM PP1

Model	PP1	PP2
Architecture	3D CNN + Bi directional LSTM	3DCNN + Bi directional GRU
Trained Dataset	500 silent speaking videos of 1 speaker	34 000 silent speaking videos of 34 speakers
Framework	Tensorflow	Pytorch
Word Prediction Accuracy	75%	**90 % for Adults** **85% for Children**
Optimal For	Adults	Adults and Children

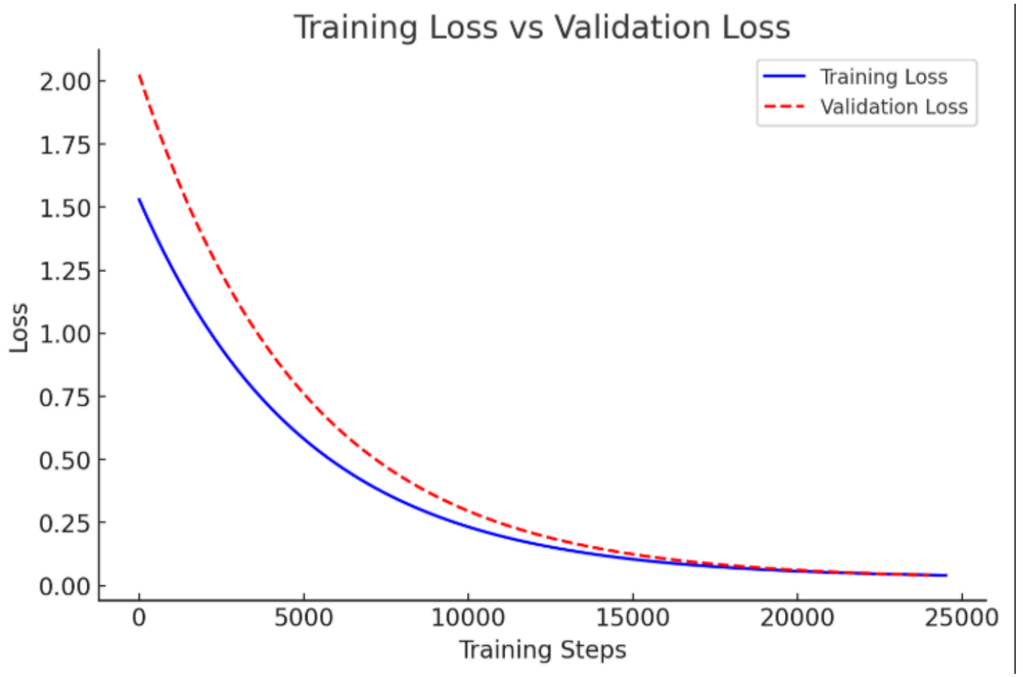
OBJECTIVE 01

MODEL RESULT EVALUATION

Old Model



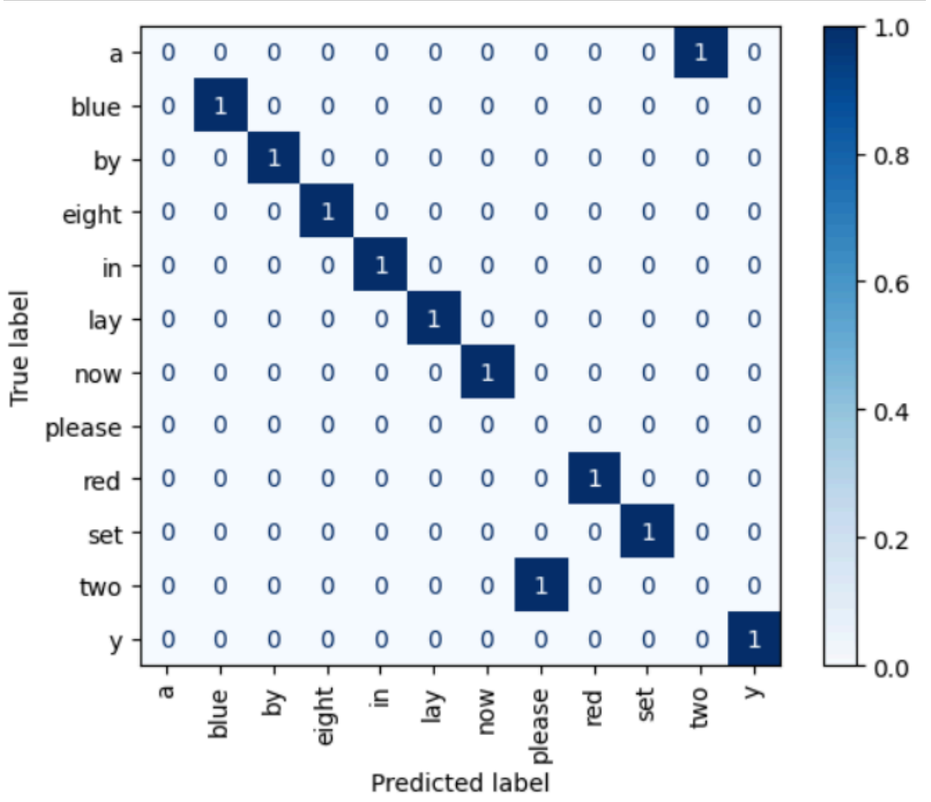
New Model



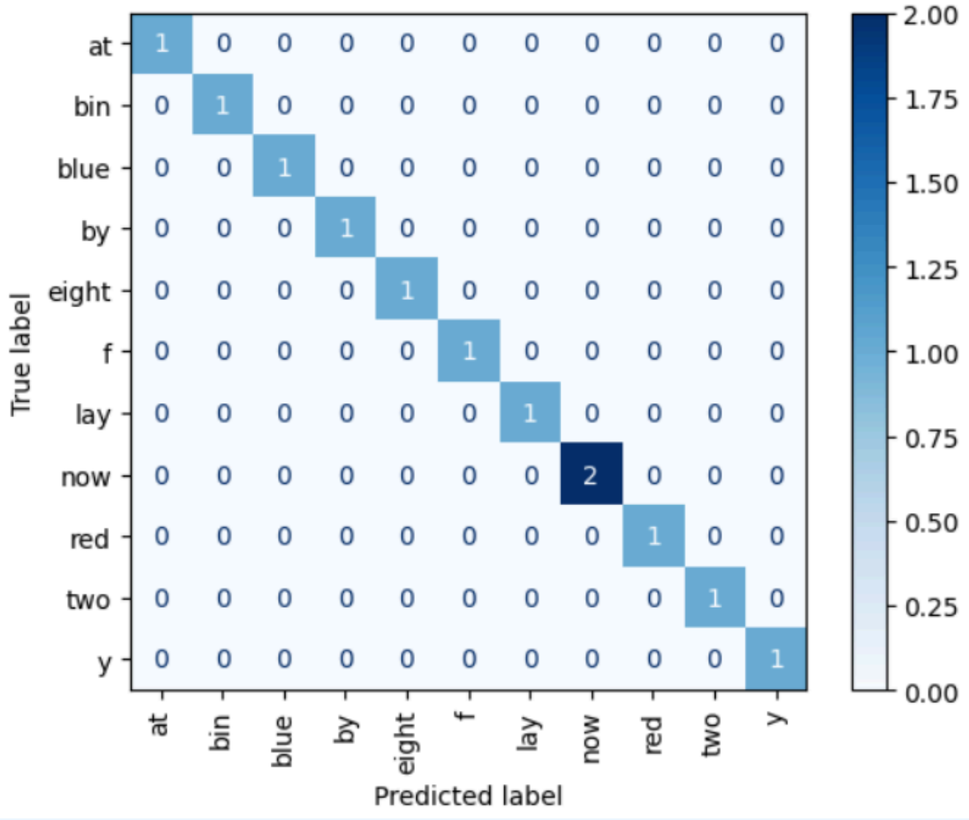
OBJECTIVE 01

MODEL RESULT EVALUATION

Old Model



New Model



OBJECTIVE 01

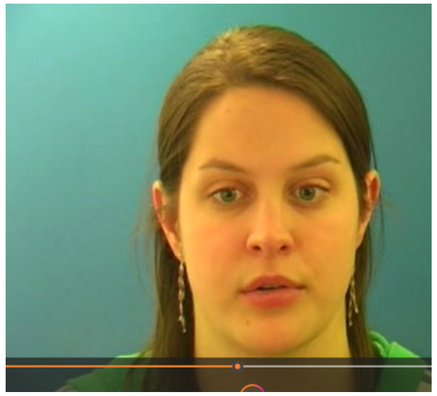
PREDICTIONS FROM THE NEW MODEL

Expected Text

Set white with V six now

Result

```
!python inference.py --input_video setWhiteTest.mpg
SET WHITE WITH V SIX NOW
```

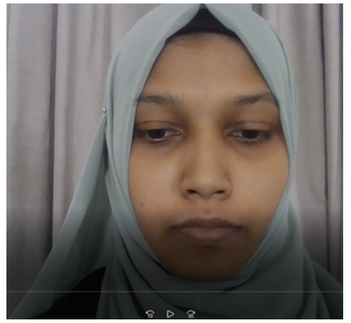


Expected Text

Place red in A zero now

Result

```
!python inference.py --input_video femaleSpeaker.mpg
PLACE RED IN A ZERO NOW
```

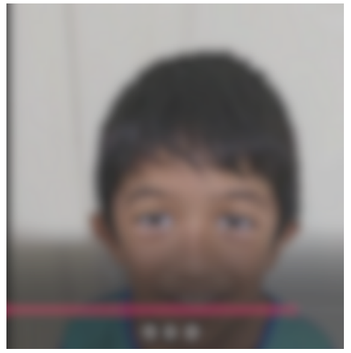


Expected Text

Set white with V six now

Result

```
!python inference.py --input_video UmiraV1.mp4
PLACE WHUTE IN U SIX NOW
```

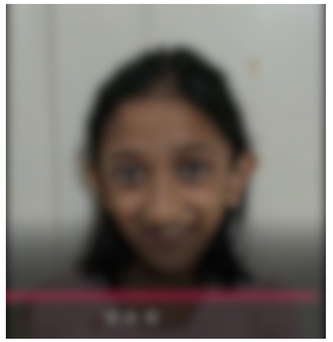


Expected Text

Set white at D six now

Result

```
!python inference.py --input_video YusufEditedV1.mp4
SET RED AT D SIX NGAIN
```



Expected Text

Bin blue at F two now

Result

```
[3] !python inference.py --input_video MaryamEditedV1.mp4
SET GREEN IN Z NINE AGAIN
```

OBJECTIVE 02

Identifying and evaluating the child’s current level of speech and language.

OVERVIEW

*Objective 02 - 50% completed by PP1

*Objective 02 - 100% completed by PP2

OBJECTIVE 02

Identifying and evaluating the child's current level of speech and language.

EVALUATING SPEECH

VIDEO PROCESSING

Trained Lip Reading Model

Lip_Reading_Text

Output_text VS Actual text

LEVENSHTEIN ALGORITHM

OVERALL ACCURACY

AUDIO PROCESSING

DIRECT TRANSCRIPTION

Whisper Model [1]

Whisper_Text

PHONEME ALIGNMENT

Wave2Vec model [2]

Wave2Vec_text

OBJECTIVE 02

Identifying and evaluating the child's current level of speech and language.

```python
# Function to calculate accuracy using Levenshtein distance
def calculate_accuracy(actual_text, predicted_text):
    return Levenshtein.ratio(actual_text.lower().strip(), predicted_text.lower().strip()) * 100
```

Function to calculate accuracy for each text using Leveinshtein Algorithm

```python
audio_path = extract_audio(video_path)
whisper_text = whisper_transcription(audio_path)
wav2vec_text = extract_wav2vec_text(audio_path)

lip_reading_text = None
lip_reading_accuracy = None

# # If mode is "sentence", use lip-reading
if mode.lower() == "sentence":
    video_tensor = load_video(video_path, device)
    coords_tensor = generate_lip_coordinates("samples", detector, predictor)

    with torch.no_grad():
        pred = lip_model(video_tensor[None, ...].to(device), coords_tensor[None, ...].to(device))

    lip_reading_text = ctc_decode(pred[0])[-1]  # Get final output text
    lip_reading_accuracy = calculate_accuracy(expected_text, lip_reading_text)

# Calculate accuracy for audio models
whisper_accuracy = calculate_accuracy(expected_text, whisper_text)
wav2vec_accuracy = calculate_accuracy(expected_text, wav2vec_text)
```

Calculating Accuracies


```
expected_text": "set white with v six now",
whisper_text": " Set white with V6 now.",
whisper_accuracy": 86.95652173913044,
wav2vec_text": "SET WHITE WITH V SIX NOW",
wav2vec_accuracy": 100.0,
lip_reading_text": "SET WHITE WITH V SIX NOW",
lip_reading_accuracy": 100.0,
overall_accuracy": 95.65217391304348,
```


Output

OBJECTIVE 03 & 4

Identifying the areas where the child is finding difficult to pronounce and suggest new words

OVERVIEW

*Objective 03 - 100% completed by PP2

*Objective 04 - 100% completed by PP2

OBJECTIVE 03 & 04

Identifying the areas where the child is finding difficult to pronounce and suggesting new words

Threshold Value = 70

Overall Accuracy < 70% => Needs Improvement

Metaphone Algorithm is used to generate similar phonetic indexes.

nltk corpus words is used to get the list of English words.

Filter words with similar phonetic encodings.

How to ensure that the generated words are child friendly ?

Dolch Sight Words List which is a most commonly used set of words for each age category is used for additional filtering.

OBJECTIVE 03 & 04

Identifying the areas where the child is finding difficult to pronounce and suggesting new words

```python
def suggest_similar_words(target_word, num_suggestions=5):
    """Suggests child-friendly words with similar pronunciation using Metaphone encoding.
    target_encoding = doublemetaphone(target_word)[0]  # Get primary Metaphone encoding

    # Find words with the same Metaphone encoding
    similar_words = [
        word for word in dolch_words  # Use predefined Dolch words list
        if doublemetaphone(word)[0] == target_encoding and word != target_word
        and len(word) <= 6  # Limit word length for young children
        and syllapy.count(word) <= 2  # Ensure easy pronunciation
    ]

    # Return up to `num_suggestions` words
    return similar_words[:num_suggestions]
```


```json
{
  "expected_text": "bat",
  "mode": "word",
  "whisper_text": " Apple.",
  "whisper_accuracy": 22.22222222222222,
  "wav2vec_text": "HE APPLE",
  "wav2vec_accuracy": 18.181818181818176,
  "lip_reading_text": "HE APPLE",
  "lip_reading_accuracy": 18.181818181818176,
  "overall_accuracy": 20.2020202020202,
  "feedback": "Needs more practice!",
  "similar_words": [
    "bad",
    "boat",
    "bed"
  ]
}
```

OBJECTIVE 05

Developing an interactive flash card game with flash cards with words and sentences in different categories.

OVERVIEW

*Objective 05 - 20% completed by PP1

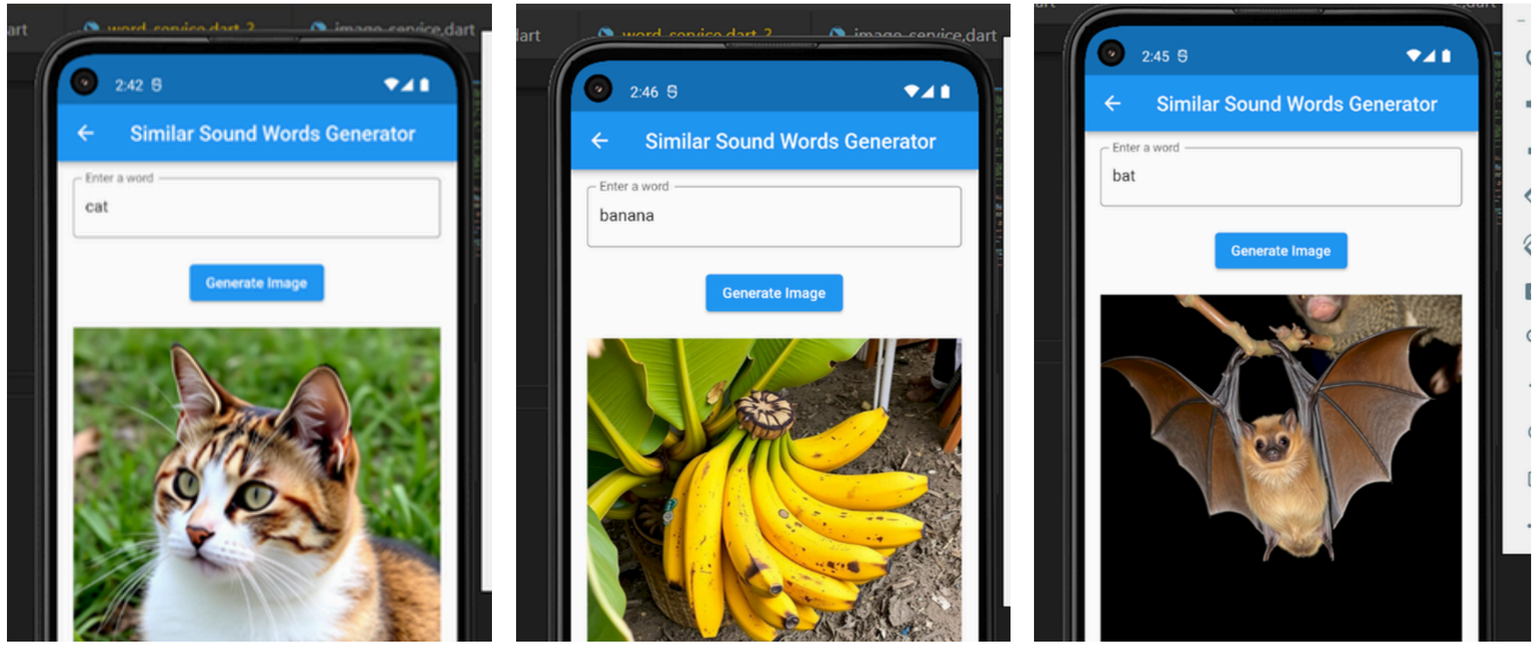
*Objective 05 - 100% completed by PP2

OBJECTIVE 05

Developing an interactive flash card game with flash cards with words and sentences in different categories.

2.2 Generating images for texts : Outputs

Outputs obtained from the Flux model by Blackforest AI [4].



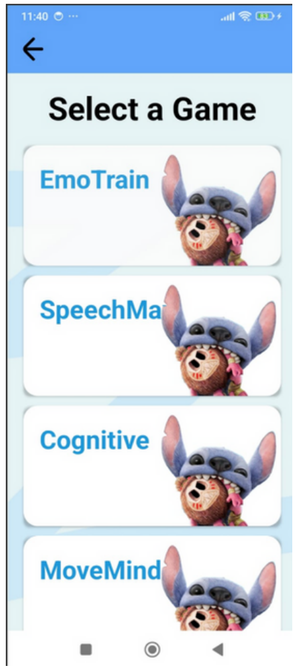
Why real-life images ?

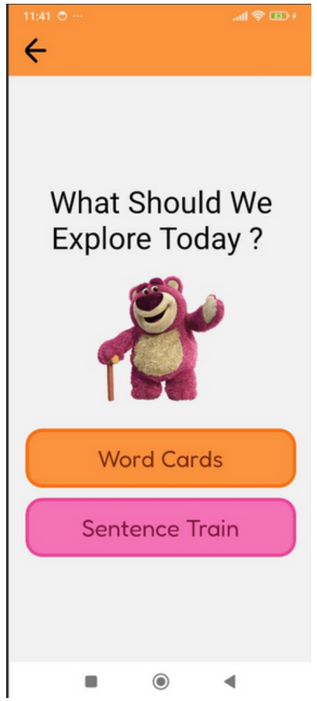
"Children with ASD achieved significantly greater retention accuracy when learning from photographs rather than cartoons [7] ."

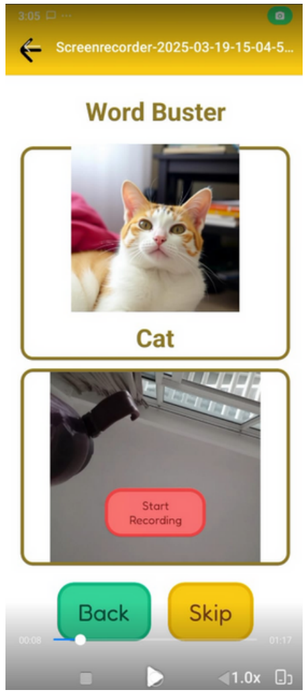
OBJECTIVE 05

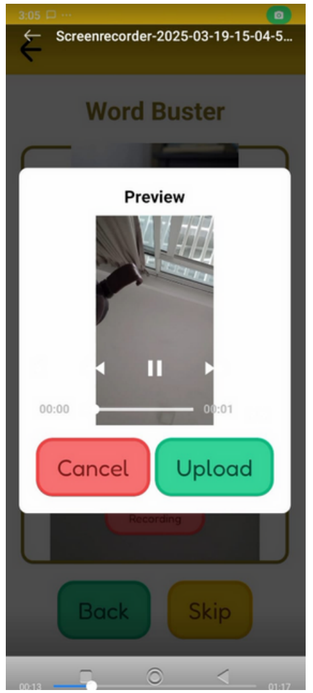
Developing an interactive flash card game with flash cards with words and sentences in different categories.

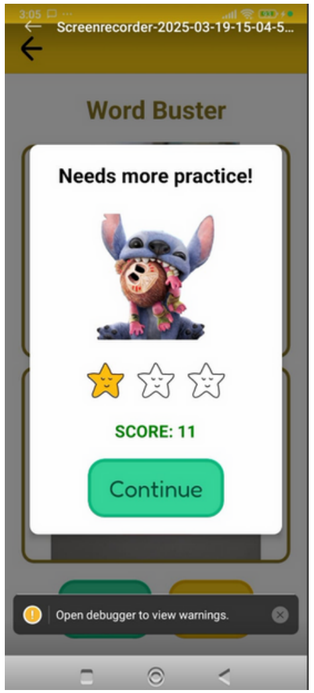
Flash Card Game











OBJECTIVE COMPLETION STATUS

	Status
Lip pattern Identification when pronouncing a word or sentence	100 %
Training the LLM to suggest new words using gen AI for the identified difficulty areas.	100%
Employing a pipeline to identify areas where the child is finding difficult to pronounce.	100%
Identifying and evaluating the child's current level of speech and language.	100%
Developing an interactive flash card game with flash cards with words in different categories.	100%

100%

REFERENCES

[1] P. H. Perera, "Reasons for which parents seek help for their children with autism," in Management of Children with Special Needs : Manual for Primary Health Care Workers in Sri Lanka, Colombo, Sri Lanka., Ministry of Health Sri Lanka, p. 13.

[2] M. Santiputri, E. B. Sembiring, N. Z. Janah and M. K. Mufida, "Abangmanis: Speech Theraphy for Autism Monitoring Mobile Application," 2019 2nd International Conference on Applied Engineering (ICAE), Batam, Indonesia, 2019, pp. 1-6, doi: 10.1109/ICAE47758.2019.9221735. keywords: {Autism;speech therapy;monitoring;multimedia;mobile},

[3] "stability.ai," [Online]. Available: https://stability.ai/.

[4] "replicate.com," [Online]. Available: https://replicate.com/stability-ai/sdxl.

[5] "github.com," [Online]. Available: https://github.com/black-forest-labs/flux.

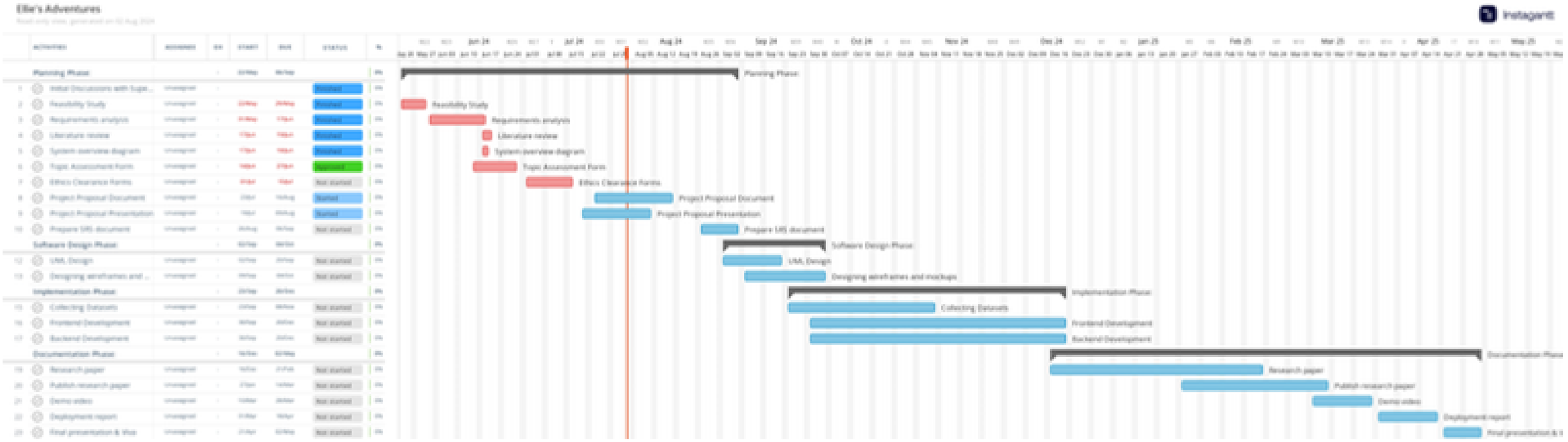
[6] "huggingface.co," [Online]. Available: https://huggingface.co/black-forest-labs/FLUX.1-schnell.

[7] C. H. Cheriece K Carter, "Are Children With Autism More Likely to Retain Object Names When Learning From Colour Photographs or Black-and-White Cartoons?," PubMedCentral, 2021.

SOFTWARE FEATURES

- CI/CD Pipeline
- Unit Testing
- API documentation using Swagger
- Cloud Deployment
- Stack-base code structure
- Sensory friendly UIs
- Logging
- Scripting

GHANTT CHART

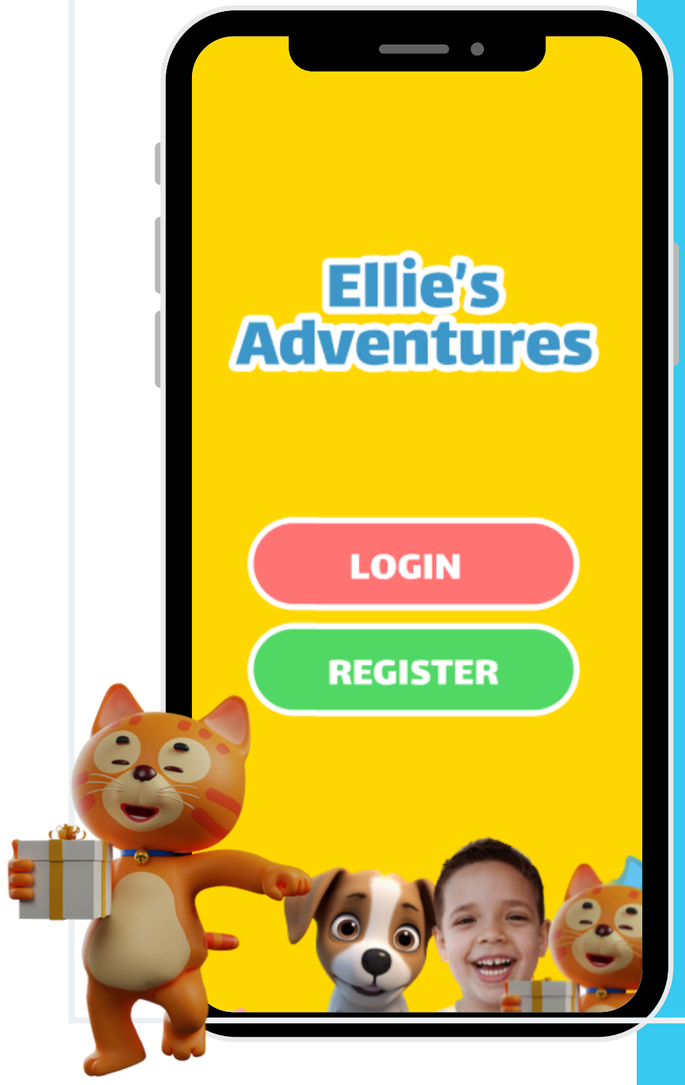


BUDGET

Initiative	Amount (lkr)
Travelling cost for data collection	30 000
Cost of deployment to cloud	25 000
Cost of storage and database	7 000
ML model training and deployment	35 000
Cost of hosting in play store	8 000
ccost of hosting in app store	33 000 / year

COMMERCIALIZATION

Free Package - LKR 0.00	Premium Package - LKR 4000
• Limited real-time feedback • Limited GenAI features • 5 free genAI credits daily • 5 free exercises daily • No real-time AI features • Limits on customizing the 3d avatar • Limits on voice selection for the avatar	• Unlimited access to the application • More personalized feedbacks



REFERENCES

[1] https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders, Health Organization," [Online]. Available: "World https://www.who.int/news-room/fact sheets/detail/autism-spectrum-disorders.

[2] A. H. Vanniarachchy, "Autism in Sri Lanka," Ceylon Today, [Online]. Available: https://ceylontoday.lk/2024/07/20/autism-in-sri-lanka/.

[3] P. H. Perera, "Reasons for which parents seek help for their children with autism," in Management of Children with Special Needs : Manual for Primary Health Care Workers in Sri Lanka, Colombo, Sri Lanka., Ministry of Health Sri Lanka, p. 13.

[4] C. J. Zampella, L. A. L. Wang, M. Haley, A. G. Hutchinson and A. d.Marchena, “Motor Skill Differences in Autism Spectrum Disorder: a Clinically Focused Review,” Current psychiatry reports, 2021.

03

THANK YOU!

03 DEMONSTRATION