Review - 1 Internship Module

Custom Behavior & Sentiment Model for Lost Child Detection System (Body Tracking)

FACULTY MENTOR : Dr Anuradha Yenkikar

INDUSTRY MENTOR: Kaushal Sharma

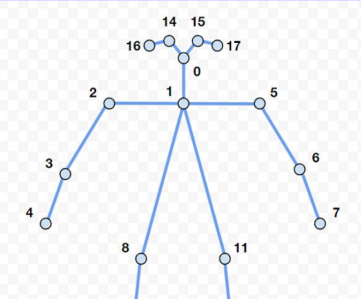
-Vedant Mukhekar (22211607)

CDAC, PUNE

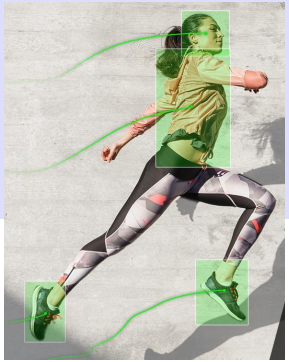
Contents of Presentation

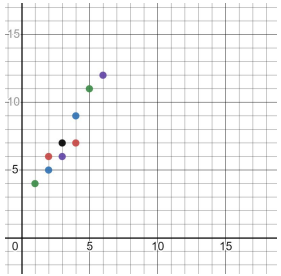
SR NO	CONTENT	SLIDE NO
1	PROBLEM STATEMENT	3
2	INTRODUCTION	4
3	LITERATURE SURVEY	6
4	MODEL CAPABILITIES	7
5	METHODOLOGY	8
6	TECH STACK	9
7	EDA	10
8	RESULTS	12
9	CONCLUSION & FUTURE SCOPE	13
10	THANK YOU	14

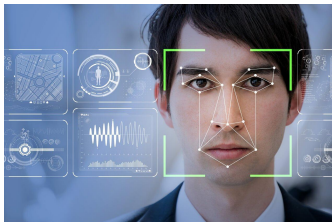
Problem Statement (Traditional Body Tracking System)



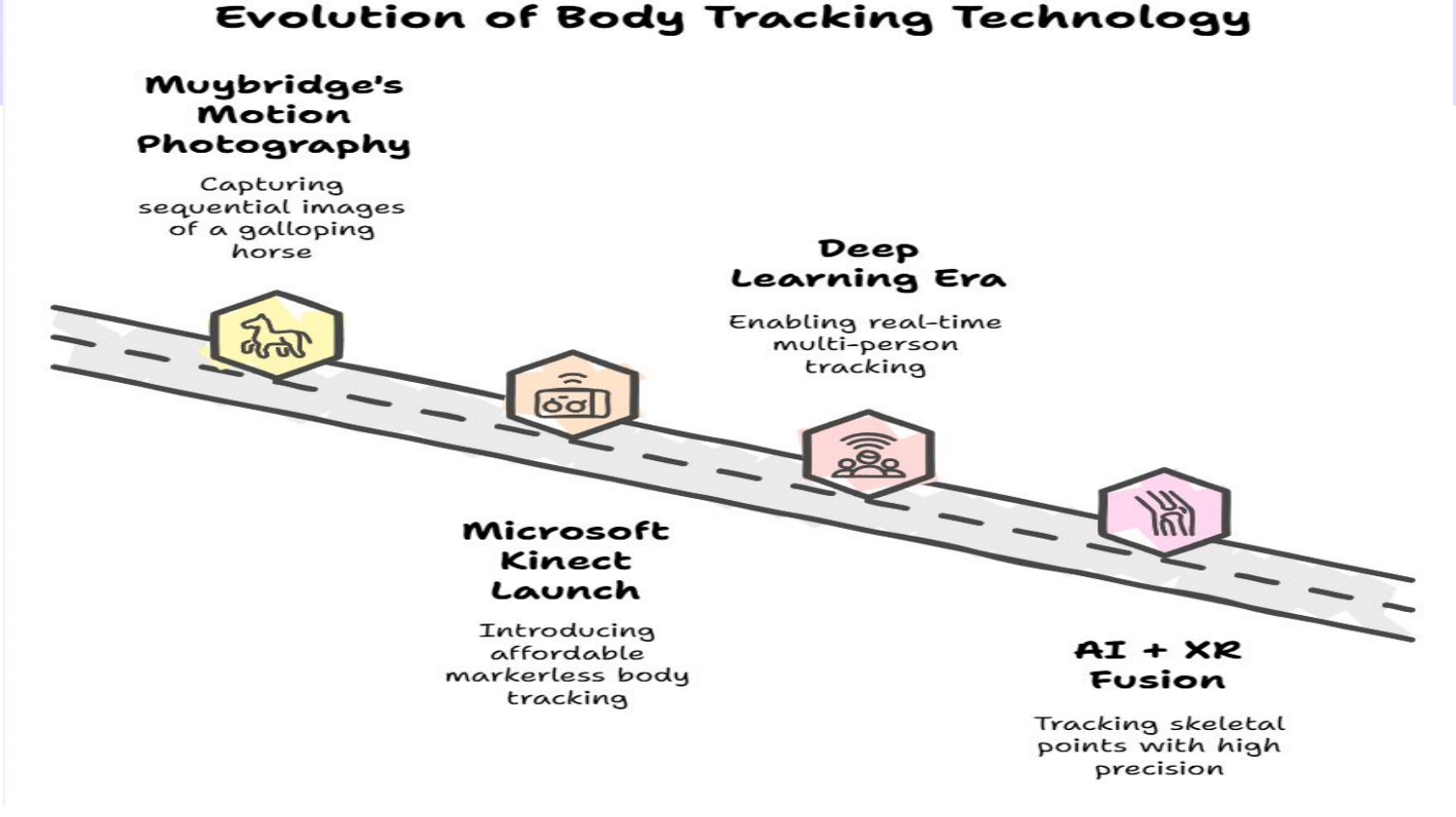
Characteristic	Challenge
Facial Context	No facial landmark analysis
Action Frequency	No behavior frequency mapping
Prediction Timeline	No behavioral evolution insight
Movement Intensity	No physical activity quantification
Dimensionality	Low data points per frame
Sentiment Recognition	No subtle cue interpretation
Output Labels	Generic, non-customizable labels







Introduction



Introduction

Custom Model

Personalized system trained on unique behaviors. Going beyond basic gesture recognition.

Enhanced Fusion

Integrating facial points with pose and hand landmarks. Capturing emotional and physical cues.

Real-Time Insights

live behavior predictions. Visual dashboards for comprehensive behavioral intelligence.

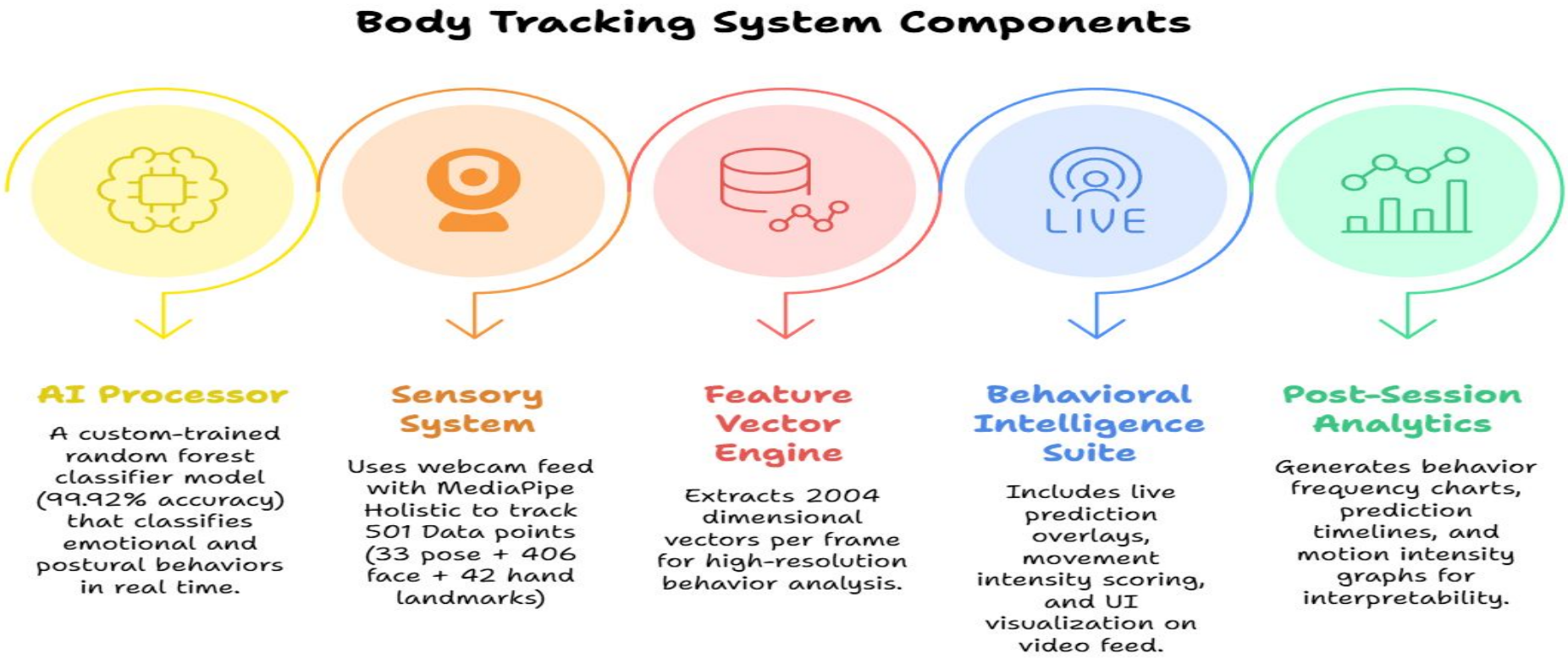
Literature Survey

.

Sr No	Title	Author	Techniques Used	Applications
1	Missing Child Tracking System Using Deep Learning Methodologies	S. Gurusubramani et al.	Deep Learning, Iris Recognition	Proposed a missing child tracking framework using biometric-based deep learning, particularly iris recognition, improving identity verification in law enforcement scenarios.
2	Crowdsourced Children Monitoring and Finding with Holding Up Detection Based on IoT	L.-W. Chen et al.	iBeacon, IoT, Sensor fusion, Feature Extraction	System tracks children using smartphones and wearable IoT devices. "Holding up" gestures (e.g., raising hand) detected to alert for lost status. Uses Bluetooth beacon for proximity-based tracking.
3	Automatic Temporal Segment Detection and Affect Recognition from Face and Body Display	H. Gunes, M. Piccardi	Temporal Segment Detection, Fusion of Face & Body cues, Phase Synchronization	Recognizes emotions (affect) using a combination of facial and body motion over time; detects temporal segments of emotional expression.
4	Sign Language to Text Conversion Using Deep Learning Techniques	P. Sonsare et al.	YOLO-NAS, RNNs, Transformers, MediaPipe	Real-time recognition of sign language gestures using YOLO-NAS for localization and deep models for gesture-text conversion. Demonstrated high accuracy across Indian, American, and Japanese sign languages.
5	Revolutionizing Emotional Intelligence: AI-Driven Mood and Pose Detection through Facial Data	H. Vanpariya et al.	HAAR Cascade, MediaPipe, Mood Detection Engine	A facial expression recognition pipeline using MediaPipe and HAAR cascades to detect mood in real-time video. Highlights usability in improving user experience and smart interfaces.
6	Design and Implementation of Live Stream System Based on Kurento	P. Li, C. Liu, P. Zhang	Kurento Media Server, WebRTC, Low Latency Protocols	Built a low-latency video streaming system using Kurento for surveillance and real-time communication—suitable for smart monitoring in lost person detection systems.
7	Sensor Fusion for 3D Human Body Tracking with an Articulated 3D Body Model	S. Knoop, S. Vacek, R. Dillmann	Sensor Fusion (Camera + Magnetic Sensors), 3D Human Pose Estimation	Proposed real-time tracking of full-body 3D pose using multi-sensor fusion. Well-suited for robotics and human motion capture. Potential for child body gesture tracking .

8	Estimating Emotional Intensity from Body Poses for Human-Robot Interaction	Mingfei Sun, Yiqing Mou, Hongwen Xie, Meng Xia, Michelle Wong, Xiaojuan Ma	Local joint transformation descriptors; LSTM-RNN temporal modeling of body pose	Real-time estimation of emotional intensity from body movements—relevant to interpreting distress behavior
9	MediaPipe Hands: On-device Real-time Hand Tracking	Fan Zhang et al. (Google Research)	MediaPipe palm detection, hand landmark model, real-time on-device inference	Core hand-tracking foundation; demonstrates on-device performance and accuracy
10	Feasibility of 3D Body Tracking from Monocular 2D Video Feeds in Musculoskeletal Telerehabilitation	(MDPI Sensors, 2024)	MediaPipe Pose, ROM calculation; coordinate alignment; monocular-to-3D tracking evaluation	Validates MediaPipe Pose accuracy for joint tracking in real-world scenarios, showing low error rates
11	Optimizing Hand Region Detection in MediaPipe Holistic Full-Body Pose Estimation	Amit Moryossef	ROI refinement; enriched feature set using z-dimension; improved IoU for hands	Enhances accuracy of holistic tracking—addresses limitations in landmark detection
12	Facial Landmark Detection Evaluation on MOBIO Database	Na Zhang	Manually labeled 22-point landmarks on mobile video frames; benchmarked landmark localization methods	Evaluates mobile-based facial landmark detection, relevant for emotion detection in challenging real-world conditions
13	LUVLi Face Alignment: Estimating Landmarks' Location, Uncertainty, and Visibility Likelihood	Abhinav Kumar et al.	Joint estimation of 68 landmarks with uncertainties and occlusion model; deep network and LUVLi loss	Improves robustness in facial landmark localization under occlusions/head poses
14	MaskFace: multi-task face and landmark detector	Dmitry Yashunin, Tamir Baydasov & Roman Vlasov	Combined face detection and landmark localization via RoIAlign and keypoint head (Mask R-CNN style)	State-of-the-art performance on AFLW, 300W, WIDER FACE for both detection and landmark tasks
15	A Detailed Look At CNN-based Approaches in Facial Landmark Detection	Chih-Fan Hsu et al.	Regression-based vs heatmap (pixel-wise classification) CNN architectures; hybrid loss function; evaluation across multiple datasets (300W, LFPW, AFW, COFW)	Provides systematic comparison of CNN methods for accurate landmark localization
16	Towards Accurate Facial Landmark Detection via Cascaded Transformers	Hui Li et al.	Cascaded transformer-based coordinate regression using deformable attention; landmark interrelationships exploited	Achieves new SOTA results on benchmarks under large pose and occlusion
17	Facial Landmark Detection: a Literature Survey	Yue Wu & Qiang Ji	Comprehensive review of holistic, CLM, regression-based & deep learning methods; covers datasets (300W, Menpo, 300-VW)	Offers background on all major landmark detection paradigms and dataset challenges
18	System and method to determine true facial feature measurements of a face in a 2D image	Peyush Bansal (Lenskart Solutions Pvt. Ltd.)	Patent describing measurement of facial features from 2D image—landmark-based distance metrics for accurate face mapping	Commercial facial mapping used in virtual try-on systems; relevant to precise facial point mapping in real-world deployment

Model Capabilities



Model Capabilities

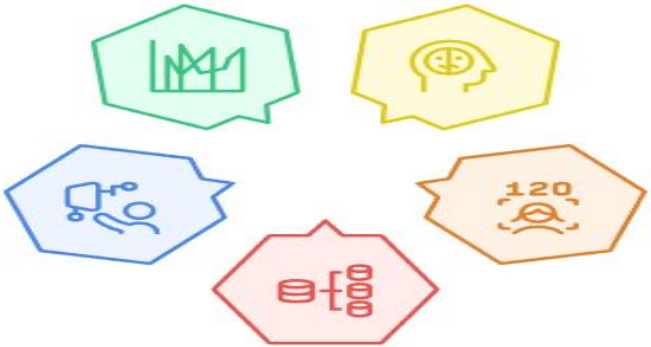
Application in Child Detection System

Forensic Graphs

Matplotlib charts log label frequency, movement intensity, and behavioral changes over time.

Live Prediction Engine

Real-time inference with overlayed labels and motion scoring via OpenCV + MediaPipe.



Feature Depth

Dense feature vectors enhance detection of abnormal child behaviors vs normal crowd activity.

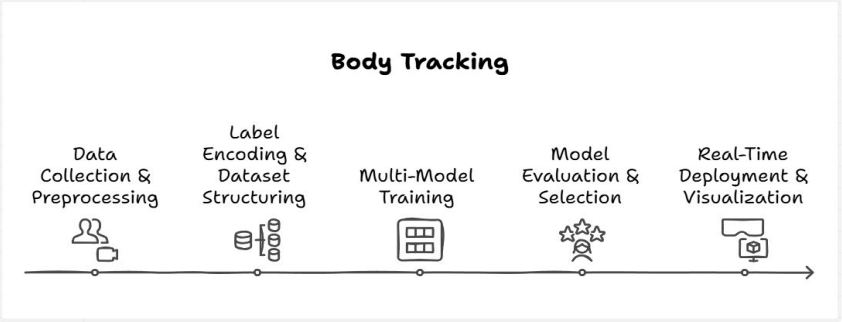
Emotion Classification

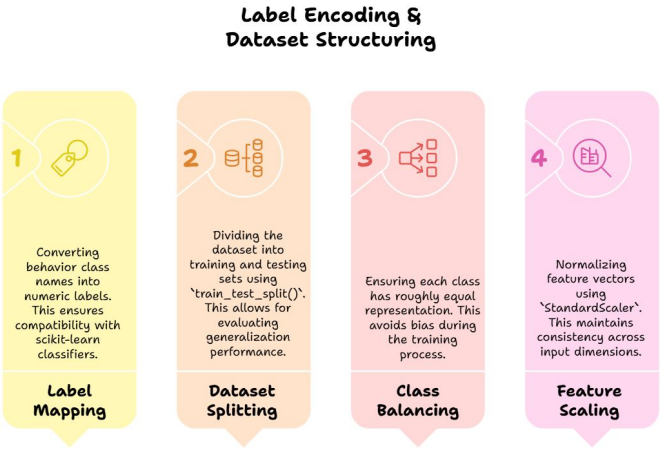
Uses Random Forest to detect distress behaviors like “fear_1” and “melancholy” from 2004-dimensional vectors.

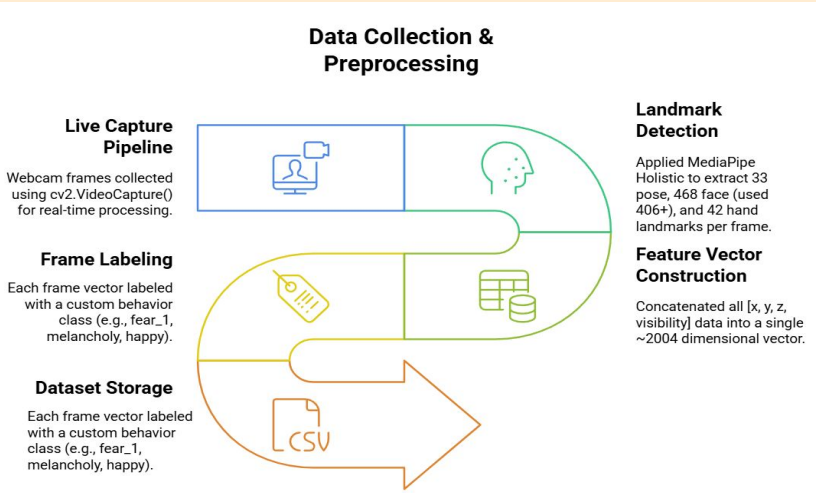
Landmark Precision

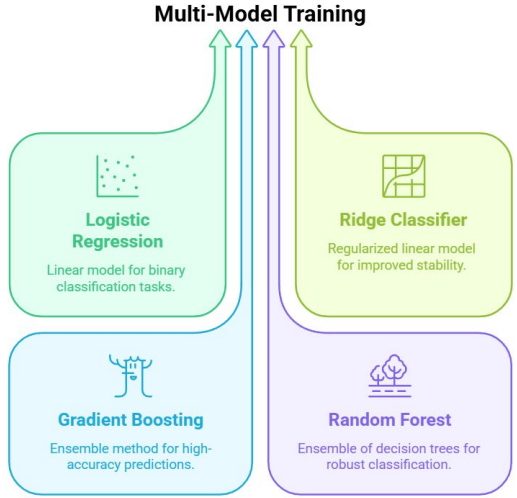
Tracks 501 points/frame (33 pose, 406 face, 42 hand) to capture fine emotional posture shifts.

Methodology

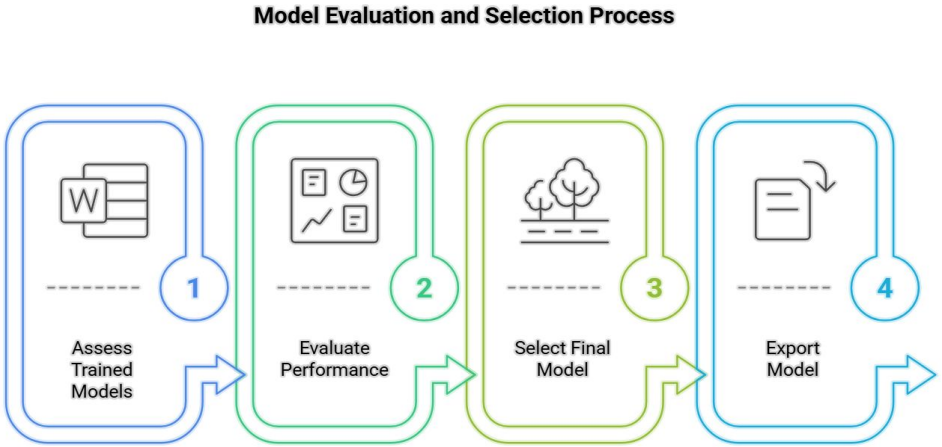


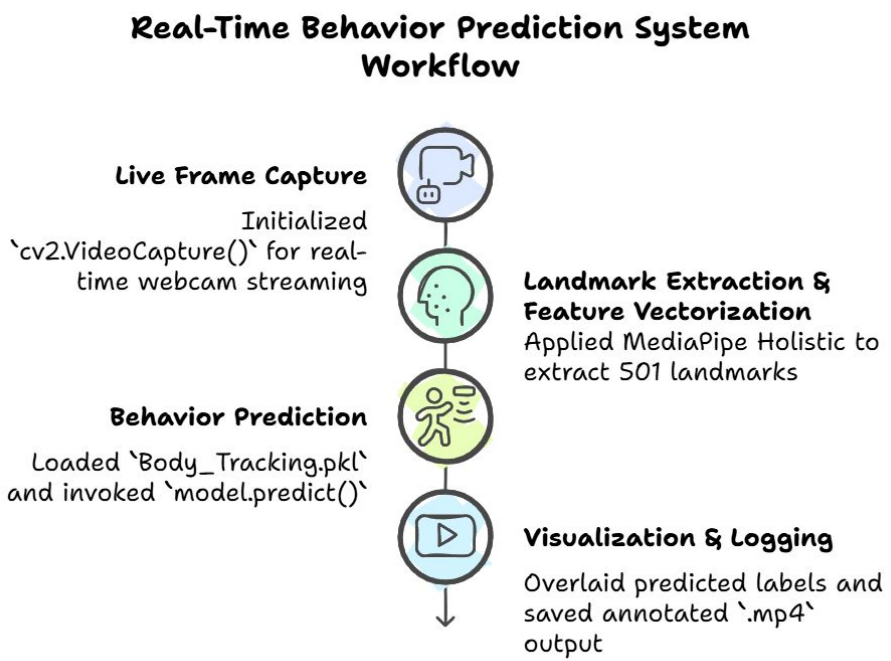






Methodology

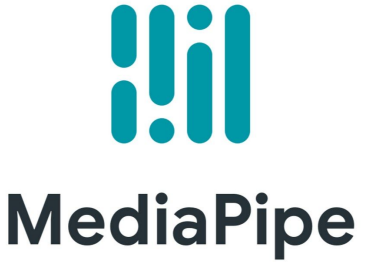




Tech Stack


NumPy


Joblib


MediaPipe

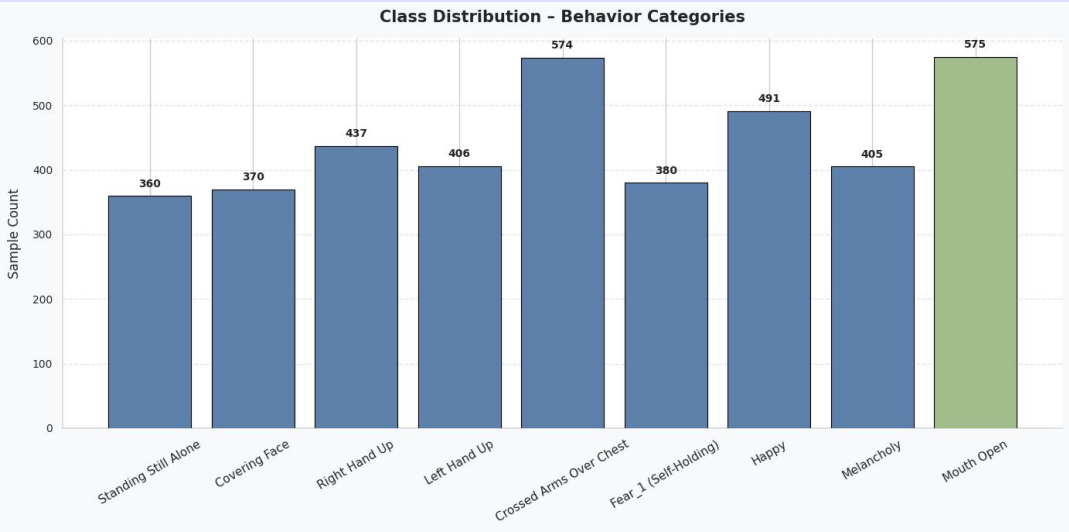

scikit
learn

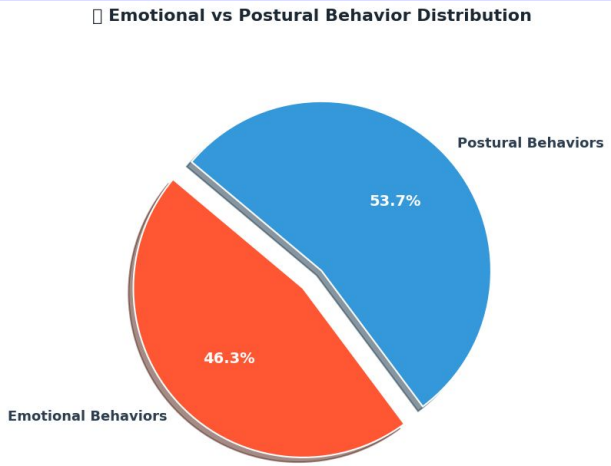

OpenCV

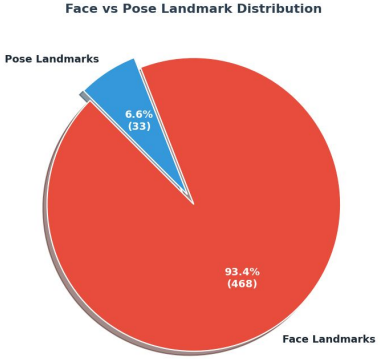

matplotlib

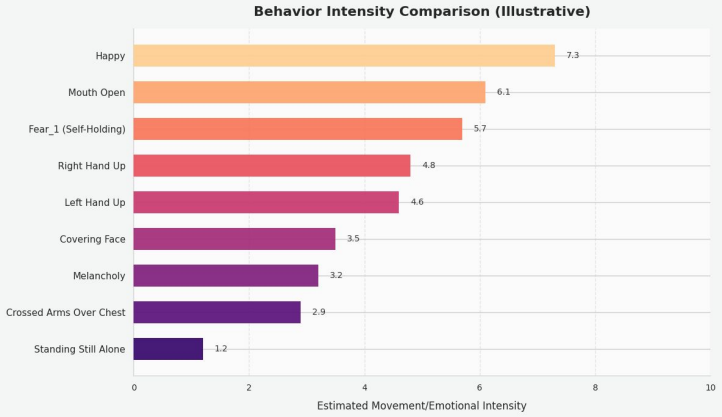

Google
Text-to-Speech

EDA

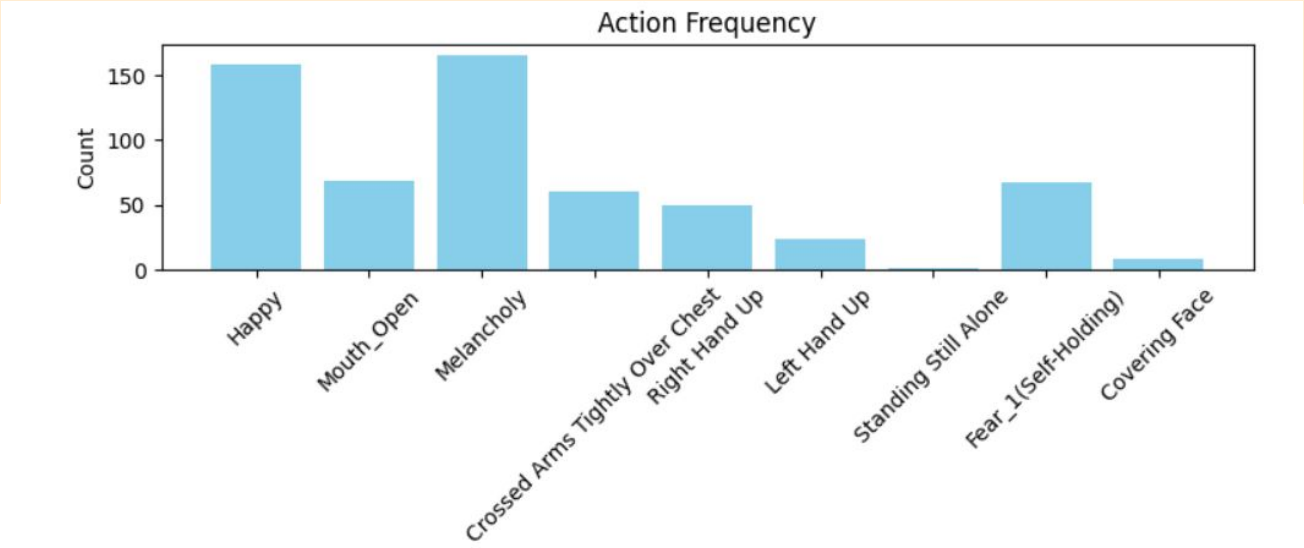

Class Distribution – Behavior Categories
Sample Count
600
500
400
300
200
100
0
360
370
437
406
574
380
491
405
575
Standing Still Alone
Covering Face
Right Hand Up
Left Hand Up
Crossed Arms Over Chest
Fear_1 (Self-Holding)
Happy
Melancholy
Mouth Open

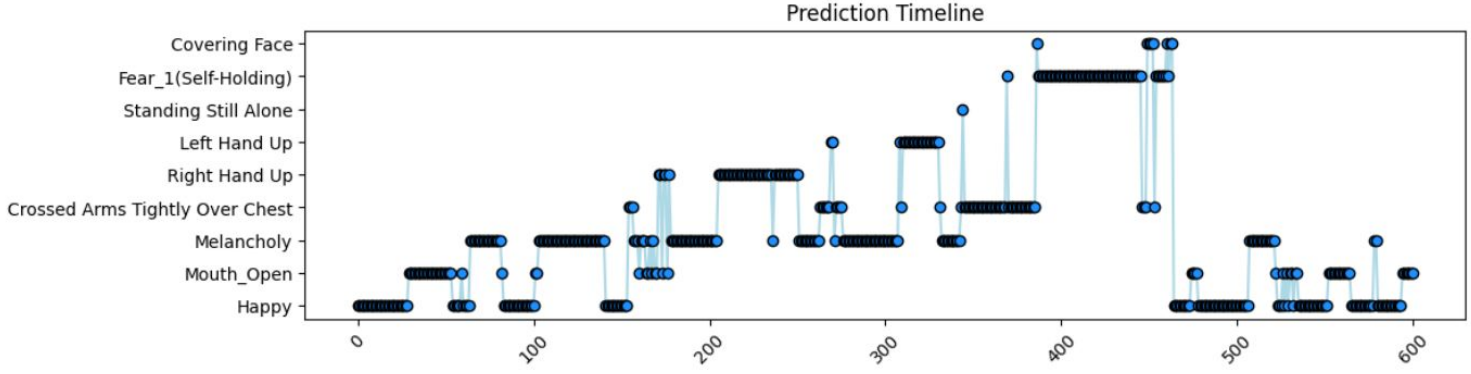

Emotional vs Postural Behavior Distribution
Postural Behaviors
53.7%
46.3%
Emotional Behaviors


Face vs Pose Landmark Distribution
Pose Landmarks
6.6%
(33)
93.4%
(468)
Face Landmarks

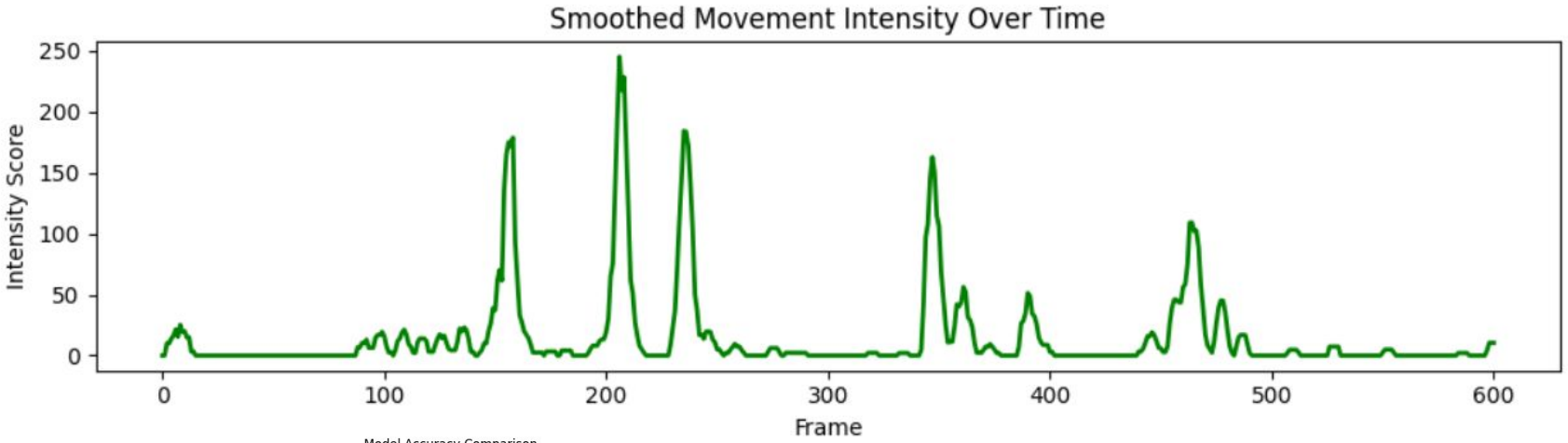

Behavior Intensity Comparison (Illustrative)
Happy
7.3
Mouth Open
6.1
Fear_1 (Self-Holding)
5.7
Right Hand Up
4.8
Left Hand Up
4.6
Covering Face
3.5
Melancholy
3.2
Crossed Arms Over Chest
2.9
Standing Still Alone
1.2
0
2
4
6
8
10
Estimated Movement/Emotional Intensity

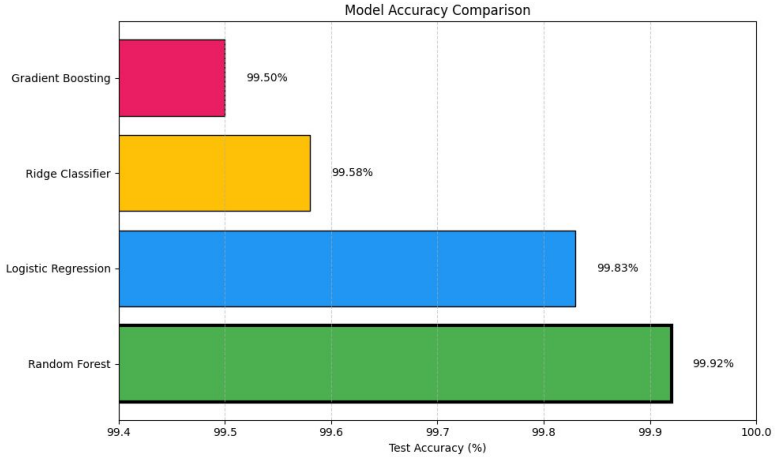
RESULTS

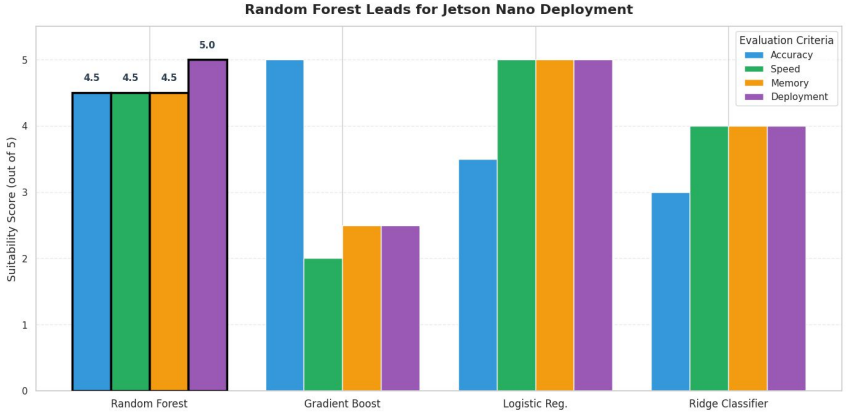

Action Frequency
Count
150
100
50
0
Happy
Mouth_Open
Melancholy
Crossed Arms Tightly Over Chest
Right Hand Up
Left Hand Up
Standing Still Alone
Fear_1(Self-Holding)
Covering Face


Prediction Timeline
Covering Face
Fear_1(Self-Holding)
Standing Still Alone
Left Hand Up
Right Hand Up
Crossed Arms Tightly Over Chest
Melancholy
Mouth_Open
Happy
0
100
200
300
400
500
600

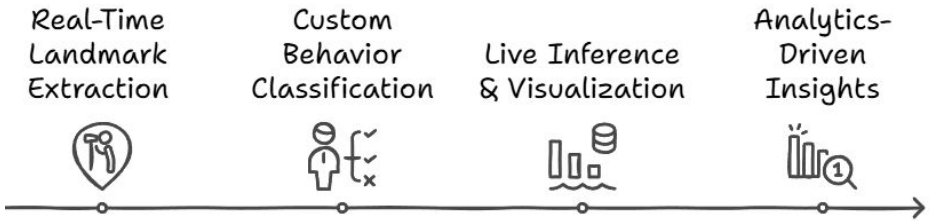
RESULTS

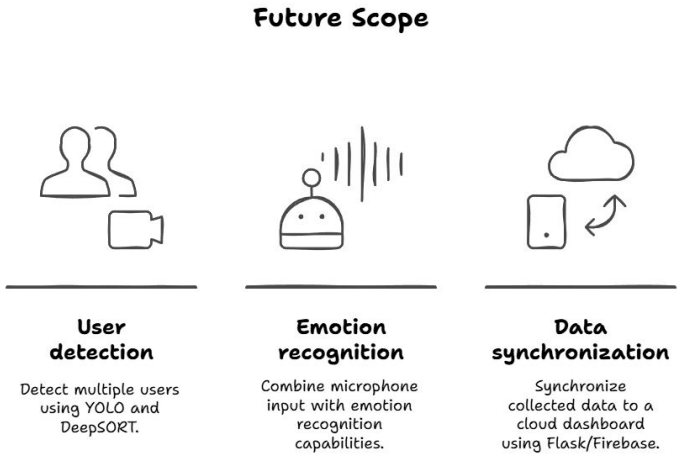

Smoothed Movement Intensity Over Time
Intensity Score
250
200
150
100
50
0
0
100
200
300
400
500
600
Frame


Model Accuracy Comparison
Gradient Boosting
99.50%
Ridge Classifier
99.58%
Logistic Regression
99.83%
Random Forest
99.92%
99.4
99.5
99.6
99.7
99.8
99.9
100.0
Test Accuracy (%)


Random Forest Leads for Jetson Nano Deployment
Evaluation Criteria
Accuracy
Speed
Memory
Deployment
Suitability Score (out of 5)
4.5
4.5
4.5
5.0
0
1
2
3
4
5
Random Forest
Gradient Boost
Logistic Reg.
Ridge Classifier

CONCLUSION & FUTURE SCOPE





THANK YOU