Archives available at journals.mriindia.com

International Journal of Recent Advances in Engineering and Technology

ISSN: 2347 - 2812

Volume 14 Issue 02s, 2025

Mood Sync: Personalized Music and Driver Safety through Facial Emotion Recognition

[1]Mr. Vaibhav U. Bhosale, [2]Atharva A. Chinke, [3]Shreeya A. Shete, [4]Avishkar V. Bhujbal, [5]Shreyash M. Taralekar

[1]Project Guide, Dept. of Computer Science and Eng Karmaveer Bhaurao Patil College of Engineering,Satara, Maharashtra, India.

[2,3,4,5] Dept. of Computer Science and Eng Karmaveer Bhaurao Patil College of Engineering,Satara, Maharashtra, India.

Email: [1]vaibhav.bhosale297@gmail.com, [2]atharvachinke799@gmail.com, [3]shreeyashete06@gmail.com, [4]avishkar.033@gmail.com, [5]shreyastaralekar27@gmail.com

Peer Review Information

Submission: 21 Oct 2025

Revision: 18 Nov 2025

Acceptance: 05 Dec 2025

Keywords

Deep Learning models, frameworks, drowsiness detection, emotion recognition, integrated system, fatigue signs, emotional detection, music recommendation, dlib library, real-time monitoring, AWS Rekognition, AWS Polly, AWS S3 buckets, driver safety, personalized music

Abstract

With the latest advancements in Deep Learning models and frameworks, we can tackle more complex problems than ever before. In this paper, we focus on two key areas: detecting drowsiness and recognizing emotions. Our goal is to create a system that can understand a driver's emotional and physical state, and respond appropriately. By alerting the driver when signs of fatigue are detected and suggesting music based on their current emotions, we aim to enhance their driving experience.

For drowsiness detection, we use the dlib library and a facial landmark shape predictor to monitor the driver's eye conditions in real time. If the eyelids stay closed for a short period, an alert is triggered to wake the driver. Additionally, we incorporate AWS Rekognition to improve facial emotion detection, AWS Polly to generate audio alerts, and AWS S3 buckets to efficiently store and manage data. This integrated approach not only ensures driver safety but also personalizes their journey with music that suits their mood.

Introduction

In today’s fast-paced world, ensuring driver safety and enhancing the driving experience are more important than ever. Emotion-Driven Tunes is a cutting-edge system that uses facial recognition technology to achieve these goals. By analyzing a driver's facial expressions, this system can detect emotions in real time and respond accordingly, creating a safer and more enjoyable journey for all.[3]

Our project is focused on developing a standalone hardware application that is capable of detecting and classifying a driver's emotions. Based on these detected emotions, it plays personalized music from Spotify through an internet connection. Additionally, the system detects drowsiness and triggers an alarm to alert the driver, ensuring their safety. Personalization is key in our approach—both the music playlist and the alarm can be tailored to the user's preferences, providing a customized and enhanced driving experience.

Music plays a significant role in making car rides more enjoyable. It can elevate the mood, reduce stress, and even help maintain focus. Our system's ability to recommend music based on the driver's emotions ensures that the music always complements their mood, making for a more pleasant journey. This system is highly compatible with various types of vehicles, from personal cars to commercial trucks, ensuring that all drivers can benefit from enhanced safety and entertainment.[2] Imagine driving home after a long day at work and feeling a bit stressed. The system recognizes your emotions and starts playing soothing music to help you relax. On a long night drive, it can detect when you're feeling drowsy and instantly play an upbeat playlist to keep you alert, while simultaneously triggering an alarm to ensure your safety. [2]

Sleep plays a pivotal role in how our bodies function [1]. The state of sleepiness, which is the struggle to stay awake during activities, impacts this process significantly [2]. Sleep deprivation, characterized by an insufficient amount of sleep either in terms of duration or quality, can result from various factors including circadian rhythm disorders, deliberate sleep avoidance, or involuntary insomnia [3]. This lack of sleep impairs brain function, leading to slower reaction times and diminished decision-making abilities.

Benefits of Creating Playlists and Listening to Music According to Our Mood

- Music Can Influence Our Mood:

Ever noticed how an upbeat song can instantly lift your spirits? Music has this incredible power to change our mood. Lively and energetic tunes can make us feel more positive and ready to take on the day, while calming melodies can help us unwind and find peace. Crafting a playlist that matches how you're feeling can really elevate your emotional state and make any moment more enjoyable.

- Music Affects Mental Health in a Good Way:

Music is more than just entertainment; it's a form of therapy. Listening to your favorite tracks can help ease anxiety and lift the fog of depression. It's like having a good friend who understands what you're going through. Music gives us a safe space to process and express our emotions, promoting a sense of well-being and mental balance.

- Reduce Stress Levels:

Stressed out? Music can come to the rescue. It has this magical ability to lower cortisol levels, which are the hormones that make us feel stressed. When you listen to tunes that resonate with you, it helps melt away the tension and anxiety.

Related Work

A. Existing Systems

1. Rule-Based Systems:

Imagine a system that reads your facial expressions and decides how you're feeling based on some pre-set rules. That's basically what rule-based systems do. They analyze facial features and expressions to classify emotions using predefined rules and algorithms. For example, a smile might be mapped to happiness, while a frown could mean sadness. These systems rely on a database of facial landmarks and their corresponding emotions to interpret expressions. They work well in controlled settings but can struggle with the complexity and unpredictability of real-world conditions.

2. Steering Wheel Movement:

Another way to keep drivers safe is by looking at how they handle the steering wheel. Drowsy drivers often show less consistent steering behavior—like frequent lane drifting, overcorrecting, or making jerky movements. By monitoring the steering wheel's movements, the system can spot patterns that suggest the driver is getting sleepy. Sensors track the steering wheel's position and movement over time, providing real-time feedback on the driver's state. However, this method might not account for other factors that can affect steering behavior, such as road conditions or the driver's unique style.

3. Vehicle-Based Measures:

This approach uses data from the vehicle itself to assess the driver's condition. It looks at things like steering wheel movements, lane position, speed, and braking patterns. By combining data from various sensors and systems within the vehicle, this approach gives a comprehensive view of the driver's behavior. For instance, if a driver keeps drifting out of their lane or brakes erratically, the system can guess that they might be drowsy or distracted. These measures often work alongside other systems to provide a well-rounded assessment of driver safety. However, they might need advanced hardware and software integration within the vehicle.[2][5]

B. Review of Existing System /Enhanced System

Our system addresses the limitations of existing systems and offers a more comprehensive and personalized approach to driver safety and entertainment.

- Standalone Hardware Application:

Think of our system as a super smart co-pilot! It's a standalone hardware application that can detect and classify a driver's emotions. This means you don't need complicated vehicle integration, making it super accessible and compatible with various types of vehicles.

- Emotion-Based Music Recommendation:

Imagine cruising down the road and your favorite tunes start playing, perfectly matching your mood. Based on the emotions it detects, our system plays personalized music from Spotify through an internet connection. Music is essential for making car rides more enjoyable— it can boost your mood, reduce stress can be seen in Fig.1.

- Drowsiness Detection and Alert System:

Safety first! Our system doesn't just guess if you're sleepy—it directly monitors your physical and emotional states. When it detects drowsiness, it triggers an alarm to wake you up, ensuring you stay safe on the road. Unlike methods that only watch your steering, our approach is more accurate and timelier.

- Personalization:

Personalization is at the heart of what we do. Both the music playlist and the alarm can be tailored to your preferences, providing a customized and enhanced driving experience. Whether it's your favorite upbeat tracks to keep you energized or a specific alarm sound that really gets your attention, we've got you covered. This not only improves safety but also makes those long drives more enjoyable and less monotonous.

Emotion Recognition And Music Recommendation

Emotion Recognition for Safer Driving

1. Recognizing Facial Emotions:

Imagine a system that can read the driver's face to sense their emotions—whether they're happy, stressed, or fatigued. This can be achieved using advanced machine learning models like Convolutional Neural Networks (CNNs). These models study facial features to identify different emotions. To make the detection even sharper, combining facial expressions with other inputs, like voice or movement patterns, can help. For instance, researchers have found that blending facial details with motion data can achieve impressive accuracy levels. Real-time analysis is key here, and tools like OpenCV (for extracting facial features) and TensorFlow or PyTorch (for building the models) are ideal choices.[4]

2. Detecting Drowsiness:

Sleepy drivers are one of the biggest risks on the road, so detecting signs of drowsiness is vital. This can be done by observing cues like how long someone's eyes stay closed, head movements, or frequent yawns. Technologies like YOLO (You Only Look Once) or Vision Transformers (ViT) can help monitor these signs effectively. When combined with emotion recognition, this creates a powerful system for driver safety.[3]

3. Safety Alerts:

If the system notices that the driver is drowsy or experiencing negative emotions, it could immediately step in to help. For instance, it might sound an alarm or play a voice message through the car's speakers. Libraries like gTTS (Google Text-to-Speech) could even deliver these alerts as spoken words.

1. Matching Music to Emotions:

Why not let the driver's mood decide the playlist? If the system detects stress, it could suggest soothing music, or if the driver seems fatigued, it could opt for something lively and energetic. By tapping into Spotify's API, the system can create and play mood-specific playlists, all based on emotion recognition.[6]

2. Smart Personalization:

Personalization takes this a step further. By learning about the driver's music preferences and linking them to their emotional states, the system can refine its recommendations over time. Using machine learning techniques like collaborative filtering or decision trees, it could tailor playlists that feel just right for any mood.

3. Hardware to Make It Happen:

This could all be packed into a sleek, standalone device equipped with a camera to capture facial cues, a microphone to analyze voice, and a speaker to play music or alerts. With internet connectivity, the system could integrate with Spotify and even receive cloud-based updates to stay ahead.

Important Extras to Consider

- Privacy First: Protecting user data is essential, so robust encryption and anonymization methods are a must.
- Smooth and Fast: The system needs to work in real time without any noticeable delays to keep the driving experience uninterrupted.

Fig.1. Emotions

Material And Methods

A. Proposed Methodology

1. Data Collection
   - Use a camera built into the hardware device to capture images or video frames of the driver's face.
   - "We used EAR = (||p2 - p6|| + ||p3 - p5||) / (2 * ||p1 - p4||) to monitor eye openness. A threshold EAR < 0.25 over multiple frames triggered drowsiness alerts."
   - Gather pre-labeled datasets like FER2013 for identifying emotions (e.g., happy, sad, angry) and others specifically tagged for detecting drowsiness (e.g., alert vs. drowsy states).[1]
2. Preparing the Data
   - Resize and normalize the images to ensure consistent dimensions and pixel values.
   - Convert images to grayscale for simpler, faster analysis without losing key facial features.
   - Apply noise-reduction techniques.
   - Use techniques like flipping, rotating, and zooming to expand the dataset and make the system more versatile.
3. Facial Detection
   - Utilize Amazon Rekognition to quickly and accurately identify the driver's face in captured images or video.
   - Rekognition will also help pinpoint critical facial landmarks (like the eyes and mouth), which are essential for both emotion detection and drowsiness.
4. Extracting Features
   - Calculate the Eye Aspect Ratio (EAR) from facial landmarks (identified through Rekognition) to detect signs of drowsiness, such as closed or nearly closed eyes.
   - Use Rekognition's tools to identify specific features of the driver's face for emotion recognition.
   - Store all processed images and data securely in Amazon S3, and manage access with IAM Policies to ensure security and privacy.

5. Emotion and Drowsiness Analysis
   - Emotion Recognition: Analyze facial features using Rekognition to classify emotions like happiness, sadness, or anger as shown in Fig. 1.
   - Drowsiness Detection: Monitor eye movement data and thresholds to identify when driver becomes drowsy.

6. Personalized Music Recommendations
   - Connect to Spotify's API to suggest playlists tailored to the driver's emotions. For example, calming music might play when sadness is detected, or upbeat tracks when drowsiness is observed.
   - The system relies on an emotion-to-music mapping model to match emotions with suitable music genres or tracks.
   - Custom alarms for drowsiness are also generated via Polly to ensure the system aligns with the driver's preferences.
7. Hardware Integration
   - Build a standalone hardware device that combines a camera, processor, and internet connectivity.
   - Use Amazon Lambda to integrate all AWS services (Rekognition, Polly, S3) seamlessly and efficiently.
   - This device will act as a self-contained solution for monitoring driver emotions, recommending personalized music, and ensuring driver safety with drowsiness alert performance as shown in Fig.7, Fig.8, Fig.9, Fig.10, Fig.11, Fig.12.

B. Technology Stack

1. Deep Learning Architecture & AWS Integration
   - Convolutional Neural Networks (CNNs): Our solution relies on CNNs to extract spatial hierarchies from images. These neural networks, built with layers that include convolutional, pooling, and fully connected stages, are essential for detecting subtle patterns in facial expressions.
   - We fine-tune pre-trained models such as

VGG16, ResNet, and Inception to customize them for emotion recognition and drowsiness detection.

Cloud-Powered Deep Learning:

- Amazon Rekognition leverages state-of-the-art deep learning frameworks to perform facial analysis at scale. With the help of the Boto3 client, our system communicates with Rekognition to analyze images in real time.
- Custom layers are integrated into our models to focus particularly on driver-specific features, such as eye movements that indicate drowsiness.[1]

2. Facial Landmark Detection

Amazon Rekognition:

- This service detects and maps facial landmarks (e.g., eyes, nose, mouth) automatically, providing high-accuracy facial geometry analysis.
- Using these landmarks, we calculate the Eye Aspect Ratio (EAR) to determine the degree of eye closure—a critical measure for detecting drowsiness.

Data Security & Management:

- All captured images and processed data are securely stored in Amazon S3, with access strictly managed via IAM policies.
- The Boto3 client streamlines our interactions with Rekognition and S3, ensuring that data handling is both efficient and secure.

3. Emotion Classification

Emotion Detection via Amazon Rekognition:

- Rekognition's pre-trained deep neural networks classify facial expressions into emotions like happiness, sadness, anger, or surprise, which directly informs our music recommendation engine.

Extended Training & Inference:

- In addition to Rekognition, our custom CNN models—fine-tuned on datasets like FER2013— complement the detection pipeline for unique, context-specific requirements.
- Real-time inference is coordinated through AWS Lambda, enabling serverless, on-demand processing of live video feeds.

Data Flow & Security:

- Classification results are transferred securely through the Boto3 client and stored in Amazon S3 following strict IAM guidelines, ensuring real-time data analysis and secure storage.

4. Real-Time Processing

Orchestrated Workflow with AWS Lambda:

- AWS Lambda functions coordinate the entire processing pipeline—capturing frames, sending them to Amazon Rekognition for facial and emotion analysis, and managing the return of these results.
- Techniques like model quantization and hardware acceleration (e.g., using GPUs) work in tandem with Lambda to minimize latency.[9]

Boto3-Driven Data Handling:

- The Boto3 client facilitates seamless API communications between the local device and AWS services, ensuring that real-time decisions are accurate and timely, and the model architecture can be seen in Fig.*2.*

5. Personalized Music & Safety Alerts

Spotify API Integration:

- The system connects with Spotify to provide personalized music recommendations. Depending on the driver's detected emotional state—be it calm, upbeat, or even stressed—playlists are dynamically curated.

Amazon Polly for Audible Alerts:

- When drowsiness is detected, Amazon Polly converts text-based warnings into clear, real-time voice alerts, helping to keep the driver alert and safe.

Integration via AWS Services:

- Boto3 is used to manage all of our API calls, ensuring that services like Polly and our custom endpoints communicate flawlessly across the cloud ecosystem.

6. Amazon Rekognition Integration

Centralized Facial Analysis:

- Amazon Rekognition does more than just landmark detection—it continuously evaluates facial expressions to determine a wide range of emotions and identifies key indicators of drowsiness.

API-Driven Workflow:

- Direct calls via the Boto3 client let our system access Rekognition's powerful analysis capabilities, processing images and video streams in real time.

Security & Data Handling:

- All processed data, including emotion metrics and facial analysis results, are securely transmitted using IAM policies and stored in Amazon S3.

C. System Architecture

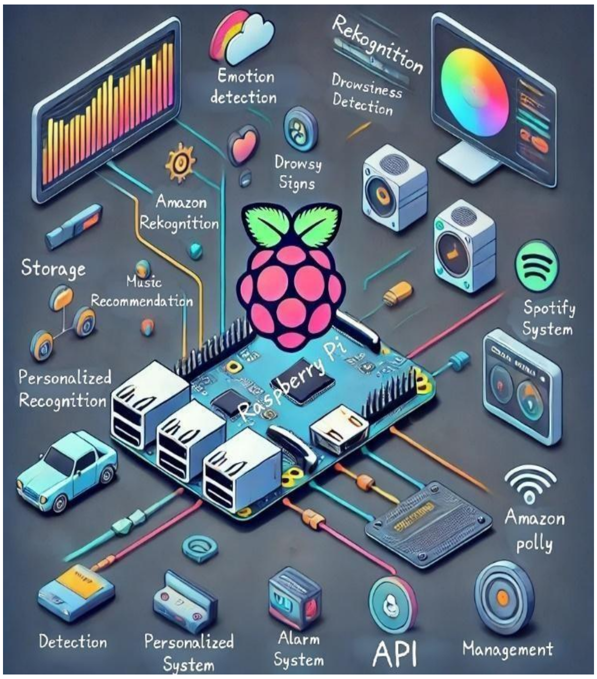

Fig.2. Architecture

Real-Time Alerts & Personalization:

- Results from Rekognition feed directly into our Lambda-managed workflow, triggering immediate alerts via Polly and ensuring the music recommendations are always aligned with the driver's current state.

7. Standalone Hardware Integration

Device Components:

- The hardware features a driver-facing camera, an on-board processor, and robust internet connectivity designed specifically for real-time facial analysis and cloud communications.[1]

Cloud Connectivity & Workflow Orchestration:

- Our device interfaces with AWS services like Rekognition, Polly, and S3 through AWS Lambda functions invoked via the Boto3 client, ensuring smooth, end-to-end integration.

Comprehensive In-Vehicle Solution:

- The combined functionalities—emotion recognition, drowsiness detection, personalized music playback, and audible safety alerts—are seamlessly integrated into a single device, providing both an enjoyable and secure driving experience.

Access Control & Data Security:

- Strict IAM protocols govern every interaction, ensuring that all stored data in Amazon S3 and exchanges between services are protected, enabling a trusted user experience, and the model architecture can be seen in Fig.2.

D. Objectives

1. Enhance Driver Safety

We want every driver to feel secure and supported behind the wheel. To achieve that:

- Real-Time Drowsiness Detection:

Our system continuously checks the driver's facial cues using Amazon Rekognition. If it notices signs of fatigue, it immediately triggers a gentle audio alert via Amazon Polly. This timely reminder encourages drivers to take a break when needed.[9]

- Nonintrusive Alerts:

Safety warnings and music recommendations are designed to be subtle and unobtrusive—ensuring that drivers remain focused on the road without unnecessary distractions.

2. Real-Time Performance and Efficiency

Speed and responsiveness are key to safety and enjoyment:

- Low Latency Processing:

By leveraging AWS Lambda and GPU acceleration, we ensure that video frames are processed in near real time, keeping any feedback almost instantaneous.

- Efficient Algorithm Deployment:

Our face detection, emotion recognition, and drowsiness analysis algorithms are finely tuned for speed and accuracy, delivering fast results without compromising on detail.[8]

- Continuous Monitoring:

The system keeps a constant eye on the driver's state, dynamically updating both music suggestions and safety alerts as new information is captured.

3. Robustness Across Environmental Conditions

Whether the weather changes or lighting shifts dramatically, our system is built to perform reliably:

- Adaptive to Lighting Variations:

Designed to handle everything from bright sunlight to deep nighttime conditions (even when weather isn't cooperating), we use advanced image preprocessing techniques to maintain performance.

- Handling Facial Variations:

Whether a driver is wearing sunglasses, a hat, or viewed from various angles, our setup accurately recognizes facial expressions and detects signs of drowsiness as per the actual output Fig.13.

4. High Accuracy in Emotion Recognition and Drowsiness Detection

Precision is key when your safety depends on it:

- Advanced Deep Learning Models:

We fine-tune top models like VGG16, ResNet, or Inception (with the extra power of Amazon Rekognition) to capture even the subtlest facial cues and moments of fatigue.

- Tailored Feature Extraction:

Using custom layers and calculating metrics like the Eye Aspect Ratio (EAR), our system focuses on critical features—specifically those that signal a drop in alertness.

- Continuous, Data-Driven Improvement:

Through regular real-world testing and user feedback, we refine our models to adapt to changing conditions and ensure top-notch accuracy over time [5]

5. Personalized Music Recommendation

Your ride should always have the perfect soundtrack:

- Emotion-Based Playlists:

We map detected emotions to mood-specific tracks using models inspired by Russell's Circumplex Model. The Spotify API then fetches personalized playlists that resonate with how you're feeling at that very moment.

- Customizable User Experience:

Drivers can tailor their music preferences and alarm sounds to their liking, ensuring that both the tunes and safety alerts match their unique style and mood as per propsed in *Fig.3.*

- Seamless Integration with Safety Alerts:

While the music adapts to the driver's emotional state, clear safety notifications are always prioritized—keeping you aware without interrupting your drive.

- Personalized commands for processing

We can communicate with the system through commands, without making physical contact with the system.

The basic commands are as follows:

Diagnose, Play music, Stop music, Increase volume, Decrease volume, Standby.

6. Data Privacy & Security

Your privacy is as important as your safety:

- Secure Data Handling:

All captured images and processed data are securely stored in Amazon S3, with strict access controls enforced by IAM policies.[7]

- Anonymization and Compliance:

We anonymize facial data to protect your identity, ensuring our system complies with privacy regulations while still delivering top-notch functionality.

Implementation Setup

- The system was implemented on a Raspberry Pi 5 running Raspbian OS (64-bit), equipped with 8GB RAM for efficient multitasking and processing. A Logitech USB webcam was used to capture real-time video of the driver's face.
- Frames were preprocessed locally and then encoded using base64 before being sent to AWS Lambda via the Boto3 client for analysis.
- The Pi connects to the internet using an onboard WiFi module, enabling seamless communication with cloud services.
- The deployment pipeline follows a modular approach:

(1) frame capture → (2) Lambda invocation → (3) emotion and drowsiness detection via Rekognition → (4) Spotify playlist control or alert via Polly → (5) response handling on device. This lightweight cloud-integrated architecture ensures near real-time responsiveness with minimal latency and reliable performance in dynamic driving environments.

Result And Decisions

Proposed Model

Fig.3.Proposed Model

Model Distribution:

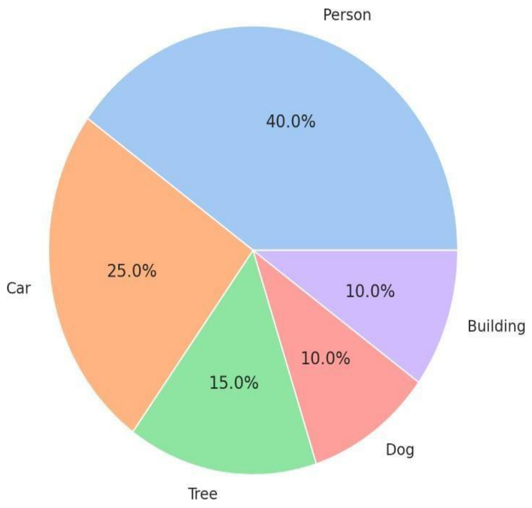

Fig.4.Model Distribution chart

- Above chart provides a clear visual representation of how the AWS Rekognition model identifies and categorizes objects, which is crucial for evaluating its effectiveness and focus areas.
- The pie chart titled "Label Detection Distribution" illustrates the performance of the AWS Rekognition model in detecting various labels as per above Fig.4.

1. Person: AWS Rekognition excels at identifying individuals in images and videos. It can detect faces, analyze facial attributes (like emotions, age range, and facial hairs.
2. Car: The model can identify vehicles, including cars, in various contexts. It can distinguish between different types of vehicles and is often used in traffic monitoring, parking management, and smart city solutions.
3. Tree: Rekognition can detect natural elements like trees, which is useful in environmental monitoring, forestry management, and even landscape photography analysis.
4. Building: The model can recognize architectural structures, which is helpful in urban planning, real estate, and tourism- related applications.
5. Dog: AWS Rekognition can identify animals, including dogs, and even specific breeds in some cases. This feature is often used in wildlife monitoring, pet identification, and content moderation.

AWS Recognition Emotions Accuracy:
The bar graph showcases the Amazon Rekognition model's accuracy in detecting emotions. It evaluates emotions such as Happy, Sad, Angry, Surprised, Disgusted, Fearful, Neutral, and Confused, with accuracy levels ranging from 70% to 100%. Each emotion is represented by unique colors or line styles.

4.

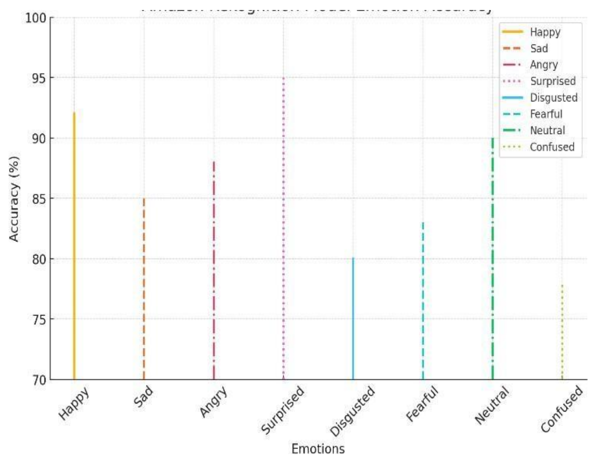

Fig.5.Emotions Accuracy

1. Driver Safety First Safety lies at the heart of our system. By using Amazon Rekognition, it monitors the driver's facial expressions to identify early signs of fatigue or drowsiness. If drowsiness is detected, Amazon Polly triggers an instant audio alert, encouraging the driver to take a break.

The system's alerts are subtle—think soft chimes or calm voice prompts—ensuring safety without distracting the driver. Additionally, backup mechanisms ensure that any detection failure doesn't compromise safety, as the system will still prompt the driver to assess their condition.

2. Real-Time Responsiveness Timely feedback is critical. Powered by AWS Lambda and optimized with GPU acceleration, the system processes video frames within moments to provide instant feedback.

All detection algorithms—whether for facial cues, emotions, or drowsiness—are tailored to deliver fast and accurate results and accuracy shown in Fig.5. Continuous tracking keeps recommendations and alerts aligned with the driver's real-time condition.

3. Built for Real-World Conditions The system adapts to varying environments, handling bright daylight, nighttime, or even bad weather with ease. Using preprocessing techniques, it remains reliable regardless of changes in lighting.

Accuracy Chart (testing vs testing) and (training vs testing):

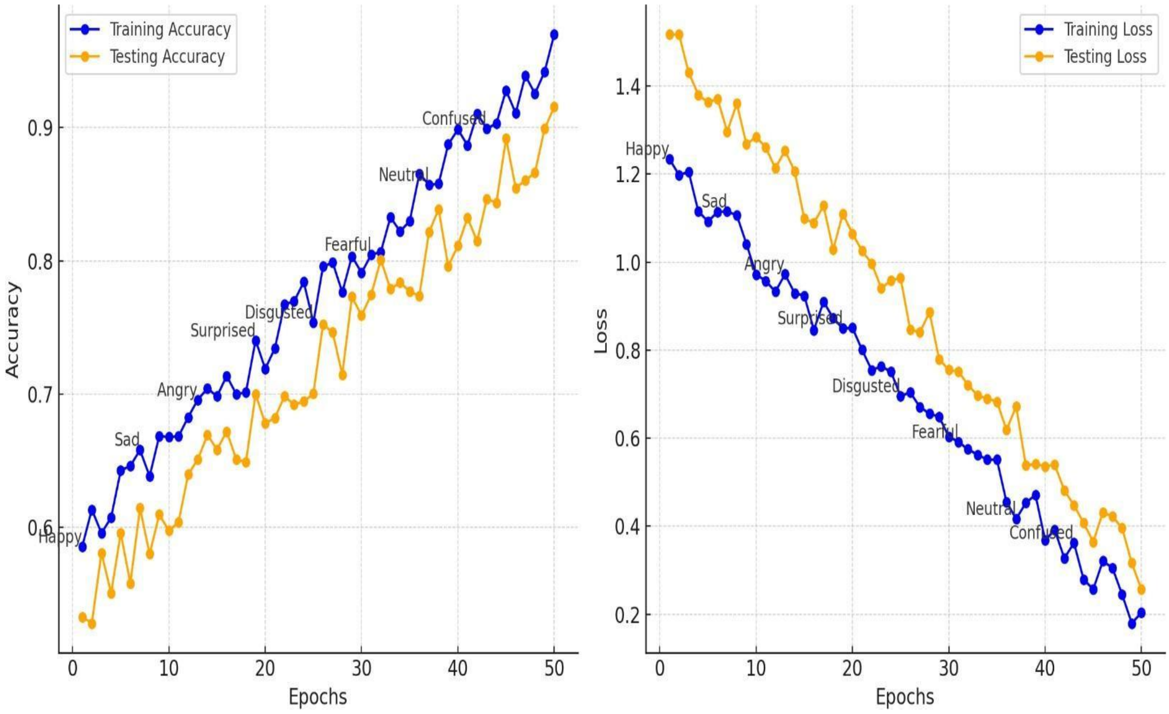

Fig.6.Accuracy

The graphs provide a visual representation of the performance of a model over 50 training epochs, with two key metrics being tracked: accuracy and loss. each graph conveys in simpler terms, as per accuracy in Fig.6.

- Training vs. Testing Accuracy (Left Graph):

This graph displays how well the model is learning to correctly classify data (accuracy) during training and testing over time (epochs):

- Blue Line (Training Accuracy): It begins at around 0.6 (60%) and steadily climbs, reaching approximately 0.95 (95%) by the 50 th epoch. This shows that the model improves its understanding of the training data as it progresses.
- Orange Line (Testing Accuracy): Starting at roughly 0.55 (55%), it increases more slowly, ultimately reaching about 0.85 (85%) at epoch 50. This indicates the model is getting better at generalizing to new, unseen data, though it doesn't match the training accuracy.

- Training vs. Testing Loss (Right Graph):

This graph focuses on "loss," a measure of how far off the model's predictions are from the actual data, over the same epochs:

- Blue Line (Training Loss): It starts at a higher value near 1.4 and continuously decreases, ending close to 0.1 by the 50th epoch. This suggests the model is learning effectively from the training data, with fewer mistakes over time.
- Orange Line (Testing Loss): Similarly, it begins at about 1.4 and decreases to approximately 0.3 at the 50th epoch. The higher final value compared to training loss highlights a small gap in performance when tested on unseen data.

- Explanation:
  - The graphs collectively show that the model is learning well, with increasing accuracy and decreasing loss, both for training and testing datasets.
  - The gap between training and testing performance (higher accuracy and lower loss for training) suggests a degree of overfitting, but the testing results are still impressive as shown in Fig.6.

RESULT TABLE

Table 1: System Performance Matrices on Raspberry Pi 5

FEATURE	SCENARIO	DESCRIPTION
HardwarePerformance (Raspberry Pi 5)	Overall Performance of all integrated components when deployed on the Raspberry Pi 5 Metric: CPU /GPU usage, power Consumption, thermal performance, and real-time resource allocation	The Raspberry Pi 5 is expected to efficiently run all pre-processing tasks while offloading heavy image analysis to AWS Lambda, operating within acceptable thermal and power limits in a portable device
Error Handling and Robustness	Errors such as malformed imaged data, network outage, or unexpected lambda responses Metric: Appropriate logging, fallback behaviors, and stability without system crashes.	In the event of any errors (e.g., invalid image payload, API, invocation issues) the system should log detailed error messages and gracefully handle failures
Drowsiness Alert Trigger	A Subject's eyes remain closed for consecutive frames. Metric: A drowsiness counter increments; upon exceeding the threshold, SNS(email/SMS) alerts are sent and an audio alert is played.	When a sufficient number of frames (e.g. more than 4)indicates closed eyes, the system should trigger an alert by sending SNS notification and playing an audible warning.
Emotion Detection	The Pi processes frames to determine the dominant emotion(e.g., "happy" or "sad") from the detected face. Metric: Accurate identification and labeling of emotions	The system must extract the emotions details from the "Emotions" "array" returned by lambda and label the detected emotion.
Voice Command Recognition	The user issues spoken commands (e.g., "pause music"," resume music")via a microphone connected to the Pi. Metric: Commands are accurately interpreted and actions are triggered accordingly.	The integrated voice recognition system (using SpeechRecognition) should capture and correctly parse commands, leading to the expected actions (i.e. pausing/resuming Spotify, toggling emotion analysis).
IAM/Rekognition Permissions	Metric: The lambda function attempts to call rekognition: DetectFaces on the provided image. Scenario: Successful detection without	The Lambda's execution role should possess the AmazonRekognition ReadOnlyAccess policy (or a custom inline policy
Face Detection Accuracy	Scenario: The Pi continuously captures video frames and sends them to Lambda for analysis. Metric: Correct detection of facial features like eyelid status and emotion attributes.	Each captured frame is expected to be successfully encoded(JPEG+base64) sent to Lambada, and then analyzed for facial features.
Lambada Invocation	Input: A valid base64-encoded JPEG image is sent from the Raspberry Pi to FaceDetection Lambda via boto3. Metric: Response JSON includes a Face Details array.	The pi sends the image payload over a stable network connection, and the AWS Lambda function processes the image and returns a JSON response containing FaceDetails (e.g. Eyes open, Emotions).

Overall System Latency	<u>Scenario</u>: Measure the end-to-end delay from video capture on the Pi to Lambada invocation, processing, and response display. <u>Metric</u>: Round-trip latency (in seconds).	The entire process (capturing video frame, encoding, invoking Lambda, processing the result, and triggering subsequent actions) should operate in near-real-time, with delays of 1-3 seconds, to maintain the interactivity.
Emotions Detection and Spotify Control	<u>Metric</u>: Correct emotion detection and automated playlist playback. The PI processes frames to determine the dominant emotion (e.g. "happy" or "sad") from the detected face and selects a corresponding Spotify.	The system must extract emotion details from the "Emotions" array returned by the lambda, determine the most common emotion, and (if permitted) initiate playlist playback via Spotify audio output.

Raspberry PI

Performance Graphs:

CPU Usage:

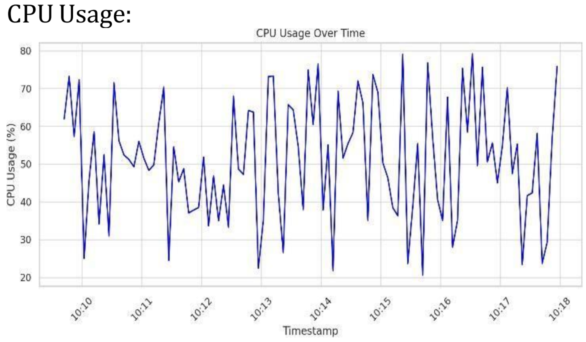

Fig.7.CPU Usage

GPU Usage:

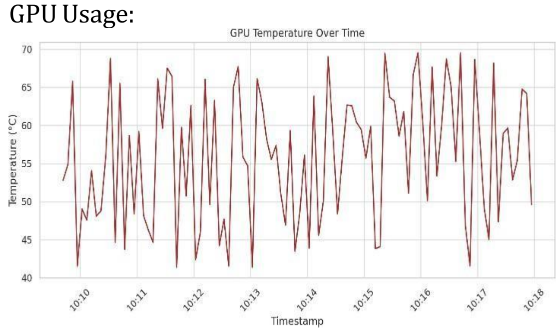

Fig.8.GPU Usage

l. Disk Read:

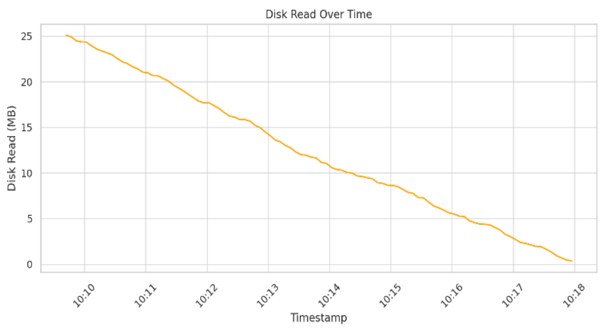

Fig.9. Disk Read

Disk Write:

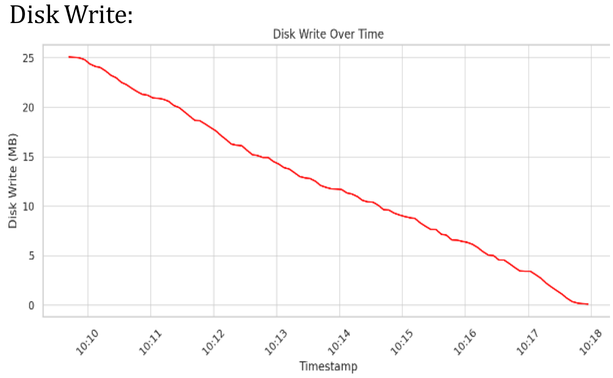

Fig.10. Disk Write

Internet Latency

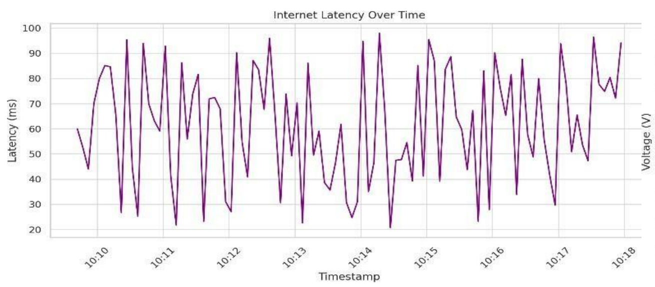

Fig.11. Internet Latency

5. Voltage

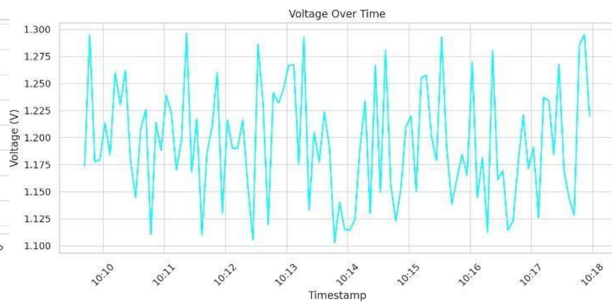

Fig.12. Voltage

ACTUAL RESULT:

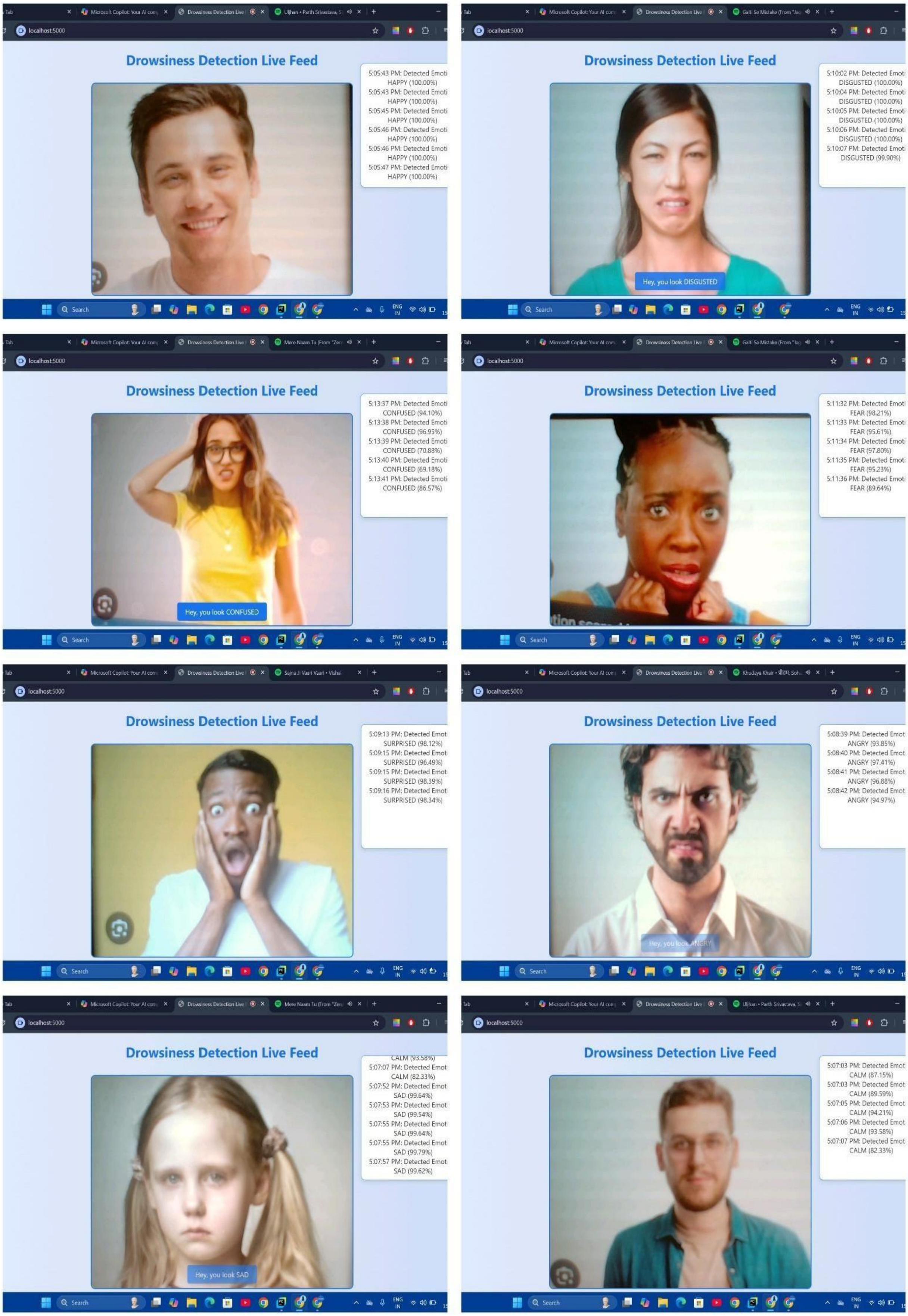

Fig.13

Conclusion

The Emotion-Driven Tunes project brilliantly demonstrates the power of integrating facial recognition technology with emotion-based music personalization to enhance driver safety and overall driving experience. By continuously monitoring the driver's facial expressions, the system can detect signs of fatigue, stress, or distraction and respond with appropriate alerts or interventions to ensure road safety.[10]

Acknowledgment

We wholeheartedly thank everyone who helped bring this paper to life. Your insights, guidance, and unwavering support have not only enriched this work but have also been a driving force behind our personal and professional growth. We truly appreciate every one of you for sharing your expertise and for walking this journey alongside us.

References

T. Lahari, S. Garaga, P. V. L. Manasa, and S. K. Sri, "Emotion Based Music Recommendation," 2024.
This module accurately and swiftly detects faces, providing the emotion detection model with clear, reliable data to interpret facial expressions effectively.

A. Phaneendra, M. Muduli, and S. L. Reddy, "Emuse: An Emotion Based Music," 2022.
Merge a power of computer vision with advanced machine learning to understand facial expressions and tailor music recommendations that truly resonate with your mood.

P. A. Álvarez, A. Guiu, J. R. Beltrán, J. García de Quiros, and
S. Baldassarri,
"DJ-Running: An Emotion-based System for Recommending Spotify Songs to Runners,"
Department of Computer Science and Systems Engineering, University of Zaragoza, 2019.
Alongside the main system, a mobile application is proposed to track a runner's activities while seamlessly integrating with Spotify for tailored song recommendations.

P. Phaneendra, M. Muduli, and S. L. Reddy,
"Emuse – An Emotion Based Music Recommendation System," 2022.
Music has an incredible ability to lift our spirits and change how we feel; this system offers a natural remedy to ease stress and transform a bad day.

D. Vaishnavi Reddy, B. Somaditya, O. Sathvika, C. Ram Prakash, and K. Srikanth,
"Emotion Based Music Recommendation System," 2022.
Traditional recommendation systems rely on fixed algorithms that cannot adjust to the natural shifts in our moods and tastes, often leaving users with suboptimal suggestions.

S. Wang, H. Shuai, L. Zhu, and Q. Liu,
"Emotion Based Music Recommendation System using CNN Expression Complementary Disentanglement Network for Facial Expression Recognition," 2020.
Integrating facial expression detection into the music player can transform the listening experience by matching songs exactly to the user's current mood.

P. Jayaprakash,
"Emotion Aware Smart Music Recommended System by Using CNN," 2022.
Extensive research in computer vision and machine learning is unlocking new possibilities for creating more intuitive, human- centric technologies.

S. Katkuri, M. Chegoor, K. C. Sreedhar, and M. Sathyanarayana,
"Emotion Based Music Recommendation System," 2023.
This work introduces a method for categorizing music into various emotional types such as happy, sad, or angry, allowing for distinct mood-based recommendations.

T. Lahariw, S. Garaga, P. V. L. Manasa, and S. K. Sri, "Emotion Based Music Recommendation System," 2024.
This paper presents an innovative system that uses real-time emotion recognition—with webcam analysis integrated with Spotify and YouTube—to create a dynamic, mood-based listening experience.

A. Jaiswal, A. K. Raju, and S. Deb,
"Facial Emotion Detection Using Deep Learning," 2020.
Advancements in AI and deep learning now enable facial recognition systems to detect emotions such as anger, happiness, and sadness, paving the way for emotion-sensiti