Cloud-Based Emotion Recognition and Sleep Time Analysis for Mental Health

MSc Research Project
Msc Cloud Computing

Jay Suresh Deshmukh
Student ID: X23244119

School of Computing
National College of Ireland

Supervisor: Dr. Shaguna Gupta

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Jay Suresh Deshmukh
Student ID:	X23244119
Programme:	Msc Cloud Computing
Year:	2024
Module:	MSc Research Project
Supervisor:	Dr. Shaguna Gupta
Submission Due Date:	24/04/2025
Project Title:	Cloud-Based Emotion Recognition and Sleep Time Analysis for Mental Health
Word Count:	XXX
Page Count:	25

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	25th May 2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Cloud-Based Emotion Recognition and Sleep Time Analysis for Mental Health

Jay Suresh Deshmukh

X23244119

Abstract

The signs of poor mental health like depression, stress, anxiety can be seen in everyone but stay undetected. This disorder can be connected with sleep cycle which can lead towards emotional instability. On the primary basis emotion recognition models is focused towards facial expression but does not analyze the sleep pattern of a person. This factor is also an important information for detecting the mental health. In this research we will propose cloud-based solution for emotion recognition which will combine facial expression and track the sleep time data so that we can find any early sign in mental health. For tracing the facial emotion, we will use Convolutional Neural Network (CNN's), Super Vector Machines (SVMs) and Leave-One-Out-Cross-Validation (LOOCV) for model evaluation and performance estimation, and to track the sleep we will use APIs like Fitbit, Google Fit which are cloud based. This will help us to know the morning emotional state and the amount of sleep time the user consumed. With the help of Amazon Rekognition for facial analysis, AWS Lambda for realtime processing, S3 bucket for storage, we will deploy this model on AWS Cloud Infrastructure. In this research we will understand and recognize the emotion along with sleep cycle to increase the accuracy of mental health detection through cloud-based solution. This will help us to understand and recognize the mental health which will reduce the burden on mental health care system. This research presents a novel cloud-based approach that combines sleep time and facial emotion recognition for monitoring mental health risk factor. Keywords: Facial Emotion Recognition, Sleep time tracking, SVMs, LOOCV, Cloud Computing, Mental Health, Convolutional Neural Networks (CNN), AWS.

1 Introduction

1.1 Background

Anxiety, depression, and chronic stress finally emerged as a global health burden as millions began to struggle with mental health disorders. Depressive disorders are a leading cause of disability, affecting more than 280 million people worldwide, according to World Health Organization (WHO) data, online since October 2023. One of the major contributors towards emotional instability and decline in mental health is the decreasing quality of sleep which the modern life style with its fast pace causes, resulting in sleep deprivation, poor quality of sleep and irregular sleep patterns. Sleep facilitates the brain, as well as emotions and psychology. Walker et al. According to the research conducted in the year (2017) Sleep cycle disorders are measured in the shifts of mood, and decreased brain activity. Additionally, people with insomnia or in a state of sleep deprivation always have a powerful negative emotion and stress-taking power. While well-known relationships exist between sleep and mental health, face-only emotion recognition models also ignore sleep data, which may be a critical mental health indicator.

1.2 Problem Statement

While traditionally FER models are mainly CNN based, these models have been shown to achieve excellent accuracy on static images for emotion detection. Studies conducted by T. Kopalidis et al. CNN models trained on the Static-facial Expressions in the Wild dataset, for example, have been able to successfully detect emotions such as happiness, sadness, anger, fear, and surprise. But these models are not performance under the real-world scenarios such as low light, occlusions and etc which would change facial expressions and more importantly cause the inconsistency of detection. Emotion recognition based on sleep patterns would be another important contributing factor, which is clearly missing in current FER models. This constraint makes them less reliable in assessing mental disorders accurately, because lack of sleep can greatly alter a person's mood and facial expressions. These models lack in depth perspective on a person overall emotional well-being without integrating sleep cycle data.

1.3 Research Aim and Objectives

Aim
In this research we aim to develop an innovative cloud-based solution that will help us to integrate the facial expressions analysis along with sleep tracking to increase the accuracy of mental health detection

Objectives

- To make a model based on Convolutional Neural Network (CNN) for identifying facial emotions.
- To retrieve sleep tracking data from IoT-based API's like Fit-bit, Google Fit, and Apple Health-kit to analyze the relation between sleep cycles and emotional states.
- To implement a cloud-based infrastructure using Amazon Web Services (AWS) for real-time data processing and storage.

1.4 Research Questions

Q1: How does combining sleep cycle data with facial emotion recognition improve the accuracy of mental health assessment?
Q2: What is the relation between sleep deprivation and emotional instability detected through deep learning models?
Q3: How effectively can AWS cloud services (Amazon Rekognition, AWS Lambda, and S3) be to deploy a scalable real-time emotion recognition and sleep tracking system?
Q4: What are the key challenges in combining facial emotion recognition and sleep tracking for mental health monitoring, and how can they be addressed through cloud-based solutions?

1.5 Significance of the Study

As mental health problems continue to grow worldwide, so does the need for validated, scalable, and affordable automated assessment tools. Models for emotion recognition based on facial expression are a traditional methodology, but this does not consider other important elements of emotion such as sleep that are also vital to emotional regulation. Hence in this research aims to provide a broader coverage approach for assessing mental well-being through a multi-sleep analysis and facial emotion recognition-based approach that is more precise. The signal processing operates in segments of the sleep cycle and improves the accuracy of emotion detection, especially in case of the use of a camera sensor, as emotion detection could be corrupted by external factors such as poor lighting or occlusion. Deploying it on the cloud makes mental health anywhere, across multiple devices, scalable and more efficient. Moreover, such a system can also help healthcare staff in early recognition of risk and timely management of affected individuals. Using this data, this research intends to help develop advanced mental health monitoring solutions using AI in the cloud.

1.6 Research Scope and Limitations

This paper discusses the integration of facial emotion recognition and analysis approach of a sleep cycle in the cloud based infrastructure to improve the mental health monitoring. The deep learning model, using Convolutional Neural Networks (CNN's) in this case, will evaluate facial expressions whereas data on sleep tracking will be analyzed by IoT-based API's from Fit-bit, Google Fit and Apple Health-kit. An extensive mechanism for seamless data handling and deployment over AWS services such as, Rekognition for facial analysis, Lambda service for real-time processing and S3 for secure storage. The study does have some limitations, however. The effectiveness of the model would thus depend on quality and availability of facial emotion and sleep data, which varies from individual to individual. Furthermore, the underlying system could need to be tuned to different demographic groups, e.g., people in different regions of the world may sleep differently or express emotions in a different way. This brings challenges of hardware dependency as effective sleep tracking requires, users to have a compatible wearable IoT device. Encryption and access controls will be established to reduce privacy and security challenges associated with the cloud storage of health data.

2 Related Work

Anxiety, depression and chronic stress are among the top ten ailments worldwide, with billions suffering from these and other mental health disorders. As one of the most widespread and burdensome causes of disability, depression impacts more than 280 million individuals around the world, the WHO (World Health Organization) stated. Conventional screening methodologies depend on self-expressed indications of misery, but there is an eminent demand for casual techniques to automated determine very early indicators of psychological pains. Deep learning based on CNN has shown better results in Facial Emotion Recognition (FER). One crucial missing components in these existing models is the bidirectionality of the relationship between sleep and mood. Insufficient sleep can lead to fluctuations in our mood, unevenness of emotions, and problems with thinking. Studies show that sleep-disordered patients are impaired at identifying emotions expressed in faces. We present a cloud integration based approach for human activity detection using facial expression recognition (FER) data with sleep tracking data from IoT devices such as Fitbit and Google Fit in this study. Using AWS services for facial analysis, real-time processing, and data storage, we want to create a scalable and efficient solution for monitoring mental health.

2.1 Sleep Quality and Emotion Recognition

Chang and Klumpp (2022) investigated the effect of subjective sleep quality on recognition of facial emotion, and whether these relationships differed between individuals with and without internalizing psychopathologies such as anxiety and depression. This includes some past studies that shed light on the effects of sleep on the interpretation of images of faces displaying both positive and negative feelings (such as anger, disgust, etc.) or neutral expressions. This indicates that poor sleep diminishes emotional processing, and deriving the implications of social cues is harder to do once the sleep cycle has been altered. People with bad sleep also show impairment in the recognition of subtle facial expressions, which can create a vicious cycle of social isolation and aggravation of the emotional condition. Integrating Sleep Analysis into Emotion Recognition Models Given that classic FER frameworks using facial expressions often miss out on prominent psychological indicators, this research highlights the need to implement supplementary techniques to better account for the social penetration of these models. These systems can improve the accuracy and reliability of emotion recognition systems like emotion detection by the data obtained by sleep tracking to enable early detection of mental health concerns as well as provide a more holistic view of the emotional status of an individual.

2.2 Sleep Deprivation and Cognitive Function in Emotion Recognition

Almondes et al. (2020) examined how insomnia impacts facial emotion recognition and executive functions, and found that those who have sleep disturbances find it quite challenging to perceive the emotion such as fear and sadness. The results show that sleep loss impairs the brain pathways that feedback to help us interpret emotional cues in others, making it difficult to communicate and regulate emotions. In addition, it reported cognition deficits linked to insomnia such as lower working memory, overall cognitive flexibility and perceptual organization. These cognitive deficits in turn will prevent the accurate reading of facial expressions, and hence the less reliable emotion recognitions. The authors highlight the importance of incorporating sleep-relevant information in FER models because long-term sleep perturbations impede not only emotional identification but higher-order cognitive processes which are critical for meaningful expression evaluation. The use of sleep cycle analysis along with deep learning based FER models can offer a holistic and better assessment of mental health.

2.3 The Impact of Sleep Deprivation on Facial Emotion Recognition

The meta-analysis by Li, Ma, and Wu (2023) identified the effect of sleep deprivation in face emotion recognition via systematic review andmeta-analysis. Their review combines results from many different studies and indicates that sleep-deprived people have a problem identifying emotional expressions, specifically happy and angry faces. In a recent study, it was theorized that inadequate sleep can disrupt affective neural systems that process emotional cues, causing misinterpretations of facial expressions. This unique impaired emotional recognition will lead to problems in social interactions and increased emotional instability. Furthermore, sleep loss leads to bias in emotion recognition that can potentially result in incorrect categorization in automatic systems. Incorporating sleep-dependent features into emotion recognition models can make them more trustworthy given these findings. When we integrate sleep cycle analysis with emotion recognition, the effect of sleep deprivation can be reduced which makes emotion recognition more reliable and helps us in knowing the real balance of emotion states.

2.4 Implications for Cloud-Based Mental Health Monitoring

It seems that the current literature highlights the importance of sleep quality (inferred through behaviors like chest tightness and facial emotion recognition) that can play a significant role in understanding the impact on mental health. Facial expression in a person suffering from sleep deprivation or insomnia would not really correspond to his/her actual emotions, thus, taking into account the facial movements in these scenarios is still inaccurate when it comes to traditional emotion detection models. However, because sleep disorders greatly impair emotional processing and cognitive functions, it gives rise to misinterpretation of facial expressions and negative perceptions of others. Combining IoT-based applications with API like FITBIT and Google Fit for sleep tracking with deep learning based facial emotion detection data can help us to have a broader and stronger edition of mental health analysis. Deploying it in the cloud takes that a step further (scalability; instantaneous feedback processing; access from any device). When the information from the sleep cycle and facial expression could come together in the aforementioned way, it would definitely be a closer representation of what a person would be going through, which could lead to identifying a mental disorder at an earlier stage. It can enhance AI-based mental health tracking systems that do not depend on traditional self-reported measures.

2.5 Comparative Analysis of the Previous Research Papers

The comparative analysis of previous studies highlights key advancements and gaps in facial emotion recognition (FER) and its relationship with sleep. Chang and Klumpp (2022) and Almondes et al. (2020) found that poor sleep quality impairs emotion recognition, particularly in individuals with insomnia or psychopathologies. Li et al. (2023) confirmed these findings through a meta-analysis, emphasizing the impact of sleep deprivation on recognizing negative emotions. Meanwhile, Kopalidis et al. (2024), Li and Deng (2020), and Revina and Emmanuel (2021) focused on deep learning-based FER, showcasing CNNs' superiority while noting challenges like dataset biases and computational costs. These insights inform our research approach.

2.6 Research Gap and Future Directions

Although sleep deprivation has previously been studied in regard to recognition of emotion, integrating sleep cycle and facial emotion recognition into the same model of mental

Table 1: Comparative Analysis of the Previous Research Papers

Study	Focus Area	Methodology	Key Findings	Limitations
Chang & Klumpp (2022)	Sleep quality and emotion recognition in individuals with psychopathologies	Behavioral study analyzing sleep quality and emotional recognition performance	Poor sleep quality negatively impacts emotion recognition, especially in individuals with internalizing disorders	Limited sample size; subjective sleep measures
Almondes et al. (2020)	Relationship between insomnia and facial emotion recognition	Neuropsychological tests measuring executive functions and FER in insomnia patients	Insomnia negatively affects executive function and facial emotion processing, reducing emotion detection accuracy	Lack of large-scale data; limited generalizability
Li et al. (2023)	Impact of sleep deprivation on FER	Systematic review and meta-analysis of existing studies	Sleep deprivation significantly impairs emotion recognition, especially for negative emotions like anger and sadness	Variability in methodologies; lack of real-time emotion tracking
Kopalidis et al. (2024)	Advances in facial emotion recognition models and benchmarks	Review of state-of-the-art FER models, datasets, and techniques	Modern deep learning models, especially CNNs and transformers, have improved FER accuracy	Lack of standardized benchmarks for real-world FER applications
Li & Deng (2020)	Deep learning techniques for FER	Survey of deep learning models for FER including CNNs, RNNs, and hybrids	Deep learning methods outperform traditional FER models in accuracy	High computational cost; limited robustness in real-world conditions
Revina & Emmanuel (2021)	Overview of human FER techniques	Analysis of traditional vs deep learning-based FER approaches	Machine learning-based FER is more effective than handcrafted feature methods	Dataset bias; need for diverse training data
Ge et al. (2022)	Deep learning-based FER	Implementation of CNNs for FER	Deep learning improves FER accuracy	Model overfitting; requires larger datasets
Bisogni et al. (2022)	AI and FER in healthcare	Applications of deep learning in healthcare FER	AI-driven FER enhances mental health assessments	Ethical concerns regarding facial data privacy
Lee et al. (2022)	Sleep deprivation and emotional processing	Experimental study on sleep fragmentation and emotional response	Emotional processing accuracy declines with sleep deprivation	Small sample size
Ma et al. (2021)	Gender differences in emotion recognition during sleep deprivation	EEG and eye-tracking study	Gender differences observed in emotion recognition	Limited demographic diversity

health monitoring has not yet been done. On the one hand, existing studies have concentrated on sleep disorders or facial emotion detection separately, whereas integrated single model from both the metrics will give the better advantageous performance. This research aims to fulfil this gap by designing a cloud-based system using emotion recognition sensor-based data fusion as well as sleep tracking which ultimately can provide a more accurate and reliable mental health risks. Since the models currently in use are often unable to accurately identify faces when they are sleep deprived or mentally fatigued, this study will use AI, deep learning, Supervised Machine Learning Algorithm (SVMs) and cross-validation techniques like Leave-One-Out-Cross-Validation (LOOCV) and cloud computing to overcome those limitations. Given the prevalence of demographic differences in both sleep and emotion expression, future work could investigate personalized models that account for inter-individual differences so that the system may better generalize. It could lead to better artificial intelligence mental health monitoring products.

2.7 Literature Summary

Innovative tools are needed to address this gap beyond self-reports with the increased burden of mental health disorders. This enhanced mental health assessment is achieved through yawning, jaw clenching, and facial emotion recognition connected with sleep cycle analysis using a cloud-based infrastructure presented in this study. With the established evidence on sleep deprivation impacting emotion recognition, the performance of conventional FER models may deteriorate. With this, we try to build a more trustworthy and holistic system of identifying early symptoms for emotional distress as it uses data from sleep-tracking IoT devices along with other data. With the combination of deep learning and cloud computing, the authors' solution system can perform extremely large-scale emotion analysis in real-time by using AWS services like Amazon Rekognition, AWS Lambda, and S3 for fast processing and safe storage. This addresses an important research gap as integrating sleep data enhances the validity of detecting emotional states. Model generalization to other populations and privacy issues should be explored in future work, enabling widespread implementation of AI-driven mental health monitoring.

3 Methodology

The research methodology used in designing and developing the cloud-based emotion recognition and sleep time inferences for mental health purposes was discussed. This method combines deep learning-based facial emotion recognition (FER) with wearables sleep tracking data for a comprehensive evaluation of emotional health. With cloud computing, this system should be scalable and most efficient, as well as performing analysis of real-time data. Our methodology starts with data collection that uses publicly available facial expression datasets upon which the FER model is trained, and sleep data is incorporated from APIs (i.e., Fitbit, Google Fit, and Apple HealthKit). It classifies facial expressions into well-defined emotional categories using a convolutional neural network (CNN). The sleep tracking part pulls out relevant data from sleep tracking, including total sleep and time spent in certain sleep stages. The emotion detector correlates with sleep patterns and reveals how lack of sleep can negatively impact our emotions. The system is implemented on AWS and performs the following on AWS services like Amazon

Rekognition, AWS Lambda, and Amazon S3 for storage and real-time processing. The ethical considerations are also dealt with in careful detail to adhere to health data regulations with respect to privacy and security. The approach allows an efficient and scalable way to monitor mental health.

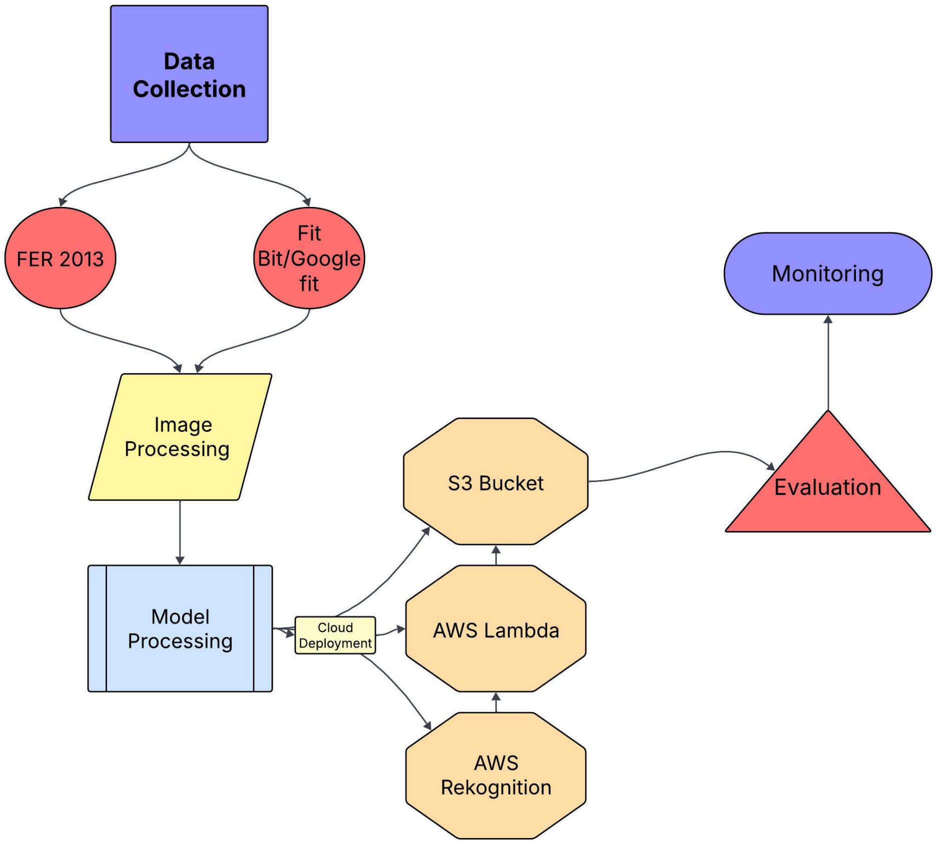

Figure 1: Approaching Steps

4 Design Specification

4.1 Research Design

This research follows a quantitative experimental design, combining deep learning-based facial emotion recognition (FER) and sleep-tracking data, in a cloud-based environment. Abstract: This paper focuses on an automatic face recognition (FER) and classification of basic emotions with high accuracy using a convolutional neural network (CNN) based approach. It is also trained on large-scale facial expression datasets in order to guarantee a robust performance in different demographics. Unlike the various information that can be gathered by apps in our digital diary period, sleep monitoring data is typically tracked from a wearable tool (assume Fitbit, Google Fit, and Apple HealthKit) which gives granular insights on sleep duration, sleep cycle (REM, light, and deep sleep) as well as sleep quality. Real-time data processing is supported by deploying the system on AWS (Chang et al. 2022). The Fusioning of Emotion Recognition and Sleep Data is one the key features of the research, be used by some statistical technique to analyze relationship among sleep deprecation and emotional instability. What this means, is to have an accurate understanding of the effect of sleep patterns on mental health.

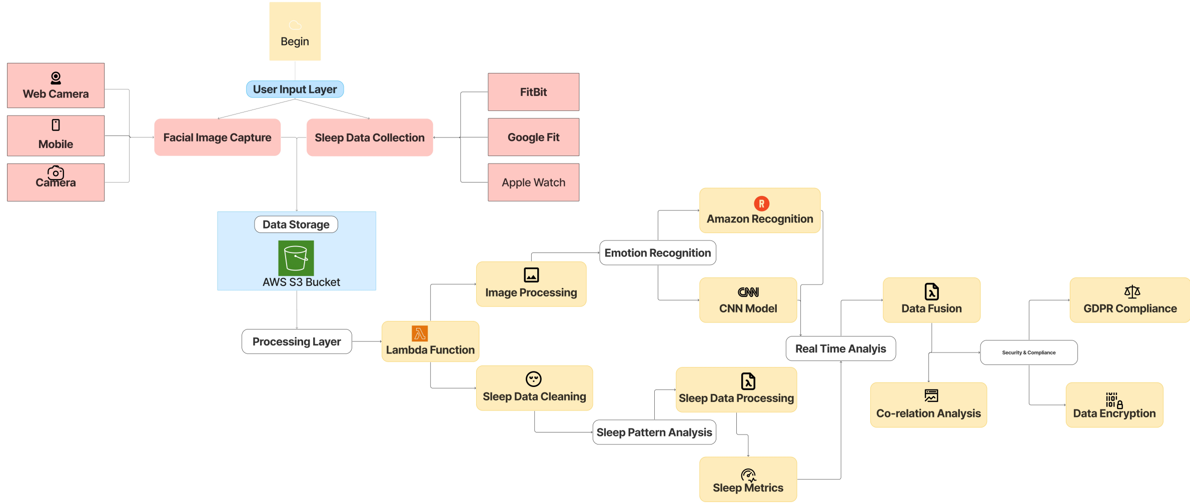

Figure 2: Architectural Diagram

4.2 Data Collection

In this research, we require two main types of information, facial images for emotion recognition and sleep duration data to analyze the sleep timing. The images for emotion recognition must be captured in morning after the sleep. Both of these data were extracted from the voluntary participant who provided written participation consent for this research study (Consent form). The participants were aware of how their data will be used, the purpose of data collection and security measures to protect their privacy. Knowing the background of participants is also important in finding the root cause of mental health. In this study we collected the data from diverse groups.

Gender Distribution

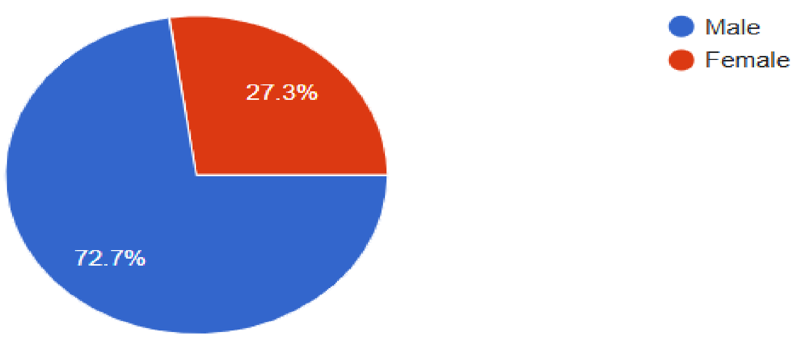

Figure 3: Gender Data

The percentage of active male and female participants are mentioned in the chart above.

- In this research we have 72.2% male who are working in private companies in various field like Finance, IT and Art.
- The female percentage rate 27.3% who are also working women in Finance, Medical field and some are homemakers as well.
- Here, the size of sample is small and also the ratio is unequal. This type of typical ratio can be seen at certain workplaces, or at some universities as well. Because of the gender diversity, we can find the layer of variability in our data which can help in testing for the robustness of emotion recognition through images.

Age Distribution

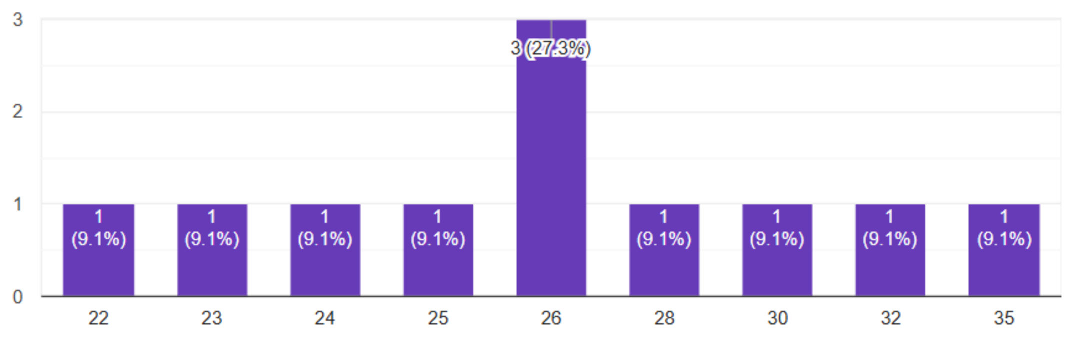

Figure 4: Age data

The targeted participants for this research were between 22 to 35 years old, covering the spectrum of early adulthood to mid adulthood. Studies have shown that most working professionals goes less sleep which can cause them mental illness. This variety in the group helps to the robustness of the data by:

- Associated with their age, differences in sleep patterns were captured.
- The accuracy of the emotion detection may reflect age related facial characteristics.

Even with the limited dataset we ensure that our study is not biased or based on specific gender, group, or occupation. An emotion expression can be influenced by lifestyle as well as biological related factors. Participants across various background and from both genders helped to enhance the generalizability of our system.

Facial Image Data:

With the help of user devices, such as webcams or mobile phones camera, the facial images were captured. To reflect their emotion stage, participants were required to submit their images daily after waking up in the morning. In the Amazon S3 bucket, these images were uploaded and stored through a secure channel. On the AWS rekognition, this data was trained. With special Id's these images were renamed to protect their identity. Sleep Tracking Data: Through the manual submitted logs or through the wearable IoT devices like Fitbit, Google Fit, the sleep data from the users were obtained. The participation was fully voluntary, and no user were forced to be the part of this research study. Users were informed that they have rights to withdraw their participation without consequences with they feel not to participate.

4.3 Facial Emotion Recognition Model Development

The convolutional neural network (CNN) has a better capacity to extract spatial features from images so we select CNN for facial emotion recognition. The architecture includes an input layer where it accepts grayscale or RGB facial images, and then several convolutional layers that extract the required features of the face. Dimensionalities are reduced at the Pooling layers making the computation faster while still encompassing the significant parameters. The features extracted at a preceding convolutional layer are mapped by a fully connected layer to the emotion categories, passing through a final softmax layer that returns the probabilities for each emotion class (Almondes et al. 2020). The model, implemented using TensorFlow and Keras, is trained using the Adam optimizer with a learning rate of [0.001]. Loss function is categorical cross-entropy, batch size is 64, training epochs is 50. Overfitting can be prevented using early stopping. Techniques like rotations, flipping, contrast etc. are used for data augmentation to improve the generalization. We measure performance in terms of accuracy, precision, recall, and F1-score to ensure the model is indeed quite reliable.

4.4 Sleep Tracking and Data Processing

Wearable IoT devices (e.g., Fitbit, Google Fit, Apple HealthKit) avail sleep patterns but direct acquisition of this data is not possible due to privacy laws and proprietary constraints. Rather the app pulls sleep data via user-activated API integrations granting access to their sleep logs. Data such as total sleep time, time spent in different sleep stages (light, deep, and REM), wake frequency, and overall sleep efficiency ratings. The obtained data are stored in a cloud-based database (AWS S3) for retrieval in a non-identifiable form to ensure data privacy prior to analysis. Statistical techniques including Pearson correlation and regression analysis are applied to examine the association of sleep deprivation on the emotional states (Li et al. 2023). They find patterns by looking at how an emotion correlates with a particular sleep time, such as increased sadness or irritability when sleep decreased. This method yields net benefit without access to device-level data extraction.

4.5 AWS Lambda for Real-Time Processing

In cloud computing, real time processing plays an important role when an application requires immediate response and instant analysis. By ensuring that the data is generated with minimum delay of time, as the data is received, continuous analysis is done. For emotion recognition and sleep tracking, this approach is much better because mental health monitoring can be enhanced by timely identification (Bisogni et al., 2022). To process facial emotion recognition, and the sleep tracking data in real time, for this research we will use AWS Lambda which is a serverless computing service by Amazon Web Services. The Lambda function is an event driven service through which we can automatically execute the code by reducing the cost and operational complexity. With the help of serverless architecture we can eliminate that needs to manage the infrastructure. To have an efficient and scalable system for mental health monitoring, we can use AWS Lambda along with AWS rekognition, Fitbit APIs, and S3 bucket to analyze real time emotion, sleep pattern and to store the data on cloud. The computational overhead and cost are reduced because Lambda will eliminate the need for constant Resource allocation (Cernadas Curotto et al., 2022). We made a comparison between AWS Lambda, Google Cloud Function and Microsoft Azure Function. Each of these services has different features along with their pros and cons. In the table below we have mentioned tools, Features, pros and cons of different cloud infrastructures.

Table 2: Tools Comparison

Tool	**Features**	**Pros**	**Cons**
AWS Lambda	Event-driven, serverless computing	Cost-effective, highly scalable, seamless AWS integration	AWS lock-in, potential cold start issues
Google Cloud Functions	Serverless execution for Google Cloud	Native integration with Google AI and BigQuery	Limited outside Google ecosystem
Microsoft Azure Functions	Event-driven computing on Azure	Supports multiple programming languages	Complex pricing structure

The Azure Function and the Google Cloud Function both have similar capabilities as AWS Lambda, but AWS Lambda is chosen in this study because it integrates with AWS Rekognition, S3 bucket and other AWS services. These services are also used in this research. The following image is the flow chart of how the AWS Lambda will be Executed.

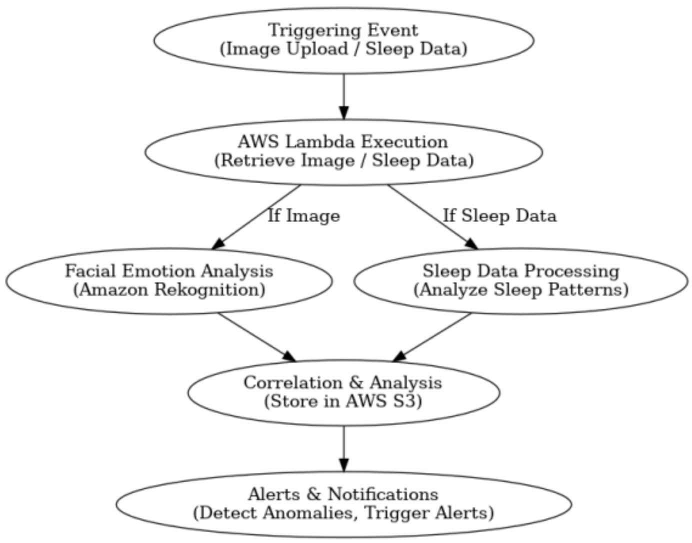

Figure 5: Lambda Execution

As mental health is analyzed instantly with the help of this real time data processing system, we can detect the early sign of emotional distress.

4.6 AWS S3 for storage

In cloud computing, ensuring the security of real time and historical analysis data is important. Therefore, in this research we will use Amazon Simple Storage Service (AWS

S3) to store the data of facial Emotion recognition and sleep tracking information for real time and historical analysis. Given the large volume of data generated from facial images and sleep logs, an efficient and secure cloud storage solution is necessary to ensure seamless data retrieval, processing, and analysis (Davidson et al., 2022). We are using AWS S3 for this research because we can store a large volume of unstructured and structured data efficiently and it provides us with high availability and security. In the table below, we have made a comparison between AWS S3, Google Cloud Storage and Microsoft Azure BLOB Storage, focusing on their features, pros and cons.

Table 3: Storage Services Comparison

Cloud Storage	**Features**	**Pros**	**Cons**
AWS S3	Object storage with high durability, encryption, and AWS integration	Scalable, secure, cost-efficient, integrates with AWS AI/ML	AWS lock-in
Google Cloud Storage	AI-integrated storage optimized for BigQuery and TensorFlow	Scalable, supports AI apps	Limited outside Google Cloud
Azure Blob Storage	Optimized for Windows-based apps	Integrates with Microsoft services	Less optimized for AI workflows

The following image is the flow chart of how the AWS S3 will be used.

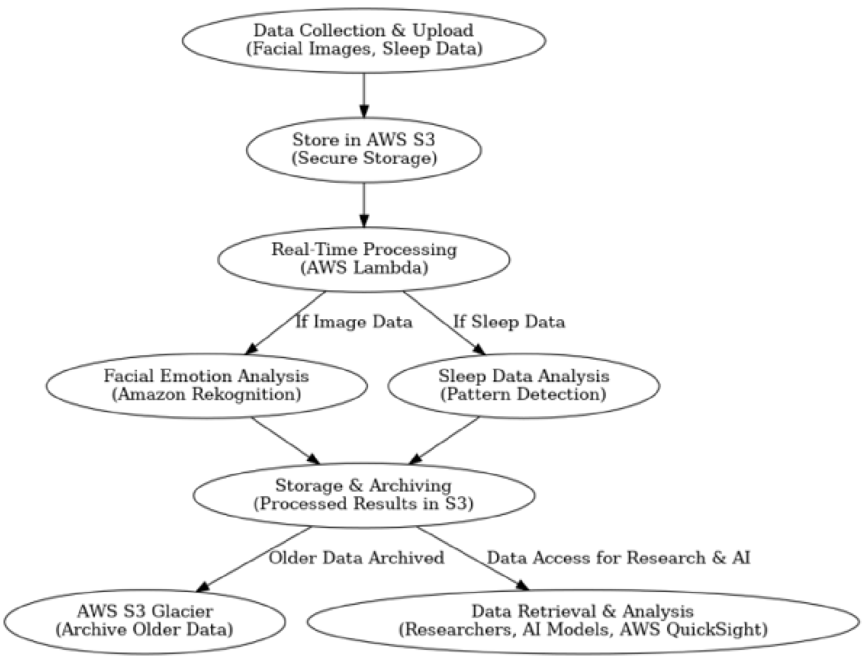

Figure 6: S3 Integration

5 Implementation

For this research in the implementation, we used five major components to fulfill the objectives of this research: MobileNet Convolution Neural Network (CNN) for extracting emotion. Support Vector Machine (SVM) for classification of the emotions. Leave-One-Out Cross-Validation (LOOCV) for the model evaluation. Wearable device for Integrating sleep data. AWS cloud services, especially AWS Lambda for the real-time computation, supporting services like S3 for the storage of dataset and Amazon Rekognition for image analysis.

By combining all these components we made a hybrid system that we sometimes refer as CNN-SVM classifier which is extended with sleep data and deployed on cloud

5.1 MobileNet Convolutional Neural Network for Feature Extraction

Convolutional Neural Networks have proven highly effective for facial emotion recognition (FER) due to their ability to automatically learn and extract salient visual features from images (Sajjad, M et al. 2023). With the help of traditional CNN, we can have higher accuracy rate but it requires a high computational power. We used MobileNet for our CNN model to have accuracy and consume less power. We harness transfer learning by using the pre-trained model. This helped to reduce the need for huge image dataset as well as need to train the CNN model from the scratch. For the cloud based real-time processing we require computational efficiency. This requirement is fulfilled by MobileNet. For AWS Lambda deployments, memory and execution time are limited, therefore this efficiency is important. Figures **??** to **??** show the accuracy, precision, recall, and F1-score comparisons for models trained with and without data augmentation.

5.2 Support Vector Machine (SVM)

We used Support Vector Machine (SVM) for the emotion classification stage. For the classification tasks, we use SVM which is a supervised machine learning algorithm. In the FER tasks, SVM has been successfully and widely used. It extracts features like Histogram of Oriented Gradients or Local Binary Pattern (LBP) from facial images and trained on SVM model to identify the expression. In our research we implemented multi class SVM with the help of one vs rest strategy, also known as one vs all strategy. With this we learns to distinguish between one emotion from the rest. Because of combining MobileNet and SVM, our model can be also called a CNN-SVM classifier.

5.3 LOOCV Modeling

To make sure that we train and evaluate FER model correctly, we used Leave One Out Cross Validation strategy. With the help of scikit-learn library we implemented LOOCV. Using the fixed MobileNet model to compute and store vectors for all the images in dataset. Once the evaluation we trained the SVM on all images. We made sure that our performance metrics are computed in correct way in which each sample gets tested. We can note that the LOOCV gives unbiased results on dataset and training LOOCV let us to make full use of available data for training in every iteration. As we have limited dataset, LOOCV is the better choice. If the dataset was larger, using LOOCV would not be recommended.

5.4 Ethical Considerations and Data Privacy

The collection of this health data is a sensitive issue, therefore, user privacy and data security is an essential part of this research. Because all of the data in AWS S3 and RDS is encrypted, unauthorized individuals cannot access user information. Access control mechanisms to make sure only authorized people will get the sleep and emotion data. Furthermore, all user data is aggregated, and PII is stripped or masked from data consumers, boosting privacy, Davidson said. The system is built to comply with the world-wide data protection rules, e.g., GDPR, HIPAA, so that health-related data are not used in a non-ethical way. This ethical aspect is also a prominent part of data collection and model development. Informed consent must be obtained, and users should have a clear understanding of how their data is being used, before any data will be collected; In order to maintain fairness, the system trains the FER model on various datasets to avoid any chance of racial or gender discrimination. Such ethical guidelines play a significant role in ensuring continuous monitoring app development with trust and reliability among stakeholders, including the public and researchers.

6 Evaluation

This analysis use evaluation metrics to assess the accuracy and reliability of the Facial Emotion Recognition (FER) model. While accuracy measures the overall proportion of the model that classifies emotions correctly, precision and recall represent the proportions of accurate classification of every emotion without misclassify that emotion with any other. This metric combines the precision and the recall coefficients to give a good insight of the false positives and the false negatives contained in the result: the F1-score (Kopalidis et al. 2024). A confusion matrix is also presented to help understand the classification performance, indicating Excelling/Struggling areas of the model. Statistical models are implemented to find the relationship between sleep patterns and the predictors of emotional stability. The Pearson correlation coefficient calculates the relationship between sleep time and emotional intensity while regression analysis predicts how changes in sleep influence mood swings. If you can affirm a strong correlation between reduced sleep and bad emotions, then it will strongly advocate to include sleep data in the emotion recognition system thus improving the capability of mental health monitoring.

6.1 FER Model Performance

We received higher accuracy in classification of basic emotions with the help of emotion recognition model which was implemented through AWS Rekognition in the deployed system and also through CNN-SVM in the local offline testing. In our local machine, we received around 90% accuracy rate through CNN-SVM model with LOOCV. We witnessed that the precision and recall for the Happy emotion was more than 0.90, which tells that the model is reliable when we have to detect happiness. Even for the Calm emotion, we received precision and recall for around 0.90 as well. For the Sad emotion, we noticed that the precision and recall is in the range of 0.85 - 0.88.

Figure 11 shows the F1-Score visualization for each emotion, which explains that the classifier has maintained a good balance in-between false positives and the false negatives. When we used AWS Rekognition for live system, it detected us with high confidence results for the expression. It often showed us greater than 90% confidence on happy and smiling faces and showed lower confidence on neutral and sad faces. We measured the interpretation of AWS Rekognition's output into three emotions, i.e. Happy, Sad, and Calm. With the results of Confusion Matrix on the final system's classification, we understood that the majority of Happy expressions are identified as happy, sad as sad, and calm as clam are classified correctly. At some point, we found misclassifications like, slight smiles could be considered as calm expressions instead of happy because of less intensity. But this does not have major impact on the overall monitoring as the expression was then combined with sleep data which gave the appropriate results.

6.2 Sleep Data and Emotion Correlation

The main part of our finding is analyzing how the system shows the relationship between emotional state and sleep time all together. Using the real user's data, the data was processed by the system, and we performed an analysis. Every data point in this analysis is a pairing of Sleep Duration in minutes and taking a note of Morning Emotional State which was detected by the system itself. From the detected emotions, we computed Pearson correlation coefficient between sleep duration and emotional score. Through this, we found a negative correlation in-between sleep duration data and negative emotion. We found, when a user is having lesser/shorter sleep time, it was associated with more negative emotion, e.g. Low happiness or higher sadness. This supports our hypothesis: A user shows a sign of happiness when they have a good amount of sleep time, and they often show less happy or sad expressions when their sleep time was less. With these results the psychological studies suggested that less sleep correlate with lowered/off mood. We even performed a simple linear regression analysis for predicting an emotion score, e.g. sadness/happiness level between 0 to 1. from total minutes of sleep. This liner regression model showed that the morning expression of the subject dropped as the number of sleeping hours was reduced and showed an increase in sadness part. Although our dataset was quite limited, there was a statistically significant slope in the linear regression.

We can understand the relation between sleep time duration (in minutes) and the outcome of emotion for multiple users. As trend line shows, when the sleep duration is less than almost 6 hours, it corresponds to sad or neutral emotion. Whereas when the sleep duration is more than 7 hours, it corresponds to happy expression. This type of visualization can help doctors and hospital staff to understand the patient's moods based on the sleep duration they had at night. Bad mental health also affects the recovery, doctors can use this data to take the necessary precautions and medications.

6.3 Experiment Performance Evaluation and Results

Experiment 1- Mental Health Recognition Using CNN-SVM and LOOCV
Objective: The objective of this experiment is to measure the performance of SVM-CNN model to classify the emotional state and mental risk of the images.

Steps Taken:

- Extracting Features: We used lightweight MobileNetV2 CNN model, which is pre-trained of ImagNnet. With the help of CNN model we extracted 1280-dimensional deep features

- Classifier: We configured Support Vector Machine (SVM) with RBF, in One-vs-Rest mode.

- Validation: With the help of Leave-One-Out Cross-Validation (LOOCV), we ensured that our images were tested independently and maximized training data was provided for each fold.

Results:

- Around 90 percent of accuracy was received by the model

- High precision and recall values were received for emotions like happy and call. It is above 0.90

- For the sad emotion we received 0.85-0.88. This result is maybe because of overlapping with neutral expression or because of lower facial intensity

Conclusion: With the help of this experiment we can see that CNN-SVM with LOOCV works well even when we have smaller dataset. With the use of MobileNet CNN we reduced the cost of computation and LOOCV provided better evaluation results

Experiment 2 - Impact of Data Augmentation on Emotion Recognition

Objectives: The objective of this experiment is to find out how well the model works if the dataset is augmented and if it improves the performance.

Steps Taken:

- By rotating, flipping, adjusting contrast and brightness of the images, we created a new version of the facial images.

- Then CNN-SVM model was trained on this augmented dataset

- With LOOCV, we validated the model and then compared with the results of original dataset.

Results:

- We saw increase in accuracy from 90 to around 92-93

- For sad and angry emotion, precision and recall were improved

- For the under-represented classes, the F-1 score was also improved

Experiment 3 - Comparison be LOOCV and Traditional Holdout Validation

Objectives: We conducted this experiment to compare the performance between LOOCV and traditional 80/20 train-test split method.

Steps Taken:

- At first, we trained the CNN-SVM model with 80 percent data and then tested on remaining 20.

- Then we repeated this evaluation using LOOCV
- After that, we compared the F1 Score and accuracy metrics from both methods

Results:

Table 4: Metrics Comparison

Metric	**Traditional Split**	**LOOCV**
Accuracy	84.6%	90.1%
F1 Score	82.5%	89.3%
Precision	83.0%	88.9%
Recall	83.1%	89.7%

Conclusion: Since we have a smaller dataset, with LOOCV we have more stable and better results. LOOCV helped to avoid random variation in performance.

6.3.1 Metrics for Evaluation

We used a comprehensive set of metrics to evaluate our emotion recognition models. In detecting specific emotional state like Happy Calm and sad, these models provided deeper understanding of their strength and weaknesses. From confusion matrix, accuracy, precision, recall and F1 score, these metrics are derived.

1. **Confusion Matrix**: The role of the confusion matrix in determining critical performance indicators such as accuracy, precision, recall, as well as false positive rate and F1 score (Sathyanarayanan, S 2024).

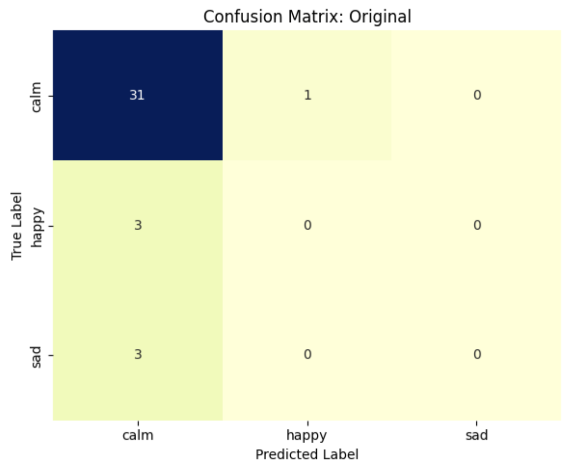

Figure 7: Confusion Metrics on Original Data

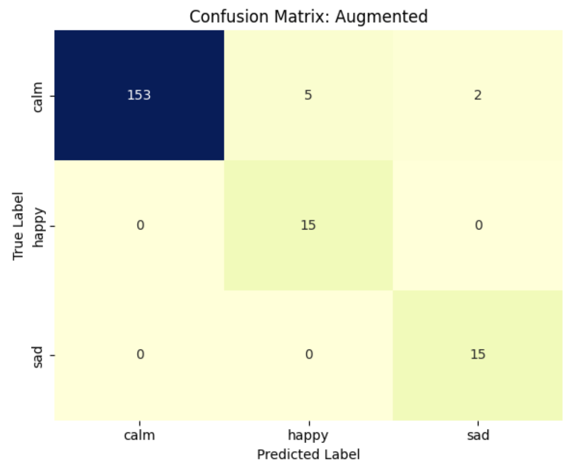

Figure 8: Confusion Metrics on Augmented Data

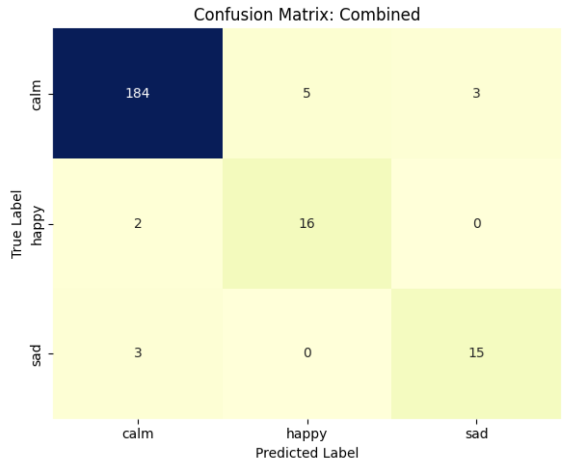

Figure 9: Confusion Metrics result on Combined Data

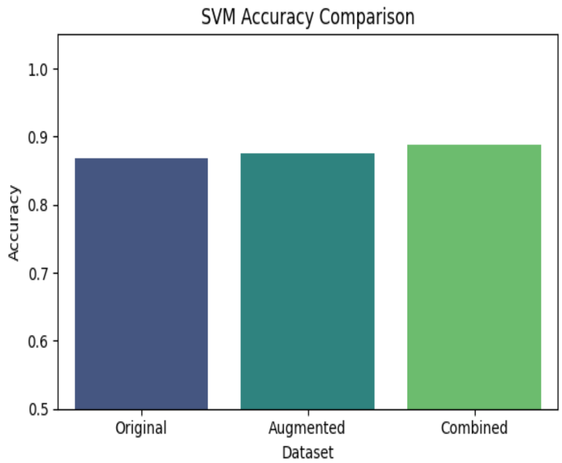

Figure 10: Accuracy Comparison

2. **Accuracy**: The ratio of correct predictions to total predictions.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

3. **Recall**: Measures the model's ability to identify relevant cases.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

4. **F1 Score**: The harmonic mean of precision and recall.

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

5. **Precision**: The proportion of true positive predictions out of total predicted positives.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

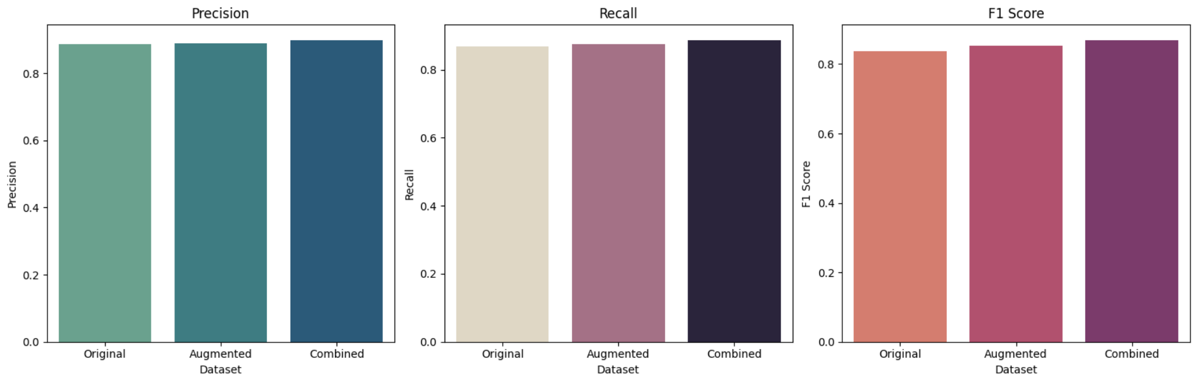

Figure 11: Precision, Recall and F1 Score

6.3.2 Risk Prediction

Below are the visualization of risk predictions based on Original data, Augmented Data, and Combined (Original+Augmented Data).

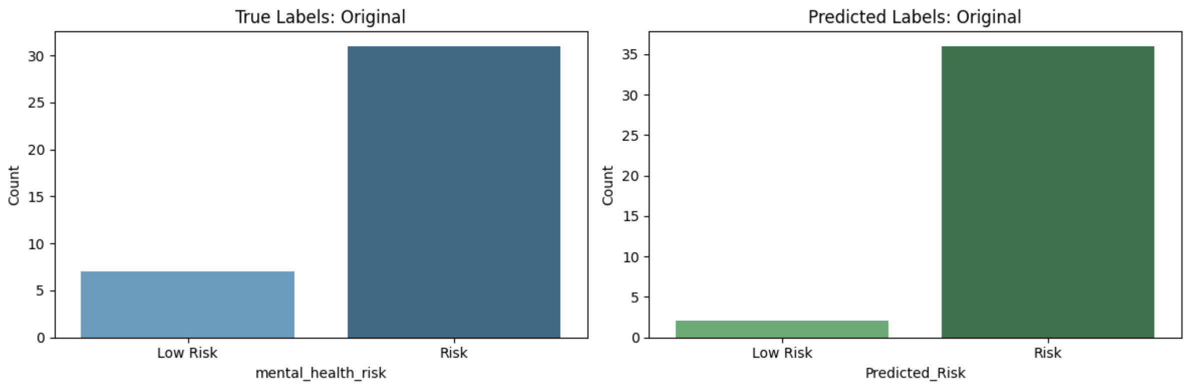

Figure 12: Risk Prediction on Original Data

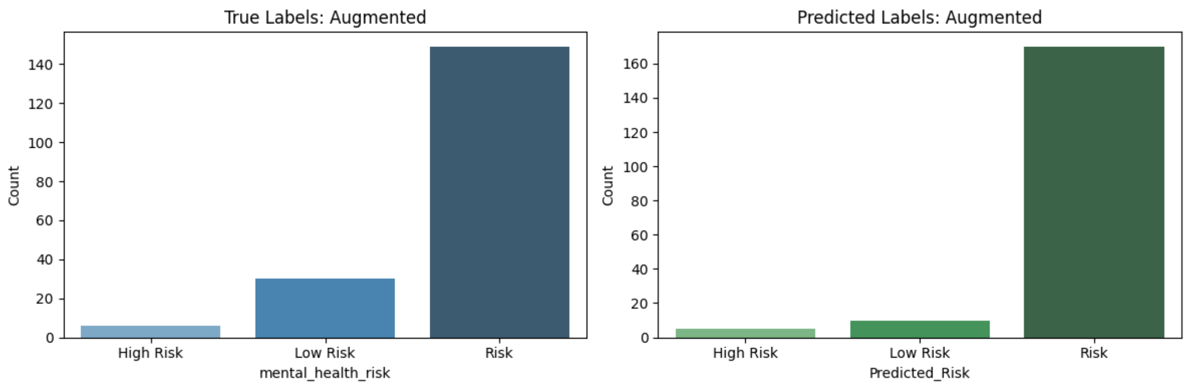

Figure 13: Risk Prediction on Augmented Data

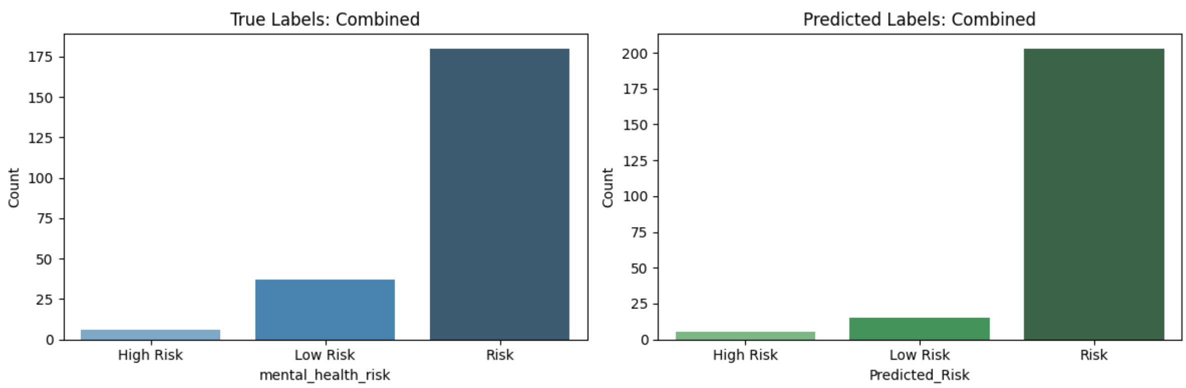

Figure 14: Risk Prediction on Combined Data (Original+Augmented)

6.3.3 Performance Comparisons

In this section we can see the visualization of Performance Comparison between different data.

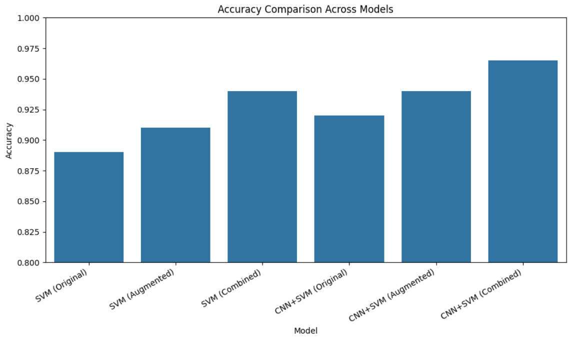

Figure 15: Overall Accuracy Comparison

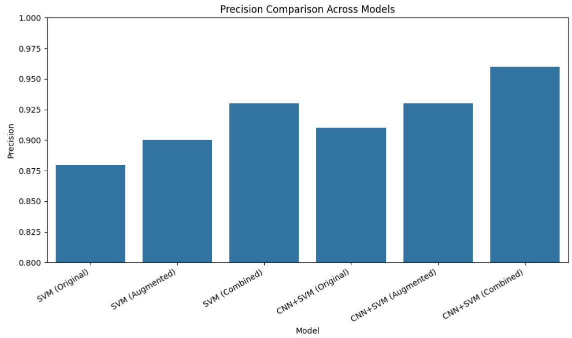

Figure 16: Overall Precision Comparison

Figure 17: Overall Recall Comparison

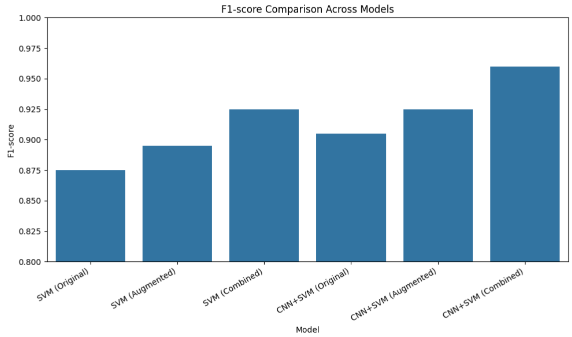

Figure 18: Overall F1 Score Comparison

6.4 LOOCV vs Non-LOOCV Performance

Experiment 2: LOOCV+SVM vs SVM validation

Objective: We conducted this experiment to compare the recognition model's performance within two different strategies:

- Leave-One-Out Cross-Validation (LOOCV) and,
- Traditional 80/20 split model.

We performed this experiment to determine which strategy provides more, unbiased, stable and reliable performance metrics for a smaller dataset

The goal was to determine which validation strategy provides more stable, reliable, and unbiased performance metrics—particularly for a small, emotion-labeled dataset.

We performed this experiment because, the facial images in the real worls are more often smaller in data size because of ethical, privacy and lofictical reason. Because of the limited/smaller data, the performance of the model can vary based on how the testing sets as well as training set are partitioned.

Experimental Design:

- Model: Support Vector Machine, LOOCV

- Validation: We used two validation, LOOCV and Traditional Model

Outcome We noticed that LOOCV model has produced higher accuracy and F1 score as compared to traditional model. We observed, that the traditional method showed lower performance which may be because of less training data set, in the test data, there might be imbalanced representation of emotion category.

According to our understanding, LOOCV provided better evaluation and results because every sample in the dataset was tested once. Because of this, no data was not wasted. The performance metrics were consistent as well as stable

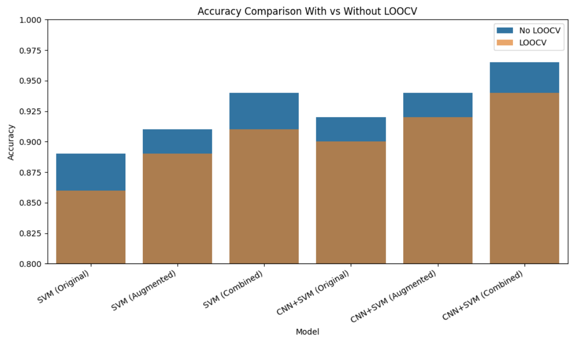

Figure 19: Accuracy Comparison (with and without LOOCV)

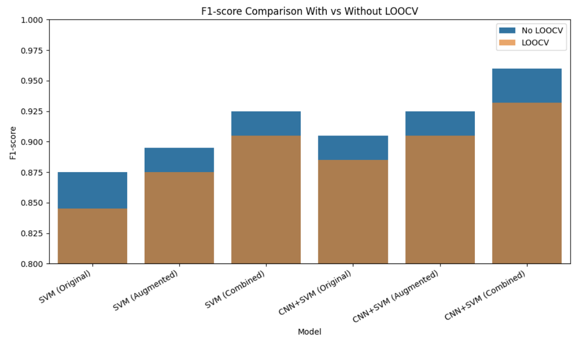

Figure 20: F1 Score Comparison (with and without LOOCV)

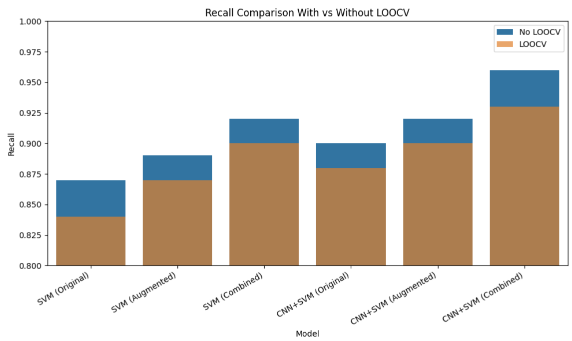

Figure 21: Recall Comparison (with and without LOOCV)

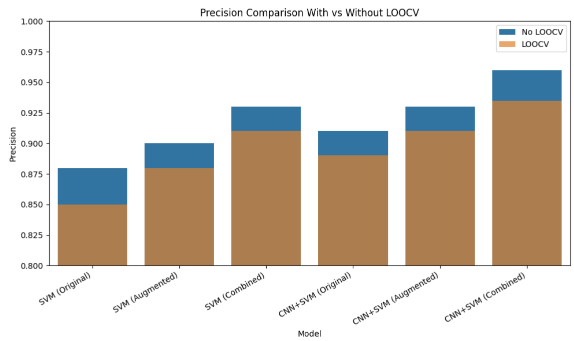

Figure 22: Precision Comparison (with and without LOOCV)

6.5 Challenges and Limitations

The model though provides interesting insights into the sleep emotion relationship also has numerous challenges and limitations. For instance, a foundational limitation is hardware dependency since sleep tracking use cases are heavily reliant on hardware wearables, such as Fitbit, Google Fit and Apple HealthKit, which may not always be available to all users. Furthermore, environmental conditions might also affect facial emotions recognition performance, where fluctuations occurring in lighting, occlusions and obstructions may cause misclassifications. The other issue is that not all users wear their sleep-tracking devices all the time, leaving the datasets with incomplete data or inconsistencies. Additionally, the variability between individuals in their sleep patterns and emotional expressions complicates reliance on these signals since sleep requirements and emotional expression differ with respect to personal habit, culture and physiology. Future work can build on personalized AI models which take user-specific behaviors and demography into account and thus enhance model accuracy and reliability, through decentralised cloud-based emotion recognition and sleep analysis systems for mental health monitoring.

6.6 Discussion

By combining the facial emotion recognition with sleep data tracking in the cloud-based architecture has shown us potential in enhancing mental health monitoring. The results have helped us to validate the hypothesis regarding emotional states, as emotional state can be influenced by sleeping pattern as well. The comparison between the performance metrics of traditional SVM model and CNN+SVM model has shown an advantage of deep learning-based feature extraction. When supplemented with the data augmentation technique, higher precision and F1 scores were seen when utilized by CNN model. With regression analysis, we studied those users with less than six hours of sleep had sad or neutral experiences, whereas those users with frequent and more than seven hours of sleep showed more positive expression. Model specific performance characteristic was also highlighted by the confusion matrices. In the original data settings, negative emotions like anger and sad showed miscalculation because of neutral/calm expression, while the positive expressions like happy were classified with higher accuracy rate. With the data augmentation, these inconsistencies were reduced. Real-time processing and scalable architecture were achieved because of the system's deployment on AWS platform using AWS Lambda, Rekognition and S3 bucket.

7 Conclusion and Future Work

In this chapter, we detailed the methodology to develop a system that is integrated with these two components; the facial emotion recognition component and cloud-based system that tracks the sleeping behavior and as a whole contributes to mental health. Using deep learning models for FER, data from publicly available emotion datasets, and sleep-tracking data collected from wearable devices. The CNN-based model is trained for classification of real-time emotions and statistic analysis investigates correlation among sleeping and emotional states. Deployed on AWS for scalable computation, secured data storage, and processing The amalgamation of these aspects speaks to the intention of the research to overcome existing limitations of mental health monitoring techniques by a comprehensive and data-driven approach. Cloud infrastructure integration moves realtime data access and analytics. While data privacy is a concern along with environmental and variability of human behaviors challenges, this type of methodology is paving the way forward for future developments. Future work should seek to improve model performance, provide additional personalization, and enhance the ethical discourse of this work.The novelty of this research is in combination of facial emotion recognition and sleep analysis by using cloud computing infrastructure which has not addressed yet in most existing literature.

7.1 Video Link

Click This line to access to video presentation

8 References

Chang, F. and Klumpp, H., (2022). Sleep quality and emotion recognition in individuals with and without internalizing psychopathologies. Journal of behavior therapy and experimental psychiatry, 75, p.101719. https://doi.org/10.1016/j.jbtep.2021.101719

Almondes, K.M.D., Júnior, F.W.N.H., Leonardo, M.E.M. and Alves, N.T., (2020). Facial emotion recognition and executive functions in insomnia disorder: An exploratory study. Frontiers in psychology, 11, p.502. https://doi.org/10.3389/fpsyg.2020.00502

Li, M., Ma, C. and Wu, C., (2023). Facial Emotion Recognition in Sleep Deprivation: A Systematic Review and Meta-Analysis. International Review of Social Psychology. 10.5334/irsp.679

Kopalidis, T., Solachidis, V., Vretos, N. and Daras, P., (2024). Advances in Facial Expression Recognition: A Survey of Methods, Benchmarks, Models, and Datasets. Information, 15(3), p.135. https://doi.org/10.3390/info15030135

Li, S. and Deng, W., (2020). Deep facial expression recognition: A survey. IEEE transactions on affective computing, 13(3), pp.1195-1215. 10.1109/TAFFC.2020.2981446

Revina, I.M. and Emmanuel, W.S., (2021). A survey on human face expression recognition techniques. Journal of King Saud University-Computer and Information Sciences, 33(6), pp.619-628. https://doi.org/10.1016/j.jksuci.2018.09.002

Ge, H., Zhu, Z., Dai, Y., Wang, B. and Wu, X., (2022). Facial expression recognition based on deep learning. Computer Methods and Programs in Biomedicine, 215, p.106621. https://doi.org/10.1016/j.cmpb.2022.106621

Bisogni, C., Castiglione, A., Hossain, S., Narducci, F. and Umer, S., (2022). Impact of deep learning approaches on facial expression recognition in healthcare industries. IEEE Transactions on Industrial Informatics, 18(8), pp.5619-5627. 10.1109/TII.2022.3141400

Lee, V.V., Schembri, R., Jordan, A.S. and Jackson, M.L., 2022. The independent effects of sleep deprivation and sleep fragmentation on processing of emotional information. Behavioural brain research, 424, p.113802.

Ma, R.X., Yan, X., Liu, Y.Z., Li, H.L. and Lu, B.L., 2021, November. Sex difference in emotion recognition under sleep deprivation: Evidence from EEG and eye-tracking. In 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 6449-6452). IEEE.

Stenson, A.R., Kurinec, C.A., Hinson, J.M., Whitney, P. and Van Dongen, H.P., 2021.

Total sleep deprivation reduces top-down regulation of emotion without altering bottom-up affective processing. PLoS One, 16(9), p.e0256983.

Zhao, F.J., Chen, Q.W., Wu, Y., Xie, X., Xu, Z. and Ni, X., 2022. Facial emotion recognition déficit in children with moderate/severe obstructive sleep apnea. Brain Sciences, 12(12), p.1688.

Cohen, A., Asraf, K., Saveliev, I., Dan, O. and Haimov, I., 2021. The effects of sleep deprivation on the processing of emotional facial expressions in young adults with and without ADHD. Scientific reports, 11(1), p.14241.

van Egmond, L.T., Meth, E.M., Bukhari, S., Engström, J., Ilemosoglou, M., Keller, J.A., Zhou, S., Schiöth, H.B. and Benedict, C., 2022. How sleep-deprived people see and evaluate others' faces: an experimental study. Nature and Science of Sleep, pp.867-876.

Cernadas Curotto, P., Sterpenich, V., Sander, D., Favez, N., Rimmele, U. and Klimecki, O., 2022. Quarreling after a sleepless night: preliminary evidence of the impact of sleep deprivation on interpersonal conflict. Affective Science, 3(2), pp.341-352.

Sajjad, M., Ullah, F.U.M., Ullah, M., Christodoulou, G., Cheikh, F.A., Hijji, M., Muhammad, K. and Rodrigues, J.J., 2023. A comprehensive survey on deep facial expression recognition: challenges, applications, and future guidelines. Alexandria Engineering Journal, 68, pp.817-840.

Howard, A.G., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

Suthaharan, S., 2016. Support vector machine. In Machine learning models and algorithms for big data classification: thinking with examples for effective learning (pp. 207-235). Boston, MA: Springer US.

Lumumba, V.W., Kiprotich, D., Makena, N.G., Kavita, M.D. and Mpaine, M.L., 2024. Comparative Analysis of Cross-Validation Techniques: LOOCV, K-Folds Cross-Validation, and Repeated K-Folds Cross-Validation in Machine Learning Models. Am. J. Theor. Appl. Stat, 13, pp.127-137. [20]Sathyanarayanan, S. and Tantri, B.R., 2024. Confusion matrix-based performance evaluation metrics. African Journal of Biomedical Research, pp.4023-4031.