JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# Mental Health Monitoring Using Facial Recognition

[1]Aakash Kishor Kale, [2]Aditya Salish, [3]Siddhesh Navnath Sabale, [4]Siddharth Anand Sancheti,

[5]Shubhangi Chavan

[1]Department of Computer Engineering,
[1]Pillai College of Engineering, New Panvel, Mumbai, Maharashtra, India

***Abstract :*** Artificial intelligence and machine learning have the power to revolutionize the field of mental health recognition. By analyzing large sets of data. Al-based systems can detect patterns and make predictions that can help diagnose and treat mental health issues. Al and machine learning used in mental health recognition is a great success. For example, machine learning algorithms have been used to predict suicide risk by analyzing social media data, and natural language processing techniques have been used to analyze speech patterns and detect depression. In addition to recognizing mental health issues. Al can also be used to develop more effective treatment options. By analyzing large sets of data, Al can identify patterns in patient behavior and tailor treatment plans to individual needs. This could potentially reduce the need for trial and-e for approaches and lead to more successful treatment outcomes. However, it's important to note that Al is not a replacement for human expertise and care. While Al can assist healthcare professionals in making more informed decisions, it should not be relied upon as the sole basis for diagnosis or treatment. Additionally, ethical considerations must be taken into account when developing and deploying AI systems in the mental health domain.

## I. INTRODUCTION

Artificial Intelligence is a technique for building systems that mimic human behaviour or decision-making. Machine Learning is a subset of AI that uses data to solve tasks. These solvers are trained models of data that learn based on the information provided to them. ML algorithms use our data to learn and automatically solve predictive tasks. AI consists of many subfields that use techniques to mimic specific behaviours we associate with natural human intelligence. Other subfields are building intelligent systems that replicate human behaviours, such as Robotics, Computer Vision and Pattern Recognition.

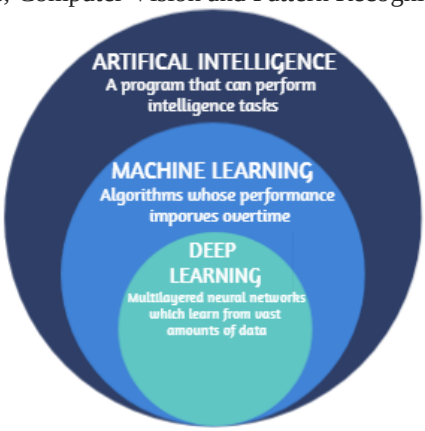


**Fig 1: The figure shows inter-relation between Artificial intelligence, Machine Learning and Deep Learning**.

The main objective of the implemented system is to aid in the early detection and diagnosis of mental health disorders using Facial recognition technology that has the potential to identify subtle changes in facial expressions that may indicate the presence of a mental health disorder. The goal of the system is to provide a standardized approach and also improve the accuracy and reliability of mental health detection. Overall, the objective of using facial recognition technology in mental health detection is to improve the identification and treatment of mental health disorders, ultimately leading to better mental health outcomes for individuals. The current scope of the project focuses on building core competencies in computer vision, deep learning, and GUI development to deliver real-time facial emotion classification. It detects faces in video streams using a pre trained model, crops the faces, and feeds them into a CNN to categorize one of 7 basic emotions. The results are displayed live through bounding boxes and text overlays on the video. Some application logic is implemented to log the emotion sequence over time and provide analysis and personalized song recommendations based on the dominant mood. However, there is substantial room to expand the sophistication and applicability of the system. Some areas for future work include expanding the emotion categories recognized, training the models on larger and more diverse facial expression datasets, optimizing and quantizing models for edge deployment, packaging the application for non-technical end users, enhancing the GUI and visualizations, customizing for specific real-world use cases in mental health, education, automotive safety, etc. Additional sensory modalities like voice, gestures and posture could also be integrated for richer emotion analysis. In summary, the current project delivers a proof-of-concept while future work can significantly enhance the real-world viability and impact across multiple domains.

## II. CRITICAL ANALYSIS

In this chapter the relevant techniques in literature are reviewed. It describes various techniques used in the work. The techniques in this category are adapted to the individual needs, interests and preferences of the user or society. Various techniques in this category are listed here. These techniques have various advantages and are used extensively in literature.

| SN | Title of the Literature paper | Proposed system | Advantages and Disadvantages |
|---|---|---|---|
| 1. | "A Survey on Wearable Sensors for Mental Health Monitoring" by Nuno Gomes, Matilde Pato, André Ribeiro Lourenço, and Nuno Datia, 2023 [1] | The proposed system uses wearable sensors for monitoring mental health. It aims to track physiological and behavioral parameters, such as heart rate, breathing patterns, sleep quality, physical activity, and social interactions. | Advantages: Wearable sensors provide continuous and real-time monitoring, which can be more effective than periodic assessments or self-reports. Disadvantages: There are concerns about privacy and security of the data collected by wearable sensors, particularly with regard to sensitive health information. |
| 2. | "Diagnosis of Depressive Disorder Model on Facial Expression Based on Fast RCNN", National Library of Medicine, Young-Shin Lee, 2022 [2] | The system proposed uses a deep learning method that recognizes vector-based information to assist in the diagnosis of depressive disorder that can be devised by checking the position change of the eyes and lips based on pictorial data. | Advantages: The system uses real time emotional evaluation using active participation from the user. Disadvantages: The disadvantage that comes with the system is hardware failure and unavailability of users to take pictures for the diagnosis. |
| 3. | "Detection of Types of Mental Illness through the Social Network Using Ensembled Deep Learning Model", Hindawi Computational Intelligence and Neuroscience, by Syed Nasrullah. 2022 [3] | The proposed system uses a deep learning model based on ensemble learning techniques to classify mental illnesses. | Advantages: The system uses CNN and RNN which gives it more accuracy and precision giving it advantage over existing systems. Disadvantages: The system however does not give real time analysis of patients' data which is crucial for mental health monitoring |
| 4. | "Prediction of Mental Health using Machine Learning", JETIR, by Dr.J.Arokia Renjit, Adlin Sajeesha M.J, Sangavai V.D, Sree Devi D.S, 2022 [4] | The proposed system uses machine learning algorithms to collect and pre-process datasets to map data into graphs to check quality of data and predict the outcome with higher accuracy. | Advantages: The system gives higher accuracy results in case of public tests and is able to predict mental illness accordingly. Disadvantage: This system also lacks the real time analysis of the patient's condition and is limited to specific dataset. |
| 5. | "Mental Health Analysis in Social Media Posts: A Survey", Archives of Computational Methods in Engineering, by Muskan Garg, 2022 [5] | The proposed survey suggests a method using multimedia features extracted from social media and posts to classify potential diseases into different categories using classification algorithms. | Advantages: The system predicts mental health, especially suicidal ideation which helps to take appropriate measures in such cases. Disadvantages: This system also lacks in the real time analysis of the patient's condition as the data extracted can be from older posts. |
| 6. | "Mental Health Prediction Using Deep Learning" by U. Sairam, Santhosh Voruganti, 2022 [6] | The proposed system uses genetic algorithm, classification and machine learning techniques to build semi-automated system to identify patient's mental illness | Advantages: The advantage this system gives is day-to-day feedback and guidance for nurturing employee's health and the system uses algorithm like MLP and SVM to give more accurate results Disadvantages: The main disadvantage is that the system does |

| | | | |
|---|---|---|---|
| | | | not follow a real time approach to get the status of mental health. |
| 7. | "Assessment of Anxiety, Depression, and Stress using Machine Learning models", Elsevier Publications by Prince Kumar, Shruti Garg, Ashwani Garg, 2020 [7] | This system focused on the prediction of five severity levels of anxiety, depression and stress using eight different machine learning models. | Advantages: The hybrid approach that the system approaches improve the accuracy of the outcomes. Neural networks perform better than all other methods. Disadvantages: Whereas the hybrid approach came with a disadvantage of having computational time exceptionally higher. |

I. Summary of Literature Survey

## III. Implemented System

Facial recognition technology has been used to aid in the recognition of mental health issues. This technology can detect facial features, such as expressions, and analyze them to identify patterns and signs of mental health issues. For example, it can detect changes in mood, such as sadness or anxiety, and track them over time. This data can then be used to personalize treatment plans and monitor progress. However, there are concerns about the accuracy of facial recognition technology and the potential for it to be misused. It is important to ensure that these systems are developed and deployed ethically, with appropriate safeguards in place to protect patients' privacy and rights.

### *A. Existing System*

The existing system implements deep learning algorithm that uses R-CNN for the extraction of data from facial images, and a region proposition network (RPN) is learned from them. The system involves participation from the user or the patient itself by asking the user to take pictures regularly to track the changes in eyebrow, eye and lip movements for the diagnosis of the depressive disorder. The system also uses KakaoTalk chatbot to perform deep learning that can determine the picture. The system then classifies these images based on the features extracted from them and runs a diagnosis of depressive disorder to detect Mental Health of the user.


**Fig 2: Existing System Architecture**

The existing system implements deep learning algorithm that uses R-CNN for the extraction of data from facial images, and a region

### *B. Implemented System*

The implemented system uses various methods such as facial recognition, surveys by using Artificial intelligence and machine learning algorithms like SVM, NLP and Random Forest Algorithm. The system uses R-CNN algorithm through which a feature map is extracted by sequentially applying convolution, linear rectification, and max pooling processes from facial data provided as input. The system will then recognize the emotions based on the extracted features and then classify as and track them. The system will have a continuous feed of video data to track the condition of the user in real time through which the data can be analyzed with the data stating depressing situations. The system will then detect potential mental health issues the user is facing and provide therapy such as recommendations of songs that help to relax the user.

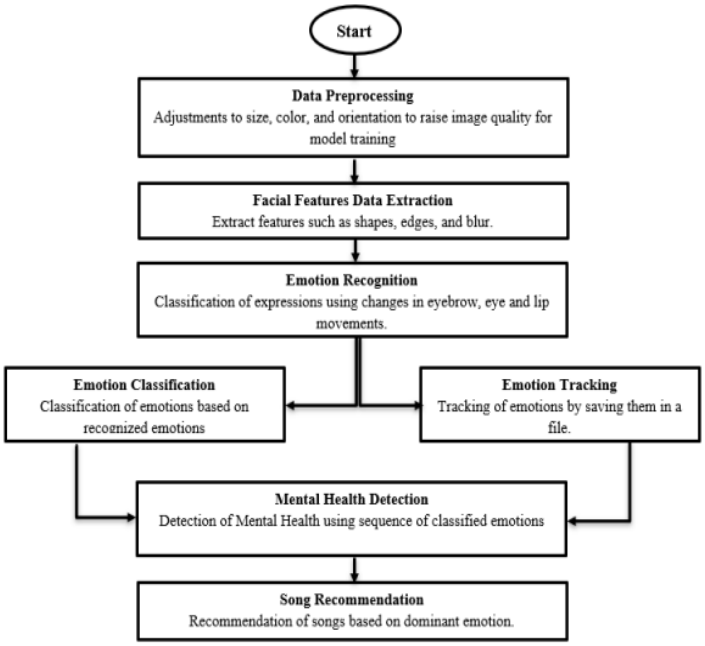


**Fig 3: Implemented System Architecture.**

R-CNN (Region based Convolutional Neural Networks) is a popular deep learning model that can be used for object detection and segmentation. Here's a possible implementation method for mental health recognition using facial recognition with R-CNN. Collect a dataset of facial images with labels indicating whether the person has a mental health issue or not. Preprocess the dataset by resizing the images, normalizing the pixel values, and augmenting the data (e.g. flipping, rotating, adding noise). Train an R-CNN model on the preprocessed dataset using a deep learning framework like TensorFlow, OpenCV or PyTorch. Fine-tune the model by adjusting the hyperparameters, selecting the appropriate loss function, and optimizing the learning rate. Evaluate the model on a separate test set by calculating metrics like precision, recall, and F1 score. 12. Deploy the model in a real-world setting by integrating it into a mental health screening system or a mobile app that can analyze facial images in real-time. Continuously monitor and validate the performance of the model to ensure its accuracy and reliability. It's important to note that this implementation method is just a general guideline, and there are many variations and modifications that can be made depending on the specific use case and requirements. Additionally, ethical considerations must be taken into account when developing and deploying facial recognition systems for mental health recognition.

For Face Detection it uses a pre trained RFB-320 face detector model from Caffe framework Implements helper functions like generate_priors(), convert_locations_to_boxes() etc to interpret the model outputs. It detects faces in each video frame and extracts face regions. For Emotion Classification. It Leverages a pre trained ONNX model file for emotion prediction. It resizes the cropped face to 64x64 pixels as input. The Model has 7 output nodes corresponding to the emotion classes. It uses OpenCV DNN module to run inference on the ONNX model. The GUI is built using Tkinter for Python, it contains buttons to start video, view analysis, and exit. Click handlers defined to trigger different functions Analysis. Logs sequence of emotions predicted over time in a text file. It Calculates frequency distribution of emotions. It Matches most frequent emotion to related content. It uses Spotify library to retrieve song recommendations. It Queries different premade playlists based on detected emotions.

### *C. Methodology and Techniques*

Collecting data for this type of system would involve gathering a diverse set of images and videos of individuals with different mental health conditions. This data could be obtained through partnerships with mental health clinics or through online platforms that allow individuals to share their stories and experiences. The data should be properly anonymized to protect individuals' privacy. The next step involves using machine learning algorithms to identify facial expressions and other biometric data associated with different mental health conditions. This process can involve a range of techniques, including facial landmark detection, emotion recognition, and micro-expression analysis. The resulting data can be used to create a feature vector that represents the individual's mental health state. Once the features have been extracted, they can be used to train a deep learning model that can recognize patterns and classify individuals into different mental health conditions. This model could be a Region based convolutional neural network. (R-CNN) or a recurrent neural network (RNN) that has been trained on a large dataset of annotated images and videos. With a trained model in place, the system can begin analyzing real-time facial expressions of individuals through a camera or other input device. The system can then use the trained model to identify potential mental health conditions in the individual. This could involve a threshold-based approach where the system makes a decision based on a specific level of confidence or a continuous approach where the system provides a probability distribution over the different mental health conditions. Based on the identified mental health condition, the system can recommend songs or playlists that are known to have a positive impact on mental health. This recommendation can be generatzd using natural language processing and sentiment analysis of lyrics, as well as other musical features such as tempo, rhythm, and instrumentation. The system could also consider the individual's preferences and past listening behavior to personalize the recommendations. Facial recognition is a promising technique for mental health recognition. Some different techniques for mental health recognition using facial recognition include the following Facial expression analysis involves analyzing facial expressions to identify emotions, such as happiness, sadness, or anxiety, that are associated with various mental health conditions. Certain mental health conditions, such as depression or anxiety, can affect skin color. By analyzing changes in

skin color, facial recognition technology can detect these conditions. • Eye tracking involves tracking eye movements to identify patterns associated with mental health conditions, such as social anxiety or ADHD. Thermal imaging can detect changes in body temperature associated with mental health conditions, such as depression or anxiety. • 3D facial recognition involves analyzing the dimensions of a person's face to detect subtle changes that may be associated with mental health conditions.

*D. Applications*

Following could be some applications of the implemented system:

- The emotion analysis over time could be useful for monitoring mental health conditions like depression, bipolar disorder, etc. Patients and clinicians can track emotional patterns.
- The system could be applied in marketing research to gauge customer reactions and sentiment from facial expressions in surveys, interviews, etc.
- It could be used as an educational tool to teach students about emotions, mental health, psychology, machine learning, etc.
- The real-time emotion recognition could enable interactive art installations that respond to the viewer's facial expressions.
- The personalized song recommendations based on detected emotions could be expanded into recommender systems for various products and services.
- The system could enable new modes of emotional and expressive interaction with computers and devices.
- Detecting driver emotion and fatigue could help improve safety in autonomous vehicles.

## IV. Result And Discussion

Here are some datasets commonly used in facial recognition with their types and number of items:

- FER-2013: The data consists of 48x48 pixel gray scale images of faces. The faces have been automatically registered so that the face is more or less centered and occupies about the same amount of space in each image. The task is to categorize each face based on the emotion shown in the facial expression into one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral). The training set consists of 28,709 examples and the public test set consists of 3,589 examples.
- Labeled Faces in the Wild (LFW): A public benchmark for face verification, also used for face recognition. Consists of more than 13,000 images of faces collected from the web.
- CelebA: A large scale face attributes dataset with more than 200,000 celebrity images, each with 40 attribute annotations.
- YouTube Faces: A dataset of face videos that contains over 3,000 identities over 10,000 videos.
- MORPH: A longitudinal face dataset with over 55,000 images of 13,000 individuals.
- AffectNet: A database of facial expressions, valence and arousal annotations. Contains over 1 million facial images.

There are several datasets used for mental health recognition, including AffectNet, FER2013, and EmoReact. AffectNet contains over one million facial images with emotion labels and has been used for emotion recognition res earch. FER2013 is a dataset of over 35,000 facial images labeled with one of seven emotions: anger, disgust, fear, happiness, sadness, surprise or neutral. EmoReact is a more recent dataset that contains over 100,000 videos and has been used for emotion recognition and affective computing research. These datasets are used in conjunction with machine learning algorithms to recognize and predict mental health conditions.

Evaluating a facial recognition model for mental health detection is a complex task, as it involves sensitive and nuanced aspects of mental health. It's essential to consider both technical and ethical aspects of model evaluation. Here are some key evaluation metrics and considerations for such a model:

1. Accuracy: Accuracy is a fundamental metric that measures the proportion of correctly classified cases. In the context of mental health detection, this would be the model's ability to correctly identify individuals with mental health issues
2. Sensitivity (True Positive Rate): Sensitivity measures the proportion of true positive cases correctly identified by the model. In the mental health context, this metric assesses the model's ability to detect individuals with mental health issues when they actually have them.
3. Specificity (True Negative Rate): Specificity measures the proportion of true negative cases correctly identified by the model. In this context, it assesses the model's ability to correctly identify individuals without mental health issues as such.
4. Precision: Precision measures the proportion of correctly identified positive cases out of all the positive cases identified by the model. In mental health detection, it assesses how many of those identified as having mental health issues actually have them.
5. F1 Score: The F1 score is the harmonic mean of precision and sensitivity, providing a balanced measure of a model's performance. It's especially useful when there's an imbalance in the dataset.
6. Area Under the Receiver Operating Characteristic (ROCAUC): ROC-AUC measures the model's ability to distinguish between individuals with and without mental health issues. It is particularly important when you're dealing with imbalanced datasets.
7. Mean Squared Error (MSE): If the model outputs a continuous mental health score, MSE can be used to measure the discrepancy between predicted and actual mental health scores
8. Root Mean Squared Error (RMSE): Similar to MSE, RMSE is appropriate when the model predicts continuous mental health scores. It provides a more interpretable error metric compared to MSE. 19
9. Confusion Matrix: A confusion matrix provides a detailed breakdown of true positives, true negatives, false positives, and false negatives, offering insights into the model's performance across different classes.

10. Ethical Considerations: Beyond technical metrics, it's crucial to evaluate the model's ethical implications. Ensure that the model does not introduce bias, discriminate against any group, or violate privacy and consent norms. Assess the potential for harm and the interpretability of the model's decisions.
11. Cross-Validation: Use cross-validation techniques to assess the model's generalization to unseen data, which is essential to ensure its real-world applicability.
12. Human Evaluation: Incorporate expert evaluations or user feedback to validate the model's predictions and gauge its real-world performance.
13. Interpretable Features: Consider using techniques that provide insight into which facial features are driving the model's decisions. This can help in understanding the model's decision-making process.

Based on these factors, we can start the process of emotion analysis. The following steps must be followed:

1. The user is required to start the application.
2. On clicking start the camera will start, and the users are to then record their face for a desired amount of time. After the recording they are to press Q on the keyboard
3. On clicking Analysis they will be shown a pop up card based on their detected emotion.

**Fig 4: Input Sample.**

**Fig 5a: Output Sample.**


**Fig 5b: Output Sample.**

## V. CONCLUSION

The project implements a facial emotion recognition system using computer vision and deep learning techniques. It captures video from a webcam or file and detects faces using a pretrained RFB320 face detection model. For each detected face, the system crops the face region from the frame, resizes it to 64x64 pixels, and feeds it into a pre trained ONNX deep learning model to predict one of 7 emotions - neutral, happiness, surprise, sadness, anger, disgust, or fear. The predicted emotion and bounding box are displayed on

the face in the video stream. The sequence of predicted emotions over time is logged to a text file. The project includes a GUI with Start, Analysis, and Exit buttons. The Analysis screen shows statistics on the frequency of each emotion and suggests related mental health conditions and treatments based on the most frequent emotion. It also retrieves a song recommendation from Spotify using the Spotipy library and displays it along with the health information in a formatted webview.

## REFERENCES

**[1]** Nuno Gomes, Matilde Pato, André Ribeiro Lourenço, and Nuno Datia, "A Survey on Wearable Sensors for Mental Health Monitoring", 2023

**[2]** Young-Shin Lee, "Diagnosis of Depressive Disorder Model on Facial Expression Based on Fast R-CNN ", National Library of Medicine, 2022

**[3]** Syed Nasrullah, "Detection of Types of Mental Illness through the Social Network Using Ensembled Deep Learning Model", Hindawi Computational Intelligence and Neuroscience, 2022

**[4]** Dr.J.Arokia Renjit, Adlin Sajeesha M.J, Sangavai V.D, Sree Devi D.S, "Prediction of Mental Health using Machine Learning", JETIR, 2022

**[5]** Muskan Garg, "Mental Health Analysis in Social Media Posts: A Survey", Archives of Computational Methods in Engineering, 2022

**[6]** U. Sairam, Santhosh Voruganti, "Mental Health Prediction Using Deep Learning", 2022

**[7]** Prince Kumar, Shruti Garg, Ashwani Garg, "Assessment of Anxiety, Depression, and Stress using Machine Learning models", Elsevier Publications, 2020

**[8]** Jung Yuchae, Yong Ik. Multimedia Tools and Applications, 76 (9) (2020).

**[9]** Norizam, Sulaiman. Determination and classification of human stress index using the nonparametric analysis of EEG signals. Diss. UniversitiTeknologi MARA, 2020.

**[10]** Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms (2020).

**[11]** Mental Stress Level Prediction and Classification based on Machine Learning (2020).

**[12]** M Srividya, S Mohanavalli, and N Bhalaji. Behavioral modeling for mental health using machine learning algorithms. Journal of medical systems, 2018.

**[13]** Verslus A., Verkuil B., Spinhoven P., Ploeg M.M., Brosschot J.F. Changing Mental Health and Positive Psychological WellBeing Using Ecological Momentary Interventions. A Systematic Review and Meta-analysis. J. Med. Internet Res. 2016 [14] Enrique Garcia-Ceja, "Mental health monitoring with multimodal sensing and machine learning: A survey", Elsevier publications, 2018