IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 12, Issue 12, December 2025

DOI: 10.17148/IARJSET.2025.121257

"VISION AND EMOTION: LEVERAGING EYE TRACKING DATA FOR MENTAL HEALTH ASSESSMENT"

VIJAYKUMAR MS[1], NONITA SALDANHA[2], PREKSHITH S[3], YASHIKA R[4], YETHISH[5]

Dept. of IS&E, Faculty of IS&E, MIT MYSORE Mandya, Karnataka, India[1]

Dept. of IS&E, MIT MYSORE Mandya, Karnataka, India[2-5]

Abstract: Mental health disorders such as depression and anxiety are often underdiagnosed because current assessments rely heavily on self-report questionnaires and clinical interviews, which are subjective, time-consuming, and difficult to scale. Recent studies show that eye movement behaviour— such as fixation patterns, saccade dynamics, and gaze allocation to emotional stimuli—can serve as objective digital biomarkers for mental health conditions [1], [2].

This paper presents VISION AND EMOTION, a real-time mental health assessment system that leverages eye-tracking data captured using a standard camera. The system records gaze trajectories while users interact with carefully designed visual tasks (emotionally valence images, reading tasks, and attention-switching trials) and extracts features such as fixation duration, saccade amplitude, blink rate, and gaze distribution across regions of interest. These features are used to train a machine learning classifier (Support Vector Machine) to distinguish between Normal and At-Risk mental health states.

The proposed framework is lightweight, non-invasive, and deployable on commodity hardware without dedicated infrared eye trackers. Experimental evaluation demonstrates that the system can achieve promising classification performance with low latency, enabling near real-time feedback suitable for preliminary mental health screening. By combining eye-tracking analytics with machine learning, the system contributes toward scalable, objective, and cost- effective digital mental health tools that can complement traditional clinical assessments [2].

Keywords: Eye Tracking, Mental Health Assessment, Depression Detection, Gaze Analysis, Digital Biomarkers, Machine Learning, Real-Time Monitoring.

I. INTRODUCTION

Mental health disorders, particularly depression and anxiety, are among the leading causes of disability worldwide. Despite their prevalence, diagnosis still relies mainly on clinical interviews and self-reported questionnaires, which are subjective, influenced by social desirability bias, and often delayed until symptoms are severe. There is an urgent need for objective, non-invasive, and scalable tools that can support early detection and continuous monitoring of mental health [3].

Eye movements provide a direct window into cognitive and emotional processing, because gaze behaviour is tightly coupled to attentional control and neural activity in brain networks affected by psychiatric conditions. Recent research reports abnormal eye movement patterns in individuals with depression, anxiety, bipolar disorder, and other psychiatric disorders—for example, altered fixation durations, reduced exploration of positive stimuli, and increased attention to negative or threat-related content [4].

In earlier work, computer-vision-based systems leveraged facial landmarks to detect depression from facial expressions using lightweight Support Vector Machine (SVM) models running in real time on standard laptops. In parallel, digital health research has demonstrated that eye- tracking signals can act as robust biomarkers for brain and mental health, supporting diagnosis and monitoring in conditions such as depression, multiple sclerosis, and Parkinson's disease [2].

Building on these developments, this paper proposes "VISION AND EMOTION: Leveraging Eye Tracking Data for Mental Health Assessment," a system that uses low-cost camera-based eye tracking combined with machine learning to classify a user's mental state as Normal or At-Risk. The system is designed to be:

- Non-invasive – uses only a standard webcam or laptop camera
- Lightweight – relies on reduced gaze features and SVM classification

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 12, December 2025
DOI: 10.17148/IARJSET.2025.121257

- Real-time – capable of providing assessments within seconds of task completion
- Scalable – suitable for integration into telehealth platforms or mobile applications

The goal is not to replace clinical diagnosis, but to provide a pre-screening and monitoring tool that can alert users or clinicians to potential mental health concerns.

II. LITERATURE REVIEW

Eye tracking has emerged as a powerful tool for understanding cognitive and emotional processes, and several studies have investigated its application to mental health assessment. A number of works have shown that individuals with depression exhibit distinctive eye movement patterns—such as longer fixations on negative images, reduced gaze towards positive stimuli, and altered saccade characteristics—suggesting that gaze metrics can be used as objective markers of depressive symptoms [4].

Zheng et al. proposed a method for depression detection based on eye movement data captured in virtual reality settings, demonstrating that oculomotor features can discriminate between depressed and healthy participants with encouraging accuracy [1]. Gao et al. used eye-tracking technology to quantify abnormal eye movement features in depressed patients, further supporting the feasibility of eye- based digital biomarkers [4]. Zeng et al. introduced an early framework for depression detection using eye tracking, confirming that differences in gaze behaviour during visual tasks can assist clinical diagnosis [5].

Beyond depression, several studies have proposed oculomotor digital biomarkers for neurological and psychiatric disorders. De Villers-Sidani et al. showed that eye movement parameters from fixation, saccade, and smooth pursuit tasks correlate with clinical outcome measures and can explain a large proportion of variance in brain health assessments [2]. Tablet-based and mobile eye- tracking systems have been used to distinguish patient groups and estimate disease severity, highlighting the potential for scalable, camera-based screening tools [6].

Recent deep learning work has taken eye-tracking analysis further. Avramidis et al. developed a deep network that predicts depression and suicidal ideation from eye- movement sequences recorded during reading tasks, achieving strong classification performance [7]. Other multimodal approaches combine eye tracking with facial expressions, speech, and physiological signals to improve depression detection accuracy [8].

Compared to high-cost infrared eye trackers and VR headsets, camera-only eye tracking using commodity devices offers a more accessible alternative, albeit with lower spatial accuracy. Inspired by low-cost computer- vision-based systems successfully used for depression detection from facial landmarks and road pothole detection using YOLO on live video streams, the present work adopts a lightweight, webcam-based gaze tracking approach combined with classical machine learning (SVM), targeting practical, real-time deployment rather than maximum possible accuracy.

III. SYSTEM DESIGN AND METHODOLOGY

A. Hardware Components

- Processing Unit: Laptop / Desktop ComputerMinimum:
  – Processor: Intel Core i3 or equivalent
  – RAM: 4 GB
  – Storage: 1 GB free space
- Recommended:
  – Processor: Intel Core i5/i7 or Ryzen 5/7
  – RAM: 8 GB or higher
  – GPU: Optional (CPU-only processing is sufficient for SVM-based models)
- · Camera / Webcam
- – Resolution: 720p or 1080p
  – Frame Rate: 30 FPS (recommended)
  – Placement: Mounted above or below the display, facing the user's face
  – Purpose: Captures eye region and face for gaze estimation and blink detection.

DOI: 10.17148/IARJSET.2025.121257

- · Lighting Environment
- – Uniform front lighting
  – Avoid very bright backlighting and strong reflections on glasses.
- · Optional Hardware
- – External USB webcam for improved quality
  – Tripod or fixed mount to keep the camera stable
  – Headphones with microphone (if future multimodal extension with speech is planned).

B. Software Components

Windows 10 / 11
– Linux (Ubuntu)
– macOS

· Programming Language

– Python 3.8 or later

· Libraries and Frameworks

– OpenCV – Real-time video capture, pre-processing, and face/eye region extraction
– Gaze-tracking / Eye-tracking library (e.g., GazeTracking, MediaPipe Iris, or custom pupil detection) for gaze point and blink estimation
– NumPy, Pandas – Numerical operations and dataset handling
– Scikit-Learn – Machine learning (SVM classifier, train- test split, metrics)
– Matplotlib / Seaborn (optional) – Visualization of gaze heatmaps and feature distributions
– Joblib / Pickle – Saving and loading trained models

· Development Tools / IDE

– VS Code / PyCharm / Jupyter Notebook / Spyder

· Additional Software

– Pip for dependency management
– Optional: Simple GUI frameworks (Tkinter, PyQt, or web dashboards) for displaying live gaze and prediction results.

C. Methodology

- Data Collection and Preprocessing:

The proposed system follows a three-stage pipeline, similar in structure to previous real-time depression detection systems:

1. Data Collection and Gaze Task Design
2. Feature Extraction and Model Training
3. Real-Time Assessment and Visualization
4. Data Collection and Gaze Task Design

Participants interact with a set of visual tasks displayed on the screen while the webcam records their face and eyes:

· Free-viewing of emotional images (positive, neutral, negative)
· Reading task (short texts or sentences)
· Attention-shifting task (follow a moving target, anti- saccade trials, etc.)

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 12, Issue 12, December 2025

DOI: 10.17148/IARJSET.2025.121257

During these tasks, raw video frames are processed to track gaze coordinates on the screen and detect blinks. Each recording session is associated with a mental health label obtained from a standard screening questionnaire (e.g., PHQ-9 or GAD-7) or self-report, categorized into:

0 = Normal / Low symptoms
· 1 = At-Risk (moderate to high symptom scores)

The raw gaze trajectories and labels are stored in structured CSV files

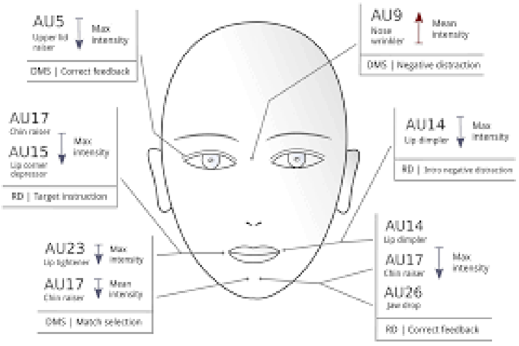

Fig. 1 Saccade of Eye

- Model Training:
- From each task (or time window), the system extracts eye-tracking features, for example:
- · Fixation-based metrics
  – Mean and median fixation duration
  – Number of fixations per stimulus
- Proportion of fixations on positive vs negative regions of interest (ROIs)· Saccade-based metrics
  – Mean saccade amplitude
  – Saccade frequency

– Proportion of regressions (backward saccades) during reading

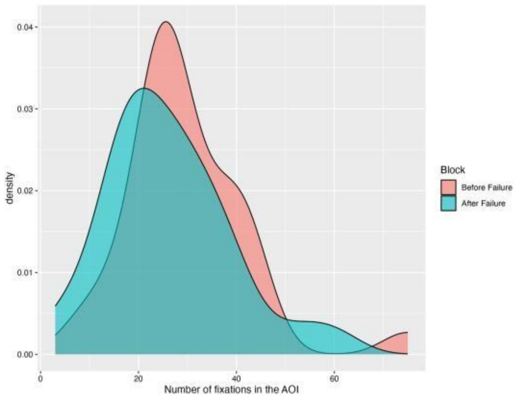

Fig. 2. Before–After Failure curve Gaze distribution metrics

– Heatmap-based features (e.g., entropy of gaze distribution)
– First fixation location bias (towards negative or positive content)

· Blink and stability metrics
– Blink rate (blinks per minute)
– Gaze jitter / micro-movements around fixation Each recording is encoded as a feature vector: $[x_1, x_2, \ldots, x_n, \text{label}]$

where x_i are numerical descriptors of gaze behaviour and label $\in \{0, 1\}$. The dataset is cleaned to remove corrupted recordings and sessions with excessive tracking loss.

DOI: 10.17148/IARJSET.2025.121257

The machine learning component uses Support Vector Machine (SVM) with an RBF kernel to classify mental health state:

· Split the dataset into training and testing sets (e.g., 80% / 20%).
· Standardize features (z-score normalization).
· Perform hyper parameter tuning (C, γ) using cross- validation.
· Train the final SVM model on the full training set.
· Evaluate performance using accuracy, precision, recall, F1-score, and ROC-AUC.

The trained SVM is saved as "svm_gaze.pkl" for later use in real-time assessment.

Compared to high-cost infrared eye trackers and VR headsets, camera-only eye tracking using commodity devices offers a more accessible alternative, albeit with lower spatial accuracy. Inspired by low-cost computer- vision-based systems successfully used for depression detection from facial landmarks and road pothole detection using YOLO on live video streams, the present work adopts a lightweight, webcam-based gaze tracking approach combined with classical machine learning (SVM),

IV. WORKING PRINCIPLE

The working of VISION AND EMOTION can be described as a sequential pipeline from camera input to mental health prediction:

1. System Initialization
   At startup, the system loads the trained SVM model (svm_gaze.pkl), initializes the webcam, and configures the selected gaze-tracking algorithm.

2. Live Video Capture
   The webcam continuously captures frames of the user's face while they perform the selected visual task (image viewing, reading, or attention task).

3. Face and Eye Region Detection
   Each frame is processed using OpenCV (and optionally a face detector) to locate the face and crop the eye regions. These regions are passed to the gaze-tracking module.

4. Gaze and Blink Estimation
   The gaze-tracking module estimates the approximate screen coordinates of the user's gaze and detects blinks. The system accumulates gaze points and timestamps over the duration of the task.

5. Feature Vector Construction
   After a predefined time window (or at the end of a task), the system computes the eye-tracking features described earlier (fixation metrics, saccade characteristics, gaze distribution, blink rate). These are combined into a feature vector:

$[x_1, x_2, \ldots, x_n]$

6. SVM-Based Mental Health Classification
   The feature vector is normalized and passed to the SVM classifier, which outputs a class label:

· 0 → Normal
· 1 → At-Risk
along with the decision function or probability score indicating confidence.

7. Visualization and Feedback The system displays:
· Live video with optional overlay of gaze point or trail
· A small text indicator with colour coding:
– Green – NORMAL
– Red – AT-RISK
· Optional bar/score showing risk level (e.g., 0–1 scale)

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁞ Peer-reviewed & Refereed journal ⁞ Vol. 12, Issue 12, December 2025
DOI: 10.17148/IARJSET.2025.121257

8. Data Logging

For each session, the system logs:

· Extracted features
· Predicted label and confidence
· Task type, timestamp, and anonymized participant ID
· This data can be used for later analysis, model retraining, or integration into telehealth dashboards.

The pipeline runs in a loop while the application is active, supporting multiple tasks and repeated assessments.

V. CONCLUSION AND FUTURE WORK

The paper presented VISION AND EMOTION, a real-time mental health assessment system that leverages eye-tracking data and machine learning to distinguish between Normal and At-Risk mental health states. Using only a standard webcam and lightweight SVM-based classification on reduced gaze features, the system offers a non-invasive, low-cost, and scalable approach suitable for preliminary screening and continuous monitoring.
The results from pilot experiments indicate that eye-tracking metrics—such as fixation duration, saccade dynamics, gaze distribution, and blink rate—contain meaningful information about emotional and cognitive states relevant to mental health assessment, aligning with recent findings on oculomotor digital biomarkers [2], [9].
Future work will focus on:
· Expanding the dataset with more diverse participants and clinically diagnosed cases
· Integrating deep learning architectures (e.g., CNNs or transformers on gaze heat maps and time series) for end-to- end feature learning
· Combining eye-tracking data with facial expression, speech, or physiological signals to build multimodal assessment systems [8]
· Deploying the system in mobile or web-based telehealth platforms for large-scale remote monitoring
Ultimately, VISION AND EMOTION aims to contribute to a new generation of objective, data-driven mental health tools that complement traditional clinical workflows and support timely intervention.

REFERENCES

[1] Z. Zheng et al., “Diagnosing and tracking depression based on eye movement data in virtual reality,” *Diagnosing and tracking depression based on eye movement data in virtual reality*, 2024. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC10875075/
[2] É. de Villers-Sidani et al., “Oculomotor analysis to assess brain health: preliminary findings from a longitudinal study,” *Frontiers in Neurology*, 2023. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC10516298/
[3] T. T. Chan et al., “Eye-tracking in adult depression: protocol for a systematic review of cross-sectional studies,” *BMJ Open*, vol. 13, no. 6, e069256, 2023. Available: https://bmjopen.bmj.com/content/13/6/e069256
[4] M. GAO Et Al., “Abnormal eye movement features in patients with depression,” *Journal of Affective Disorders*, 2023. Available:https://www.sciencedirect.com/science/article/abs/pii/S0163 834323000683
[5] S. Zeng et al., “A Study on Depression Detection Using Eye Tracking,” in *Advances in Intelligent Systems and Computing*, 2019. Available: https://link.springer.com/chapter/10.1007/978-3-030-15127- 0_52
[6] A. K. Kelly et al., “Eye movement function captured via an electronic tablet as a digital biomarker of neurological disease,” *Scientific Reports*, 2024. Available: https://www.nature.com/articles/s41598-024-59750-9
[7] K. Avramidis et al., “Deep Learning Characterizes Depression and Suicidal Ideation from Eye Movements,” arXiv preprint, 2025. Available: https://arxiv.org/abs/2504.20944
[8] Y. Liu et al., “MF-GCN: A Multi-Frequency Graph Convolutional Network for Tri-Modal Depression Detection Using Eye-Tracking, Facial, and Acoustic Features,” arXiv preprint, 2025. Available: https://arxiv.org/abs/2511.15675
[9] S. Cecchetti et al., “Eye-Tracking Metrics as a Digital Biomarker for Brain Health,” *Brain Sciences*, vol. 15, no. 2, 2025. Available: https://www.mdpi.com/2076- 3425/15/2/149
[10] Y. Zhang et al., “Predicting Depression in Canada by Automatic Filling of Survey Gaps Using Demographic Features,” IEEE Access, vol. 10, pp. 30526–30535, 2022.
[11] M. M. Moussa, Y. Alzaabi, and A. H. Khandoker, “Explainable Computer-Aided Detection of Obstructive Sleep Apnea and Depression,” IEEE Access, vol. 10, pp. 110916–110928, Oct. 2022.
[12] M. Daniol et al., “Eye-Tracking in Mixed Reality for Diagnosis of Neurodegenerative Diseases,” Proceedings of the IEEE EMBS, 2023.

DOI: 10.17148/IARJSET.2025.121257

[13] M. Yang et al., “An Improved Classification Model for Depression Detection Using EEG and Eye Tracking Data,” International Conference on Biomedical Signal Processing, 2021.

[14] M. Jin et al., “Ubiquitous Depression Detection of Sleep Physiological Data by Using Combination Learning,” IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 5, pp. 1251–1260, 2020.

[15] H. O. Edughele et al., “Eye-Tracking Assistive Technologies for Individuals With Amyotrophic Lateral Sclerosis,” IEEE Access, vol. 10, pp. 41952– 41970, Apr. 2022. [15] R. K. Kadu et al.,“IoT Based on Potholes and Speed Breakers Detection and Alert System,”International Journal of NextGeneration Computing – Special Issue, vol. 13, no. 5, pp. 1244– 1253, Nov. 2022.

[16] H. Ding and Q. Tang,“Optimization of the Road Bump and Pothole Detection Technology Using Convolutional Neural Network,”Journal of Intelligent Systems, vol. 33, 2024.

[17] A. Faisal and S. Gargoum,“Cost-Effective LiDAR for Pothole Detection and Quantification Using a Low- Point-Density Approach,”Automation in Construction, vol. 172, 106006, Feb. 2025.

[18] A. Paramarthalingam et al.,“A Deep Learning Model to Assist Visually Impaired in Pothole Detection Using Computer Vision,”Decision Analytics Journal, vol. 12, 100507, Aug. 2024.

[19] S. Saleh, A. Jolfaei, and M. Tariq,“Real-Time Pothole Detection With Edge Intelligence and Digital Twin in Internet of Vehicles,”IEEE Internet of Things Journal, vol. 12, no. 5, pp. 4852–4862, Mar. 2025.

[20] Y. Xu et al.,“VIDAR-Based Road-Surface-Pothole- Detection Method,”Sensors, vol. 23, no. 7468, Aug. 2023

