© April 2025 | IJIRT | Volume 11 Issue 11 | ISSN: 2349-6002

Facial Micro-Expression Analysis Using Eye Landmark Detection

Hemant Kumar Bhardwaj[1], Nitin Kumar[2], Suaib Saifi[3], Varun Tyagi[4], Aman Goswami[5]
[1]*Assistant Professor, R.D. Engineering College, Uttar Pradesh, India*
[2,3,4,5] *Department of Computer Science & Engineering, R.D. Engineering College, Uttar Pradesh, India*

***Abstract*—This paper addresses the challenge of facial micro-expression analysis in image sequences, with a focus on face landmark identification in the eye region. Various modern approaches for this task are examined, highlighting their key characteristics. The Spontaneous Actions and Micro-Movements (SAMM) dataset is selected for experiments due to its use of the Facial Action Coding System (FACS) for precise facial expression labelling. The study compares landmark detection using the widely known Active Shape Model (ASM) with a newly developed approach. The suggested approach enables efficient and accurate eye landmark detection with minimal computational resources. Key implementation stages of this method are outlined, demonstrating its advantages over ASM in terms of accuracy and simplicity.**

***Index Terms*—Active Shape Model (ASM), emotion recognition, eye detection, Facial Action Coding System (FACS), facial landmark detection, Micro-expressions, Spontaneous Actions and Micro-Movements (SAMM) dataset.**

I. INTRODUCTION

Facial expressions play a crucial role in human communication, conveying emotions across different cultures. Paul Ekman, a pioneer in the psychology of emotions, shown that facial expressions are universally recognized, regardless of cultural background [1]. Extensive Study in this domain have demonstrated that micro-expressions, which are involuntary and fleeting facial movements, cannot be consciously controlled or suppressed. Consequently, individuals trained in micro-expression recognition can gain authentic insights into a person's emotional state in real time. Since facial behavior is a key aspect of nonverbal communication [2], its analysis has been extensively applied to improve human-computer interaction (HCI) [3,4]. One of the most reliable methods for automatic facial analysis is smile detection. For instance, Daniel Modaff et al. developed an automated classification method that categorizes reactions to media content based on over 1,500 mimic responses collected online. Recent advancements in automated facial behavior analysis have also contributed to medical research, helping identify conditions such as depression and post-traumatic stress disorder (PTSD) [5]. Furthermore, progress in micro-expression analysis technology, sensor technology, and signal processing has led to applications in road safety systems, enabling real-time monitoring of driver fatigue and drowsiness [6,7]. The intersection of emotion recognition and learning has also gained attention across multiple disciplines, including psychology [8,9], education [10-12], neurobiology [13], and informatics [14-16]. The integration of computer vision and artificial intelligence (AI) has made it possible to develop interactive learning systems that assess a student's emotional state and tailor the teaching process accordingly.This paper focuses on the analysis of facial micro-expressions, which reveal concealed emotions and potential deception. Since micro-expressions are extremely subtle and last only a fraction of a second, detecting them with the naked eye is difficult. Therefore, computer vision applications for automatic micro-expression detection have gained significant interest. In recent years, machine learning and AI have enabled the recognition of psychological states and social signals through facial expression analysis. One of the earliest and most prominent applications of micro-expression detection is deception detection, which has been utilized in criminology, law enforcement, medicine, and cybersecurity. Other notable applications include business communication, border security, HCI, medical diagnosis, the automobile industry, education, and entertainment.

With the rapid development of computer vision technology, micro-expression detection Systems are increasingly optimizing their efficiency, reducing reliance on complex optical equipment while maintaining high accuracy. A fundamental step in micro-expression analysis is facial landmark detection, particularly in the eye region, it offers critical indicators for interpreting subtle emotional states. This task can be effectively performed using computer vision techniques, minimizing the need for specialized hardware and human intervention.

II. RELATED WORK

Detection of facial landmarks is a controlled or automated process that relies on manually annotated images for training. The process typically initiates with facial recognition, where an image is analyzed and a rectangular bounding box is generated using a face detection algorithm [17,18]. This bounding box is then used to initialize the positions of facial landmarks, which serve as important points of reference for further study.

Facial landmark detection techniques can be roughly divided into three primary methods:

1. Holistic methods
2. Constrained Local Model (CLM) methods
3. Regression-based methods [19]

Among the holistic approaches, the most well-known technique is the Active Appearance Model (AAM) [20], while the Active Shape Model (ASM) [21] is widely recognized as the leading CLM-based method. Both AAM and ASM are considered classic techniques for facial landmark detection. Comparative studies show that ASM offers several advantages over AAM, particularly in precise contour localization and robustness against illumination variations. Due to these strengths, applications needing high-precision facial shaping are better suited for ASM.

Since the development of the traditional ASM, several enhanced variations have emerged to improve accuracy, reliability, and computational efficiency. These methods typically predict facial landmark positions by training on manually annotated images. However, some alternative approaches detect facial landmarks independently, without requiring prior training. While these methods eliminate the need for initialization, they introduce the challenge of ambiguity, as multiple positions may closely resemble the target landmark, leading to detection inconsistencies.

With the rise of deep learning, numerous neural network-based approaches for facial landmark detection have been explored. Current deep learning models make use of large datasets to automatically learn facial structures, significantly improving detection accuracy and generalization. One example of such technology is OpenFace®, an advanced open-source software that detects facial landmarks, head poses, facial expressions, and gaze directions. However, implementing deep learning-based solutions often requires complex computational resources and significant training, which limits their use for light-duty tasks.

In contrast, this paper proposes a straightforward method that works for eye landmark detection, leveraging the geometric properties of the eye region rather than complex neural networks. Despite its simplicity, the suggested approach shows positive results and turns out to be a workable solution for detecting micro-movements and facial micro-expressions.

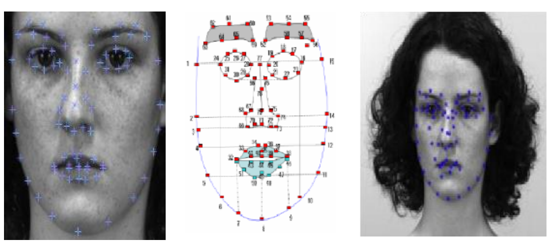
III. DATASET

For our experimental study on eye landmark detection, we selected the Spontaneous Actions and Micro-Movements (SAMM) dataset [22]. The SAMM dataset is among the most varied and high-frame-rate datasets, capturing micro-facial movements at 200 frames per second. Unlike many other datasets, SAMM does not classify facial movements based on emotional labels; instead, it utilizes the Facial Action Coding System (FACS) to categorize facial expressions based on muscle activity rather than subjective emotional interpretation. This approach provides an objective framework for analyzing micro-expressions without relying on assumptions about emotional states.

The FACS system annotates facial expressions by decomposing them into Action Units (AUs), each corresponding to a specific muscle movement. A trained FACS coder examines facial movements in video sequences and labels them accordingly. This method ensures that facial micro-movements are analyzed based on muscle behavior rather than inferred emotions, making it highly compatible with the SAMM dataset.

For experimental implementation, we used the MATLAB® environment, which provides specialized toolboxes and built-in functions for rapid prototyping and testing of landmark detection techniques. Initially, we employed the Active Shape Model (ASM) to detect facial landmarks, as it is designed to automatically identify structural points on facial features such as the eyes, nose, lips, mouth, and eyebrows.

The ASM learning stage involves generating a statistical face model from a training set containing manually annotated landmarks (Fig. 1a). The facial landmark structure used in our experiments consists of 79 key points (Fig. 1b). After training, the averaged facial model is aligned to a test image, and landmark positions are further refined. However, as shown in Fig. 1c, the ASM method sometimes produces inaccurate eye landmark placements, which do not precisely align with the expected structure. Given the importance of eye landmarks in micro-expression analysis, these inaccuracies highlight the limitations of ASM and motivate the need for a more precise and adaptable detection approach.



a b

c

Figure 1. The facial landmarks: a – the manually marked face; b – the layout of the facial landmarks; c – the result of the ASM method.

IV. EXPERIMENTAL RESULTS

A. Eye Landmark Detection Approach

Traditional facial landmark detection techniques often rely on holistic models, where landmarks are estimated based on the entire facial structure rather than focusing on specific facial features. While these approaches provide a general representation, they frequently result in positioning inaccuracies, particularly in localized facial regions such as the eyes. In contrast, neural network-based methods offer higher precision, but their effectiveness is highly dependent on the chosen architecture, requiring significant computational power for training and real-time execution.

To overcome these challenges, we adopted a Facial Action Coding System (FACS)-based approach, which provides a structured framework for micro-expression analysis. FACS defines 26 key facial regions, where subtle muscle movements occur, with the eye region playing a critical role in micro-expression detection. By leveraging FACS, we achieved more accurate and consistent eye landmark detection, independent of illumination changes, facial variations, and head movements.

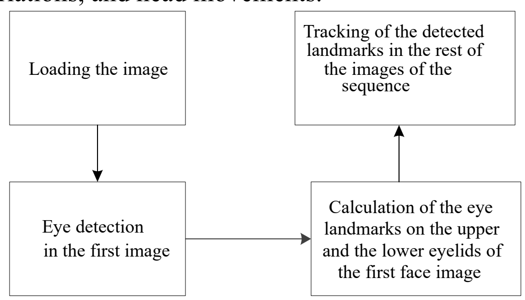


Figure 2. The main stages of the eye landmarks detection technology.

B. Proposed Eye Landmark Detection Methodology

Our proposed eye landmark detection system follows a multi-stage process, as illustrated in Fig. 2. The Viola-Jones algorithm, integrated into MATLAB's Computer Vision Toolbox, was chosen for initial face and eye detection due to its efficiency in real-time processing [17]. The system begins by loading an image sequence and applying a face detection algorithm to identify the facial region of interest (ROI). Once the face is detected, the eye region is extracted for further processing.

1) Eye Landmark Detection Steps

The detection process consists of the following stages:

1. Local Binary Pattern (LBP) Calculation – The LBP operator is utilized to analyze texture features within the eye region (Fig. 3a).
2. Outer and Inner Corner Detection – The horizontal extreme points with the highest intensity values are identified as the inner and outer eye corners (Fig. 3b).
3. Pupil Detection – The following steps make up the detecting procedure.:

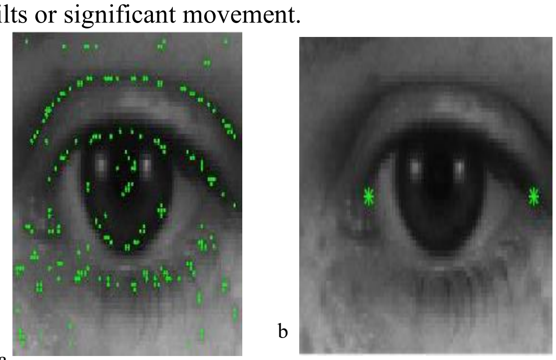
- A median filter is applied to remove noise.

- The grayscale image is converted to a binary image.
- The image is inverted to enhance contrast (Fig. 4a).
- Objects within the image are analyzed, and circular regions are detected to locate the pupil center (Fig. 4b).

4. Eye Center Calculation – The midpoint between the inner and outer eye corners is computed, and a perpendicular reference line is drawn through this point to improve accuracy (Fig. 5b).
5. Upper Eyelid Landmark Detection –
   - The highest intensity point above the pupil is identified as the upper eyelid peak (Fig. 6a).
   - Additional landmarks are positioned at 30°, 60°, 120°, and 150° angles relative to the vertical midline, ensuring uniform distribution across the upper eyelid contour (Fig. 6b).
6. Lower Eyelid Landmark Detection –
   - The lowest intensity point below the pupil is identified as the lower eyelid peak.
   - Similar angular divisions are applied to distribute lower eyelid landmarks symmetrically.
7. Landmark Mapping onto the Facial Image – Once detected, the eye landmarks are mapped onto the original facial image, ensuring precise alignment.

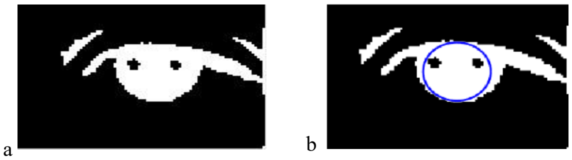
C. Tracking of Eye Landmarks Across Image Sequences

Once the landmarks are detected in the first frame, tracking is performed using the Kanade-Lucas-Tomasi (KLT) algorithm, which maintains landmark stability across subsequent frames. While cascade object detectors could be applied in each frame, they demand a lot of processing power and may fail in cases of head tilts or significant movement.



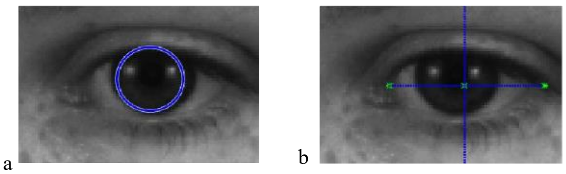
a b

Figure 3. The found points: a – the calculated points with the most intensities; b – the found corners of the eye.



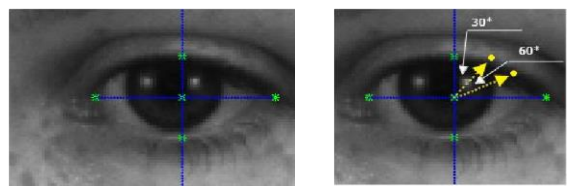
a b

Figure 4. The inverted binary image of the eye: a – after transformation into the binary image; b – with the circle describing the pupil of the eye



a b

Figure 5. The detection of the pupil and the center of the eye: a – the found pupil of the eye; b – the center of the eye.



a b

Figure 6. The detection of the landmarks of the eyelids: a – the upper and lower eyelid landmarks detection; b – the first sector landmarks detection.

D. Results of Eye Landmark Detection

The accuracy of our eye landmark detection system is demonstrated in Fig. 7a, where the detected landmarks are superimposed onto the facial image. The final processed output is shown in Fig. 7b, confirming the precision, stability, and efficiency of our proposed approach.

a b

Figure 7. The detection results: a – the detected landmarks on the upper and the lower eyelids of the right eye; b – the result of the application of the developed eye landmarks detection technology.

V. CONCLUSION

The proposed eye landmark detection technology offers greater precision in identifying upper and lower eyelid landmarks compared to the Active Shape Model (ASM) method. Unlike ASM, this approach eliminates the need for a learning phase and doesn't need pictures to be manually annotated for training, making it a more efficient and user-friendly solution. Additionally, the system is designed to function with grayscale images, further reducing computational complexity.

However, since the method relies on the Viola-Jones algorithm, it has limitations when the subject is not directly facing the camera, which may result in the non-detection of one eye. Future research will focus on expanding the dataset and exploring alternative detection techniques to enhance robustness and accuracy across a wider range of head positions and facial variations.

REFERENCES

[1] Paul Ekman, Emotion in the Human Face, 2nd Edition, Malor Books, 2013, 456 p.

[2] Fernando De la Torre and Jeffrey. F. Cohn, "Guide to Visual Analysis of Humans: Looking at People. Facial Expression Analysis," 2011, pp. 1-31.

[3] Zhihong Zeng, Maja Pantic, Glenn I. Roisman, and Thomas S. Huang, "A Survey of Affect Recognition Methods: Audio, Visual and Spontaneous Expressions," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, pp. 126-133.

[4] Marina Stewart Bartlett, Gwen Littlewort, Ian Fasel, and Javier R. Movellan, "Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction," Computer Vision and Pattern Recognition Workshop, 2003, 5 p.

[5] Hamdi Dibeklioglu, Zakia Hammal, Ying Yang, and Jeffrey F. Cohn, "Multimodal Detection of Depression in Clinical Interviews," ICMI '15: Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, 2015, pp. 307-310.

[6] Thomas A. Ranney, "Driver Distraction: A Review of the Current Stateof-Knowledge," National Highway Safety Administration, 2008, 25 p.

[7] Carlos Busso and Jinesh Jain, "Advances in Multimodal Tracking of Driver Distraction," A Novel Way to Start Speech Dialogs in Cars by Talk-and-Push (TAP), 2011, pp. 253-270.

[8] Charles S. Carver "Negative Affects Deriving from the Behavioral Approach System," Emotion (Washington, D.C.), 2004, pp. 3-22.

[9] Edward L. Deci and Richard M. Ryan, Handbook of Self-Determination Research, University Rochester Press, 2004, 480 p.

[10] Elizabeth A. Linnenbrink and Paul R. Pintrich, "Multiple Goals, Multiple Contexts: The Dynamic Interplay Between Personal Goals and Contextual Goal Stresses," in Simone Volet and Sanna Järvelä (Eds.), Advances in Learning and Instruction Series. Motivation in Learning Contexts: Theoretical Advances and Methodological Implications, Pergamon Press, 2001, pp. 251-269.

[11] Debra K. Meyer and Julianne C. Turner, "Discovering Emotion in Classroom Motivation Research," Educational Psychologist, vol. 37 (2), 2002, pp. 107-114.

[12] Debra K. Meyer and Julianne C. Turner, "Re-Conceptualizing Emotion and Motivation to Learn in Classroom Contexts," Educational Psychology Review, 18(4), 2006, pp. 377-390.

[13] Antonio Damasio, Descartes' Error: Emotion, Reason, and the Human Brain, Penguin Books; Reprint edition, 2005, pp. 336 p.

[14] Barry Kort, Rob Reilly, and Rosalind Picard, "An Affective Model of Interplay between Emotions and Learning: Reengineering Educational Pedagogy – Building a Learning Companion," International Conference on Advanced Learning Technologies (ICALT), IEEE Press, Madison, 2001, pp. 43-46.

[15] Jack Mostow and Gregory Aist, "Evaluating Tutors that Listen: An Overview of Project LISTEN," in Smart Machines in Education, MIT/AAAI Press, 2001, pp. 169-234.

[16] Gregory Aist, Barry Kort, Rob Reilly, Jack Mostow, and Rosalind Picard, "Experimentally Augmenting an Intelligent Tutoring System with Human-Supplied Capabilities: Adding Human-Provided Emotional Scaffolding to an Automated Reading Tutor that Listens," ICMI '02:

Proceedings of the 4th IEEE International Conference on Multimodal Interfaces, 2002, pp. 483-490.

[17] Paul A. Viola and Michael J. Jones, “Robust Real-Time Face Detection,” International Journal of Computer Vision, 2004, pp.137-154. [18] Anna D. Sergeeva, Alexander V. Savin, Victoria A. Sablina and Olga V. Melnik, “Emotion Recognition from Micro-Expressions: Search for the Face and Eyes,” 8th Mediterranean Conference on Embedded Computing (MECO) Proceedings. Budva, Montenegro, 2019, pp. 632-635.

[18] Yue Wu, Qiang Ji, “Facial Landmark Detection: A Literature Survey,” International Journal of Computer Vision, Volume 127, Issue 2, February 2019, pp.115-142.

[19] Timothy F. Cootes, Gareth J. Edwards, and Christopher J. Taylor, “Active Appearance Models,” IEEE Trans. PAMI 23, 2001, pp. 681-685.

[20] Timothy F. Cootes, Christopher J. Taylor, David H. Cooper, and Jim Graham, “Active Shape Models – Their Training and Application,” Computer Vision, Graphics and Image Understanding, 1995, pp. 38–59.

[21] Adrian K. Davison, Cliff Lansley, Nicholas Costen, Kevin Tan, Moi Hoon Yap, “SAMM: A Spontaneous Micro-Facial Movement Dataset,” in IEEE Transactions on Affective Computing, Volume 9, No. 1, 2018, pp. 116-129.