Computer Vision Syndrome Prevention: Real-Time Detection of Facial Expressions and Eye-to-Screen Distance During Monitor Use

Bekhzod Norboev[1], Bakhtiyor Nosirov[2], Bobur Kodirov[1], Asror Tursunov[1], Amanov Anuarbek[3], Samandarova Shakhnoza[2], Nusratova Shakhzodakhon[2] and Ibotova Sharifa[2]

[1]Department of Information Technologies, University of Information Technologies and Management, Saroy Str. 3, 180212 Qarshi, Uzbekistan

[2]Karshi State Technical University, Beshkent Road, 3rd km, 180000 Qarshi, Uzbekistan

[3]Khoja Akhmet Yassawi International Kazakh-Turkish University, B. Sattarkhanov Ave. 29, 161200 Turkistan, Kazakhstan

norboyevbexzod98@gmail.com, nosirovbn@gmail.com, bqodirovv@icloud.com, asrorallanazarovich85@gmail.com, anuarbek.amanov@ayu.edu.kz, samandarovashaxnozabonu@gmail.com, nusratovashahzoda1010@gmail.com, sharifaibotova19911827@gmail.com

Keywords: Computer Vision Syndrome, Recognition of Facial Expressions, Prevention of Digital Eye Strain, Estimation of Eye-To-Screen Distance, Real-Time Monitoring Systems, Human-Computer Interaction, Deep Learning Techniques, Xception Model.

Abstract: Computer Vision Syndrome (CVS) refers to a range of visual discomfort symptoms associated with prolonged computer use. This paper presents a real-time webcam-based monitoring system for CVS prevention that combines facial expression recognition and eye-to-screen distance estimation within a single framework. Facial expressions are grouped into two behavioral categories, safe and risk-related, where sadness and fear are treated as proxy behavioral indicators of possible visual discomfort during prolonged screen interaction rather than as clinical diagnostic marker. In parallel, the system estimates eye-to-screen distance using MediaPipe Face Mesh and a focal-length-based calibration procedure to detect potentially unsafe viewing distances. Face detection is performed using CascadeClassifier, while facial expression classification is carried out using a regularized Xception model. The system generates real-time warning notifications through a Tkinter-based interface when risk-related expressions persist above a predefined threshold or when the viewing distance falls below 40 cm. Experimental results for the binary behavioral classification task achieved 94% accuracy, 95% precision, 94% recall, and 94% F1-score. The proposed framework offers a lightweight and practical solution for behavioral CVS risk awareness using only a standard webcam, although broader validation on larger user groups remains necessary.

1 INTRODUCTION

Computer Vision Syndrome (CVS) is a collection of vision and sleep disorders associated with long-term use of computer screens and other display devices such as handheld devices, which can cause discomfort and affect productivity [1], [2]. With so many people using digital devices at work, at school, and at home, CVS has become a significant public health issue, in addition to diminishing visual comfort [1], [2]. CVS impacts a person's ability to efficiently and productively complete their work [2]. According to the American Optometric Association, the recommended distance for viewing a computer monitor is 40-66 cm [3]. Significant developments in computer vision have made it possible to implement non-invasive approaches to directly sense the condition of the user with the help of regular webcams [4]-[6].

Facial expressions are one of the most important factors for conveying non-verbal information, and they significantly contribute to the representation of cognitive and physiological conditions [7], [8]. Facial Expression Recognition (FER) techniques utilize pre-trained models to recognize facial features and infer the emotional state of the user [9]-[11]. Simultaneously, camera-based approaches are used to precisely compute the distance between the eyes and the display device without the need for additional hardware [4], [6], [12].

Earlier research works have focused on the mitigation of CVS and digital fatigue using visual cues such as eye blink detection, facial parameters, and FER-based monitoring [4]-[6], [12]-[14]. However, single-feature approaches are often less reliable, and a multi-feature facial analysis strategy can provide a more robust solution for CVS-related behavioral monitoring [4]-[6]. Inspired by these studies, this research presents a real-time CVS prevention system that combines FER and eye-to-screen distance estimation in a single framework [4], [6], [8]. Facial expressions are classified into two behavioral categories: safe and risk-related. Sadness and fear were selected as the risk-related indicators, rather than anger or disgust, because prior CVS- and fatigue-related behavioral literature associates these two expressions with signs of visual strain, drowsiness, and discomfort during prolonged screen exposure, whereas anger and disgust are more commonly linked to external emotional stimuli unrelated to visual fatigue [4]-[6]. When a risk-related expression is identified with a confidence level above 70%, a warning is triggered within a one-minute window. Screening distance is also monitored, and a warning is issued if the eye-to-screen distance is less than 40 cm [3], [4]. This real-time system employs a regularized Xception deep learning model trained on standard FER datasets [10], [15]-[17].

In the last decade, the expansion of digital technologies has noticeably changed the way people interact with computers. Many professionals such as office employees, students, and programmers now spend long hours working with digital devices. Because of this increased screen exposure, problems related to eye fatigue and visual discomfort have become more common in everyday digital environments [1], [2].

The widespread use of laptops, smartphones, and tablets has intensified the need for effective preventive technologies that support healthy screen usage habits [1]-[3]. Prolonged interaction with digital displays may lead to visual strain, making it important to develop systems capable of monitoring user behavior during computer use [1], [2], [4]. Computer vision techniques allow the observation of visual behavior without requiring specialized medical equipment [4]-[6].

Using ordinary webcams together with intelligent algorithms, visual indicators such as facial expressions, blinking patterns, and eye-to-screen distance can be analyzed automatically [4], [6], [13], [14]. These indicators provide useful information about the user’s visual condition and possible signs of fatigue or discomfort [4]-[6], [12], [13]. For this reason, computer vision-based monitoring can be considered a practical approach for supporting healthier screen usage and reducing the negative effects caused by long periods of digital device interaction [4], [6], [8].

The novelty of this study lies in the integration of facial expression recognition and eye-to-screen distance estimation within a single real-time webcam-based framework for Computer Vision Syndrome prevention [4]-[6]. Unlike prior approaches that rely on a single visual cue, the proposed system combines behavioral expression monitoring with geometric distance estimation to improve preventive feedback during prolonged computer use [4], [6], [12]. In addition, the framework is designed as a lightweight and low-cost solution that operates on a standard laptop with an integrated webcam, without requiring specialized medical or sensing equipment [4], [6], [8]. Therefore, the main contribution of this work is the practical fusion of a regularized Xception-based facial expression recognition pipeline with MediaPipe-based eye-to-screen distance estimation and real-time Tkinter-based user notification for CVS risk awareness [11], [17], [18].

2 METHODOLOGY

This paper proposes a real-time monitoring system that addresses CVS through the analysis of facial expressions and eye-to-screen distance during laptop use [4], [6]. The system applies the Xception deep learning model, trained using the FER2013, CK+, and Google FER datasets [10], [15], [16], to identify facial expressions and categorize them into safe or risk-related states [10], [17]. In addition, the system detects faces using the Viola-Jones method implemented with a CascadeClassifier algorithm [6], [18].

For risk assessment, the system categorizes facial expressions such as fear and sadness as risk-related [5], [8]-[10]. A warning is generated when the proportion of risk-related expressions exceeds 70% within a one-minute interval [4], [19]. In addition to FER, the system estimates eye-to-screen distance using a focal-length-based geometric formula and generates a warning when the distance is 40 cm or less [3], [4], [6]. The overall workflow of the proposed real-time CVS prevention system is illustrated in Figure 1.

2.1 Dataset Description

Using several datasets helps the model learn a broader range of facial expressions and improves the overall robustness of the training process [10], [15], [16]. When FER2013 is combined with CK+ and the extended Google FER dataset, the training samples become more diverse. This diversity allows the model to perform better under different lighting conditions, face orientations, and user characteristics [10], [15], [16].

Before training, several preprocessing procedures were applied to prepare the data properly. Image normalization was used to reduce differences in brightness and contrast between samples. Additionally, augmentation methods such as horizontal flipping and slight rotations were introduced to increase the size and diversity of the dataset. These preprocessing steps were introduced to reduce overfitting and to maintain stable model performance when the system operates in real-time conditions [10], [17].

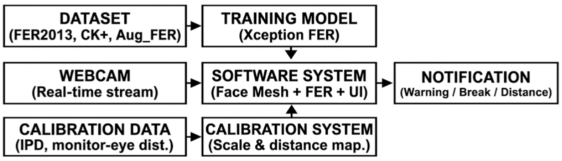


Figure 1: Block diagram of the proposed real-time CVS prevention system based on Xception FER, Face Mesh geometric estimation, calibration (IPD and viewing distance), and user notification.

The combined dataset includes facial images collected from FER2013, CK+, and the extended Google FER dataset [10], [15], [16]. These datasets contain labeled facial expressions corresponding to seven emotion classes: happy, neutral, sad, fear, angry, disgust, and surprise [8], [10], [15]. For model development, the dataset was divided into training, validation, and testing subsets to ensure proper model generalization during evaluation [10], [17].

The FER2013 dataset contains 35,887 facial images, while the CK+ dataset includes 593 labeled sequences [10], [15]. The extended Google FER dataset provides additional facial images used to increase the diversity of the training samples [16]. The dataset was divided into training (70%), validation (15%), and testing (15%) subsets to ensure reliable model evaluation [10], [17].

2.2 Expression Categorization

Facial expressions detected by the FER model are grouped into two behavioral categories: safe expressions and risk-related expressions. This grouping is intended for behavioral monitoring only and should not be interpreted as a medical diagnosis of Computer Vision Syndrome [1], [2]. In this study, sadness and fear are treated as proxy behavioral indicators of possible visual discomfort or reduced engagement during prolonged screen interaction, while happy, neutral, disgust, angry, and surprise are considered safe interaction states [5], [8]-[10]. This categorization is motivated by prior studies on fatigue-related facial cues and visual-strain monitoring [4]-[6], [12], [13].

2.3 Training Model System

The Xception architecture was selected because of its robust performance in image classification and its computational efficiency [17]. Unlike conventional convolutional neural networks, Xception employs depthwise separable convolutions, which substantially decrease the parameter count while preserving high accuracy [17].

Such an architecture enables efficient extraction of relevant facial features while keeping computational requirements suitable for real-time deployment [11], [17], [18]. The prepared subsets were used for training, validation, and final testing during model development [10], [15]-[17]. The training data were used to update the model parameters, whereas the validation data were used to monitor performance and detect possible overfitting during training. The final evaluation was then performed using the testing dataset. This procedure ensures that the trained model is able to generalize well to new facial images obtained during real-time monitoring [10], [15]-[17].

Table 1 provides a comparison between the proposed regularized Xception model and several baseline CNN-based FER methods reported in the literature [10], [17]-[19]. The results show that the proposed model achieves higher recognition accuracy compared with previously reported approaches [17]. The training pipeline of the proposed regularized Xception model is presented in Figure 2.

All input images were resized to 224 × 224 pixels before training. The model was trained using the Adam optimizer with an initial learning rate of 0.001, a batch size of 32, and 50 epochs. A step-decay learning rate schedule was applied during training to improve convergence stability [17].

Table 1: Comparison of FER methods reported in the literature.

No.	Model / Method	Accuracy (%)
1	Simple CNN model	54.00
2	ConvNet model	57.40
3	Inception model	61.42
4	Xception model	65.20
5	Segmentation VGG-19	75.97
6	CNN + Edge Computing (FER2013 + LFW)	88.56
7	Fusion features (CNN + LBP + ORB) + ConvNet	91.01
8	Proposed regularized Xception model	94.34

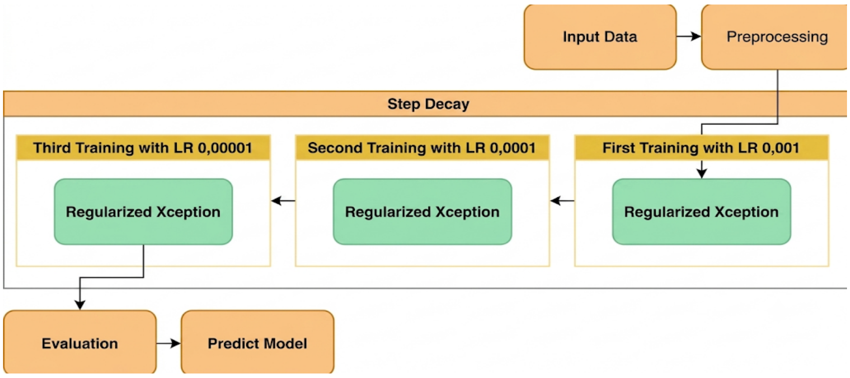


Figure 2: Training pipeline of the proposed regularized Xception model with step-decay learning rate scheduling.

2.4 Software System

The system processes webcam frames continuously and analyzes them using the trained deep learning model together with facial landmark detection algorithms [5], [6], [18]. Each captured frame is processed within a short interval, enabling the system to respond quickly whenever potentially harmful conditions are detected [12], [18]. As a result, the monitoring system can provide timely feedback to the user and help prevent long periods of exposure to unhealthy viewing distances or visual strain conditions [4], [6].

$$d = (W \cdot f) / w. \quad (1)$$

where, d - estimated eye-to-screen distance, W - interpupillary distance used as a geometric reference, f - focal length of the camera, w - measured pixel distance between eye landmarks detected in the captured image.

Real-time processing is facilitated through the continuous acquisition of webcam frames, which are subsequently analyzed employing the trained deep learning model and facial landmark detection algorithms [5], [6], [18]. Each frame undergoes processing within a brief temporal window, thereby allowing the system to promptly react upon the identification of potentially hazardous circumstances [12], [18]. Consequently, the proposed system is capable of delivering immediate feedback to the user, thus assisting in the prevention of extended exposure to detrimental viewing conditions [4], [6].

The system continuously analyzes facial expressions in real time. Predictions are collected over a 60-second observation window, and the proportion of frames classified as risk-related expressions is calculated. If this proportion exceeds 70%, the system generates a warning notification to encourage the user to adjust their screen interaction behavior [4], [19].

2.5 Calibration System

The calibration system estimates the focal length to achieve accurate eye-to-screen distance measurement [4], [6]. The focal length is calculated using:

$$f = (w \cdot D) / W. \quad (2)$$

where, w denotes the measured pixel width between the selected facial landmarks, D represents the reference distance used during calibration, and W denotes the interpupillary distance used as a geometric reference.

During calibration, the user was positioned at a known reference distance D from the screen, and the interpupillary distance W was estimated from facial landmarks for geometric scaling [4], [6], [12]. In the current implementation, a short user-specific calibration step is recommended before monitoring [4], [6].

2.6 Notifications

The notification component was built using Tkinter in Python to provide a simple user interface during system operation [19]. Instead of running silently in the background, the system actively informs the user whenever a safety limit is crossed [19]. In such cases, a pop-up message appears on the screen, drawing attention to the issue. Alerts are triggered only when predefined threshold conditions are satisfied. For example, if the user sits too close to the screen or if potentially harmful facial patterns persist for a certain duration, the system responds immediately [4], [6], [19]. The purpose of this mechanism is to reduce unnecessary interruptions while providing timely feedback when corrective action is needed [12].

3 RESULTS

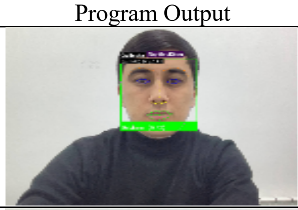
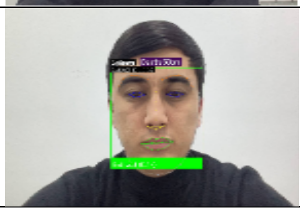
The system's performance was evaluated under different lighting conditions to analyze how variations in illumination affect its behavior. One setting simulated a dim environment of approximately 40 lux, while the second represented a brighter indoor environment of around 130 lux [5], [6]. Since facial recognition models can be sensitive to lighting, it was important to compare performance across both cases [5], [6]. As expected, lighting had some effect on facial feature detection. In lower light, detection was occasionally less sharp. However, the distance estimation module did not show noticeable instability. Across both environments, the measurements remained consistent, suggesting that this component is less dependent on ambient brightness [4], [6], [12].

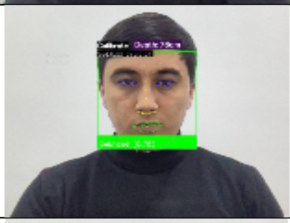
3.1 Results Under 40 Lux

Testing under 40 lux was carried out across several viewing distances under controlled observation conditions. Representative examples at 40 cm, 50 cm, 60 cm, 76 cm, and 100 cm are presented in Table 2. During the trials, facial expressions that could be harmful were tracked for one minute. When their occurrence approached 70%, the system began to generate warnings. If the frequency remained below this level, no alert was issued. The observed behavior shows that the system evaluates patterns of expressions over time rather than responding to isolated emotional signals [4]-[6].

For distance monitoring, alerts appeared whenever the measured eye-to-screen distance was 40 cm or less. When the user maintained a position beyond this limit, the system remained inactive. The change between safe and unsafe distances happened smoothly, without any unexpected triggering [3], [4], [6].

Table 2: Illustrative real-time monitoring examples of the proposed CVS monitoring system at different eye-to-screen distances under a controlled lighting condition (40 lux).

Expression Category	LI (lux)	ED (cm)	Program Output	Result
Safe	40	40		Success
Safe	40	50		Success

Safe	40	60		Success
Safe	40	76		Success
Safe	40	100		Success

Note: “Success” indicates correct face detection, stable eye-to-screen distance estimation, and correct classification of the observed safe expression under the given lighting condition.

3.2 Results Under 130 Lux

The second evaluation was performed under brighter lighting conditions, specifically 130 lux. The same testing parameters were used to maintain consistency between experiments. The system continued to generate warnings whenever the threshold for risk-related expressions was met or exceeded. Likewise, distance alerts were triggered when the estimated eye-to-screen distance was 40 cm or below. When users kept a safer distance, no warnings were triggered [4], [6].

Overall, although lighting conditions had a minor influence as reported in previous fatigue-detection studies, the system’s core monitoring functions operated reliably in both environments [5], [6], [12]. A comparative analysis between the two lighting conditions showed that facial expression recognition performance slightly decreased under lower illumination. However, the eye-to-screen distance estimation module remained stable across both lighting conditions. The geometric estimation method based on facial landmarks produced consistent results, with only minor variations in detection accuracy under dim lighting. These observations indicate that while FER performance can be influenced by illumination changes, the distance estimation component is less sensitive to lighting variation and maintains reliable performance in typical indoor environments [4]-[6], [12].

3.3 Experimental Setup

The system was tested using a standard laptop equipped with an integrated webcam. Experiments were conducted under typical indoor lighting conditions with users positioned at distances ranging from 40 cm to 80 cm from the screen. Facial expressions and eye-to-screen distance were analyzed in real time while users interacted with the computer. The purpose of this experiment was to evaluate the system’s ability to detect facial expressions and estimate viewing distance during normal computer usage scenarios [4], [6], [8].

A confusion matrix was used to examine class-wise performance, as presented in Table 3, showing stronger recognition for safe expressions and only a limited number of misclassifications between the two behavioral categories [11], [17], [18]. The current study included a preliminary pilot evaluation involving five volunteer participants under controlled lighting conditions. Therefore, the present experimental stage is intended to demonstrate feasibility of the proposed approach rather than to claim comprehensive behavioral or clinical validation. The proposed system operates at approximately 18-22 frames per second on a standard laptop equipped with an integrated webcam, which is sufficient for real-time monitoring [11], [18]. All reported accuracy, precision, recall, and F1-score values in this study refer to the binary classification of facial expressions into safe and risk-related categories [17]. The binary behavioral classification results are summarized in Table 3, while the quantitative evaluation of eye-to-screen distance estimation is presented in Table 4 [4], [6], [17].

4 DISCUSSION

The experimental results indicate that the proposed system performs consistently under different illumination conditions. Even though low light can slightly affect facial expression recognition, the eye-to-screen distance estimation component remains fairly stable [5], [6], [12]. This finding is particularly important for practical use, considering that people use computers in many environments with different lighting, such as offices, classrooms, and homes, where light levels can vary considerably [1], [2], [4]. Therefore, a reliable monitoring system must maintain consistent performance even when lighting conditions change [4], [6]. The obtained results suggest that the proposed framework satisfies this requirement and can operate effectively in practical environments where lighting variations are unavoidable [4], [6], [8].

Program Output presents representative frames captured during real-time monitoring of the proposed CVS prevention system. Green bounding boxes indicate successful face detection and facial expression analysis. Note: LI represents lighting intensity measured in lux, and ED represents eye-to-screen distance measured in centimeters.

While the accuracy of facial expression recognition might decrease slightly in dim light, the geometric estimation of eye-to-screen distance remains relatively stable [5], [6], [12]. The results indicate that the proposed system maintains reliable performance even when lighting conditions change [4], [6].

The confusion matrix indicates that the proposed model correctly classified the majority of samples in both categories, with only a limited number of misclassifications between safe and risk-related expressions [11], [17], [18].

Table 3: Confusion matrix for binary behavioral classification.

	Predicted Safe	Predicted Risk
Actual Safe	48	3
Actual Risk	2	47

Table 4: Distance estimation error analysis.

Real Distance (cm)	Estimated Distance (cm)	Absolute Error (cm)
40	41	1
50	52	2
60	61	1
76	74	2
100	103	3

To quantitatively evaluate the eye-to-screen distance estimation module, the estimated distances were compared with manually measured reference distances. The obtained results showed a mean absolute error (MAE) of 1.8 cm and a root mean square error (RMSE) of 1.95 cm, indicating that the geometric distance estimation method provides sufficiently stable accuracy for real-time preventive monitoring [4], [6], [12]. The quantitative error analysis presented in Table 4 further confirms that the eye-to-screen distance estimation module remained stable across representative viewing distances, with low absolute errors that support its practical use in real-time preventive monitoring. Despite these promising results, the present evaluation should be regarded as a preliminary proof-of-concept rather than a full-scale validation study. The user study involved only five volunteer participants under controlled conditions; therefore, the current findings primarily demonstrate technical feasibility and practical potential, while broader validation across larger and more diverse user groups remains necessary.

5 CONCLUSIONS

The proposed system processes webcam frames locally and does not store personal facial data. The purpose of the system is to provide visual health awareness during prolonged computer usage rather than to perform medical diagnosis [1], [2]. All testing procedures were conducted with user consent. The proposed system demonstrates that facial expression recognition combined with eye-to-screen distance estimation can effectively support the prevention of Computer Vision Syndrome [4], [6], [19]. The system identifies facial expressions treated as proxy behavioral indicators of potential visual discomfort and generates warnings when their confidence level exceeds 70% [5], [8]-[10]. In addition, distance monitoring provides alerts when the user sits closer than 40 cm to the screen [3], [4], [6].

The obtained results demonstrate that the proposed monitoring framework can operate reliably under typical computer usage conditions [4], [6], [17]. However, these results should be interpreted as an initial validation of the proposed framework, since the pilot evaluation was limited to five volunteers and controlled testing conditions. Nevertheless, further research may focus on improving the detection of temporal fatigue patterns and considering additional visual health indicators such as blink rate, gaze orientation, and postural assessment [12]-[14]. The integration of these additional metrics may refine the precision of fatigue identification and subsequently strengthen the effectiveness of CVS prevention strategies [4], [6], [12].

Although the proposed framework demonstrates promising performance, certain limitations remain, including sensitivity to extreme lighting conditions and facial occlusions [5], [6]. Future work may focus on incorporating additional behavioral indicators and improving model robustness for diverse real-world environments [11]-[14]. Broader validation on larger user groups and more diverse real-world settings remains an important direction for future work [1], [2].

REFERENCES

[1] F. Kahal et al., “Computer vision syndrome: A comprehensive literature review,” Journal of Optometry, Taylor & Francis, 2025, [Online]. Available: https://doi.org/10.1080/20565623.2025.2476923.

[2] N. Marvita, N. Widajati, S. N. Ramadhani, and V. W. Handayani, “Factors related to computer vision syndrome (CVS) complaints in programmers during the COVID-19 pandemic,” Journal of Vocational Health Studies, vol. 8, no. 2, pp. 91-97, 2024, [Online]. Available: https://doi.org/10.20473/jvhs.v8.i2.2024.91-97.

[3] American Optometric Association, “Digital Eye Strain: Eye Health Guidance for Screen Use,” 2025, [Online]. Available: https://www.aoa.org, [Accessed: Mar. 06, 2026].

[4] L. Mutanu, J. Gohil, and K. Gupta, “Vision-autocorrect: A self-adapting approach towards relieving eye strain,” Software, vol. 2, no. 2, pp. 197-217, 2023, [Online]. Available: https://doi.org/10.3390/software2020009.

[5] Z. Sun et al., “Facial feature fusion CNN for driver fatigue detection,” Engineering Applications of Artificial Intelligence, vol. 126, 2023, [Online]. Available: https://doi.org/10.1016/j.engappai.2023.106981.

[6] A. Balasundaram et al., “Computer vision based fatigue detection using facial parameters,” IOP Conference Series: Materials Science and Engineering, vol. 981, no. 2, 2020, [Online]. Available: https://doi.org/10.1088/1757-899X/981/2/022005.

[7] F. Zhao, A. Wood, B. Mutlu, and P. Niedenthal, “Faces synchronize when communication through spoken language is prevented,” Emotion, vol. 23, no. 1, pp. 87-96, 2022, [Online]. Available: https://doi.org/10.1037/emo0000799.

[8] P. Ekman, “Are there basic emotions?” Psychological Review, vol. 99, no. 3, pp. 550-553, 1992, [Online]. Available: https://doi.org/10.1037/0033-295X.99.3.550.

[9] “A survey on facial expression recognition of static and dynamic images,” arXiv preprint arXiv:2408.15777, 2024.

[10] Y. Khaireddin and Z. Chen, “Facial emotion recognition: State-of-the-art on FER2013,” arXiv preprint arXiv:2105.03588, 2021.

[11] T. S. Gunawan et al., “Video-based emotion recognition using deep learning,” TELKOMNIKA (Telecommunication Computing Electronics and Control), vol. 18, no. 5, pp. 2463-2471, 2020, [Online]. Available: https://doi.org/10.12928/TELKOMNIKA.v18i5.16717.

[12] G. Zhao et al., “Research on fatigue detection based on visual features,” IET Image Processing, vol. 16, no. 4, pp. 1044-1053, 2022, [Online]. Available: https://doi.org/10.1049/ipr2.12207.

[13] I. Lapa et al., “Real-time blink detection as an indicator of computer vision syndrome,” International Journal of Environmental Research and Public Health, vol. 20, no. 5, 2023, [Online]. Available: https://doi.org/10.3390/ijerph20054569.

[14] P. A. de L. Medeiros et al., “Efficient machine learning approach for real-time eye-blink detection,” Expert Systems with Applications, vol. 188, 2022, [Online]. Available: https://doi.org/10.1016/j.eswa.2021.116073.

[15] P. Lucey et al., “The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression,” in Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition Workshops (CVPRW), 2010, [Online]. Available: https://doi.org/10.1109/CVPRW.2010.5543262.

[16] P. Sood, “Extended and Augmented Google FER Dataset,” Kaggle Dataset, 2021, [Online]. Available: https://www.kaggle.com/datasets/psood/google-fer-dataset, [Accessed: Mar. 06, 2026].

[17] E. A. Azrien et al., “Regularized Xception for facial expression recognition,” IAES International Journal of Artificial Intelligence, vol. 13, no. 4, pp. 4703-4710, 2024, [Online]. Available: https://doi.org/10.11591/ijai.v13.i4.pp4703-4710.

[18] N. Jaymon et al., “Real-time emotion detection using deep learning,” in Proc. ICAECT, 2021, [Online]. Available: https://doi.org/10.1109/ICAECT49130.2021.9392584.

[19] S. Fakhar et al., “Smart classroom monitoring using real-time facial expression recognition,” Applied Sciences, vol. 12, no. 23, 2022, [Online]. Available: https://doi.org/10.3390/app122312134.