

Blink Identification Using Video Magnification for Temporal Analysis of Eyelid Motion

Maria L. C. A. Cernadas¹, Victor H. R. Cardoso¹, João C. W. A. Costa¹ and Moisés F. Silva²

Abstract—This work proposes a blink identification method based on video magnification and evaluates its performance against the Eye Aspect Ratio (EAR), a widely used geometric landmark-based approach. The methods were assessed using 221 manually annotated blink events collected from two subjects performing fixation and reading tasks at three acquisition distances. The proposed method achieved an overall F1-score of 0.997, with 100% recall in nearly all scenarios, whereas EAR reached an F1-score of 0.971 and showed systematic failures in detecting partial blinks and estimating event duration. The results indicate that representing blinking as a continuous temporal signal improves identification performance while preserving morphological information relevant to subtle blink patterns.

Keywords—Blink Identification, Video Magnification, Hilbert Transform, Partial Blink, Non-contact Monitoring

I. INTRODUCTION

Eye blinking plays a fundamental role in ocular protection, lubrication of the eye surface, and coordination with eye movements. The blink cycle consists of distinct phases, including eyelid closure, occlusion, reopening, and the interblink interval [1]. Under normal conditions, blink duration ranges from 100 to 400 ms, with average frequencies between 10 and 20 blinks per minute [2], [3]. Blink frequency also varies according to cognitive demand: reading tasks are commonly associated with reduced blink frequency, whereas conversational conditions tend to increase blinking activity [4].

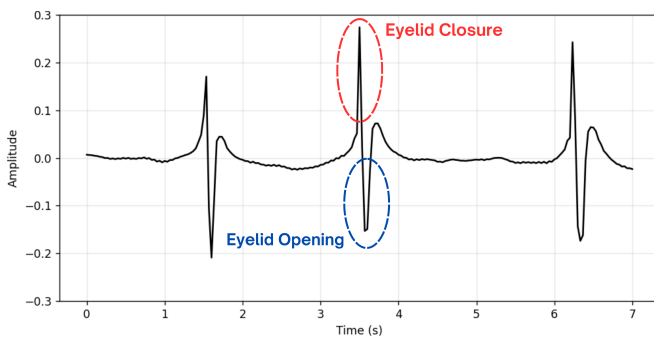


Fig. 1: Magnified modal signal showing three complete blink cycles.

Fig. 1 illustrates the characteristic morphology of eyelid motion represented by the magnified modal signal. Each cycle exhibits a positive deflection associated with eyelid closure, followed by a negative deflection during reopening. This

pattern forms the temporal signature of a blink and serves as a reference for event detection. Different blink types have been described according to eyelid closure amplitude [5]. A complete blink occurs when full eyelid closure is achieved, whereas a partial blink corresponds to incomplete closure, producing a shallower signal valley and indicating interruption before total occlusion [1]. Su et al. [6] further reported a double-valley morphology in severe partial blinks, characterized by two consecutive inflections without returning to baseline between them.

Changes in blinking behavior have been associated with neurological and ophthalmological conditions. Blink frequency tends to decrease in Parkinson's disease and may vary in other neurological disorders [7], [8]. Partial blinking is also recognized as a clinical indicator of dry eye disease, since incomplete eyelid closure compromises tear film stability [1], [6]. Several techniques have been proposed for automatic blink detection. Landmark-based approaches, such as the Eye Aspect Ratio (EAR), estimate geometric relationships around the eyelids but remain sensitive to head pose, illumination, and tracking errors [2]. Deep learning approaches can learn robust spatiotemporal features, but their generalization depends on diverse annotated data and they still face class imbalance, environmental interference, real-time, and device constraints [9]. Electroencephalography (EEG)-based techniques require contact sensors and specialized acquisition systems [8]. Video magnification instead provides a non-contact continuous signal that preserves blink morphology.

This work proposes a video-magnification blink identification method, wherein identification comprises event detection and complete/partial classification; detection refers specifically to event localization. Performance is evaluated against the Eye Aspect Ratio (EAR) using fixation and reading tasks, detection metrics, and event duration.

II. REGION OF INTEREST EXTRACTION

Initially, the video is subjected to face detection and tracking using MediaPipe Face Detection [10], supporting independent processing of multiple subjects. Template matching maintains the facial bounding box between frames, with automatic re-detection below a predefined matching threshold. The region of interest (ROI) is expanded, restricted to facial content, and resized to a fixed subject-specific spatial size; every temporal sample is retained. No frame-wise photometric normalization, geometric eye alignment, or explicit head-pose compensation is applied. Thus, both methods receive the same stabilized ROI, but follow independent signal extraction and detection pipelines.

¹Applied Electromagnetism Laboratory, Federal University of Pará, Belém-PA, Brazil. E-mail: maria.cernadas@itec.ufpa.br; ²Bandelier Technologies, Santa Fe, NM 87505, USA.

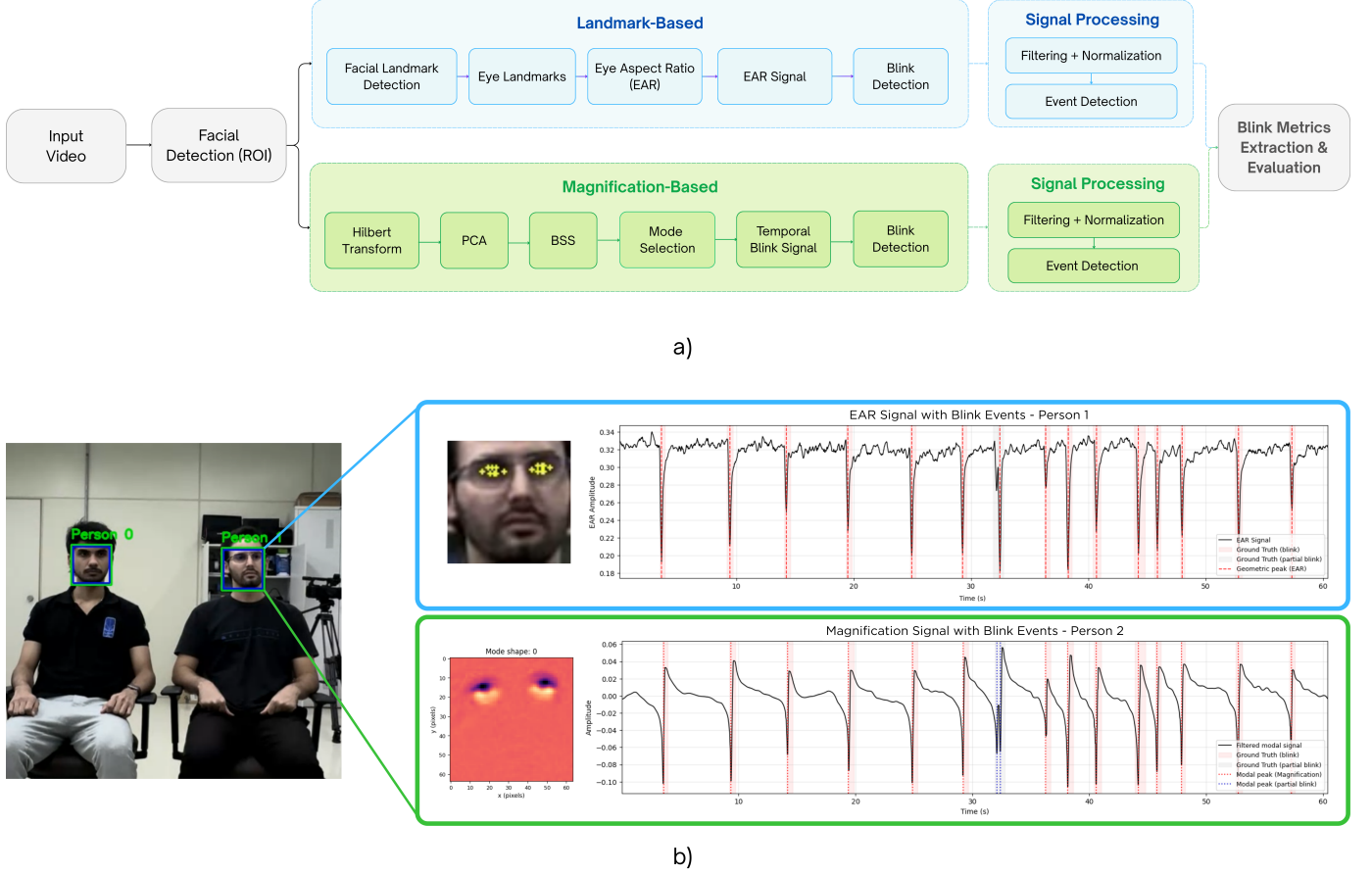


Fig. 2: Overview of the evaluated methods: (a) workflow from video input to metric extraction, with independent pipelines for the proposed magnification-based method and the comparative EAR approach; (b) EAR and magnification signals for Subject 1 during fixation at 1 m, highlighting ocular landmarks and the eyelid mode shape.

III. PROPOSED METHOD: VIDEO MAGNIFICATION-BASED IDENTIFICATION

A. Modal Signal Extraction

In the proposed video magnification-based approach, the video is modeled as a set of spatially distributed temporal series, in which each pixel within the ocular region provides an intensity measurement at a fixed position over time. As the eyelid moves, image structures shift, causing different regions to pass through the same pixel across consecutive frames. This behavior can be represented as a time-dependent spatial variation, expressed as $I(x + \delta(t))$, where $I(x)$ denotes the pixel intensity at position x , and $\delta(t)$ describes the displacement associated with eyelid motion [11], [12]. To analyze these variations, the Hilbert transform is applied to construct the analytic signal:

$$I_a(t) = I(t) + j\mathcal{H}\{I(t)\}, \quad (1)$$

where $\mathcal{H}\{\cdot\}$ denotes the Hilbert transform and j represents the imaginary unit. This transformation produces a complex signal representation whose real and imaginary components are combined into a spatiotemporal matrix, where each row corresponds to a pixel and each column to a time instant, preserving both motion amplitude and phase [13]. Since

videos contain a large number of signals, Principal Component Analysis (PCA) is employed to reduce dimensionality by projecting the data onto orthogonal components ordered according to variance. Blind Source Separation (BSS) is subsequently performed using the Complexity Pursuit algorithm, decomposing the reduced signal into statistically independent components and separating eyelid motion from noise and residual variations [13], [12]. The decomposition produces a set of independent components, each associated with a spatial pattern referred to as a *mode shape*, describing its contribution across pixels within the ocular region [12]. In the experiments, 10 principal components were considered, although this parameter can be adjusted depending on algorithm configuration. Selection of the component associated with eyelid motion is based on low-frequency spectral energy concentration [8], characteristic modal signal morphology, and spatial correspondence between the *mode shape* and the eyelid region.

B. Signal Processing

The modal signal corresponding to the selected component after PCA and BSS contains visual and dynamic noise that may compromise preservation of blink morphology. Therefore, Total Variation Denoising (TV Denoising, $\lambda = 0.02$) is applied

to suppress residual oscillations while preserving abrupt transitions associated with eyelid closure and reopening. The filtered signal is subsequently normalized using the robust z-score expression in (4), establishing a common scale for comparison with the EAR signal.

C. Blink Identification: Detection and Classification

In the magnification-based method, classification combines signal morphology, relative amplitude, and temporal context through sequential stages. Candidate events are initially identified using peak detection in the closure signal and subjected to individual morphological validation. Each event is retained only when exhibiting a consistent decrease before the peak and recovery afterward, corresponding to the expected behavior of a physiological blink. Validated events are then classified as complete or partial blinks based on relative amplitude with respect to the signal pattern within the same video, event duration, and the presence of a double-valley morphology corresponding to the pattern described by Su et al. [6] during the closure phase of severe partial blinks.

Events exhibiting this morphology are labeled as two independent consecutive partial blinks according to the consecutive blink criterion proposed by Lee et al. [14]. Classification as a partial blink is confirmed only when the event belongs to a temporally adjacent pair with compatible characteristics. An example of this pattern can be observed in Fig. 2(b), where two consecutive partial blink events are detected in agreement with the established ground truth.

IV. COMPARATIVE METHOD: EYE ASPECT RATIO (EAR)

A. EAR Signal Extraction

The landmark-based approach begins by detecting facial points within the ocular region extracted from the ROI of each person, as illustrated in Fig. 2(b) – EAR Signal. Detection was performed using MediaPipe FaceMesh [10], which estimates the coordinates of 478 facial landmarks in normalized image space. Among these landmarks, eight points per eye are selected: one horizontal pair corresponding to the medial and lateral eye corners and three vertical pairs distributed along the upper and lower eyelid margins. Based on these coordinates, the Eye Aspect Ratio (EAR) is computed following the formulation proposed by Soukupová and Čech [2]:

$$\text{EAR} = \frac{|p_2 - p_6| + |p_3 - p_5| + |p_7 - p_8|}{2|p_1 - p_4|} \quad (2)$$

where p_1 and p_4 denote horizontally opposite eye corners, while p_2-p_3 , p_5-p_6 , and p_7-p_8 define three pairs of points located on the upper and lower eyelid margins, extending the original formulation that considered two vertical pairs [2]. The index is computed independently for each eye and averaged, since blinking occurs simultaneously in both eyes.

B. Signal Processing

The EAR signal is extracted frame by frame, and detection failures produce missing values, which are handled through linear interpolation between neighboring valid samples. The

resulting signal is smoothed using a moving average filter with a five-sample window to attenuate oscillations introduced by tracking instability without compromising the temporal resolution required for blink detection. For event detection, the closure signal is defined as:

$$\text{closedness}[t] = \text{median}(\text{EAR}) - \text{EAR}[t] \quad (3)$$

capturing deviations below the baseline associated with eyelid closure events, where the median represents the fully open-eye reference level [2], [3]. The signal is subsequently normalized using robust z-score:

$$z[t] = \frac{x[t] - \tilde{x}}{1.4826 \cdot \text{MAD}(x)} \quad (4)$$

where $\tilde{x}$ denotes the median, MAD is the median absolute deviation, and the factor 1.4826 ensures consistency with the standard deviation estimator under normal distributions.

C. Blink Detection

In the EAR method, detection operates on the normalized closure signal using a peak detector constrained by minimum amplitude and prominence criteria, considering only events with durations compatible with the physiological blink interval, typically between 100 and 400 ms [2], [3]. A post-detection filtering stage removes candidate events whose amplitude relationship and temporal proximity to stronger events suggest a higher probability of representing tracking artifacts rather than true blinks.

V. EXPERIMENTAL PROTOCOL AND EVALUATION METRICS

A. Data Acquisition

Data were collected non-invasively using the rear camera of a smartphone at 30 frames per second with high-definition (HD) resolution (1280×720 pixels), under controlled conditions with stable illumination and a neutral background.

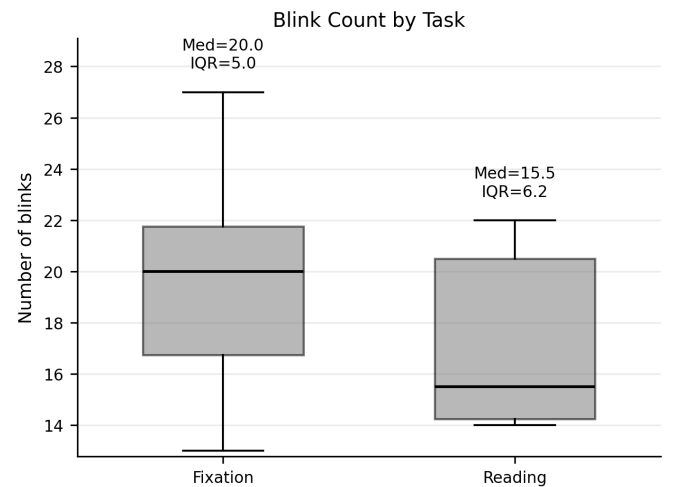


Fig. 3: Distribution of blink counts according to cognitive task.

TABLE I: Blink detection performance by participant and condition.

Person	Task	Dist.	n_{GT}	EAR						Magnification					
				TP	FP	FN	Prec.	Rec.	F1	TP	FP	FN	Prec.	Rec.	F1
Person 0	Fixation	1 m	19	19	0	0	1.000	1.000	1.000	19	0	0	1.000	1.000	1.000
		2 m	21	21	0	0	1.000	1.000	1.000	21	0	0	1.000	1.000	1.000
		3 m	22	22	0	0	1.000	1.000	1.000	22	0	0	1.000	1.000	1.000
	Reading	1 m	16	16	0	0	1.000	1.000	1.000	16	0	0	1.000	1.000	1.000
		2 m	15	14	0	1	1.000	0.933	0.966	15	0	0	1.000	1.000	1.000
		3 m	22	22	1	0	0.957	1.000	0.978	22	0	0	1.000	1.000	1.000
	Overall	–	115	114	1	1	0.991	0.991	0.991	115	0	0	1.000	1.000	1.000
Person 1	Fixation	1 m	16	15	0	1	1.000	0.938	0.968	16	0	0	1.000	1.000	1.000
		2 m	27	25	0	2	1.000	0.926	0.962	27	0	0	1.000	1.000	1.000
		3 m	13	12	1	1	0.923	0.923	0.923	13	0	0	1.000	1.000	1.000
	Reading	1 m	22	19	0	3	1.000	0.864	0.927	22	0	0	1.000	1.000	1.000
		2 m	14	13	1	1	0.929	0.929	0.929	14	1	0	0.933	1.000	0.966
		3 m	14	14	0	0	1.000	1.000	1.000	14	0	0	1.000	1.000	1.000
	Overall	–	106	98	2	8	0.980	0.925	0.951	106	1	0	0.995	1.000	0.997
Overall		–	221	212	3	9	0.986	0.959	0.971	221	1	0	0.995	1.000	0.997

* *Abbreviations:* Dist., acquisition distance; n_{GT} , number of ground-truth blink events; EAR, Eye Aspect Ratio; GT, ground truth; TP, true positive; FP, false positive; FN, false negative; Prec., precision; Rec., recall; F1, harmonic mean of precision and recall.

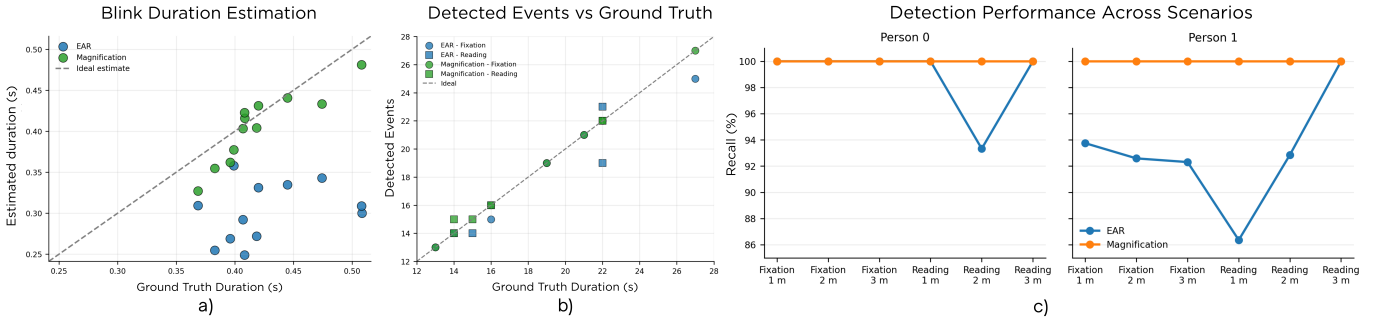


Fig. 4: Quantitative comparison of the evaluated methods: (a) estimated blink duration relative to ground truth; (b) detected events across tasks and acquisition distances; (c) recall by subject across tasks and acquisition distances.

The experiment involved two healthy adults with no history of ocular disorders, positioned at 1, 2, and 3 m from the camera. Each 60-s recording contained both subjects performing either static fixation or text reading. The six recordings yielded 12 subject-specific sequences and 221 annotated blink events. All frames were retained at the native 30 fps, without temporal subsampling. As shown in Fig. 3, the median count decreased from 20 events during fixation (interquartile range, IQR = 5.0) to 15.5 during reading (IQR = 6.2), consistently with task-related differences [4]. Eyeglasses, ocular occlusion, large pose changes, and illumination variation were not evaluated as experimental factors.

B. Evaluation Metrics

Ground truth was produced by frame-wise visual inspection of the raw videos at 30 fps, before automated evaluation. Onset was the first visible eyelid-closure frame and offset the return to the open-eye baseline; events were labeled complete or partial according to full or incomplete closure. Matching used a ± 150 ms tolerance between the detected peak

and the annotated event midpoint. For magnification, a true positive required agreement in timing and class. Precision (P), recall (R), and F1-score were micro-averaged [10]; 95% confidence intervals (CIs) for proportions used the Wilson method, and paired ground-truth detections were compared with the two-sided exact McNemar test. Event duration was also evaluated.

$$P = \frac{TP}{TP + FP}, \quad R = \frac{TP}{TP + FN}, \quad F1 = \frac{2PR}{P + R} \quad (5)$$

VI. RESULTS AND DISCUSSION

Table I presents the event-level results. Magnification achieved precision 0.995 (95% CI: 0.975–0.999), recall 1.000 (0.983–1.000), and F1-score 0.997; EAR achieved precision 0.986 (0.960–0.995), recall 0.959 (0.924–0.978), and F1-score 0.971. In an exploratory event-level paired comparison, nine of the 221 reference events were detected only by magnification and none only by EAR (two-sided exact McNemar $p = 0.0039$). Magnification produced one false positive for

person 1 during reading at 2 m, caused by slight head motion captured by the eyelid mode.

EAR achieved perfect detection only for fixation scenarios of person 0, accumulating 3 false positives and 9 false negatives overall. Person 1 presented 6 partial blink events, all detected by magnification but missed by EAR as independent events. As shown in Fig. 2(b), partial blinks produce only small fluctuations in the EAR signal and become incorporated into a single larger valley. Since EAR is computed frame by frame, temporally adjacent events may be merged into one detection. The worst performance occurred in the reading task at 1 m for person 1 (recall = 0.864), where baseline displacement prevented detection of two events, while another low-amplitude event was discarded as noise.

Fig. 4(b) confirms this behavior through detected event counts. The magnification-based method follows the ideal line across all scenarios, whereas EAR deviates more noticeably, particularly during reading. Fig. 4(a) further shows that EAR tends to underestimate event duration, while magnification remains closer to the ideal diagonal with lower dispersion.

Fig. 4(c) shows stable recall for magnification, whereas EAR decreases for person 1. The isolated magnification failure indicates susceptibility to non-blink motion when modal separation is insufficient.

For T frames, P ROI pixels, and $k = 10$ retained components, the Hilbert stage costs $O(PT \log T)$ and stores $O(PT)$ samples. The current batch implementation computes full PCA separately for the real and imaginary matrices, dominating at approximately $O(PT^2)$ when $P \gg T$; BSS then costs $O(Tk^2 + k^3)$. Real-time operation was not benchmarked and is therefore not claimed. Hardware-limited deployment would require ROI downsampling, truncated or incremental PCA, temporal windows, and device-specific profiling.

These findings constitute a proof of concept under controlled acquisition, not population-level validation. With only two healthy participants, the metrics may reflect subject and setup characteristics; event-level significance does not establish subject-level generalization. Robustness to eyeglasses, occlusion, pose and illumination changes, broader demographics, and clinical populations remains to be tested. Recent deep-learning methods offer a relevant future benchmark, although they introduce training-data and computational requirements [9].

VII. CONCLUSION

This work evaluated a video-magnification blink identification method against EAR on 221 manually annotated events. Under controlled proof-of-concept conditions, magnification achieved F1-score 0.997, reduced failures on partial blinks, and preserved event duration more consistently than EAR. The results support the feasibility of continuous eyelid-motion analysis, but clinical or population-level conclusions require broader validation.

Future work will expand demographic and clinical diversity, vary pose, illumination and occlusion, compare deep-learning baselines, and profile optimized implementations on resource-limited hardware.

ACKNOWLEDGMENTS

This study was financed in part by the Brazilian funding agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Additional support was provided by the Federal University of Pará (UFPA), the High Performance Computing and Artificial Intelligence Center (CCAD-IA), the Amazon Sustainability Institute for Science and Innovation (ISACI), and the Guamá Foundation. The authors also acknowledge the Brazilian Innovation Agency (FINEP) for the financial support and infrastructure provided for this research.

REFERENCES

- [1] Y. Jie, R. Sella, J. Feng, M. L. Gomez, and N. A. Afshari, "Evaluation of incomplete blinking as a measurement of dry eye disease," *Ocular Surface*, 2019.
- [2] T. Soukupova and J. Cech, "Real-time eye blink detection using facial landmarks," in *21st Computer Vision Winter Workshop*, 2016.
- [3] S. Al-gawwam and M. Benaissa, "Robust eye blink detection based on eye landmarks and Savitzky-Golay filtering," *Information*, vol. 9, no. 9, p. 93, 2018.
- [4] E. Fitzpatrick, N. Hohl, P. Silburn, C. O'Gorman, and S. A. Broadley, "Case-control study of blink rate in Parkinson's disease under different conditions," *Journal of Neurology*, vol. 259, no. 4, pp. 739–744, 2012.
- [5] M. Nyström, R. Andersson, D. C. Niehorster, R. S. Hessels, and I. T. C. Hooge, "What is a blink? Classifying and characterizing blinks in eye openness signals," *Behavior Research Methods*, vol. 56, pp. 3280–3299, 2024.
- [6] Y. Su, Q. Liang, G. Su, N. Wang, C. Baudouin, and A. Labbé, "Spontaneous eye blink patterns in dry eye: Clinical correlations," *Investigative Ophthalmology & Visual Science*, vol. 59, no. 12, pp. 5149–5156, 2018.
- [7] A. Esteban, A. Traba, and J. Prieto, "Eyelid movements in health and disease. The supranuclear impairment of the palpebral motility," *Journal of the Neurological Sciences*, 2004, received 12 September 2003; accepted 13 January 2004.
- [8] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "A new EEG acquisition protocol for biometric identification using eye blinking signals," *I.J. Intelligent Systems and Applications*, vol. 6, pp. 48–54, 2015.
- [9] J. Xiong, W. Dai, Q. Wang, X. Dong, B. Ye, and J. Yang, "A review of deep learning in blink detection," *PeerJ Computer Science*, vol. 11, p. e2594, 2025.
- [10] P. Kietly *et al.*, "Event-based multi-task facial landmark and blink detection," *IEEE Access*, vol. 13, pp. 45 604–45 620, 2025.
- [11] Y. Yang, C. Dorn, T. Mancini, Z. Talken, G. Kenyon, C. Farrar, and D. Mascarenas, "Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification," *Mechanical Systems and Signal Processing*, vol. 85, pp. 567–590, 2017.
- [12] M. F. Silva, A. Green, and D. Mascarenas, "A generalized technique for full-field blind identification of travelling waves and complex modes from video measurements with Hilbert transform," in *Data Science in Engineering, Volume 9: Proceedings of the 39th IMAC*. Springer, 2021, pp. 233–236.
- [13] C. S. Costa, G. N. Castro, Y. S. A. Alencar, V. H. R. Cardoso, J. C. W. A. Costa, and M. F. Silva, "Non-contact video-based heart rate monitoring using Hilbert motion magnification," in *Special Topics in Structural Dynamics & Experimental Techniques, Vol. 5: Proceedings of the 43rd IMAC, A Conference and Exposition on Structural Dynamics 2025*, M. Allen, B. Moaveni, and D. Epp, Eds. River Publishers, 2025.
- [14] W.-H. Lee, J. Seo, and J.-M. Hwang, "The method for quantified analysis and pattern visualization for eye blinking using high-frame-rate video," *Biomedical Engineering Letters*, vol. 15, pp. 1097–1107, 2025.