

Low-Cost Camera-Based Automated Blink Analysis in Pediatric Frequent Blinking: a Proof-of-Concept Functional Evaluation

Lin Zhu^{1\}

737051097@qq.com

The Second People's Hospital of Qujing City

Chuanqi Lin¹

The Second People's Hospital of Qujing City

Fanrong Yang¹

The Second People's Hospital of Qujing City

Chengsheng Peng²

The Second People's Hospital of Qujing City

Qiuling Li¹

The Second People's Hospital of Qujing City

Yating Zhang¹

The Second People's Hospital of Qujing City

Article

Keywords: blink analysis, eye aspect ratio, artificial intelligence, differential diagnosis, pediatric ophthalmology, functional evaluation

Posted Date: September 2nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10882548/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Frequent blinking is a common complaint in pediatric ophthalmology whose principal etiologies, dry eye disease (DED), allergic conjunctivitis (AC), and tic disorders (TD), overlap clinically but require divergent treatment. We present a low-cost, camera-based system quantifying five blink parameters and generating an etiology-oriented four-class output (normal, DED, AC, TD). Blinks were detected from the eye aspect ratio over MediaPipe FaceLandmarker landmarks, eye rubbing via hand proximity, and classification used an XGBoost classifier trained on 320 synthetic records as ethics approval for pediatric data was pending. Evaluation comprised eight module self-validation items in a headless environment, five camera-based sessions by a single adult author-participant, and boundary testing with 20 synthetic records. All self-validation items passed; all five sessions yielded complete parameter extraction and valid four-class outputs; and all boundary records were handled without failure. The system demonstrates a functionally complete, end-to-end, low-cost blink-analysis pipeline deployable on standard office hardware and reimplementable from the methods. The classifier reached perfect separation on held-out synthetic data, reflecting the constructed separability of the training set; no children or patients were involved, and clinical validation must precede any diagnostic claim.

Introduction

Clinical Background

Frequent blinking is among the most common presenting complaints in pediatric ophthalmology clinics. The principal underlying etiologies (DED, AC, and TD) share highly overlapping manifestations, including increased blink frequency, ocular redness, and itching, yet their management diverges substantially: DED is treated with lubricants and lid hygiene, AC with antihistamine agents, and TD with behavioral intervention or referral to pediatric neurology[1, 2]. Incorrect classification not only delays appropriate treatment but may also expose children to unnecessary medication.

In a large sample of 1054 pediatric patients presenting with abnormal blinking, 42.98% were ultimately diagnosed with a tic disorder, and 26.71% had been initially misdiagnosed, most frequently as allergic conjunctivitis, or had experienced a missed diagnosis[3]. In the same cohort, 74.17% of the children with tic disorders had comorbid ocular disease (predominantly AC and DED), 81.90% reported ocular symptoms, and 71.07% of the misdiagnosed or missed cases had coexisting allergic disease[3]; the three etiologies therefore overlap extensively at the clinical surface. Conventional diagnostic workups, including tear break-up time, fluorescein staining, and rating scales, are time-consuming, equipment-dependent, and poorly tolerated by young children with short attention spans[4]. These constraints make objective, quantitative blink assessment a clear need in ophthalmology clinics; extension to primary care settings, where comprehensive ocular surface testing is unavailable, is a longer-term vision.

Blink behaviour in healthy children follows a reproducible baseline: spontaneous blink rates of approximately 15–20 blinks/min and complete blink rates above 80%[21–23]. Each candidate etiology disturbs this baseline in a way that is measurable, which is the technical rationale for the parameters chosen here: DED is associated with tear film instability and compensatory high-frequency, low-completeness blinking[7, 33, 34]; in children, AC frequently coexists with DED and its itch–scratch reflex raises eye rubbing[19, 20, 36]; and TD produces rhythmic, stereotyped motor abnormalities that inflate inter-blink-interval variability[8, 11]. The five measurement dimensions implemented in the system (Table 1) follow directly from these contrasts.

Existing Approaches and Their Limitations

Prior investigations of blink behavior have largely been confined to binary comparisons (e.g., DED versus healthy controls) or to single-disease characterization[5–7]. Computer vision and deep learning methods have recently been applied to the automated detection of eye tics in video recordings of adolescents and adults with Tourette syndrome[8, 9]. In a proof-of-concept study, a supervised deep learning algorithm was trained to detect eye tic events in video clips from 11 adolescent patients, demonstrating the feasibility of automated tic quantification[8]. A complementary line of work has addressed the detection problem itself rather than the diagnostic task: a dual-embedding video vision transformer was proposed for robust blink detection under challenging imaging conditions[10], and related video-based architectures have been applied to tic disorder identification from clinical recordings[9]. These studies share an emphasis on event-level detection using deep architectures, whereas the present work adopts a landmark-based, parameter-level modelling strategy. This route difference is deliberate. Event-level deep detection presupposes frame-wise labelled training data and GPU inference, and its time-series outputs have limited interpretability. Parameter-level modelling retains only five summary features (Table 1), runs in real time on CPU, and yields outputs that clinicians can read directly. The trade-off is a dependence on stable landmark tracking and sufficiently long observation windows. Under the deployment constraints of rapid triage in ophthalmology clinics, the present study chose the parameter-level route to prioritize interpretability and low-cost operation.

Automated video-based analysis has further characterized blink abnormalities in adults with Gilles de la Tourette syndrome, revealing blink rates approximately twice those of healthy controls (34.1 versus 17.3 blinks/min)[11]. In pediatric populations, a three-group comparison of children with tic disorders, children with DED, and healthy children identified differences in meibomian gland morphology; however, that study relied on a dedicated ocular surface analyzer (Oculus Keratograph 5M) rather than on video-derived blink parameters[12]. A smartphone camera-based blink training study has demonstrated that consumer-grade cameras can support clinically meaningful blink analysis[13], supporting the viability of a low-cost hardware pathway. Automated blink assessment in children has been demonstrated with deep-learning analysis of blink characteristics and their relationship with tear film stability in long-term orthokeratology wearers[35]. Clinical interest in such camera-based tools is further reflected in a review of artificial intelligence for anterior segment disease[14] and in a machine-learning triage framework proposed for ophthalmology specialty clinics[15].

Unmet Need and Study Objective

In a targeted search of the literature (OpenAlex, title-and-abstract queries combining blink terms with machine-learning, multi-class, and paediatric terms; see Methods), **we did not identify** any published study that applied video-derived blink parameters obtained from a standard USB camera to the multi-class differential diagnosis of DED, AC, and TD in children. We emphasize that this was a targeted rather than a systematic review, and we therefore refrain from a formal novelty claim.

The objective of the present study was to develop and functionally evaluate the AI-Eye Diagnosis System, a low-cost, fully local, camera-based system **designed for** pediatric frequent blinking, with respect to three questions: (i) can the detection module correctly compute EAR and identify blink events under controlled conditions; (ii) can the pipeline extract all five parameters and produce valid four-class outputs end-to-end; and (iii) does the system behave robustly at extreme parameter values?

Two boundaries of this report must be stated at the outset. First, functional evaluation was performed by a single adult participant (one of the authors); **no children and no patients were involved**. Second, because institutional ethics approval has not yet been obtained, the classifier was trained on synthetic data; its performance therefore carries no clinical meaning. Accordingly, this report documents functional completeness of the pipeline, rather than diagnostic validity.

Results

System Implementation

The system was implemented as a standalone desktop application with a full-screen graphical interface comprising the five functional regions labelled in Fig. 1. Ancillary modules provide medical record entry, data management, batch entry with Excel export, and model retraining; a launcher dashboard unifies access to detection, data management, and retraining.

Module Self-Validation

All eight validation items passed in the headless environment (Supplementary Table S1). The open-eye and closed-eye EAR values were 0.3000 and 0.0125, respectively, a discriminative margin exceeding 20-fold. These items establish that the EAR computation module behaves correctly on static inputs.

Camera-Based Sessions

Five sessions were completed, with durations of 8.8–26.4 s. All five parameters were extracted in every session and four-class outputs were generated without failure (Table 3). The full four-class probability output and predicted class for each session are reported in Supplementary Table S2.

Table 3
Blink parameters extracted in five camera-based sessions (single adult participant; one of the authors).

Session	Date/Time	Duration (s)	Blink freq. (blinks/min)	Complete blink rate (%)	Blink amplitude	CV	Rubbing freq. (events/min)	Total blinks	Total rubbing	EAR frames
P1787733805	2026-08-26 16:43	26.4	20.47	55.56	0.1128	0.2972	4.55	9	2	396
P1787813470	2026-08-27 14:51	22.2	8.09	0.00	0.0953	0.7288	2.70	3	1	168
P1787813528	2026-08-27 14:52	15.4	7.77	100.00	0.1155	0.0000	0.00	2	0	117
P1787821956	2026-08-27 17:12	8.8	54.54	87.50	0.2179	1.1972	0.00	8	0	264
P1787826502	2026-08-27 18:28	18.3	19.66	83.33	0.1921	1.0487	0.00	6	0	523

The outputs were consistent with the parameter ranges on which the model was trained: the session with the highest rubbing frequency (4.55 events/min) fell within the AC training range (3.0–8.0) and returned the AC class; the two sessions with the highest CV (1.1972, 1.0487) exceeded the TD training range (> 0.40) and returned the TD class; and the session with zero complete blink rate returned the DED class. These three examples illustrate the expected correspondence between output classes and training ranges. Because the classifier combines all five features jointly, the classes of the remaining sessions were determined by their full feature combinations rather than by any single range.

The dependence of blink-event counts and complete-blink rates on the user-selected detection thresholds was examined over $\pm 20\%$ of the default thresholds (Supplementary Table S3). Within $\pm 10\%$ of the defaults, blink counts changed by at most two events in four of the five sessions; the two-blink session showed a 50-percentage-point swing in complete-blink rate, an artefact of the small denominator. Frame rate was also treated as an explicit evaluation dimension. The five sessions spanned 7.5–30 fps, the range typical of the intended low-cost hardware. Decimating the two 28–30 fps recordings to 15, 10, and 7.5 fps showed that at 15 fps blink counts fell by at most two events with moderate amplitude attenuation, and decimation to 7.5 fps missed more than half of the events (Supplementary Table S4). A minimum deployment frame rate of ≥ 15 fps is therefore recommended.

Model Performance on Synthetic Data

On the held-out synthetic test set (n = 80), the classifier achieved perfect separation (Table 4). Five-fold stratified cross-validation over the full synthetic dataset (n = 320) likewise returned an accuracy of 1.0000 in every fold.

Table 4
Classifier performance on the held-out synthetic test set.

Metric	Value
Test accuracy (n = 80)	1.0000
Weighted Cohen's κ	1.0000
One-vs-rest ROC AUC (all four classes)	1.0000
5-fold stratified CV accuracy (n = 320)	1.0000 ($\pm$ 0.0000)

Feature importances from the fitted model were: coefficient of variation 24.9%, eye-rubbing frequency 23.4%, blink frequency 20.4%, complete blink rate 20.2%, and blink amplitude 11.1%.

Interpretation. A perfect score is the expected outcome of the data generation procedure described in the Methods and Table 2: the classes were defined by substantially non-overlapping intervals on the most discriminative features, sampled independently with no noise model, so the classifier needed only to recover boundaries built into the data. The figures therefore document reproducibility of the training pipeline and motivate clinical validation as the necessary next step. We make no claim of model superiority: a decision rule over the engineered intervals would achieve identical separation. The XGBoost component was exercised because it is the element to be retrained on real clinical data in the planned validation. Feature importances reflect the separability engineered into the synthetic intervals rather than clinical informativeness.

Boundary Testing

All twenty synthetic boundary records were accepted, stored, and displayed without failure, and the retraining module correctly excluded all 20 records on the basis of their simulation marker, confirming that the anti-contamination mechanism functions as intended.

Data Output and Integrity

All five sessions produced the three designated output files with internally consistent content (*bl* $\in$ *k_params.csv*, 8 fields; *ear.csv*, 117–523 data rows; *rest.json*). No session terminated prematurely or produced malformed output.

Discussion

Principal findings. This study describes the development and functional evaluation of a low-cost, camera-based system for automated blink analysis, designed for pediatric frequent blinking. Three findings can be stated.

First, the combination of MediaPipe FaceLandmarker landmarks with the standard six-point EAR formulation yielded open/closed EAR values of 0.3000 and 0.0125, a margin exceeding 20-fold, and all eight headless self-validation items passed. This establishes correct behaviour of the EAR computation on static inputs.

Second, all five parameters were extracted and four-class outputs were produced without failure in all five camera-based sessions, and all output files were complete and internally consistent. This demonstrates that the pipeline, from camera acquisition through landmark detection, event detection, parameter extraction, and classification to structured output, operates end-to-end as designed.

Third, on synthetic data the classifier achieved perfect separation, documenting a reproducible training pipeline and motivating clinical validation.

Comparison with prior work. Automated video-based blink and tic analysis has been reported previously, but with a different scope (Supplementary Table S5). Relative to this body of work, the distinguishing features of the present system are its four-class, etiology-oriented output and its reliance on a consumer-grade camera with CPU-only operation. Prior studies enrolled patients and reported performance against clinically annotated events; the present report is a functional evaluation, and its scope is defined accordingly. The large paediatric cohort study documenting a 26.71% initial misdiagnosis rate characterized the clinical problem without proposing an automated solution[3], a gap the present system is designed to address.

A 2026 study applied video-derived spontaneous blink features to multi-class classification of Parkinson's disease[25], indicating growing interest in blink-feature multi-class modelling across clinical domains; the present work differs in target population, clinical task, and hardware cost.

Recent methodologically adjacent work published in this journal further illustrates the patient-driven paradigm to which the present report is complementary. A deep learning pipeline for automated tear film break-up time measurement reported strong diagnostic performance in dry eye cohorts[26]; a smartphone-photograph approach demonstrated high screening accuracy for horizontal strabismus[27]; and a blink-frequency threshold study linked reduced spontaneous blink rates to tear film instability in symptomatic young adults, identifying approximately 20 blinks/min as the lower

bound for maintaining tear film stability[28]. The present system operationalizes this value as a quantitative recommendation target in its treatment-suggestion module. This threshold, however, was derived from symptomatic young adults and is therefore not directly comparable with the paediatric reference ranges[5–7] underlying Table 2; it is presented as a reference target rather than a diagnostic cut-off.

Methodological considerations. Three design choices merit emphasis. First, the individual baseline strategy for EAR thresholding avoids fixed-threshold errors arising from variable face–camera distances and eyelid morphology, at the cost of requiring a brief initial calibration window. Second, the standard six-point EAR formulation[29] follows the geometric definition used in the literature and the measurement conventions summarized in a recent methodology review of blink dynamics[30]. Third, synthetic training data were used to exercise the pipeline end-to-end pending ethics approval.

Threshold and frame-rate sensitivity analyses are described in the Results and in Supplementary Tables S3–S4; they motivate the minimum deployment frame rate of ≥ 15 fps.

Limitations.

1. **Single adult author-participant and no pediatric data.** All camera sessions were self-tests by a single participant (one of the authors), and all boundary records were synthetic; no children or patients were involved. Because blink behaviour changes with development[5–6], performance in children cannot be inferred from the present data, and the parameter ranges in Table 2 follow adult reference values. Moreover, the verification sessions were short (8.8–26.4 s, 2–9 blinks), so the extracted parameter values are statistically unstable and carry no physiological meaning; they serve only to demonstrate the extraction pipeline. Finally, because the same author was developer, participant, and evaluator, independent assessment was not possible.

2. **Classifier trained on synthetic data.** The training data were generated from non-overlapping, independently sampled intervals, and the parameter–disease mappings in Table 1 constitute design premises to be confirmed in clinical data. The weakest of these premises is the association between eye-rubbing frequency and AC[12–13]: eye rubbing also occurs in DED, visual fatigue, and as a complex motor tic, and non-specific pruritus could raise rubbing frequency in any class. The planned validation should therefore confirm this feature's discriminative value explicitly, for example by testing whether rubbing-frequency distributions differ between AC and non-AC children after adjustment for symptom severity; if they do not, the AC interval in Table 2 should be revised.

3. **No comparison or reliability assessment.** The system was not compared against clinical diagnosis, rating scales, or alternative algorithms, and no inter-rater or test–retest reliability was assessed.

4. **Literature search was targeted, not systematic;** the absence of prior work identified in the Introduction should be read with that caveat.

Conclusions and future work

A low-cost, fully local, camera-based system for automated blink analysis has been developed and functionally evaluated: module self-validation passed all eight items; all five blink parameters were extracted and four-class outputs generated across five camera-based sessions; and boundary testing confirmed stable behaviour at extreme parameter values with correct exclusion of simulation records from retraining. Together these findings show that the pipeline functions as intended from camera acquisition to structured output; diagnostic validity must be established in the planned prospective clinical validation.

An ethics application for the planned clinical study has been submitted to the institutional review board of our centre; data collection and model training will begin upon approval. The next step is a prospective clinical study enrolling children with frequent blinking, whose diagnosis is established by comprehensive ophthalmological examination and, where indicated, standardized paediatric instruments (e.g., the Yale Global Tic Severity Scale[31]). Sample size will follow events-per-variable criteria[32] and empirical requirements for tabular clinical classification[33], and the study will be reported in accordance with TRIPOD[34] and STARD[35]. Diagnostic performance must then be evaluated against the clinical reference standard, and the model retrained on real data with nested cross-validation and calibration assessment. Subsequent work should address external validation, test–retest reliability, formal runtime benchmarking, validation under native low-frame-rate (< 15 fps) acquisition, and evaluation in younger and less cooperative children.

Methods

Parameter Provenance

Throughout Methods, every numerical value is attributed to one of four sources: (i) peer-reviewed literature, cited in square brackets; (ii) the system implementation, i.e., a value exactly as configured in the released software (detector settings, camera capture parameters, database schema, classifier hyperparameters, and similar engineering constants); (iii) an explicit design decision made by the authors for this functional evaluation; or (iv) measurements obtained during the functional evaluation itself (Table 3; Supplementary Tables S1–S4). Values in categories (ii) and (iii) are engineering choices rather than physiological parameters and are identified as such in the text.

System Overview

The AI-Eye Diagnosis System is a standalone desktop application implemented in Python 3.10+ (the present implementation runs under Python 3.14) with a Tkinter graphical user interface. The processing pipeline comprises five sequential stages (Figure 3): (1) standardized visual stimulation (age-appropriate cartoon video, default 180 s organized as three 60-s segments separated by 5-s rest periods with on-screen countdown; the three-segment structure with

short inter-segment rests was adapted from the repeated-measurement protocol of Chai et al.[28], who performed three repeated measurements separated by brief natural-blink intervals); (2) face and eye landmark detection (478 landmarks per frame via MediaPipe FaceLandmarker, Apache-2.0); (3) blink and eye-rubbing event detection (per-frame EAR with individually adapted thresholds; hand proximity via MediaPipe HandLandmarker); (4) parameter extraction; and (5) classification and reporting via an XGBoost classifier writing to local structured files. All processing is performed locally; no data are transmitted to external servers.

Hardware and Software Environment

The system was designed to operate on standard office hardware without a GPU. Minimum configuration: Intel Core i3 or equivalent, 4 GB RAM, 720p USB camera, 1366 × 768 display, Windows 10/11 (64-bit); these values define the deployment hardware class for which the system was designed and are engineering specifications rather than literature-derived parameters. Core dependencies, with the versions used in the present implementation, are OpenCV 4.10, MediaPipe (FaceLandmarker task), NumPy, pandas, SciPy, scikit-learn 1.9, XGBoost 3.3, and joblib, running under Python 3.14 with the Tkinter graphical user interface (Supplementary Fig. S1). All are open-source libraries distributed under MIT or Apache-2.0 licences; the versions listed are implementation configurations, not minimum requirements. The application is packaged as a standalone executable using PyInstaller, requiring no Python installation on the target machine.

Face and Eye Landmark Detection

Detection used the MediaPipe FaceLandmarker task, returning 478 three-dimensional facial landmarks per frame. Six landmarks were used per eye, consistent with the standard EAR formulation: left eye 33, 160, 158, 133, 153, 144; right eye 362, 385, 387, 263, 373, 380. MediaPipe was selected because it provides real-time monocular facial geometry estimation suitable for consumer hardware[36]. The detector was configured with $\min_f ace_{det_e} ction_{onf} nce = 0.5$ (the detector's default confidence setting; an implementation configuration). Because the MediaPipe Tasks C API does not support non-ASCII file paths, the model file (*face_l* and *marker.task*, 3.58 MB; model file size) is automatically copied to an ASCII temporary directory ($\%TEMP \frac{\%}{e} ye_{mod} e \frac{l}{-}$) prior to loading. If detection fails on a given frame, the detector returns a null result and the interface continues uninterrupted.

EAR Computation and Blink Detection

The eye aspect ratio was computed according to the standard six-point formulation[29]:

$$EAR = (\|p_2 - p_6\| + \|p_3 - p_5\|) / (2 \times \|p_1 - p_4\|)$$

An **individual baseline** was established as the median EAR over the initial frames of the session: the first third of the buffered EAR series, with a minimum of 10 and a maximum of 80 frames (implementation-specified configuration). A blink onset was registered when EAR fell below $0.7 \times$ baseline, and the blink was classified as complete when EAR fell below $0.5 \times$ baseline. These two threshold factors are design parameters of the system rather than literature-derived constants; their influence is quantified in the threshold sensitivity analysis below (Supplementary Table S3). This adaptive strategy was adopted to reduce fixed-threshold errors arising from variable face–camera distances and eyelid morphologies. Blink detection was performed on every acquired frame at the camera's native frame rate (30 fps target at 320 × 240; camera capture configuration) rather than on a decimated subset; this sampling strategy follows the frame-rate ablation evidence that event-time quantization error grows steeply below roughly 15 fps[26].

Eye-Rubbing Detection

Eye-rubbing behavior was detected using MediaPipe HandLandmarker, monitoring sustained proximity between fingertips and the periocular region: frames in which the fingertip–periocular distance satisfied a proximity criterion were flagged, and consecutive qualifying frames were aggregated into a single rubbing event. The proximity criterion itself is defined in the system implementation and is an engineering constant rather than a literature-derived value. Events were counted and normalized by acquisition duration, and per-frame rubbing flags were written to the ear.csv output, which accounts for the distinction between event counts (Table 3) and frame-level flags (Figure 2). Unlike the EAR module, hand-proximity detection was not part of the static self-validation battery; it was exercised end-to-end in the five camera sessions, in which rubbing events were detected at 4.55 and 2.70 events/min (Table 3).

Blink Parameter Extraction

Five parameters were computed per session (Table 1). The selection of these five dimensions was driven by the pathophysiological characteristics attributed to each etiology in the literature (Table 1, third column); we note explicitly that this mapping constitutes the **design premise** of the system rather than a finding of the present study.

Table 1. Definition and design rationale of the five extracted blink parameters.

Parameter	Unit	Definition	Design rationale (literature-based premise)
Blink frequency	blinks/min	Blink events per unit time	Elevated across etiologies; reference values vary with age and measurement conditions[5–7]
Complete blink rate	%	Proportion of blinks reaching complete eyelid closure	Reduction associated with compromised ocular surface protection and DED[9–10,31]
Blink amplitude	EAR units	Difference between open-eye EAR and minimum EAR during blink	Reflects depth of eyelid closure, quantified from the same six-point EAR formulation[29]
Coefficient of variation (CV)	–	SD of inter-blink intervals / mean	Elevation indicates rhythmic irregularity, associated with TD[15–16]
Eye-rubbing frequency	events/min	Eye-rubbing events per unit time	Elevation associated with AC[12–13]

Classification Model and Synthetic Training Data

Model. Classification used an XGBoost gradient-boosted tree classifier (*XGBoost* if *ier*) with the five parameters as features and four output classes (0 = normal, 1 = DED, 2 = AC, 3 = TD). Hyperparameters were: *objective* = $\mu < i: softprob, vm_4, \max_{d} epth = 4, \leq arn \in g_{rate} = 0.1, n_{estima} \rightarrow rs = 150, \subset samp \leq = 0.8, colsamp \leq_b ytree = 0.8, r$ and *om_{state}* = 42; these are implementation-specified values within the conventional ranges for XGBoost classifiers and were not tuned against any clinical outcome. No feature scaling was applied, as tree-based models are scale-invariant, and no class-weighting was required because the synthetic classes were balanced by construction.

Synthetic training data. Because the institutional ethics approval required for paediatric data collection had not yet been obtained, the classifier was trained on synthetic data so that the end-to-end pipeline could be implemented and exercised. The generation procedure was as follows. With *vm_{py}*. *r* and *om_{seed}*(42) (a reproducibility seed), 80 samples were drawn per class (320 in total); this sample size is a pragmatic choice made by the authors to exercise the pipeline with balanced classes and is not based on statistical power considerations. Each of the five features was sampled independently from a uniform distribution over a class-specific interval (Table 2). **Two properties of this generator are critical for interpretation:** (i) features were sampled independently, so no covariance structure or within-class correlation was modelled; and (ii) the intervals were specified a priori from the design premises in Table 1 and were substantially non-overlapping for the discriminative features, so the classes are close to linearly separable **by construction**. No measurement noise, missing data, or class-overlap model was incorporated.

Table 2. Parameter ranges used to generate synthetic training data (uniform distributions, independent sampling, n = 80 per class).

Class	Blink freq. (blinks/min)	Complete blink rate (%)	Blink amplitude	CV	Rubbing freq. (events/min)
Normal	U(12, 18)	U(88, 98)	U(0.15, 0.25)	U(0.15, 0.25)	U(0, 0.5)
Dry eye disease	U(22, 35)	U(40, 65)	U(0.12, 0.20)	U(0.25, 0.38)	U(0.5, 2.0)
Allergic conjunctivitis	U(18, 30)	U(70, 88)	U(0.13, 0.22)	U(0.30, 0.45)	U(3.0, 8.0)
Tic disorder	U(12, 38)	U(85, 98)	U(0.18, 0.30)	U(0.40, 0.60)	U(0, 1.0)

The intervals were specified by the authors (ophthalmologists) a priori on the basis of the design premises in Table 1, the literature values cited therein (normal blink-rate reference ranges[5–7]; reduced complete blink rate in DED[9,31]; elevated rubbing frequency in AC[12–13]; increased rhythm variability in TD[15–16]), the measured engineering discriminability of the system (open-eye EAR $\approx$ 0.30, closed-eye EAR $\approx$ 0.0125; Table 3), and clinical experience; individual discriminative gaps (e.g., rubbing frequency 3.0–8.0 for AC versus $<$ 2.0 for all other classes) have no direct literature values.

Training and evaluation on synthetic data. Data were split into training and held-out test sets (75%/25%, stratified by class, *r* and *om_{state}* = 42), yielding 240 training and 80 test records. Model performance on the synthetic test set was assessed using accuracy, weighted Cohen's κ , one-vs-rest ROC AUC, and 5-fold stratified cross-validation over the full synthetic dataset (*Strat if iedKFold, shuff* $\leq$ = *True*, *r* and *om_{state}* = 42). Feature importance was obtained from the fitted model's *feature_{imp}* or *tan_{ces}* attribute.

Retraining safeguard. A retraining module supports model updating with real clinical records. To prevent contamination, records whose remarks field carries a simulation marker are automatically excluded during data loading.

Treatment-Suggestion Module

The treatment-suggestion field displays category-specific advice generated from the predicted class: lid hygiene and lubricant recommendations for DED, antihistamine use and allergen avoidance for AC, behavioural intervention and paediatric referral for TD, and general eye-care advice for the normal class. In addition, when the measured blink frequency falls below 20 blinks/min, a quantitative prompt advises increasing blink rate to maintain tear film stability[28]. Suggestions are informational and are not intended to replace clinical judgement.

Data Management, Privacy, and Security

Each session writes three structured files to a per-session directory: $bl \in k_p \text{ arams. csv}$, $ear. \text{ csv}$ (per-frame timestamp, EAR, rubbing flag), and $rest. \text{ json}$. Records are stored in a local UTF-8 CSV database of 18 fields (implementation-specified database schema). Safeguards included local-only storage with no network transmission; anonymization using name abbreviations only; binding of each record to the acquisition machine identifier; a universally unique identifier per record for cross-device de-duplication; and explicit flagging of simulation records.

Functional Evaluation Design

Evaluation was conducted in three layers. **No patients and no children were involved at any stage.**

Layer 1 — Module self-validation. Eight items were evaluated in a headless environment (no camera): syntax compilation, module import integrity, detector initialization, open-eye EAR, closed-eye EAR, open/closed discriminability, blank-image robustness, and detection on a synthetic face image.

Layer 2 — Camera-based sessions. Five sessions were conducted using a USB camera under ordinary indoor lighting by a **single adult participant, one of the authors (an ophthalmologist and co-developer of the software)**, who volunteered for this functional evaluation and consented to the reporting of these details (N = 1, adult female, aged 39 years).

Sessions were terminated early relative to the 180 s default because the purpose was functional verification rather than data collection. Frame rate was treated as an explicit evaluation dimension because low-cost USB cameras exhibit variable frame rate; the sensitivity analysis is described in "Performance Metrics, Reporting, and Sensitivity Analysis". Extracted parameters, four-class probabilities, and predicted class were recorded for each session. In every session the application processed each acquired frame (the EAR time series contained one row per frame, 117–523 rows), indicating that frame throughput at least matched camera acquisition; per-frame processing latency was not formally benchmarked.

Layer 3 — Boundary testing. Twenty synthetic records (SIM-01 to SIM-20), five per class, were constructed with extreme boundary values (e.g., blink frequency 4.2–42.0 blinks/min; complete blink rate 95.0–100.0%), chosen by the authors to probe detection and classification behaviour at extreme parameter combinations. These records were assigned a dedicated machine identifier (*SIMULATED*) and UUID prefix (*SIM*) and were labeled in the remarks field as simulation data excluded from training.

Literature Search

The targeted search referenced in the Introduction was conducted in OpenAlex (queried August 2026) using title-and-abstract queries that combined blink terms (*blink*, *blinking*, *eyelid*) with machine-learning terms (*machine learning*, *deep learning*, *neural network*, *XGBoost*, *classifier*), multi-class terms (*multi-class*, *differential diagnosis*, *discriminat*), and population terms (*child*, *children*, *pediatric*). Results were screened for relevance to video-derived blink parameters in paediatric populations. This procedure was **not** a systematic review: no protocol was registered, no dual-reviewer screening was performed, and grey literature and non-English databases were not covered.

Performance Metrics, Reporting, and Sensitivity Analysis

Because no reference standard was available, clinical accuracy metrics against a diagnostic gold standard were not computed. Evaluation endpoints were: (i) completion of all module self-validation items; (ii) successful extraction of all five parameters and generation of valid four-class outputs in every session; and (iii) absence of system failure across the boundary-value range.

The future clinical validation of this prediction model should follow the **TRIPOD** reporting guidance[34] and the **STARD** guidance for diagnostic accuracy studies[35]; the present functional evaluation is reported with those endpoints in view.

Threshold sensitivity analysis. Because the blink-detection thresholds ($0.7 \times$ and $0.5 \times$ baseline) are user-selected parameters, their influence on blink-event counts and on the derived parameters was examined. For each of the five camera-based sessions, the onset threshold was varied over $\pm 20\%$ of the default (0.56 , 0.63 , 0.70 , 0.77 , and $0.84 \times$ baseline) with the completeness threshold held at $0.5 \times$ baseline, and the completeness threshold was varied over $\pm 20\%$ (0.40 , 0.45 , 0.50 , 0.55 , and $0.60 \times$ baseline) with the onset threshold held at $0.7 \times$ baseline. Event detection and parameter extraction were recomputed for every combination using the same pipeline that generated Table 3 (results in Supplementary Table S3).

Frame-rate sensitivity analysis. Because a blink lasts on the order of a few hundred milliseconds — high-speed video-oculography measured a closing phase of 76 ± 2 ms and total voluntary-blink durations of 572 ± 25 ms, while typical spontaneous blinks are shorter, approximately one-third of a second[37] — the frame rate of a low-cost camera can materially affect whether a blink is sampled with enough frames to be detected. The actual frame rate of each session was therefore derived from the recorded EAR timestamp files as $(\text{frames} - 1)/\text{duration}$; the five sessions spanned 7.5–30 fps, the range typical of the deployment hardware class. To quantify the effect of temporal sampling density on the detection logic, the two sessions recorded at 28–30 fps were decimated by time-uniform linear interpolation to 15, 10, and 7.5 fps, and event detection and parameter extraction were recomputed on each decimated series with the same pipeline that generated Table 3 (results in Supplementary Table S4). Decimation isolates the effect of frame rate on the detection logic; it does not reproduce native low-frame-rate acquisition, whose exposure and sensor characteristics differ.

Use of Large Language Models

During the preparation of this work, the authors used large language model (LLM)-assisted tools for language editing and writing support. The authors reviewed, verified, and take full responsibility for the content of this manuscript. No AI-generated content was used for the data, the analyses, or the conclusions reported here.

Declarations

Competing Interests

The authors declare no competing interests.

Author Contribution

L.Z. conceived the study, designed and implemented the system, performed the camera-based functional evaluation, and drafted the main manuscript text. C.L., F.Y., and C.P. contributed to the clinical design and to critical revision for important intellectual content. Q.L. and Y.Z. contributed to data management and data verification. Analysis and interpretation: all authors. All authors read and approved the final manuscript.

Acknowledgements

This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The synthetic data generator specification (complete specification in Table 2) and the functional-evaluation outputs analysed in this study (five camera-based sessions reported in Table 3 and Supplementary Tables S1–S4) are available from the corresponding author on reasonable request, and will be deposited in a public repository (Figshare) upon publication.

References

1. Szejko, N. et al. European clinical guidelines for Tourette syndrome and other tic disorders—version 2.0. Part I: assessment. *Eur. Child. Adolesc. Psychiatry*. **31**, 383–402. <https://doi.org/10.1007/s00787-021-01842-2> (2021).
2. Leonardi, A., Modugno, R. L. & Salami, E. Allergy and dry eye disease: a two-way street? *Ocul Immunol. Inflamm*. **29**, 1168–1176. <https://doi.org/10.1080/09273948.2020.1841804> (2021).
3. Tang, N. et al. Clinical characteristics of tic disorders in children and adolescents with the chief complaint of abnormal blinking. *Front. Psychiatry*. **16**, 1553358. <https://doi.org/10.3389/fpsy.2025.1553358> (2025).
4. Wolffsohn, J. S. et al. TFOS DEWS II Diagnostic Methodology report. *Ocul Surf*. **15**, 539–574. <https://doi.org/10.1016/j.jtos.2017.05.001> (2017).
5. Jeon, Y. J. et al. Relationship between the partial blink rate and ocular surface parameters. *Int. Ophthalmol*. **41**, 2601–2608. <https://doi.org/10.1007/s10792-021-01819-9> (2021).
6. Zheng, Q. et al. Impact of incomplete blinking analyzed using a deep learning model with the Keratograph 5M in dry eye disease. *Transl Vis. Sci. Technol*. **11**, 38. <https://doi.org/10.1167/tvst.11.3.38> (2022).
7. Wang, M. T. M., Power, B., Xue, A. L. & Craig, J. P. Blink completeness and rate in dry eye disease: an investigator-masked, prospective registry-based, cross-sectional, prognostic study. *Cont. Lens Anterior Eye*. **48**, 102369. <https://doi.org/10.1016/j.clae.2025.102369> (2025).
8. Conelea, C. A. et al. Automated quantification of eye tics using computer vision and deep learning techniques. *Mov. Disord*. <https://doi.org/10.1002/mds.29593> (2023).
9. Xu, X. et al. Diagnosing tic disorders from videos using multi-phase learning. *Comput. Electr. Eng*. **123**, 110216. <https://doi.org/10.1016/j.compeleceng.2025.110216> (2025).
10. Hong, J., Shin, J. S., Choi, J. & Ko, M. Robust eye blink detection using dual embedding video vision transformer. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp. 6362–6372 (2024). <https://doi.org/10.1109/wacv57701.2024.00625>
11. Verrel, J. et al. Eye blinking abnormalities in Tourette syndrome: blink more or blink differently? *Parkinsonism Relat. Disord*. **142**, 108121. <https://doi.org/10.1016/j.parkreldis.2025.108121> (2026).
12. Qian, S. et al. Morphological changes in the meibomian gland in children with tic disorders. *Quant. Imaging Med. Surg*. **13**, 6374–6383. <https://doi.org/10.21037/qims-22-390> (2023).
13. Xu, Z. et al. A smartphone-based blink training application for alleviating dry eye signs and symptoms. *npj Digit. Med*. **8**, 703. <https://doi.org/10.1038/s41746-025-02053-8> (2025).

14. Xu, Z. et al. Artificial intelligence for anterior segment diseases: a review of potential developments and clinical applications. *Ophthalmol. Ther.* **12**, 1439–1455. <https://doi.org/10.1007/s40123-023-00690-4> (2023).
15. Li, Y. Y. S., Vardhanabhuti, V., Tsougenis, E., Lam, W. C. & Shih, K. C. A proposed framework for machine learning-aided triage in public specialty ophthalmology clinics in Hong Kong. *Ophthalmol. Ther.* **10**, 703–713. <https://doi.org/10.1007/s40123-021-00405-7> (2021).
16. Kartynnik, Y., Ablavatski, A., Grishchenko, I. & Grundmann, M. Real-time facial surface geometry from monocular video on mobile GPUs. *arXiv:1907.06724* (2019). <https://doi.org/10.48550/arXiv.1907.06724>
17. Soukupová, T. & Čech, J. Real-time eye blink detection using facial landmarks. In *Proceedings of the 21st Computer Vision Winter Workshop (CVWW 2016)* (2016).
18. Leckman, J. F. et al. The Yale Global Tic Severity Scale. *J. Am. Acad. Child Adolesc. Psychiatry* **28**, 566–573 PMID: 2768151 (1989).
19. Yang, B. et al. Quantitative evaluation of lipid layer thickness and blinking in children with allergic conjunctivitis. *Graefes Arch. Clin. Exp. Ophthalmol.* **259**, 3479–3488. <https://doi.org/10.1007/s00417-021-05199-9> (2021).
20. Niu, X., Qin, A., Zhang, W., Xu, M. & Yang, Y. The effects of eye rubbing on corneal topography and biomechanics in children with allergic conjunctivitis. *Transl. Vis. Sci. Technol.* **14**, 18. <https://doi.org/10.1167/tvst.14.10.18> (2025).
21. Zametkin, A. J., Stevens, J. R. & Pittman, R. S. Ontogeny of spontaneous blinking and of habituation of the blink reflex. *Ann. Neurol.* **5**, 453–457. <https://doi.org/10.1002/ana.410050509> (1979).
22. Lavezzo, M. M., Schellini, S. A. & Padovani, C. R. Eye blink in newborn and preschool-age children. *Acta Ophthalmol.* <https://doi.org/10.1111/j.1600-0420.2007.00969.x> (2007).
23. Doughty, M. J. & Naase, T. Further analysis of the human spontaneous eye blink rate by a cluster analysis-based approach to categorize individuals with 'normal' versus 'frequent' eye blink activity. *Eye Contact Lens.* **32**, 294–299. <https://doi.org/10.1097/01.icl.0000224359.32709.4d> (2006).
24. Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. M. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* **350**, g7594. <https://doi.org/10.1136/bmj.g7594> (2015).
25. Hafsi, S. et al. Spontaneous blink features extracted from video for multi-class classification of Parkinson's disease. *Proceedings of the 2026 Symposium on Eye Tracking Research and Applications* (2026). <https://doi.org/10.1145/3797246.3803044>
26. Oganov, A., Yazdanpanah, G., Jabbehdari, S., Belamkar, A. & Pflugfelder, S. C. Dry eye disease and blinking behaviors: a narrative review of methodologies for measuring blink dynamics and inducing blink response. *Ocul Surf.* **29**, 166–174. <https://doi.org/10.1016/j.jtos.2023.05.011> (2023).
27. Riley, R. D. et al. Calculating the sample size required for developing a clinical prediction model. *BMJ* **368**, m441 (2020). <https://doi.org/10.1136/bmj.m441>
28. Silvey, S. & Liu, J. Sample size requirements for popular classification algorithms in tabular clinical data: empirical study. *J. Med. Internet Res.* **26**, e60231. <https://doi.org/10.2196/60231> (2024).
29. El Barche, F. Z. et al. Automated tear film break-up time measurement for dry eye diagnosis using deep learning. *Sci. Rep.* **14**, 11723. <https://doi.org/10.1038/s41598-024-62636-5> (2024).
30. Chawuthai, R. et al. Utilizing deep learning from mobile phone photos for early detection of horizontal strabismus: a screening approach. *Sci. Rep.* **16**, 18278. <https://doi.org/10.1038/s41598-026-48893-6> (2026).
31. Chai, Y. et al. Identification of a blink frequency threshold for maintaining tear film stability in young adults with dry eye symptoms. *Sci. Rep.* **15**, 42321. <https://doi.org/10.1038/s41598-025-26424-z> (2025).
32. Kwon, K. A. et al. High-speed camera characterization of voluntary eye blinking kinematics. *J. R. Soc. Interface.* **10**, 20130227. <https://doi.org/10.1098/rsif.2013.0227> (2013).
33. Jie, Y., Sella, R., Feng, J., Gomez-Jenkins, M. & Afshari, N. A. Evaluation of incomplete blinking as a measurement of dry eye disease. *Ocul Surf.* **17**, 440–446. <https://doi.org/10.1016/j.jtos.2019.05.007> (2019).
34. Wang, M. T. M. et al. Impact of blinking on ocular surface and tear film parameters. *Ocul Surf.* **16**, 424–429. <https://doi.org/10.1016/j.jtos.2018.06.001> (2018).
35. Wu, Y. et al. Blinking characteristics analyzed by a deep learning model and the relationship with tear film stability in children with long-term use of orthokeratology. *Front. Cell. Dev. Biol.* **12**, 1517240. <https://doi.org/10.3389/fcell.2024.1517240> (2025).
36. Chen, L. et al. High incidence of dry eye in young children with allergic conjunctivitis in Southwest China. *Acta Ophthalmol.* **94**, e727–e730. <https://doi.org/10.1111/aos.13093> (2016).
37. Bossuyt, P. M. et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* **351**, h5527. <https://doi.org/10.1136/bmj.h5527> (2015).

Figures

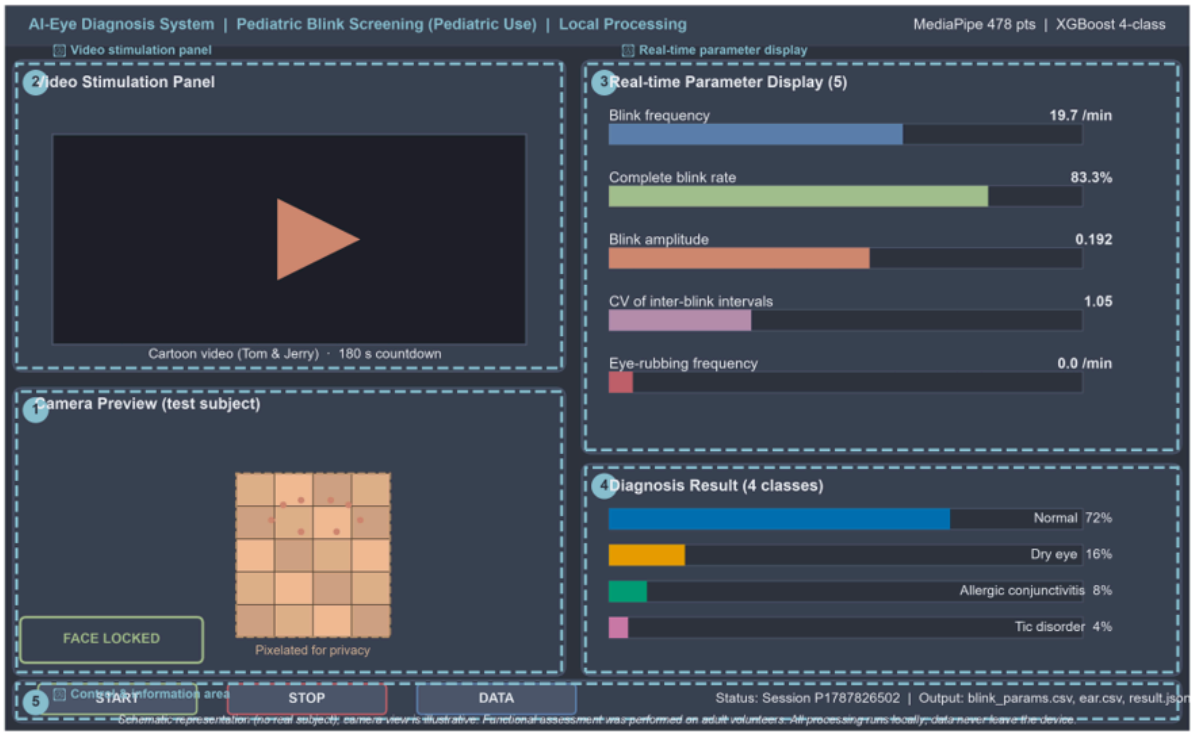


Figure 1

Graphical user interface of the AI-Eye Diagnosis System. 1 Camera preview with landmark overlay and pixelated facial region (identity protection); 2 video stimulation panel; 3 real-time parameter display; 4 diagnosis result panel with four-class probability output (normal, DED, AC, TD); 5 control and information area (start/countdown/treatment suggestion field). The interface is presented in English; the title bar states the intended pediatric use of the system.

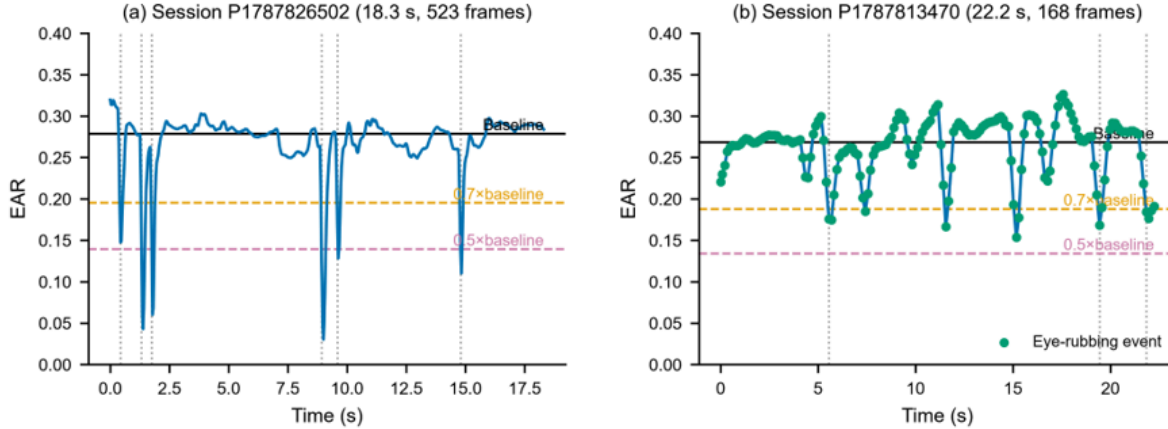


Figure 2

Representative EAR time series from two camera-based sessions, showing the individually adapted baseline (solid line) with the $0.7 \times \text{baseline}$ onset threshold and $0.5 \times \text{baseline}$ completeness threshold. Vertical dashed lines mark blink onsets detected with the system's event logic; green dots mark frames classified as eye-rubbing. (a) Session P1787826502 (18.3 s, 523 frames; six blinks detected); (b) Session P1787813470 (22.2 s, 168 frames; three blinks per the system's blink statistics reported in Table 3; the continuous green band indicates that most frames of this short session were classified as eye-rubbing).

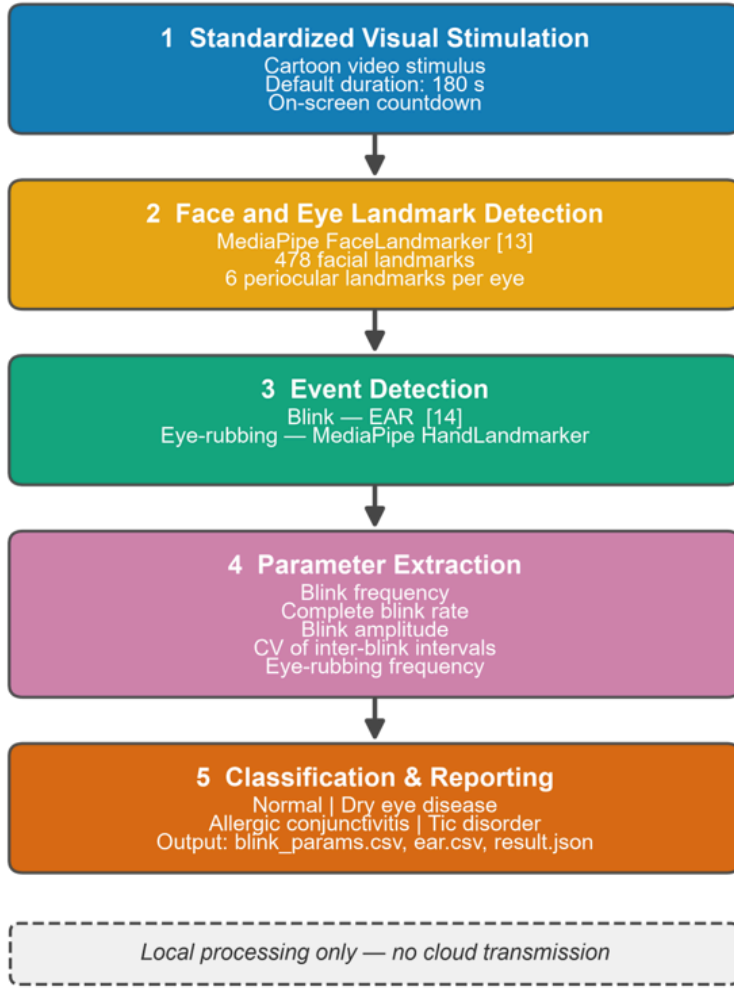


Figure 3

System architecture and processing workflow of the AI-Eye Diagnosis System. The pipeline comprises five sequential stages: (1) standardized visual stimulation; (2) face and eye landmark detection via MediaPipe FaceLandmarker[36]; (3) blink and eye-rubbing event detection using the EAR algorithm[29] and MediaPipe HandLandmarker; (4) extraction of five blink parameters (blink frequency, complete blink rate, blink amplitude, coefficient of variation, eye-rubbing frequency); and (5) classification (XGBoost, four classes: normal, DED, AC, TD) and local reporting.

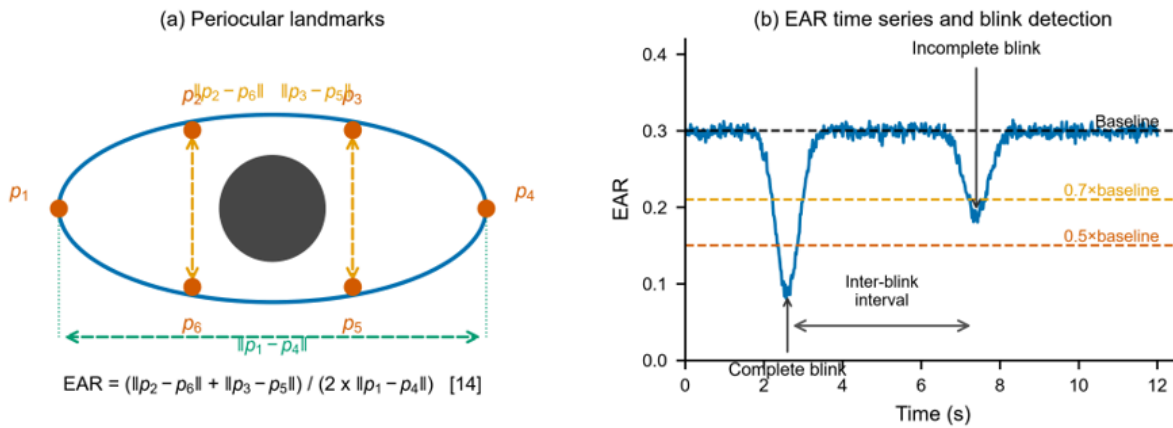


Figure 4

(a) Periocular landmarks used for EAR computation: the six landmarks p_1 – p_6 per eye and the Euclidean distances $\|p_2 - p_6\|$, $\|p_3 - p_5\|$, and $\|p_1 - p_4\|$; $EAR = (\|p_2 - p_6\| + \|p_3 - p_5\|) / (2 \times \|p_1 - p_4\|)$ [29]. (b) EAR time series and blink detection: the individually adapted baseline (solid line) with the $0.7 \times \text{baseline}$ onset threshold (orange) and $0.5 \times \text{baseline}$ completeness threshold (vermillion). A complete blink is defined by an EAR deflection reaching below $0.5 \times \text{baseline}$; an

incomplete blink reaches the $0.7 \times$ baseline onset threshold but not the $0.5 \times$ baseline completeness threshold. The inter-blink interval is the time between successive blink onsets.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [03SupplementaryInformationv1.26.docx](#)
- [SupplementaryFigureS1TechnologyStack.tif](#)