Automated Tear Film Break-Up Detection Using Classical Image Processing

Helin Melisa Ergezen
Masters in Computer Vision
University of Porto
Portugal
up202502605@up.pt

Gagandeep Kaur
Masters in Computer Vision
University of Santiago de Compostela
Spain
gagandeepkaur.gagandeep@rai.usc.es

***Abstract*—Tear film break-up time (BUT) is a key diagnostic indicator of dry eye disease, but manual evaluation remains subjective and time-consuming. This work presents a fully automated image-processing pipeline for fluorescein tear film break-up detection that integrates iris localization, eyelash removal, green-channel analysis, adaptive thresholding, morphological cleaning, and polynomial-based break-up estimation. The system is evaluated on eight clinical sequences comprising 90 frames with ground-truth labels. The proposed approach achieves an overall accuracy of 82.22%, with an F1-score of 0.765, demonstrating the viability of an interpretable hardware-independent tear film assessment without requiring large training datasets. The results suggest that classical, explainable image-processing methods remain competitive with machine-learning-based systems, especially in low-data clinical settings.**

***Index Terms*—Tear film break-up time, dry eye disease, fluorescein imaging, classical image processing, segmentation, ophthalmology automation.**

I. Introduction

Dry eye disease (DED) affects millions of people around the world and is often diagnosed by tear film break-up time (BUT). In the clinical workflow, fluorescein dye is applied to the ocular surface and clinicians visually observe the emergence of dark spots where the tear film becomes unstable. Despite its diagnostic importance, manual evaluation suffers from inter-observer variability and inconsistency caused by subjective interpretation, variations in illumination, and rapid tear film dynamics.

Automating BUT detection can substantially improve standardization, reduce clinical workload, and support earlier diagnosis. However, most existing systems require either specialized hardware or rely on machine learning models that require extensive annotated datasets. This work proposes a fully automated classical image-processing pipeline that requires no training data, is robust across sequences, and remains interpretable for clinicians.

II. Related Work

Computer-aided tear film analysis has attracted considerable attention as researchers attempt to overcome the limitations of subjective manual assessment. Fluorescein-based segmentation approaches represent the first automation attempts. Penedo et al. proposed a semi-automated threshold-based system that achieved moderate correlation with clinical BUT measurements [1]. Alonso-Caneiro et al. applied texture-based analysis to high-speed videokeratoscopy images, demonstrating sensitivity to tear film irregularities but reporting limited generalizability across different imaging conditions [2].

With the rise of machine learning, deep learning approaches have been applied to related ocular imaging problems. Kuo et al. employed convolutional neural networks to classify the dysfunction of the meibomian gland, achieving high precision but requiring large labeled datasets and sacrificing the interpretability as an essential component for clinical acceptance [3].

Interferometry-based methods such as LipiView II offer hardware-driven automated analysis of tear film lipid layers [4]. Although clinically valuable, such devices require specialized equipment and do not directly assess fluorescein break-up patterns, which remain the clinical gold standard.

Existing solutions highlight the feasibility of automated tear film assessment, but reveal gaps in interpretability, accessibility, and data requirements. The method proposed in this work addresses these challenges using a fully classical, explainable, hardware-independent pipeline.

III. Methodology

The complete processing workflow consists of a fully classical image-processing pipeline that starts from raw fluorescein images and outputs frame-level break-up (BU) decisions. The main stages are: (i) contrast-enhanced preprocessing of the green channel, (ii) robust iris localization via multi-cue edge detection and Hough Transform, (iii) frame alignment to a reference iris, (iv) eyelid/eyelash removal using sequence-specific masks, (v) dark-region segmentation inside the Region of Interest (ROI), and (vi) polynomial modeling of the break-up evolution curve for automatic BUT detection.

A. Preprocessing and Green Channel Extraction

For each frame, we extract the green channel, which maximizes the visibility of the fluorescein signal. Let $I_{\text{BGR}}(x, y)$ denote the input image. We take $G(x, y)$ as the green component and enhance its local contrast by adjusting the Contrast Limited Adaptive histogram (CLAHE). This operation

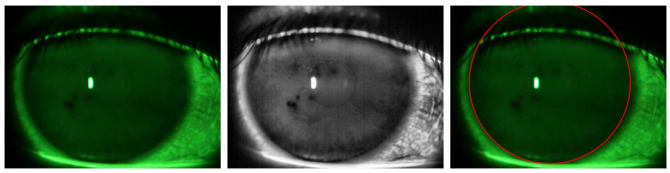

Fig. 1: Iris detection pipeline for sequence `but1`. Left: Original fluorescein frame. Center: CLAHE-enhanced green-channel image used for edge detection. Right: Final iris localisation obtained with the multi-cue Hough Transform, shown as a red circle overlaid on the green frame.

Fig. 2: Eyelid and eyelash removal in sequence `but7`. Left: Circular iris mask centred on the detected iris. Center: Binary mask implementing the sequence-specific top and bottom cuts to remove eyelid and eyelash regions. Right: Final ROI obtained by applying the cut mask to the circular iris mask, isolating the exposed corneal surface used for tear film analysis.

improves visibility of subtle intensity variations while avoiding saturation in highly illuminated areas.

A 5×5 Gaussian filter is then applied to $G(x, y)$ to suppress sensor noise and small artifacts before edge detection. This combination (green-channel projection, CLAHE, and Gaussian smoothing) provides a more homogeneous input for the subsequent iris and break-up segmentation stages.

B. Iris Detection via Multi-Cue Edge Detection

The iris detection process employs three complementary edge detection methods to ensure robust circular boundary localization:

- **Adaptive thresholding**: An adaptive Gaussian threshold is applied to $G(x, y)$, generating a binary map where the dark iris ring is highlighted relative to its local neighbourhood.
- **Otsu global thresholding**: Otsu's method is used to obtain a global threshold that maximises between-class variance, producing a complementary segmentation of dark and bright regions.
- **Canny edge detection**: Canny is applied to capture gradient-based edges, with non-maximum suppression and hysteresis providing thin, well-localised contours.

Each of the resulting binary maps is refined using morphological operations: elliptical closing and opening followed by dilation. These steps remove small holes and bridges in the iris boundary while suppressing isolated noise.

Circular iris candidates are then extracted using the circular Hough Transform over a restricted radius range (e.g. 350–450 pixels) and several inverse resolution values ($dp \in \{1.0, 1.2, 1.5\}$) and accumulator thresholds (param2 $\in \{15, 20, 30\}$). All detected circles across the three edge maps are pooled into a single candidate set.

Each candidate circle (c_x, c_y, r) is scored using two cues:

1) **Centering score**: The distance between (c_x, c_y) and the image center is penalised, favouring iris estimates located near the expected eye centre.
2) **Edge-consistency score**: Intensity values are sampled along the circle in the smoothed green channel; a low standard deviation along the contour indicates a stable, well-defined iris boundary.

A weighted combination of these scores is used to select the best iris candidate. This multi-cue scoring proves more robust than relying on the Hough accumulator alone, particularly in sequences with weak contrast or reflections.

C. Frame Alignment to a Reference Iris

Even small head or eye movements can translate the iris between frames and introduce artificial oscillations in the break-up curve. To reduce this effect, the first successfully processed frame of each sequence is used as reference. Its iris centre $(c_x^{\text{ref}}, c_y^{\text{ref}})$ and radius r^{ref} define the reference geometry.

For each subsequent frame, an iris is detected as described above, yielding $(c_x^{\text{curr}}, c_y^{\text{curr}}, r^{\text{curr}})$. The frame is then translated by

$$T = (t_x, t_y) = (c_x^{\text{ref}} - c_x^{\text{curr}},\ c_y^{\text{ref}} - c_y^{\text{curr}}),$$

using an affine warp so that the iris centre aligns with the reference frame. After alignment, iris detection is optionally re-run on the aligned frame to refine the centre and radius; if this fails, the reference parameters are reused. This alignment step stabilises the ROI across time and reduces spurious fluctuations in the measured break-up percentage.

D. Eyelid and Eyelash Removal

Eyelids and eyelashes introduce dark structures that can be mistakenly classified as break-up. To restrict the analysis to the exposed corneal region, a circular mask is first created using the detected iris radius. Then, sequence-specific vertical cuts are applied:

- A fraction of the circle radius is removed from the top (top cut) to discard upper eyelid and eyelashes.
- A (typically smaller) fraction is removed from the bottom (bottom cut) to exclude lower eyelid structures.

These percentages are stored per sequence and tuned empirically, reflecting different eyelid positions and camera setups. The resulting binary mask defines the ROI where tear film analysis is performed, as illustrated in Fig. 2. Alignment ensures that the same anatomical region is consistently analysed across frames.

E. Black Threshold Estimation in the ROI

Break-up regions appear as dark, fluorescein-depleted patches within the ROI. Instead of adopting a fixed or purely

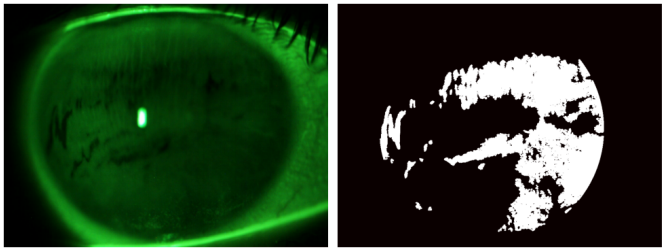

Fig. 3: Example of tear film break-up segmentation for the last frame of sequence `but2`. Left: Original fluorescein frame. Right: Corresponding break-up map inside the ROI after percentile-based black-thresholding, morphological opening and small-region removal (white = detected break-up pixels).

global threshold, we estimate a *data-driven* black threshold from the first (baseline) frame of each sequence.

Given the first masked green-frame $G_0(x, y)$ and ROI mask $M(x, y)$, we optionally erode M to exclude boundary artifacts. The set of valid pixels is

$$\mathcal{V} = \{G_0(x, y) \mid M(x, y) = 1\}.$$

The black threshold T_{black} is defined as a low percentile (e.g. 3%) of the intensity distribution over $\mathcal{V}$. This strategy adapts to global illumination and camera gain while focusing on the darkest regions that are likely to correspond to early break-up.

F. Break-Up Segmentation and Morphological Cleaning

For each aligned frame i in the sequence, the green channel $G_i(x, y)$ is masked with $M(x, y)$ and compared to the threshold T_{black}:

$$B_i(x, y) = \begin{cases} 1, & \text{if } G_i(x, y) < T_{\text{black}} \text{ and } M(x, y) = 1, \\ 0, & \text{otherwise.} \end{cases}$$

To suppress isolated noise and very small artifacts, a morphological opening with an elliptical structure element is applied to B_i. Connected components are then extracted, and components below a minimum area (e.g. 100 pixels) are discarded. The remaining binary map encodes candidate tear film break-up regions, as shown in Fig. 3.

The percentage of break-up within the ROI is computed as

$$\text{BreakUp\%}(i) = \frac{\sum_{x,y} B_i(x, y)}{\sum_{x,y} M(x, y)} \times 100, \tag{1}$$

yielding an evolution curve over time for each sequence.

G. Polynomial Curve Fitting and Adaptive BUT Detection

The raw evolution curve can still exhibit frame-to-frame fluctuations due to residual noise and subtle motion. To obtain a smoother representation, a second-degree polynomial $p(i)$ is fit to the sequence of BreakUp%(i) values using least-squares regression.

Let $p_{\min}$ and $p_{\max}$ denote the minimum and maximum of the fitted curve, and $\Delta p = p_{\max} - p_{\min}$ its dynamic range. Sequences with very small Δp are considered non-discriminative and no BUT is reported.

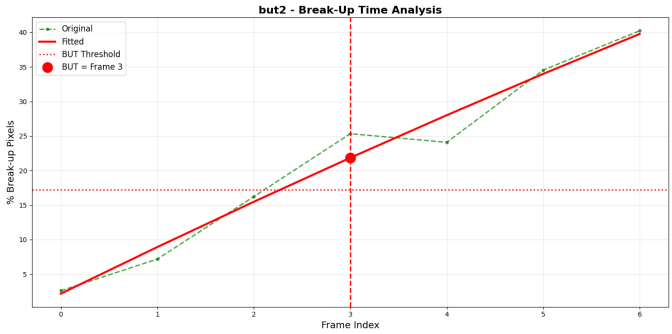

Fig. 4: Break-up evolution curve for sequence `but2`. The green dashed line shows the raw frame-wise percentage of break-up pixels, while the solid red line represents the fitted second-degree polynomial. The horizontal dotted line indicates the automatically selected adaptive threshold, and the vertical dashed line marks the detected BUT at frame index 3, highlighted by the red marker.

For the remaining sequences, the derivative of the fitted curve is used to estimate the steepness of the break-up transition. If the maximum slope exceeds a predefined threshold, the decision threshold is set closer to $p_{\min}$ (lower percentage) to detect early, abrupt break-up. For more gradual transitions, the threshold is shifted closer to $p_{\max}$, enforcing a stronger increase before labeling a frame as post-break-up. Formally, the BUT frame is defined as the first frame index i for which

$$p(i) > p_{\min} + \alpha \Delta p,$$

where the factor α is automatically adapted from the curve slope.

This adaptive strategy, illustrated in Fig. 4, allows the method to handle both sharp and slowly evolving break-up patterns with a single set of global parameters.

IV. Results

Eight fluorescein sequences (90 frames) with ground truth labels were used for evaluation. Global parameters (black threshold percentile, morphological kernel size, minimum region area, polynomial degree) remained constant across sequences; only eyelid/eyelash cut percentages required per-sequence tuning.

A. Per-Sequence Performance

Table I summarizes detection performance for each sequence. Sequences with good contrast and stable illumination (`but2`, `but5`, `but8`) achieve perfect accuracy, where the percentile-based threshold and morphological cleaning produce clear separation between pre- and post-break-up frames. Challenging sequences (`but4`, `but6`, `but7`) with partial occlusions or strong shadows show more frequent misclassifications. In `but4` and `but6` the fitted evolution curve does not exhibit a sufficiently sharp increase to cross the adaptive

TABLE I: Per-Sequence Evaluation Metrics

Seq	Total	Correct	Acc (%)	Prec	Rec	F1
but1	14	12	85.71	1.000	0.714	0.833
but2	7	7	100.00	1.000	1.000	1.000
but3	9	7	77.78	1.000	0.667	0.800
but4	13	10	76.92	0.000	0.000	0.000
but5	6	6	100.00	1.000	1.000	1.000
but6	6	5	83.33	0.000	0.000	0.000
but7	24	16	66.67	0.364	0.800	0.500
but8	11	11	100.00	1.000	1.000	1.000
Overall	**90**	**74**	**82.22**	**0.788**	**0.743**	**0.765**

threshold, so the system never labels any frame as break-up. As a result, all ground-truth BU frames become false negatives and there are no true positives, which by definition yields precision, recall and F1 equal to zero despite moderately high accuracy driven by correctly classified non-break-up frames. In contrast, `but7` exhibits high recall but moderate precision, indicating over-segmentation of illumination artifacts as break-up.

Overall, the method maintains stable behavior across heterogeneous sequences with a single global parameter set. The only per-sequence tuning is the eyelash-cut configuration, which is interpretable and directly related to visible eyelid position.

B. Overall Accuracy and Summary Metrics

Table I presents aggregated system performance. The method achieves 82.22% accuracy with precision of 0.788, recall of 0.743, and F1 score of 0.765, correctly classifying most frames while maintaining balance between false alarms and missed detections. Polynomial fitting and adaptive thresholding are crucial: without smoothing, frame-wise percentages fluctuate due to alignment errors or noise, making fixed thresholds sensitive to outliers. By modeling global trends and adapting thresholds to curve characteristics, the system becomes robust to local variations while detecting sustained break-up increases.

C. Confusion Matrix Analysis

Table II shows the confusion matrix. The method correctly identifies the most stable and unstable frames. False positives occur when eyelid shadows or reflections survive morphological cleaning, or when accumulated artifacts cause premature threshold crossing. False negatives arise in sequences with gradual break-up remaining close to the black threshold, producing modest curve increases that don't reach detection levels. Error distribution reflects known fluorescein imaging challenges: non-uniform illumination, eyelid interference, and subtle patterns. Importantly, each error type traces to specific processing stages, preserving interpretability and suggesting improvement directions.

V. Discussion

The proposed pipeline demonstrates strong performance using only classical image processing. Key advantages include:

TABLE II: Confusion Matrix

	Pred No BU	Pred BU
GT 0	48	7
GT 1	9	26

(1) no training data required, making it suitable for data-scarce clinical environments; (2) full interpretability with transparent processing stages; (3) hardware independence requiring only standard RGB cameras; (4) computational efficiency enabling real-time processing; and (5) robustness through adaptive thresholding across diverse sequences.

Primary limitations include sensitivity to iris detection failures in severe reflections, illumination variability causing local artifacts, and reduced sensitivity to gradual break-up patterns. Future work should focus on improved tracking algorithms, automated eyelid segmentation to eliminate manual parameter tuning, temporal smoothing for consistency, and large-scale clinical validation studies.

VI. Conclusion

This work presents a semi-automated, explainable, and hardware-independent method for fluorescein tear film break-up detection. Achieving 82.22% accuracy without learning-based components, the system provides a promising foundation for clinical integration, particularly in settings lacking large annotated datasets or specialized equipment. The interpretability of each processing stage facilitates clinical validation and trust, while the computational efficiency enables practical deployment. Future work will focus on addressing current limitations through improved tracking algorithms, automated ROI refinement, and large-scale clinical validation studies.

References

[1] M. G. Penedo, B. Remeseiro, L. Ramos, N. Barreira, C. García-Resúa, E. Yebra-Pimentel, and A. Mosquera, "Automatization of dry eye syndrome tests," in *Image Analysis and Modeling in Ophthalmology*, E. Y. K. Ng, U. R. Acharya, J. S. Suri, and A. Campilho, Eds. Boca Raton, FL: CRC Press, 2014, pp. 293–320.

[2] D. Alonso-Caneiro, D. H. Szczesna-Iskander, D. R. Iskander, S. A. Read, and M. J. Collins, "Application of texture analysis in tear film surface assessment based on videokeratoscopy," *Journal of Optometry*, vol. 6, no. 4, pp. 185–193, 2013.

[3] J. Wang, T. N. Yeh, R. Chakraborty, S. X. Yu, and M. C. Lin, "A deep learning approach for meibomian gland atrophy evaluation in meibography images," *Translational Vision Science & Technology*, vol. 8, no. 6, p. 37, 2019.

[4] TearScience, Inc., "LipiView® II Ocular Surface Interferometer," Traditional 510(k) Premarket Notification K152869, U.S. Food and Drug Administration, 2015.