



DEDetector: Smartphone-Based Noninvasive Screening of Dry Eye Disease

VAIBHAV GANATRA and SOUMYASIS GUN, Microsoft Research, India

PALLAVI JOSHI, ANAND BALASUBRAMANIAM, and KAUSHIK MURALI, Sankara Eye Hospital, India

NIPUN KWATRA and MOHIT JAIN, Microsoft Research, India

Dry Eye Disease (DED) is an eye condition characterized by abnormalities in tear film stability. Despite its high prevalence, screening for DED remains challenging, primarily due to the invasive nature of most diagnostic tests. The Fluorescein Break-Up Time (FBUT) test, which involves instilling fluorescein dye in the eye, is a widely used method for assessing tear film stability. In this work, we propose *DEDetector*, a low-cost, smartphone-based automated Non-Invasive Break-Up Time (NIBUT) measurement for DED screening. Utilizing a 3D-printed Placido ring attachment on a smartphone's camera, *DEDetector* projects concentric rings onto the cornea, capturing a video for subsequent analysis using our proposed video processing pipeline to identify tear film stability. We conducted a real-world evaluation on 46 eyes, comparing *DEDetector* with the traditional FBUT method. *DEDetector* achieved a sensitivity of 77.78% and specificity of 82.14%, outperforming FBUT.

CCS Concepts: • Human-centered computing → Ubiquitous and mobile computing; • Applied computing → Consumer health.

Additional Key Words and Phrases: Tear Break-up Time, TBUT, Non-invasive, NIBUT, Medical, Diagnosis, Evaluation.

ACM Reference Format:

Vaibhav Ganatra, Soumyasis Gun, Pallavi Joshi, Anand Balasubramaniam, Kaushik Murali, Nipun Kwatra, and Mohit Jain. 2024. DEDector: Smartphone-Based Noninvasive Screening of Dry Eye Disease. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 4, Article 190 (December 2024), 26 pages. <https://doi.org/10.1145/3699742>

1 Introduction

According to the Tear Film and Ocular Surface Society Dry Eye Workshop (TFOS DEWS) II [7], dry eye disease (DED) is “*a multifactorial disease of the ocular surface characterized by a loss of homeostasis of the tear film, and accompanied by ocular symptoms, in which tear film instability and hyperosmolarity, ocular surface inflammation and damage, and neurosensory abnormalities play etiological roles.*” In simpler terms, DED is a condition where the tear film on the eye’s surface becomes abnormal. There are many aspects of the tear film that could be considered abnormal, such as its stability, volume, and osmolarity, many of which are interrelated [7]. The common symptoms of DED include irritation, redness, sensation of a foreign body, blurring, tearing, and heightened light sensitivity. For more details about DED types, symptoms, diagnostics, and treatment, please refer Appendix A.

DED represents one of the most frequent ophthalmological complaints, affecting 8% of the world’s population [18]. In the U.S., studies found the prevalence of DED to increase with age (2.7% in 18-34 years old vs 18.6% in >75 years old) and a higher occurrence in women than men (8.8 vs 4.5%) [11, 39]. With the increasing air pollution and widespread penetration of digital display devices (such as computers and smartphones), even the Global

Authors' Contact Information: Vaibhav Ganatra, t-vaganatra@microsoft.com; Soumyasis Gun, Microsoft Research, India; Pallavi Joshi; Anand Balasubramaniam; Kaushik Murali, Sankara Eye Hospital, India; Nipun Kwatra, nkwater@microsoft.com; Mohit Jain, mohja@microsoft.com, Microsoft Research, India.



This work is licensed under a Creative Commons Attribution International 4.0 License.

© 2024 Copyright held by the owner/author(s).

ACM 2474-9567/2024/12-ART190

<https://doi.org/10.1145/3699742>

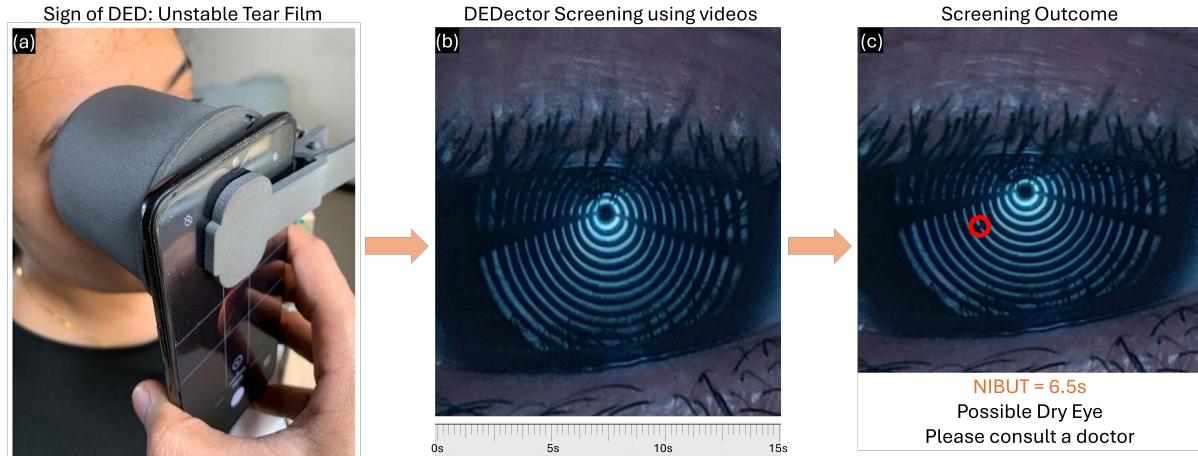


Fig. 1. *DEDetector* – a non-invasive, low-cost, automatic tool for screening Dry Eye Disease. (a) Smartphone attachment setup of SmartKC [16] (image reproduced with permission), (b) captured video with reflected Placido pattern on the tear film, and (c) tear break-up detected by *DEDetector*.

South sees dry-eye disease an emerging threat—e.g., 275 million people in India may be affected with DED by 2030 [26]. This chronic condition also constitutes a substantial economic burden for both individuals and society at large [40], as it leads to depression, anxiety and lack of performance in work, thus impacting quality of life.

Despite the high prevalence, clinical diagnosis of DED proves to be challenging, as no single gold-standard diagnostic marker has yet been established [63]. This complexity arises from the poor correlation between signs and symptoms, the multifactorial nature of the disease, and the significant variability in current metrics across time and season [15, 44]. Consequently, the TFOS DEWS II Diagnostic Methodology report [63] has outlined a diagnostic algorithm comprising of the presence of patient-reported ocular discomfort symptoms (using DEQ-5 or OSDI questionnaire) coupled with at least one indication of unstable ocular surface. These indicators comprise of tear film stability (commonly measured using tear break-up time), ocular surface staining (utilizing dyes to identify damaged tissues), and hyperosmolarity of the tear film. The TFOS DEWS II report also identifies two sub-categories of DED: evaporative (mainly from a deficient lipid layer in meibomian gland dysfunction), or aqueous deficient (a reduced tear volume). Identifying the type of DED, by examining meibomian glands, lipid thickness/dynamics, and tear volume , is essential for effective DED management

In this work, we focus on the tests used for diagnosing DED. (Note: We do not delve into the DED sub-classification methods). Among the three diagnostic indicators—tear film stability, ocular surface staining, and hyperosmolarity—tear film stability is the fundamental diagnostic criteria [54, 63]. There are various techniques to assess tear film stability: evaluation of tear break-up time, topographic and interferometric techniques, confocal microscopic methods, aberrometry, and visual function tests. In clinical practice, the most frequently employed test of tear film stability is the measurement of the tear film break-up time (TBUT) [63]. It is the interval of time that elapses between a complete blink and the appearance of the first break in the tear film. To enhance visibility of the tear film, sodium fluorescein eye drops are put in the eyes, hence it is often referred as the fluorescein break-up time (FBUT). Measuring TBUT **non-invasively**, called NIBUT, is the preferred test to assess tear film stability, and DEWS II [63] recommends “*FBUT only to be used if NIBUT not available*” due to complications with administering the fluorescein drops (details in Section 4).

Most NIBUT systems involve observing reflection of an illuminated pattern (e.g., Placido ring pattern used in corneal topography systems [16]) over the tear film (similar to Figure 1b). The tear film breaks can be identified manually or in an automated manner from a video recording of the reflection pattern (Figure 1c). Automated NIBUT measurements have become popular in both clinical practice and research [7], as they remove the subjectivity involved in manual observation and do not require medical expertise. The sensitivity and specificity of NIBUT in diagnosing DED vary based on the specific technique employed, with reported values ranging from 60-88% sensitivity and 60-95% specificity [7]. However, commercial medical devices to measure NIBUT automatically (like Keratograph 5M) are non-portable, very expensive, and require a trained technician to operate. These factors make frequent and point-of-care screening of dry eye difficult in remote areas.

Existing methods for portable and point-of-care screening of DED generally fall into two categories—smartphone apps [46, 68] that utilize patient-reported questionnaires (such as the OSDI questionnaire [50]) to assess DED symptoms, and apps/devices that analyze blink patterns to identify markers correlated with a positive DED diagnosis [9, 34, 64, 68]. Both these approaches have limitations. Questionnaires may introduce subjectivity in responses and potential language comprehension barriers (various translations and cultural adaptations of the OSDI questionnaire have been proposed [42, 49]), whereas the DEWS II Diagnostic Methodology report [63] does not include blink pattern analysis in its diagnostic criteria. Consequently, apps based on blink pattern analysis yielded poor sensitivity of 50% [46], making them ineffective as screening tools.

Motivated by these shortcomings, we propose *DEDetector*—a low-cost, portable, smartphone-based automated NIBUT measurement device for DED screening (Figure 1). On the hardware side, *DEDetector* builds upon the attachment proposed by Gairola et al. [16] comprising of a 3D-printed cone-shaped Placido ring attachment on a smartphone’s camera along with a blue-colored circular LED to help project concentric black-and-blue rings on the human eye. Videos recorded with the attachment were manually analysed to determine the type of tear film deformations associated with DED. This information was then used to develop *DEDetector*’s video processing pipeline, which analyzes the projected rings (called mires) in the captured tear film videos to identify breaks, ultimately outputting automated NIBUT readings. We evaluated *DEDetector* in real-world settings on 46 eyes from 23 patients, and flagged eyes with a NIBUT of <10 sec as dry eyes. *DEDetector* achieved a sensitivity of 77.78% and specificity of 82.14% on our dataset, in comparison to FBUT achieving a sensitivity of 72.22% and specificity of 75.0%, for DED screening. The software and hardware details for *DEDetector* is publicly available at: <https://github.com/microsoft/DEDetector>

In summary, the key contribution of our work are: (a) a novel method of using an existing low-cost smartphone-based Placido attachment [16] for assessing tear film stability, (b) detailed manual analysis to categorize the types of non-invasive mire deformations observed and their correlation with DED, (c) an end-to-end video processing pipeline that takes mire videos as input and provides NIBUT values as output, and (d) evaluation of the proposed approach on 46 distinct eyes in a real-world setting, and comparing the results with FBUT, for dry eye screening.

2 Related Work

Over the past decade, the Ubicomp community has proposed several smartphone-based diagnostic solutions for various health conditions, spanning from keratoconus [16], lung function [33], ear infection [4], to brain injury [38]. Additionally, prior works have measured vital parameters like hemoglobin level [60] and blood pressure [61]. These novel solutions often replicate the functioning of approved medical procedures and/or devices by incorporating mechanical/electrical attachments to smartphones. As our work centers on dry eye, our literature review focuses on low-cost and smartphone-based dry eye diagnostics solutions. Table 1 compares various DED screening approaches. Notably, *DEDetector* is the only smartphone-based, low-cost and portable method that follows DEWS II recommended criterion for DED screening.

Device	DEWS II Rec.	Phone	Cost	Portable	Hardware	Input	Output
DryEye-Rhythm [46]	No	Yes	Low	Yes	None	J-OSDI, Eye video (30s)	J-OSDI score, Max blink interval
EyeScore [68]	No	Yes	Low	Yes	None	J-OSDI, Eye video (60s)	(Partial) Blink Rate, "Eye Score"
DEvice [17]	No	No	-	Yes	Humidity, Temp sensor	60s reading with eye cup	Relative humidity, DED
SDE [64]	No	No	Medium	Yes	Off-the-shelf Radar	RF Signals of blink patterns	DL-based DED probability
Pauk et al. [58]	Yes	No	High	No	Slit lamp, LED light	Slit lamp exam (white light)	Manually identified NIBUT
Tomey RT-7000* [20]	Yes	No	High	No	Device	Mire videos	NIBUT
IDRA Plus* [59]	Yes	No	High	No	Device	Mire videos	NIBUT
Keratograph 5M* [36]	Yes	No	High	No	Device	Mire videos	NIBUT
<i>DEDector(ours)</i>	Yes	Yes	Low	Yes	Placido attachment, LED Light	Mire videos	NIBUT

Table 1. Comparison of various non-invasive dry eye assessment devices with *DEDector*. Note: 'DEWS II Rec.' column denotes whether the device uses a DEWS II recommended principle for diagnosing DED; 'Hardware' column denotes hardware required in addition to a smartphone; 'Cost' column denotes whether the cost of the device is *Low* ($\leq \$250$), *Medium* ($\$250-\5000), or *High* ($\geq \5000); * denotes medical devices currently used in clinics for DED diagnosis. Notably, *DEDector* is the only low-cost and portable device that uses a DEWS II recommended criterion to screen for DED.

2.1 Low-cost Dry Eye Diagnostics

Diagnostic methodologies for DED can be categorized and assessed based on cost, invasiveness of the diagnostic method, and the requirement of a skilled medical professional to conduct the test. The ideal diagnostic approach for DED should be low cost, non-invasive, and self-administrable.

The majority of low-cost clinical approaches for DED diagnosis are invasive and require a doctor to conduct the test. For instance, a box comprising 100 Schirmer's strips is economically priced at $\sim \$20$. However, the Schirmer's test involves the insertion of paper strips into the lower eyelid, constituting an uncomfortable experience for the patient [63]. Similarly, the Fluorescein Break-Up Time (FBUT) test and corneal staining test necessitate the application of dyes onto the ocular surface, making them invasive and requiring a trained medical professional for execution. Vidas Pauk et al. [58] propose a method for noninvasive examination of the lipid layer under diffuse white light using a slit lamp. However, the examination needs to be performed by a medical professional. Furthermore, none of these tests provides a digital record.

Conversely, medical instruments like the Oculus Keratograph 5M [36], Tomey RT-7000 Auto Refractor-Keratometer [20], and IDRA Plus [59] offer a non-invasive alternative to FBUT, called Non-Invasive Break-Up Time (NIBUT). However, these devices are expensive ($\geq \text{USD } 20,000$), thereby limiting their accessibility. On the positive side, these devices maintains a digital record and can be operated by medical staff with minimal training.

Non-invasive and cost-effective methods for DED diagnosis are rare. The Ocular Hygrometer – DEvice©(AI, Rome, Italy) [17] is a device reported in literature to satisfy both conditions. The device functions by measuring variations in relative humidity around a closed ocular surface. In conclusion, there exists a demand for an affordable, non-intrusive, and self-administrable digital screening solution to facilitate screening and regular monitoring of DED. Our proposed solution, *DEDector*, meets these requirements.

2.2 Smartphone-based Dry Eye Diagnostics

Existing literature on smartphone-based diagnostic solutions for DED can be broadly categorized into three approaches: questionnaire-based, eye images/videos analysis-based, and novel hardware-based.

Questionnaire-based approaches mainly rely on the OSDI questionnaire [50]. Nagino et al. [45] proposed a protocol for a non-invasive smartphone-based screening method for DED. This served as the foundation for the DryEyeRhythm [46] app, which digitized the Japanese version of OSDI (J-OSDI) [42]. The app demonstrated a strong correlation between digital and paper-based J-OSDI. Inomata et al. [25] used DryEyeRhythm to conduct a crowd sourced study on DED symptoms. Through a multidimensional correlation analysis, they identified seven

clusters of DED patients with distinct biological and behavioral characteristics. DryEyeRhythm has since been used for several crowdsourced studies, as the questionnaire can be easily self-administered [22–24].

Eye images/videos based approaches focuses on blink pattern analysis for DED diagnosis. The EyeScore app [68] utilizes blink rate and partial blink count from a 1-minute video for DED diagnosis. Similarly, DryEyeRhythm [46] also incorporates blink analysis by measuring the maximum duration for which an individual can keep their eye open, called Maximum Blink Interval (MBI), along with J-OSDI for DED diagnosis. Lower MBI has been found to be correlated with DED. Hong and Hasegawa [21] proposed a novel method to estimate Tear Meniscus Height (TMH) from the reflection of a ring light off the cornea, and Taniai et al. [55] proposed algorithmic changes in that TMH estimation method to improve its performance. However, both these methods [21, 55] have not been evaluated on their ability to screen for DED.

Novel hardware-based approaches for DED screening include solutions such as SDE [64], which uses wireless RF signals to extract fine-grained spontaneous blinking action along with unsupervised domain adaptation to screen for DED. Additionally, wearables such as DualBlink [9] and Tiger1 [34] have been proposed to monitor patients' blink rate. However, DEWS II [63] does not include blink pattern analysis in its diagnostic criteria. Tiger1 [34] ensures that the patient follows the 20-20-20 rule, whereas DualBlink [9] enforces blinks via light flashes, physical taps, or small air puffs when the blink rate dips.

A common benefit of using smartphone- and wearable-based methods is the automatic preservation of digital records. However, a major shortcoming of all prior solutions is that they do not consider or measure *tear film stability*. As per the DEWS II Report [7], an unstable tear film is one of the primary indicators of DED. Additionally, the introduction of novel hardware results in increased costs and maintenance. With *DEDetector*, our objective is to develop a low-cost screening tool, that provides reliable diagnoses for DED based on tear film stability [7].

3 Data Collection

We recruited 23 patients from a leading tertiary eye-care hospital in Bengaluru, India for our study. The hospital has a team of >15 doctors (including >3 cornea specialists), >10 optometrists, and treats over 500 patients daily. The data was collected over a period of 3 weeks between November-December 2023. The data collection protocol was approved by the hospital's Institutional Review Board.

3.1 Sample Size Calculation

The minimum sample size for our evaluation was computed using the following [5]:

$$n = \frac{r + 1}{r} \cdot \frac{SD^2 \cdot c^2}{d^2}, \quad (1)$$

where n is the number of samples in each group (DED vs Normal, in our case), r is the ratio of controls (Normal) to cases (DED) in the study (we assume a ratio of 1), SD is the population standard deviation of the variable being measured (FBUT, in our case), c is a constant (2.8 in our case) determined by the significance and power of the test (usually fixed at 5% and 80%, respectively), and d is the clinically significant difference in the mean values of the variable of interest (FBUT) between the case and control groups. We estimated SD as the standard deviation in FBUT of 50 randomly selected patients visiting the hospital in the month of September 2023, and found the value to be 5.22s. According to DEWS II [63], a difference of 5s is clinically significant between DED and Normal patients, so $d = 5$. Plugging in these values, we obtain a sample size of 17.5 or more per group, resulting in a total of at least 36 eyes. This sample size is in alignment with prior works [10, 63].

3.2 Data Collection System

We use the 3D-printed Placido attachment (see Figure 1) proposed by Gairola et al. [16] for collecting video data. The attachment consists of a conical head featuring 28 alternating opaque and empty rings, encased within a

cylindrical shell. Two 2mm supports at 0° and 180° uphold the (Placido) rings on the conical head. The cylindrical shell effectively blocks environmental light, and a set of 12 USB-powered blue LEDs are attached at the base, arranged in a ring pattern to provide illumination. The conical head is covered with a butter paper, which functions as a diffuser and ensures uniform illumination. For more details, please refer to the SmartKC paper [16]. Note: This differs from certain medical devices that use infrared light to project mires. We avoid infrared light to reduce hardware costs for illumination and sensing.

The Placido attachment is affixed over the smartphone camera, and the LED lights are powered by the smartphone's USB port. The projection of Placido rings onto the cornea creates a distinctive pattern of dark and white rings (mires), resulting from reflections off the tear film. For data capture, the attachment is positioned over the patient's eye, completely covering it, and the smartphone's default camera app is used to record a video.

Note: The Placido attachment is not a contribution of this work. We use the hardware proposed by Gairola et al. [16], without modification. However, this work is the first to utilise a smartphone-based Placido attachment for assessing tear film stability and estimating NIBUT cost-effectively.

3.3 Method

Patients typically undergo a comprehensive set of diagnostic tests, both subjective and clinical assessments, to establish a reliable clinical diagnosis for DED. In our study, following the ophthalmologist's recommendation, we used the OSDI questionnaire [50] for the subjective assessment of DED symptoms. A hospital staff member explained the OSDI questionnaire in detail to patients, often in the regional language, and collected the completed responses. Clinical tests for DED—FBUT, Schirmer's test and corneal staining test—were conducted by a senior ophthalmologist, with outcomes consolidated by hospital staff for later analysis. In addition to test results, demographic details such as age and gender were also recorded. Unfortunately, we lacked access to any device that can measure automated NIBUT (such as Keratograph 5M) due to its high cost (>USD 20,000). The patient ID was the only common identifier between the OSDI questionnaire responses, test results recorded by the ophthalmologist, and videos captured by our device.

Since tests in DED diagnosis involve alternation between blinking, eye drop instillation, and exposure to bright light, the sequence in which these tests are performed can significantly impact the outcomes [7, 13]. It is crucial to proceed from the least invasive to the most invasive test [63]. To achieve that, we adopted the following protocol:

- (1) Inclusion criteria: First, the patient is examined by the ophthalmologist under a slit lamp to identify any corneal surface abnormalities, such as keratitis or conjunctivitis. Patients with such conditions are excluded from the study. Also, only patients aged 18 and above are enrolled.
- (2) The patient is then asked to fill out the OSDI questionnaire. The OSDI questionnaire often require translation into the local language of the patient. To overcome any comprehension gaps, the ophthalmologist, during the patient examination in the preceding step, also collected a summary of the symptoms exhibited.
- (3) Videos for both eyes are then recorded. For each eye, the patient is asked to cover the other eye with their hand (similar to SmartKC instructions [16]), and the Placido attachment is placed on the patient's eye for video capture. The patient is instructed to blink slowly three times. After the third blink, the patient is asked not to blink for 15s. If the patient blinks within 15s of the last blink, instead of restarting the data capture, we continue with the same capture, now aiming for 15s from the latest blink, a common practice when examining fluorescein breaks. To prevent hand shaking and stabilize the video, the data collector's elbows rest on a table during data capture. All videos were captured by the first author.
- (4) Next, corneal staining and FBUT test is performed by the ophthalmologist.
- (5) Finally, 20 minutes after the FBUT test, Schirmer's test (without anesthetic) is performed. This gap between the FBUT test and Schirmer's test allows the eyes to return to normalcy after the fluorescein instillation.

3.4 Participants and Dataset

The final dataset consists of 23 patients (10 female, 13 male) with a total of 46 distinct eyes. The average age of the participants was 36 ± 8.56 years (20-52 years). Out of the 46 eyes, 18 (39.13%) were diagnosed positive for DED, while 28 eyes (60.87%) were normal. The mean FBUT for all participants was 10.32 ± 5.15 s. For eyes diagnosed with DED, the mean FBUT was significantly shorter at 6.67 ± 3.82 s, in comparison to normal eyes at 12.68 ± 4.51 s. The ground truth was the final dry eye diagnosis by a senior ophthalmologist. It was derived by combining patients' symptoms, the OSDI questionnaire, manual examination under the slit lamp, FBUT, corneal staining test, and Schirmer's test, in accordance with the DED diagnostic criteria used in similar studies [59, 64].

The mean OSDI score for our patients was 22.71 ± 22.03 . Ideally, individuals with DED should exhibit a higher OSDI score compared to their healthy counterparts. However, we observed the opposite trend, where the mean OSDI score for normal patients (27.24 ± 20.02) is higher than that of DED patients (15.68 ± 23.14). We attribute this discrepancy to the subjective nature of the test and potential language barriers affecting comprehension. Translation and cultural adaptation of the OSDI questionnaire in local languages is non-trivial, and considerable effort has been dedicated to obtain translated versions, such as Japanese-OSDI [42] and Brazilian-Portuguese OSDI [49]. Additionally, patients, particularly in the Global South, may face difficulties in accurately filling the questionnaire due to low literacy rates. Consequently, OSDI scores are excluded from further analysis. Due to these challenges, the ophthalmologist of our hospital relied more on symptoms reported by the patient (in the local language) during examination, rather than the OSDI questionnaire when considering the diagnosis.

The mean Schirmer's test values were 21.89 ± 11.66 , with lower values in DED patients (13.72 ± 9.78) compared to normal patients (29.25 ± 7.63), as expected. According to the DEWS II report [63], Schirmer's test values are primarily used to identify DED type: Evaporative, Aqueous Deficient, or Mixed. Since this work does not focus on DED type classification, we exclude Schirmer's test values for further analysis. The corneal staining test has a subjective outcome, which was also not used for further analysis. Although these tests were used by the ophthalmologist for the ground truth DED diagnosis, we only use the FBUT values for correlation analysis with NIBUT and the ground truth DED diagnosis for computing accuracy metrics of our method.

The final video dataset used for automated prediction of DED comprises 46 RGB videos of Placido reflections, one for each eye, collected at a frame rate of 30 frames per second (fps), with a resolution of 1920 x 1080 pixels. The videos were recorded with a OnePlus 6T smartphone, using the main camera (16MP, f/1.7 aperture), with the Placido attachment affixed on it. The videos span between 18.8-35.36s, with an average duration of 24.07 ± 3.24 s.

4 Theory of Operation

We aim to assess the utility of NIBUT estimated from videos recorded from a portable smartphone attachment for screening of DED. In this section, we discuss some key considerations of the screening tool.

How can we observe tear film break-ups on a smartphone? The non-invasive observation of the tear film involves projecting illuminated grid patterns or black-and-white concentric rings called Placido discs on the corneal surface. The tear film, with a mean thickness of $2\text{-}5.5\mu\text{m}$ [27, 62], constitutes the outermost layer of the ocular surface and serves as the primary refracting surface for incoming light [62]. Since the refractive index of the tear film (~ 1.34) [47] differs from that of the surrounding media (comprising of human cornea with a refractive index of 1.38 and air with a refractive index of 1), a portion of the incident light is reflected. The reflected light contains information about the tear film properties, which is captured and analyzed to determine tear film characteristics. This principle forms the basis of various tear film interferometry [30, 31] and reflectometry [37] methods. In the context of assessing tear film stability, a regular Placido disc reflection (called *mires*) image, as typically captured by Corneal Topographers, corresponds to an undisturbed tear film. When tear film break-up or disturbance occurs, there is a loss of reflected light from the region. This manifests into a discontinuity or irregularity in the

mire image. The NIBUT is calculated as the time elapsed from the formation of the tear film to the first observed tear film break-up, marked by distortions in the mires. A NIBUT of <10s has been found to be indicative of DED [7].

Why do we prefer NIBUT to FBUT? The TFOS DEWS II report [63] suggests using the non-invasive method NIBUT over the invasive FBUT, because the fluorescein instilled in the eye to enhance visibility of the tear film, interferes with the tear film stability by promoting tear evaporation. Consequently, the outcome of the test may not be reflective of the actual tear film stability [63]. Additionally, precise control of the volume of fluorescein administered is difficult, and it has to be performed by an ophthalmologist making the FBUT test less accessible. Finally, manual observation of the tear break-up under a slit lamp is highly subjective. On the other hand, measuring NIBUT solves these problems. Since it is non-invasive, it does not impact the tear film stability and can be conducted by a trained staff. Although manual NIBUT measurement suffers from human subjectivity, automated NIBUT devices (like IDRA Plus or Oculus Keratograph 5M) effectively addresses this concern.

What kind of mire distortion patterns are associated with DED? Our data collection setup results in the formation of black-and-blue mires onto the eye. (Note: The blue color in the mire image resembles white (Figure 2), so we refer to it as black-and-white mires instead.) The white mires are formed due to the reflection of incident light from the empty rings, while the black mires represent shadows of the opaque rings. Since the formation of black mires are not due to light reflections, any distortions observed in the black mire region may not be indicative of tear film break-ups and, consequently not suggestive of DED.

Upon manual analysis of the captured video data, we noticed a variety of tear deformation patterns. This observation aligns with a recent study [51], which categorized tear film break patterns observed during the invasive FBUT test in individuals with dry eyes into five distinct categories: line breaks, area breaks, dimple breaks, spot breaks, and random breaks. Each tear film break-up pattern is associated with distinct pathophysiology for DED. Notably, there is currently no established classification for break-up patterns observed in a non-invasive dry eye test. Hence, we conducted an extensive manual analysis of all 46 videos in our dataset, aiming to classify different types of deformation patterns observed during non-invasive examinations of the tear film and their correlation with dry eye diagnosis. Note: We did not consider pathophysiology for the break-up pattern, as it has not been well established for non-invasive tear break-up.

The manual annotation process for the videos involved two passes. In the first pass, the video segment following the blink was examined frame by frame, without any crop or zoom on the mires. Any deformation in the mires lasting for at least 0.5 seconds was noted. Ophthalmologists suggest that deformations lasting for shorter duration are challenging to observe manually, and also determining whether they indicate impaired ocular health is difficult. In the second pass, a zoomed-in version of the mire images was reviewed, using a custom visualization tool created with OpenCV [3]. This tool enabled frame-level visualization of the video with zoom on any region of interest in the frame. The two-pass approach was adopted because some breaks appear distinctly in the zoomed-out version (first pass), while others require a higher zoom level for observation (second pass). The second pass aims to identify fine-grained and small-sized deformations that may have been missed in the initial pass. To prevent unintended biases, the manual annotation was performed with the OSDI, Schirmer's, FBUT, Corneal Staining, and the final diagnosis information kept hidden from the annotator.

4.0.1 Types of Deformations Identified. We have identified four distinct types of deformations in the mires: **Line Break:** A line break appears as a linear black discontinuity in a white mire (Figure 2a). As previously mentioned, only white mires form due to the reflection of light from the tear film; thus, line breaks are exclusively observed in the white mires. We hypothesize that line breaks occur due to a tear film break-up, resulting in the loss of reflected light from the region, hence the observed discontinuity. Consequently, this break serves as an indicator of DED, as described in Section 4.0.2.

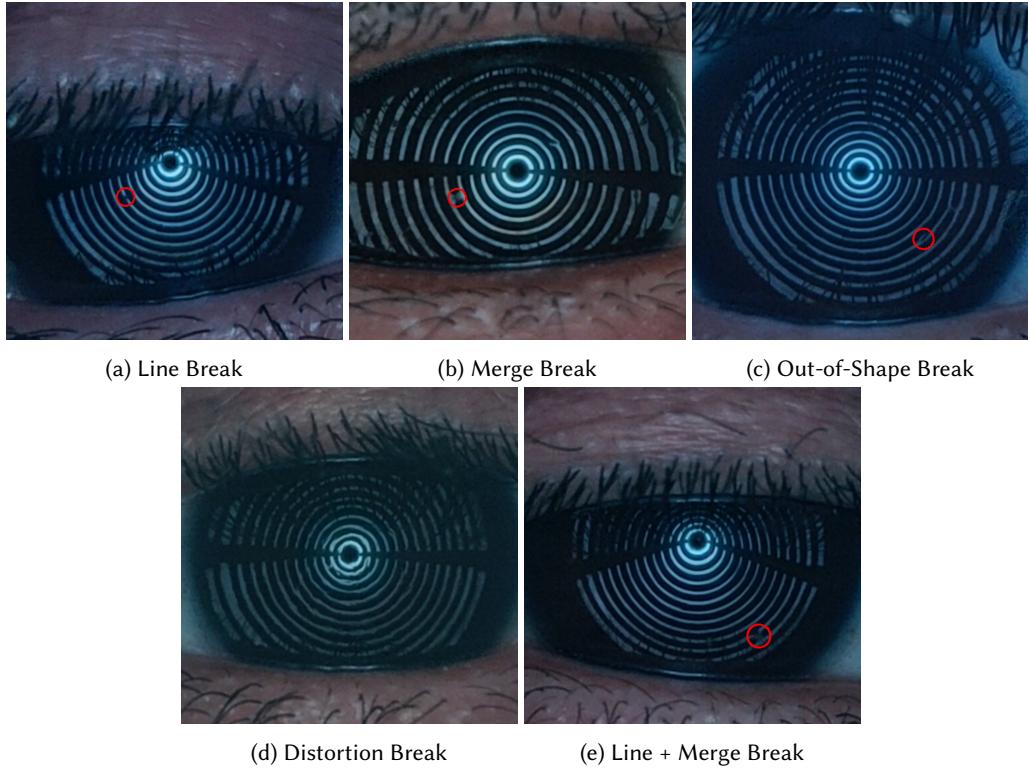


Fig. 2. (a)-(d) show the different types of breaks seen during non-invasive examination of the tear film, (e) shows the combination of line and merge break. The red circle in the figures show the location of the specific break under consideration. Note: (d) lacks a red circle as it pertains to the general shape of the mires and is not localized to a single location.

Merge Deformity: A merge deformity appears as a linear white discontinuity in a black mire (Figure 2b). Given its appearance of connecting two white mires, we term it a *merge* deformity. Previous research studies have shown that there is a constant upward movement of tears since the formation of the tear film after a blink [2]. As black mires are shadows and not formed due to reflected light, we hypothesize that merge deformities result from tear movement in the eye. Light reflected from moving tears gives the illusion of being reflected from a black mire. A previous study has shown that the upward movement of the lipid layer of the tear film is too slow to account for the required thinning observed in tear film break-ups [29]. Hence, we anticipate that these breaks are not indicative of tear film break-ups and, consequently, are not suggestive of DED. In a few videos (2), we observed a combination of a line break and a merge deformity, wherein a black discontinuity appeared in a white mire, and a white discontinuity appeared in the neighbouring black mire, as shown in Figure 2e. We hypothesize that such breaks are also caused by tear movements and, hence, are not indicative of DED.

Out-of-Shape Deformity: As the name suggests, an out-of-shape deformity is observed when mires appear deformed in their shape at specific locations. These deformations lack a specific shape definition; instead, they just appear irregular in form. An example of an out-of-shape break is shown in Figure 2c, where an extension is visible at the break location. Another criterion for classifying a break as an out-of-shape break is that it must be found in <3 mires in a frame. If such deformations are observed in three or more mires in a frame, we label it as a

Type of Deform.	% videos	Acc.	Sens.	Spec.	F1
Line	43.48	80.43	77.78	82.14	75.68
Merge	26.09	60.87	33.33	78.57	40.00
Out-of-Shape	23.91	54.35	22.22	75.00	27.59
Distortion	32.61	65.22	38.89	82.14	46.67

Table 2. Diagnostic performance of different types of NIBUT breaks. The columns labeled “% videos” shows the proportion of all videos in which the specific break is observed.

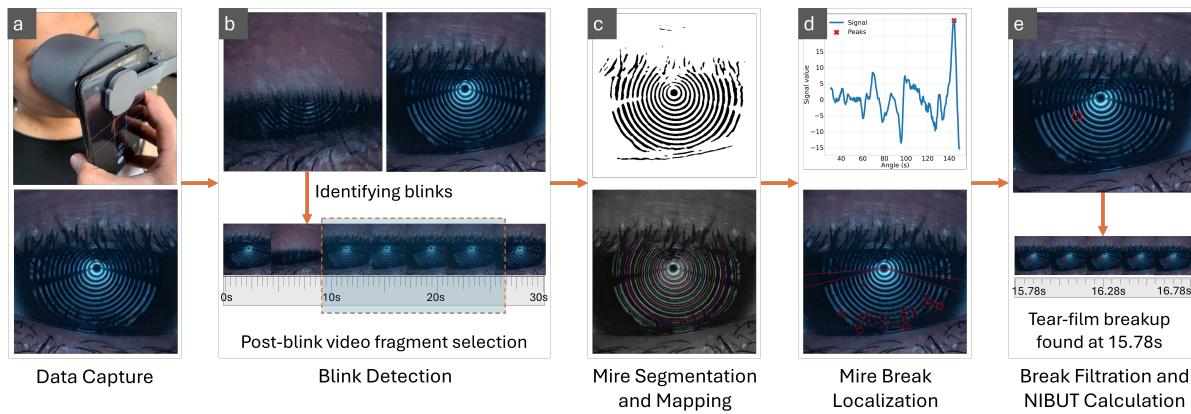


Fig. 3. Overview of the video processing pipeline for automatic detection of tear film break-ups and NIBUT calculation. Data collection setup image (a) reproduced from SmartKC [16] with permission.

distortion break, as described next. We speculate that these isolated shape deformations arise due to localized tear film dynamics, resulting in an unusual shape. For instance, a localized accumulation of tears might lead to thickening of the mires at the particular location, resulting in an irregular appearance. Once again, since these breaks do not correspond to tear film break-ups, we do not expect them to be indicative of DED.

Distortion Break: A distortion break occurs when 3 or more mires lose their typical circular shape and appear “wavy”. The boundary of the mire, instead of being smooth, exhibits “waves” or “squiggles”, as shown in Figure 2d. Unlike other breaks, a distortion break is characterized by a general distortion of mire shapes across the entire frame, instead of being localized. We speculate that this break occurs when the patient begins tearing, resulting in a distortion of the reflection patterns due to an uneven distribution of tears underneath. Hence, we do not expect a distortion break to be indicative of DED either.

We highlight that except for the linear break, there is no one-to-one correspondence between the types of breaks observed in the FBUT test [51] and the NIBUT test. This suggests that fluorescein instillation for the FBUT may alter tear film dynamics, highlighting its invasive nature.

4.0.2 Results for Types of Deformations. Table 2 shows the performance of the different types of NIBUT breaks identified through manual analysis of the 46 collected videos. Confirming our hypothesis that a line break is indicative of DED, it outperforms other break types in terms of sensitivity, specificity, accuracy, and F1-score. The other break categories exhibit poor sensitivity in diagnosing DED.

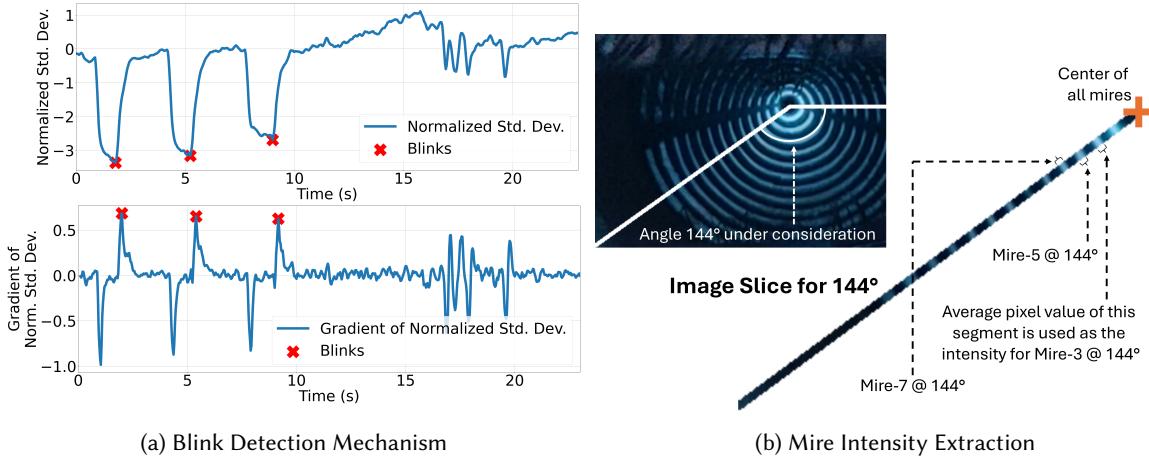


Fig. 4. (a) shows the blink detection mechanism - The upper graph shows the dips in standard deviation during a blink, but fairly stable otherwise. The lower graph shows the gradient of standard deviation - peaks of this signal correspond to the eye opening after a blink. (b) shows mire intensity extraction - Mires are segmented and numbered via a radial scanning algorithm. For each mire, average intensity is recorded along the radial segments.

5 Video Processing Pipeline

We propose a novel video-processing pipeline for automatic estimation of NIBUT from videos of Placido-disc reflections. Although automated NIBUT calculation is available in a few commercial corneal topography systems such as Keratograph 5M, IDRA Plus, and Tomey RT-7000, their software is proprietary [19, 20]. Unlike commercial systems which have head- and chin-rests and high resolution video cameras, our videos are collected with a handheld smartphone based attachment. This poses a requirement on our video-processing pipeline to be robust to camera pose variations, lighting variations, low camera resolution, etc. To the best of our knowledge, this is the first work to propose a publicly available and interpretable method for automated NIBUT assessment.

During our data collection, the patients are asked to blink 3 times before keeping the eyes open. We measure NIBUT as the time to first tear break-up after the last blink. Figure 3 shows an overview of our video processing pipeline. The pipeline first detects blinks in the input video (b). The video segment following the blinks is then analyzed to segment and number the mires (Placido disc reflection) (c). This is followed by identification of breaks in these mires (d), and finally a filtering step is performed to remove any extraneous mire breaks (e.g., due to eye-lashes) (e). NIBUT is then calculated as the time of the first sustained tear break-up from the last blink. We note that the blink detection method is motivated from prior works [1, 12] and the mire segmentation method is borrowed from SmartKC [16]. The mire break identification method and break filtration steps, however, are our novel contributions. We now describe the individual steps of the pipeline in detail:

5.1 Blink Detection

For blink detection, we use the standard deviation of a video frame's pixel values [1, 12]. When the eyes are open, the texture and color variations in the iris, sclera, iris, surrounding skin, and mire reflections lead to a high standard deviation in the pixel values. On the other hand, when the eyes are closed, the video frame mainly consists of the surrounding skin and the eyelids (also skin color), resulting in a much lower standard deviation. Finding local minima of the standard deviation can thus help in identifying the frames where eyes are closed.

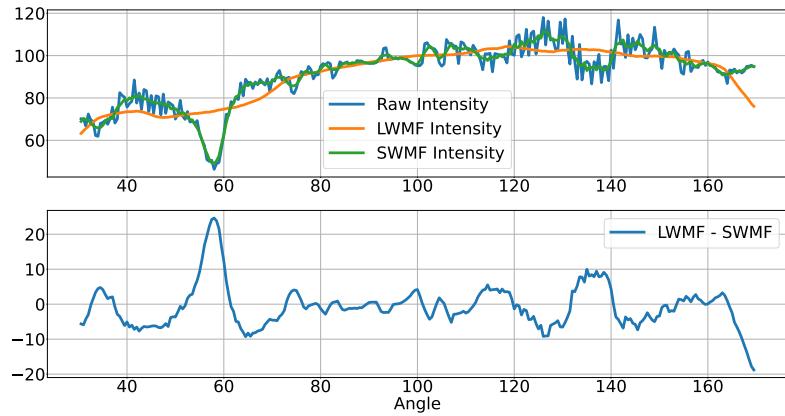


Fig. 5. Signals preparation for Break Localization

The proposed method identifies the time when the eyes are closed. However, we are interested in the timestamp when the eyes have reopened after the blink. Since our protocol mentions “**3 slow blinks**”, we observe intrinsic variation in the time patients take to open their eyes after the blink (0.1s - 2.5s), hence adding a constant value to the blink timestamps will not provide an accurate solution. To overcome this, we use the first-order difference (or gradient) of the smoothed (mean-filtered) standard deviation signal. When the eyes are open, the standard deviation will be high but stable and this gradient will be close to zero. After this, when the eyes start to close, the standard deviation will start decreasing and the gradient will sharply decrease. Next, when the eyes are opening, the standard deviation will start increasing and the gradient will sharply increase. Finally, when the eyes have opened, the standard deviation will again be high but stable, and the gradient will become close to zero. The local maxima of this gradient can thus be used to identify the time when the eye has opened after a blink (Figure 4a).

5.2 Mire Segmentation

Once we have detected the blinks, we process 11 seconds of the video fragment after the last blink. Processing the video post 11s is immaterial since our manual analysis in Section 4.0.1 found that breaks seen after 10s from the blink are not indicative of dry eye disease. To reduce computational cost, the video fragment is subsampled to consider every third frame. Since the videos are collected at a frame rate of 30fps, this sampling provides a time granularity of 0.1s. For each sampled frame, we segment and number the mires using the radial scanning algorithm proposed in [16]. This radial scanning algorithm traverses along rays at discrete angles (0° - 360°) from the center, counting the alternating black and white mires. For downstream processing, we additionally record the average grayscale pixel intensity of each radial mire segment, for every mire and angle combination (Figure 4b).

5.3 Break Localization

As shown in Figure 2a, we are interested in identifying breaks which occur in the white mire reflections. For this, we leverage the mire intensities computed in the previous step. Note, however, that the shadow from eyelashes results in a similar variation in intensity as the tear breaks. This is especially a problem in the upper half of the mire pattern, as can be seen in the images of Figure 2. Due to this, we only process the mires present in the lower half of the mire region. Our break localization then proceeds as follows.

5.3.1 Mire intensity analysis based signal generation: As shown in Figure 2a, a tear break-up will result in a dip in the mire intensity due to lack of reflection from the tear film. The radial scanning algorithm in the previous

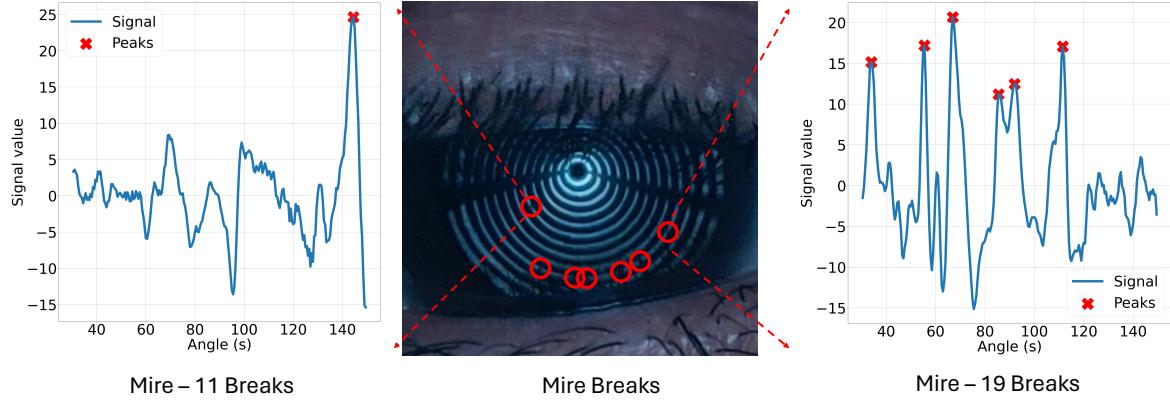


Fig. 6. Break Identification Mechanism - Peaks in the difference between large-window mean-filtered intensity and small-window mean-filtered intensity are identified as breaks.

section gives us the average intensity of each mire at a given angle. To identify these breaks, we consider one particular white mire and plot its intensity as a function of angle in Figure 5. The dip in intensity corresponds to the mire break we want to identify. However, there can be small variations in the mire intensity due to other factors as well — e.g., small errors in radial scanning algorithm, movements during video capture resulting in motion blur, frame blur due to auto-focus fluctuation, etc. Thus, we need to robustly identify the tear breaks while taking into account these sources of noise.

We consider each white mire independently and look at the mire intensity as a function of angle. First, we perform a large-window mean filtering (LWMF) to remove noise and get a baseline intensity value of the mire around each angle. We use a window size of 51 (i.e. 25 degrees clockwise and 25 degree anticlockwise) for this step. Since we choose a large window, the intensity value at each angle now gives a stable baseline representation of what the mire intensity at that angle should be if there are no breaks or variations due to noise. Thus, we can now compare the raw intensity value with this baseline to identify any sharp dips which should correspond to our tear breaks. However, as there can be small variations in intensity even without tear break-up due to the sources mentioned above, we need to apply a filtering to remove such noise. For this, we apply a small-window mean filtering (SWMF) with a window size of 5, and use this for comparison to the LWMF baseline. To summarize, for a given mire, if we denote $l(a)$ to be the LWMF intensity and $s(a)$ to be the SWMF intensity as functions of angle, then the peaks in $l(a) - s(a)$ (see Figure 5) can be used to identify tear break-ups. To identify these peaks, we first find the local maxima in this signal followed by thresholding. Note that the $l(a) - s(a)$ operation is very similar the Laplacian filtering operation used in constructing Laplacian pyramids.

5.3.2 Mire Intensity Normalization and Robust Thresholding. We note that due to the illumination setup in the Placido attachment of SmartKC [16], mires closer to the center are much brighter than the mires near the periphery. These changes in mire intensity impacts the absolute value of $l(a) - s(a)$ directly. To normalize for this, we compute the average mire intensity of each mire and normalize the intensity delta before performing peak detection. That is, if the mire's average intensity is m , we perform peak detection using $(l(a) - s(a))/m$.

Another source of changes in the peak intensity is the video sharpness. We observe that sharper videos will have more pronounced peaks, while in blurry videos the underlying smoothening will make our peaks less pronounced. Thus, our thresholding should be dependent on the underlying video sharpness. Since, our videos are captured using a handheld smartphone, there can be a large variation in the captured video sharpness. We

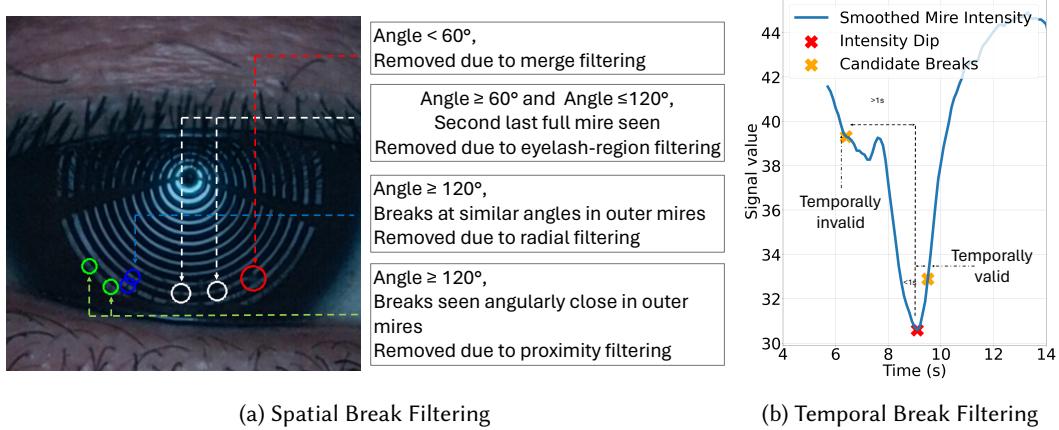


Fig. 7. Mire Break Filtering - (a) shows filtering based on spatial proximity to other breaks, (b) shows temporal filtering - a candidate break close enough to a prominent dip in intensity is temporally valid.

calculate the video sharpness using the mean variance of Laplacian of each frame, averaged over all frames of interest. We divide all the videos in our dataset into 3 sharpness categories - *Very sharp*, *Sharp* and *Somewhat Blurred*. The three categories use progressively smaller percent of the mean mire intensity as the peak threshold. The boundary values of the sharpness categories and the exact percentage values are empirically identified.

5.4 Break Filtration

Figure 6 shows the set of points which are identified as break candidates by the above method. Although the method is able to identify actual break-ups, it also detects many spurious breaks. These spurious breaks can occur due to multiple reasons – a) shadow from eyelashes, b) shadow from the Placido attachment support, c) intensity dips from tear film artifacts not related to tear break-up (see Figure 2) and d) noise during video capture. In this section we discuss methods to identify these sources of breaks and filter them from our output. Figure 7 shows the different types of breaks and the filters discussed below.

5.4.1 Eyelash Shadows: This is the most dominant reason for spurious breaks. We employ the following filters to identify and remove breaks due to eyelash shadows:

Eyelash-Region Filtering: Although we did not consider the upper-half of the mire region as it contains a lot of shadows from eyelashes, even the lower-half region has some eyelash shadows, especially near the outer mires. Due to the physiology of the eyelashes, we observe that in our samples, the eyelash shadows in the lower-half region are concentrated near the two outermost mires and in the bottom 60° sector (30° on each side of the vertical line). We thus discard this region from our calculations. This takes care of a large fraction of the spurious breaks.

Radial Filtering: Although the above region filtering removes most eyelash shadows, we may still miss a few. Again, due to the physiology of the eyelashes in the lower region, we observe that eyelash shadows are mostly formed in the radial direction. We take advantage of this and check if our break detection algorithm detects breaks on 2 consecutive white mires around the same angle. Such breaks are discarded as eyelash shadows.

Proximity Filtering: Finally, since the eyelash shadows are close to each other, they will typically form a cluster of breaks in proximity to each other. We use this observation, and ignore breaks near the outer mires, if 2 breaks on the same or consecutive white mires are seen within 20 degrees of each other.

5.4.2 Shadow from Placido Attachment Support. Due to the presence of supports in the SmartKC Placido attachment [16], light is cut off around 0° and 180° for all mires, causing intensity dips along these angles (black lines seen in Figure. 2). To avoid misidentifying these dips as tear film breaks, we locate support angles and exclude any intensity dip in that region from further analysis. Note that due to the handheld nature of our capture, the Placido attachment may not be perfectly aligned, which can result in these support angles to deviate from 0° and 180° . To automatically locate these support angles, we examine angles in a 60° window (30° clockwise and 30° counterclockwise) centered at 0° and 180° . We again go radially along the mires for each of these angles, and look for intensity dips in three or more mires. Since the location of tear film break-ups is random, the probability of multiple tear break-ups along the same radial line is very low. We thus associate these dips to be due to the support region and assume this angle to be in the shadow of the attachment support. Finally, the mean of all these angles is used as the support angle. Any candidate breaks within the neighbourhood (4.5 degrees to each side) of the estimated support regions are disregarded for further analysis.

5.4.3 Other Tear Film Artifacts. As shown in Figure 2, tear film artifacts not related to tear break-ups can also result in dips in the mire intensity. This is in particular the case with merge breaks, which sometimes causes what appears to be a line break along with a merge breaks (Figure 2e). Such line+merge breaks will be identified by our break detection algorithm. However, we observe that while the tear break related line break (Figure 2a) is localized to a (white) mire, the merge breaks are more spread out (e.g., Figures 2b,2e). As a result, although the white mire may see an intensity drop, the neighbouring black mire in such cases will simultaneously see an increase in intensity due to the white mire spreading into it. We leverage this observation, and for each break detected by our algorithm (which uses only white mires), we also look at the neighboring black mires and check for a similar peak. If we find such a peak, this break in the white mire is discarded.

5.4.4 Noise. Spurious breaks not explained by the above reasons were rare. These may be due to video noise or presence of small foreign bodies like dust particles in the setup. To handle these, we also add these filters.

Persistence Filtering: As per our discussion with ophthalmologists, tear film break-ups usually last between 1s - 5s. Breaks observed for longer than this duration are mostly due to other reasons such as speck of dust in the eye, shadow from a foreign body in attachment, video artifacts, etc. We thus remove any break candidates that persist for a longer duration (5s or more). Note that the presence of such breaks need not be contiguous and can be intermittent. The only requirement we impose is that the break is observed in sufficient number of frames ($\geq 5s \times 10 \text{ fps} = 50 \text{ frames}$). This filtering will also take care of any missed eyelash shadows which may have been missed by the previous filters, since eyelash shadows will mostly be persistent.

Temporal Filtering: Finally, we also add a temporal break detection filter. It is similar to our peak detection along the angular signal, but instead looks at the signal as a function of time. That is, for a given mire and angle, a tear break should temporally see a prominent dip in the mire intensity. We thus remove any candidate breaks, which do not see such a temporal dip within 1s of when it is detected as a spatial peak.

5.5 NIBUT Calculation

The final step is to calculate the tear break-up timestamp in the video in order to compute the non-invasive tear break-up time (NIBUT). The set of valid breaks, as obtained from the last step, are considered in the order in which they were observed in the video. For each break, we consider the video fragment of the following 1s (10 frames). If the break is seen in at least 8 of the 10 frames, we consider it to be a tear film break-up and report the time of break as the tear break-up timestamp. The traversal is conducted until we find the first tear film break-up or the list of breaks is exhausted. The difference between the tear film break-up timestamp and the blink time is reported as the NIBUT. Using a pre-configured threshold of 10s, automated prediction is obtained. If the NIBUT

Device	Threshold (s)	Sensitivity (%)	Specificity (%)
Tomey RT-7000 Keratometer [20]	5	82	60
IDRA Plus [59]	7.75	89	69
Kowa DR-1α [48]	10	69	-
Oculus Keratograph 5M [36]	6.69	80	67.59
Vidas Pauk et al. [58]*	10	87.72	88.23
DEDetector – Auto (ours)	10	77.78	78.57

Table 3. Sensitivity, specificity of automated NIBUT clinical devices compared with *DEDetector*. * Vidas Pauk et al. [58] presents results from manually identified NIBUT, whereas all other devices provide automated NIBUT.

is <10s, the eye is flagged as Dry, otherwise, it is considered as normal. The final output from the pipeline is the NIBUT as well as the automated dry eye prediction for the eye.

6 Findings

6.1 Comparison with NIBUT-based medical devices

Several studies have been conducted examining the efficacy of automated NIBUT obtained from medical devices in identifying DED [20, 48, 59]. They report NIBUTs of DED patients to be shorter compared to normal patients. A similar observation was made in the NIBUTs obtained using *DEDetector*, with the NIBUTs being smaller ($p < 0.001$) in DED patients (7.07 ± 5.15) than normal patients (13.67 ± 4.99). Since we do not have access to NIBUT measuring devices, we use the values reported in literature as a proxy for performance comparison (Table 3). We note that these results are not reflective of the performance of the respective devices on our dataset, but are a general measure of the performance of the device. We also acknowledge that DED is a multifactorial disease, and solely relying on NIBUT for a *clinical diagnosis* may lead to erroneous results. However, considering its accuracy, automated NIBUT estimation using *DEDetector* emerges as a suitable candidate for DED screening, as it is low-cost, portable, and grounded on a diagnostically sound criterion (tear film stability).

As shown in Table 3, *DEDetector* outperforms all the medical devices reporting an automated NIBUT in terms of specificity, while the sensitivity is competitive to Tomey RT-7000 Auto Refractor-Keratometer and the Oculus Keratograph 5M. The method proposed by Vidas Pauk et al. [58] outperforms *DEDetector*, but requires manual examination of the tear film to detect breaks. The DEWS II Diagnostic Methodology Report [63] recommends parallel testing by combining highly specific tests; *DEDetector* is a step in that direction.

6.2 Comparison with non-clinical DED screening methods

As mentioned earlier, DED screening methods often rely on questionnaires detailing patient's DED symptoms and on blink pattern analysis [46, 68]. In contrast, DEvice Hygrometer©(AI, Rome, Italy) [17] measures relative humidity around the eye at a given temperature to detect DED. The criteria used by these methods are not recommended by the DEWS II Diagnostic Methodology report [63] for DED diagnosis. On the other hand, *DEDetector* aligns with the DEWS II report by using NIBUT as a measure of tear film stability. Table 4 compares various non-clinical DED screening methods. We note that these are results reported in the original studies, and not on our dataset. This is because (a) the DEvice hygrometer [17] is a prototype and not commercially available, making similar humidity quantification difficult, and (b) both DryEyeRhythm [46] and EyeScore [68] apps use the OSDI questionnaire, which is unreliable in our setup due to potential language barriers. These results should be interpreted with caution, as they are based on different patient sizes and populations. For example, the EyeScore [68] app was evaluated on only 20 eyes (13 DED and 7 normal) and showed 100% accuracy, sensitivity and specificity. The small evaluation set requires larger-scale testing. Moreover, it uses a custom 'eye-score' based on

Method	Phone	Eval Set Size	Acc.	Sens.	Spec.
DryEyeRhythm [46]	Yes	82 (42 DED)	71.95	50.00	95.00
EyeScore [68]	Yes	20 (13 DED)	100.0	100.0	100.0
DEvice©(AI, Rome, Italy) [17]	No	40 (20 DED)	77.50	60.00	95.00
DEDetector - Manual (ours)	Yes	46 (18 DED)	80.43	77.78	82.14
DEDetector - Auto (ours)	Yes	46 (18 DED)	78.26	77.78	78.57

Table 4. Performance comparison of non-clinical DED screening devices compared with *DEDetector*.

Indicator	Threshold (s)	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1
FBUT	10	73.91	72.22	75.00	68.42
Manual-NIBUT	10	80.43	77.78	82.14	75.67
Auto-NIBUT	10	<u>78.26</u>	<u>77.78</u>	<u>78.57</u>	<u>73.68</u>

Table 5. Performance of various tear break-up times in screening Dry Eye Disease (DED). The numbers highlighted in bold show the best performance in diagnosis, and the underlined text highlights the second best performance.

features such as OSDI symptoms, blink and partial blink rate, contact lens usage, and patient demographics [68], with manually identified weights for these features, which can lead to overfitting. Besides EyeScore, *DEDetector* achieves the highest sensitivity and overall accuracy, which is desirable in a screening tool.

6.3 Accuracy in Screening Dry Eye Disease

6.3.1 Comparison of FBUT and NIBUT. Here, we compare the performance of FBUT and NIBUT (both manual and automated) in their ability to screen for Dry Eye Disease (DED). Predictions are obtained by thresholding FBUT / NIBUT. As per the DEWS II Diagnostic Methodology Report [63], the recommended cutoff time for FBUT is 10s, i.e., if the eye has an FBUT < 10s, it is indicative of DED. At 10s, we found FBUT to be 73.91% accurate. As stated earlier, DED is a multifactorial disease, which requires several tests and examinations to be diagnosed accurately. Using only the FBUT (which examines only one aspect, i.e., tear film stability), we are able to diagnose DED with a sensitivity of 72.22% and specificity of 75%. This result is inline with previous studies [56]. Table 5 shows the diagnostic performance of FBUT as well as NIBUT. While thresholds of 10s [32] and 12s [58] have been used in previous studies, we found a NIBUT threshold of 10s to work well in our setting. Predictions for a 10s threshold on manually annotated NIBUT obtains 80.43% accuracy, with significant improvements in both sensitivity (77.78%) and specificity (82.14%). NIBUT found in an automated manner using the proposed video-processing pipeline closely resembles the performance of manually annotated NIBUT, with an accuracy of 78.26% (sensitivity 77.78% and specificity 78.57%). It must be noted that the automated NIBUT based predictions outperform those obtained from clinically performed FBUT test. While the problems with the FBUT test have been highlighted earlier, this result demonstrates the utility of the proposed system, which is completely automatic, non-invasive and can be performed without the presence of a medical professional.

6.4 Correlation between FBUT and NIBUT

Figure 8 shows the correlation between the FBUT and manual NIBUT. We obtain a Pearson correlation coefficient of 0.406 ($p=0.005$) between the 2 break-up times. While the DEWS II Diagnostic Methodology Report [63] reports that FBUT and NIBUT are “well-correlated”, there is a large variance in the correlation values reported in the literature (e.g. $Spearman_r = 0.18$ [53] to $Spearman_r = 0.832$) [57].

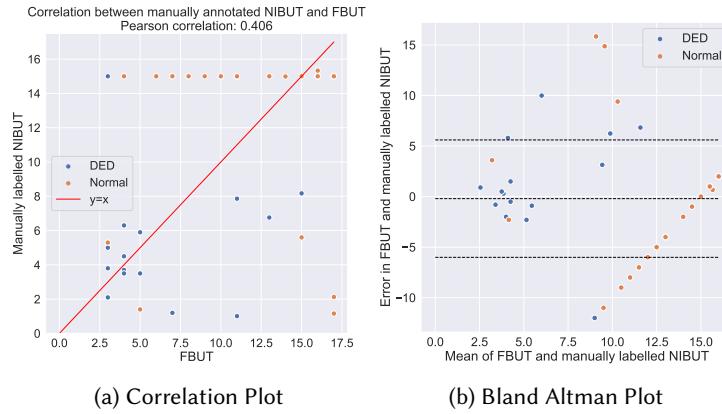


Fig. 8. Correlation between manual NIBUT and FBUT

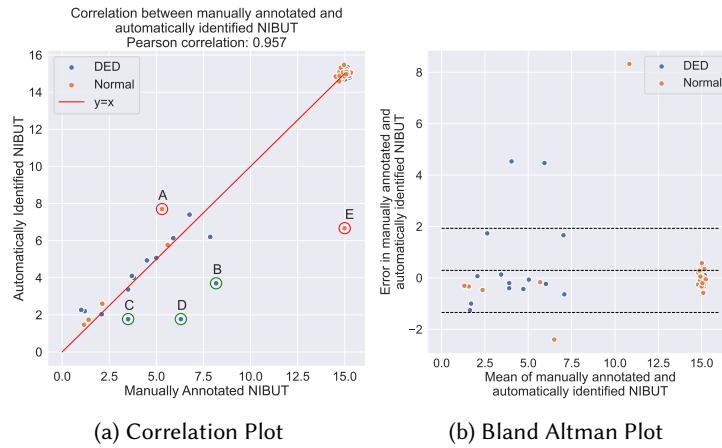


Fig. 9. Correlation between manual NIBUT and automatic NIBUT. Random noise has been added to the cluster of points at (15,15) to show the volume of points. These demonstrate the eyes for which no non-invasive tear film break-up was seen, either manually or automatically, in the first 15s. Points of disagreement between automatic and manual NIBUT have been labelled (points A-E).

Additionally, there have been studies reporting disagreement between automated NIBUT recorded using different devices [35, 52, 67]. Zeri et al. [67] compared NIBUT obtained from three devices: EasyTear View+ (Easytear, Rovereto), Polaris, and Sirius+ (CSO, Firenze) with each other, as well as with FBUT. Their findings revealed a poor correlation between NIBUT and FBUT, and also indicated that measurements from different devices were not interchangeable. Lee et al. [35] found poorly correlated ($Spearman_r = 0.061$) NIBUT values recorded using Tomey RT-7000 Auto Refractor-Keratometer and Oculus Keratograph 5M. Another study [41] compared the NIBUT obtained using 3 devices - instrument-mounted wide-field white light clinical interferometry, instrument-mounted keratoscopy and hand-held keratoscopy. The study found good correlation between the two instrument-mounted devices but poor correlation between the hand-held device and instrument-mounted ones. In summary, from the varied results of the studies, we conclude that there may be confounding factors in these

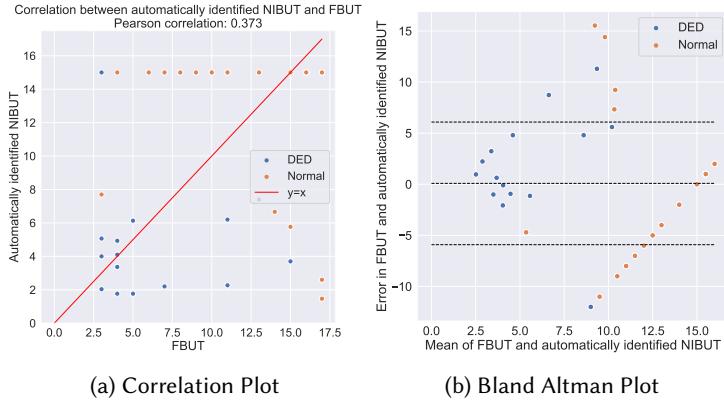


Fig. 10. Correlation between automatic NIBUT and FBUT.

devices / studies, and that no reliable conclusion can be made from the moderate correlation between FBUT and NIBUT observed in our study. The utility of the proposed method lies in the fact that it is capable of screening DED more accurately than the clinically-prevalent FBUT.

Figure 9 shows the correlation between the manually labelled NIBUT and automatic NIBUT. We obtain a very strong correlation ($Pearson_r = 0.957$), between the automated and manual NIBUT, thereby demonstrating the utility of the proposed video-processing pipeline. Finally, Figure 10 shows the correlation between the FBUT and the automatic NIBUT. We observe a similar correlation ($Pearson_r = 0.373, p < 0.05$) between FBUT and Automated NIBUT as we did for the manual NIBUT.

6.5 Limitations

In this section, we describe the limitations of the proposed method, and specific cases of disagreement between the manual annotation and video processing pipeline. These points are highlighted as A, B, C, D, and E in Figure 9a. Points B,C and D have a smaller automatically identified NIBUT than the manually labelled NIBUT. These are cases where the manual annotation missed tear film break-ups (about 6.5% of all videos). This points towards the benefits of automation, thereby removing subjectivity from the analysis. Points A and E are points where the disagreement in predictions arose because of the video-processing pipeline. The pipeline missed an earlier break for the eye in Point A, because the break was present in the upper half of the mires, and our method only consider the lower half of the mires. Point E is a normal eye, incorrectly being flagged as DED. This is because of an unusual type of mire shape deformation seen in the video, which the pipeline incorrectly flags as a tear film break. A check for the linear shape of the break was not incorporated in the pipeline in order to maintain its simplicity. Finally, mire center detection and pre-processing failed for one of the videos, which required a manual selection of the center for 1 frame (1 click in total).

In our current setup, only the data was collected using a smartphone, video processing was performed offline on a CPU. We acknowledge the importance of on-device processing and will consider it as the next step. We note that none of the video-processing pipeline steps are computationally intensive and can be executed on modern smartphones with some engineering effort.

7 Discussion

In this paper, our objective is to develop an affordable, portable, smartphone-based system for dry eye screening. *DEDetector* leverages the SmartKC attachment [16], consisting of a 3D-printed Placido attachment and an LED light

array, to record video data capturing Placido reflections on the cornea. The proposed video processing pipeline analyzes these videos, identifies breaks, and computes NIBUT. To identify relevant line breaks, our system relies on our detailed manual annotation of the captured videos. Our evaluation on 46 distinct eyes showed a strong correlation between the manual NIBUT and automated NIBUT. *DEDetector* also achieved an acceptable sensitivity of 77.78% and specificity of 82.14% for DED screening, based on ground truth provided by an ophthalmologist. We envision *DEDetector* as a valuable tool for detecting and monitoring the current state of DED. Although our results are promising, further research is needed to enhance the system's robustness.

7.1 DED Management

The primary objective of *DEDetector* is to screen DED in a low-cost, portable, and user-friendly manner. Given that dry eye is a multifactorial condition lacking a gold-standard diagnostic marker [63], coupled with the rapid fluctuation of dryness in the eyes [7], instances of both under- and over-medication are prevalent [14]. Therefore, regular monitoring of DED is crucial for effective management.

Discussing the cost here, the 3D-printed Placido attachment [16] used to project mires costs \$33. The attachment is compatible with any smartphone with a 12MP (or more) camera [16]. The OnePlus 6T smartphone used in our study costs \$350, making the total system cost \$383. This is ~50x cheaper than the Keratograph 5M, which cost upwards of \$20,000. Moreover, *DEDetector* automates the tedious process of manually identifying NIBUT from mire videos. For example, annotating our dataset's ~33,216 frames (46 videos, with mean length = 24.07s @ 30fps) requires significant human effort. *DEDetector* automates this through our proposed video processing pipeline, yielding NIBUT values strongly correlated with manual results. We believe that *DEDetector*'s fully automated nature and low cost, combined with the ubiquity of smartphones, could enable large scale screening of DED.

Once diagnosed, managing dry eye imposes a significant economic burden. In the United States, the average cost of managing DED was estimated to be \$11,302 per patient, with a societal cost of \$55.4 billion overall [66]. To mitigate the cost, there have been proposals advocating for the self-management of dry eye medication, drawing parallels with the approach adopted to manage type-1 diabetes [65]. For instance, individuals with type-1 diabetes regularly monitor their blood glucose levels, and based on these readings, they make informed decisions about insulin dosages and dietary choices to maintain their blood sugar levels within a target range. By introducing a portable and easy to use device like *DEDetector*, we aim to empower DED patients with a proactive and personalized approach to self-manage their condition. This approach not only provides a more informed means of mitigating the economic burden but also empowers DED patients to take an active role in their ongoing care.

7.2 N-in-1 Devices

The integration of multiple diagnostic functionalities into a single device is widely embraced in the medical domain. For instance, the Remidio-Visionix VX650¹ stands out by performing 24 eye exams in less than 90 seconds. Even the widely used Keratograph 5M [36] medical device for automated NIBUT based dry eye diagnosis, also serves as a corneal topographer for Keratoconus diagnosis and assists with soft contact lens fitting. The appeal of such N-in-1 devices lies in their ability to offer diverse functionalities in a single apparatus. This not only reduces costs compared to procuring multiple specialized devices but also optimizes clinic space usage. In this context, the combination of *DEDetector* and SmartKC [16] in a unified device for diagnosing DED and Keratoconus, respectively, seems advantageous. Importantly, in the US, the incidence of Keratoconus is minimal, with a prevalence rate of 0.05% [28], while dry eye is highly prevalent, with a 8.1% prevalence rate [7]. Only a small fraction of optometrists in the US have access to a corneal topographer. Optometrists equipped with a singular device like *DEDetector* for dry eye screening can leverage it for Keratoconus screening as well, potentially leading to incidental findings.

¹<https://www.remidio.com/products/remidio-visionix-vx650>

8 Conclusion

Low-cost, portable screening for DED can significantly contribute to its effective diagnosis and management. In this work, we presented *DEDetector*, a smartphone-based DED diagnostic system using a Placido ring attachment and a novel video processing pipeline. We evaluated our system on 46 distinct eyes, and achieved a sensitivity of 77.78% and a specificity of 82.14% for dry eye screening, surpassing the manual FBUT method. The evaluation also revealed a strong correlation (Pearson correlation coefficient of 0.957) between manually annotated NIBUT and automated NIBUT from the video processing pipeline. While additional research research is needed to bring *DEDetector* at par with commercial NIBUT-based medical devices, we believe that *DEDetector* has the potential to be used as an efficient and accessible dry eye screening tool.

References

- [1] Filippo Attivissimo, Vito Ivano D'Alessandro, Attilio Di Nisio, Giuliano Scarcelli, Justin Schumacher, and Anna Maria Lucia Lanzolla. 2023. Performance evaluation of image processing algorithms for eye blinking detection. *Measurement* 223 (2023), 113767. <https://doi.org/10.1016/j.measurement.2023.113767>
- [2] Robert E. Berger and Stanley Corrsin. 1974. A surface tension gradient mechanism for driving the pre-corneal tear film after a blink. *Journal of Biomechanics* 7, 3 (1974), 225–238. [https://doi.org/10.1016/0021-9290\(74\)90013-X](https://doi.org/10.1016/0021-9290(74)90013-X)
- [3] G. Bradski. 2000. The OpenCV Library. *Dr. Dobb's Journal of Software Tools* (2000).
- [4] Justin Chan, Sharat Raju, Rajalakshmi Nandakumar, Randall Bly, and Shyamnath Gollakota. 2019. Detecting middle ear fluid using smartphones. *Sci. Transl. Med.* 11, 492 (May 2019), eaav1102.
- [5] Jaykaran Charan and Tamoghna Biswas. 2013. How to calculate sample size for different study designs in medical research? *Indian J. Psychol. Med.* 35, 2 (April 2013), 121–126.
- [6] Christina Chu, Mark Rosenfield, and Joan Portello. 2014. Blink Patterns: Reading from a Computer Screen Versus Hard Copy. *Optometry and vision science : official publication of the American Academy of Optometry* 91 (01 2014). <https://doi.org/10.1097/OPX.0000000000000157>
- [7] Jennifer P Craig, Kelly K Nichols, Esen K Akpek, Barbara Caffery, Harminder S Dua, Choun-Ki Joo, Zuguo Liu, J Daniel Nelson, Jason J Nichols, Kazuo Tsubota, and Fiona Stapleton. 2017. TFOS DEWS II definition and classification report. *Ocul. Surf.* 15, 3 (July 2017), 276–283.
- [8] Lukasz Cwiklik. 2016. Tear film lipid layer: A molecular level view. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1858, 10 (2016), 2421 – 2430. <https://doi.org/10.1016/j.bbamem.2016.02.020>
- [9] Artem Dementyev and Christian Holz. 2017. DualBlink: A Wearable Device to Continuously Detect, Track, and Actuate Blinking For Alleviating Dry Eyes and Computer Vision Syndrome. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 1, Article 1 (mar 2017), 19 pages. <https://doi.org/10.1145/3053330>
- [10] Michelle Elliott, Heidi Fandrich, Trefford Simpson, and Desmond Fonn. 1998. Analysis of the repeatability of tear break-up time measurement techniques on asymptomatic subjects before, during and after contact lens wear. *Contact Lens and Anterior Eye* 21, 4 (1998), 98–103. [https://doi.org/10.1016/S1367-0484\(98\)80002-7](https://doi.org/10.1016/S1367-0484(98)80002-7)
- [11] Kimberly F Farrand, Moshe Fridman, Ipek Özer Stillman, and Debra A Schaumberg. 2017. Prevalence of diagnosed dry eye disease in the United States among adults aged 18 years and older. *Am. J. Ophthalmol.* 182 (Oct. 2017), 90–98.
- [12] A. Fogelton and W. Benesova. 2016. Eye blink detection based on motion vectors analysis. *Computer Vision and Image Understanding* 148 (2016), 23–33. <https://doi.org/10.1016/j.cviu.2016.03.011> Special issue on Assistive Computer Vision and Robotics - "Assistive Solutions for Mobility, Communication and HMI".
- [13] Gary N. Foulks. 2003. Challenges and Pitfalls in Clinical Trials of Treatments for Dry Eye. *The Ocular Surface* 1, 1 (2003), 20–30. [https://doi.org/10.1016/S1542-0124\(12\)70004-6](https://doi.org/10.1016/S1542-0124(12)70004-6)
- [14] Frederick T Fraunfelder, James J Scibba, and William D Mathers. 2012. The role of medications in causing dry eye. *Journal of ophthalmology* 2012 (2012).
- [15] G Fuentes-Páez, J M Herreras, Y Cordero, A Almaraz, M J González, and M Calonge. 2011. Lack of concordance between dry eye syndrome questionnaires and diagnostic tests. *Arch. Soc. Esp. Oftalmol.* 86, 1 (Jan. 2011), 3–7.
- [16] Siddhartha Gairola, Murtuza Bohra, Nadeem Shaheer, Navya Jayaprakash, Pallavi Joshi, Anand Balasubramaniam, Kaushik Murali, Nipun Kwatra, and Mohit Jain. 2022. SmartKC: Smartphone-Based Corneal Topographer for Keratoconus Detection. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 155 (dec 2022), 27 pages. <https://doi.org/10.1145/3494982>
- [17] Daniele Gaudenzi, Tommaso Mori, Salvatore Crugliano, Antonella Grasso, Carlo Frontini, Antonella Carducci, Siddharth Yadav, Roberto Sgrulletta, Emiliano Schena, Marco Coassini, and Antonio Di Zazzo. 2022. AS-OCT and Ocular Hygrometer as Innovative Tools in Dry Eye Disease Diagnosis. *Applied Sciences* 12, 3 (2022). <https://doi.org/10.3390/app12031647>

- [18] Johnny Gayton. 2009. Etiology, prevalence, and treatment of dry eye disease. *Clinical ophthalmology (Auckland, N.Z.)* 3 (02 2009), 405–12. <https://doi.org/10.2147/OPTH.S5555>
- [19] Tomoko Goto, Xiaodong Zheng, Stephen D Klyce, Hisashi Kataoka, Toshihiko Uno, Mike Karon, Yoshiyuki Tatematsu, Takeo Bessyo, Kazuo Tsubota, and Yuichi Ohashi. 2003. A new method for tear film stability analysis using videokeratography. *American Journal of Ophthalmology* 135, 5 (2003), 607–612. [https://doi.org/10.1016/S0002-9394\(02\)02221-3](https://doi.org/10.1016/S0002-9394(02)02221-3)
- [20] Koray Gumus, Charlene Hong Crockett, Kavita Rao, Elizabeth Yeu, Mitchell P. Weikert, Mariko Shirayama, Shigeki Hada, and Stephen C. Pflugfelder. 2011. Noninvasive Assessment of Tear Stability with the Tear Stability Analysis System in Tear Dysfunction Patients. *Investigative Ophthalmology & Visual Science* 52, 1 (01 2011), 456–461. <https://doi.org/10.1167/iovs.10-5292> arXiv:https://arvojournals.org/arvo/content_public/journal/iovs/932967/z7g00111000456.pdf
- [21] Yuting Hong and Makoto Hasegawa. 2021. Proposal of Tear Meniscus Measurement for Minor Dry-eye Detection Using Smart-phone Camera and Ring-light. In *2021 36th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)*. 1–4. <https://doi.org/10.1109/ITC-CSCC52171.2021.9501487>
- [22] Takenori Inomata, Masao Iwagami, Masahiro Nakamura, Tina Shiang, Keiichi Fujimoto, Yuichi Okumura, Nanami Iwata, Kenta Fujio, Yoshimune Hiratsuka, Satoshi Hori, Kazuo Tsubota, Reza Dana, and Akira Murakami. 2020. Association between dry eye and depressive symptoms: Large-scale crowdsourced research using the DryEyeRhythm iPhone application. *The Ocular Surface* 18, 2 (2020), 312–319. <https://doi.org/10.1016/j.jtos.2020.02.007>
- [23] Takenori Inomata, Masahiro Nakamura, Masao Iwagami, Akie Midorikawa-Inomata, Jaemyoung Sung, Keiichi Fujimoto, Yuichi Okumura, Atsuko Eguchi, Nanami Iwata, Maria Miura, Kenta Fujio, Ken Nagino, Satoshi Hori, Kazuo Tsubota, Reza Dana, and Akira Murakami. 2020. Stratification of Individual Symptoms of Contact Lens–Associated Dry Eye Using the iPhone App DryEyeRhythm: Crowdsourced Cross-Sectional Study. *J Med Internet Res* 22, 6 (26 Jun 2020), e18996. <https://doi.org/10.2196/18996>
- [24] Takenori Inomata, Masahiro Nakamura, Masao Iwagami, Tina Shiang, Yusuke Yoshimura, Keiichi Fujimoto, Yuichi Okumura, Atsuko Eguchi, Nanami Iwata, Maria Miura, Satoshi Hori, Yoshimune Hiratsuka, Miki Uchino, Kazuo Tsubota, Reza Dana, and Akira Murakami. 2019. Risk Factors for Severe Dry Eye Disease: Crowdsourced Research Using DryEyeRhythm. *Ophthalmology* 126, 5 (2019), 766–768. <https://doi.org/10.1016/j.ophtha.2018.12.013>
- [25] Takenori Inomata, Masahiro Nakamura, Jaemyoung Sung, Akie Midorikawa-Inomata, Masao Iwagami, Kenta Fujio, Yasutsugu Akasaki, Yuichi Okumura, Keiichi Fujimoto, Atsuko Eguchi, Maria Miura, Ken Nagino, Hurramhon Shokirova, Jun Zhu, Mizu Kuwahara, Kunihiko Hiroswa, Reza Dana, and Akira Murakami. 2021. Smartphone-based digital phenotyping for dry eye toward P4 medicine: a crowdsourced cross-sectional study. *npj Digital Medicine* 4, 1 (20 Dec 2021), 171. <https://doi.org/10.1038/s41746-021-00540-2>
- [26] LV Prasad Eye Institute. 2019. *Dry Eye Disease Big Data Study*. Retrieved Apr 1, 2020 from <https://www.lvpei.org/media-event/dry-eye-disease-big-data-study>
- [27] Abraham Kayal. 2021. The Physiology of Tear Film. In *Dry Eye Syndrome*, Felicia M. Ferreri (Ed.). IntechOpen, Rijeka, Chapter 2. <https://doi.org/10.5772/intechopen.98945>
- [28] Robert H Kennedy, William M Bourne, and John A Dyer. 1986. A 48-year clinical and epidemiologic study of keratoconus. *American journal of ophthalmology* 101, 3 (1986), 267–273.
- [29] P. Ewen King-Smith, Barbara A. Fink, Jason J. Nichols, Kelly K. Nichols, Richard J. Braun, and Geoffrey B. McFadden. 2009. The Contribution of Lipid Layer Movement to Tear Film Thinning and Breakup. *Investigative Ophthalmology & Visual Science* 50, 6 (06 2009), 2747–2756. <https://doi.org/10.1167/iovs.08-2459>
- [30] P Ewen King-Smith, Erich A Hinel, and Jason J Nichols. 2009. Application of a novel interferometric method to investigate the relation between lipid layer thickness and tear film thinning. *Invest Ophthalmol Vis Sci* 51, 5 (Dec. 2009), 2418–2423.
- [31] P Ewen King-Smith, Samuel H Kimball, and Jason J Nichols. 2014. Tear film interferometry and corneal surface roughness. *Invest Ophthalmol Vis Sci* 55, 4 (April 2014), 2614–2618.
- [32] Stephen R. Tonge Lakhbir S. Mengher, Anthony J. Bron and David J. Gilbert. 1985. A non-invasive instrument for clinical assessment of the pre-corneal tear film stability. *Current Eye Research* 4, 1 (1985), 1–7. <https://doi.org/10.3109/02713688508999960> arXiv:<https://doi.org/10.3109/02713688508999960> PMID: 3979089.
- [33] Eric C. Larson, Mayank Goel, Gaetano Borriello, Sonya Heltshe, Margaret Rosenfeld, and Shwetak N. Patel. 2012. SpiroSmart: using a microphone to measure lung function on a mobile phone. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (Pittsburgh, Pennsylvania) (UbiComp '12)*. Association for Computing Machinery, New York, NY, USA, 280–289. <https://doi.org/10.1145/2370216.2370261>
- [34] Euihyeok Lee, Chulhong Min, and Seungwoo Kang. 2018. Towards a Wearable Assistant to Prevent Computer Vision Syndrome. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers (Singapore, Singapore) (UbiComp '18)*. Association for Computing Machinery, New York, NY, USA, 122–125. <https://doi.org/10.1145/3267305.3267587>
- [35] Ryan Lee, Sharon Yeo, Han Tun Aung, and Louis Tong. 2016. Agreement of noninvasive tear break-up time measurement between Tomey RT-7000 Auto Refractor-Keratometer and Oculus Keratograph 5M. *Clin. Ophthalmol.* 10 (Sept. 2016), 1785–1790.

- [36] Tian Lei, Song Wenxiu, Wang Zhiqun, Zhang Yang, and Sun Xuguang. 2017. Diagnosis of dry eye disease using keratograph 5M. *Journal of Capital Medical University* 38, 1, Article 11 (2017), 5 pages. <https://doi.org/10.3969/j.issn.1006-7795.2017.01.003>
- [37] Hui Lu, Michael R Wang, Jianhua Wang, and Meixiao Shen. 2014. Tear film measurement by optical reflectometry technique. *J Biomed Opt* 19, 2 (Feb. 2014), 027001.
- [38] Alex Mariakakis, Jacob Baudin, Eric Whitmire, Vardhman Mehta, Megan A. Banks, Anthony Law, Lynn McGrath, and Shwetak N. Patel. 2017. PupilScreen: Using Smartphones to Assess Traumatic Brain Injury. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 81 (sep 2017), 27 pages. <https://doi.org/10.1145/3131896>
- [39] Paul McCann, Alison G Abraham, Adhuna Mukhopadhyay, Kanella Panagiotopoulou, Hongan Chen, Thanitsara Rittiphairoj, Darren G Gregory, Scott G Hauswirth, Cristos Ifantides, Riaz Qureshi, Su-Hsun Liu, Ian J Saldanha, and Tianjing Li. 2022. Prevalence and incidence of dry eye and meibomian gland dysfunction in the United States: A systematic review and meta-analysis. *JAMA Ophthalmol.* 140, 12 (Dec 2022), 1181–1192.
- [40] Marguerite McDonald, Dipen A Patel, Michael S Keith, and Sonya J Snedecor. 2016. Economic and humanistic burden of dry eye disease in Europe, North America, and Asia: A systematic literature review. *Ocul. Surf.* 14, 2 (April 2016), 144–167.
- [41] Kenneth J Blades Michael Tm Wang, Paul J Murphy and Jennifer P Craig. 2018. Comparison of non-invasive tear film stability measurement techniques. *Clinical and Experimental Optometry* 101, 1 (2018), 13–17. <https://doi.org/10.1111/cxo.12546> arXiv:<https://doi.org/10.1111/cxo.12546> PMID: 28503827.
- [42] Akie Midorikawa-Inomata, Takenori Inomata, Shuko Nojiri, Masahiro Nakamura, Masao Iwagami, Keiichi Fujimoto, Yuichi Okumura, Nanami Iwata, Atsuko Eguchi, Hitomi Hasegawa, Hikaru Kinouchi, Akira Murakami, and Hiroyuki Kobayashi. 2019. Reliability and validity of the Japanese version of the Ocular Surface Disease Index for dry eye disease. *BMJ Open* 9, 11 (2019). <https://doi.org/10.1136/bmjopen-2019-033940> arXiv:<https://bmjopen.bmjjournals.org/content/9/11/e033940.full.pdf>
- [43] Jun Moon, Kyoung Woo Kim, and Nam Moon. 2016. Smartphone use is a risk factor for pediatric dry eye disease according to region and age: a case control study. *BMC Ophthalmology* 16 (12 2016). <https://doi.org/10.1186/s12886-016-0364-4>
- [44] J E Moore, J E Graham, E A Goodall, D A Dartt, A Leccisotti, V E McGilligan, and T C B Moore. 2009. Concordance between common dry eye diagnostic tests. *Br. J. Ophthalmol.* 93, 1 (Jan. 2009), 66–72.
- [45] Ken Nagino, Yuichi Okumura, Masahiro Yamaguchi, Jaemyoung Sung, Masashi Nagao, Kenta Fujio, Yasutsugu Akasaki, Tianxiang Huang, Kunihiko Hirosawa, Masao Iwagami, Akie Midorikawa-Inomata, Keiichi Fujimoto, Atsuko Eguchi, Yukinobu Okajima, Koji Kakisu, Yuto Tei, Takefumi Yamaguchi, Daisuke Tomida, Masaki Fukui, Yukari Yagi-Yaguchi, Yuichi Hori, Jun Shimazaki, Shuko Nojiri, Yuki Morooka, Alan Yee, Maria Miura, Mizu Ohno, and Takenori Inomata. 2023. Diagnostic Ability of a Smartphone App for Dry Eye Disease: Protocol for a Multicenter, Open-Label, Prospective, and Cross-sectional Study. *JMIR Res Protoc* 12 (13 Mar 2023), e45218. <https://doi.org/10.2196/45218>
- [46] Yuichi Okumura, Takenori Inomata, Akie Midorikawa-Inomata, Jaemyoung Sung, Kenta Fujio, Yasutsugu Akasaki, Masahiro Nakamura, Masao Iwagami, Keiichi Fujimoto, Atsuko Eguchi, Maria Miura, Ken Nagino, Kunihiko Hirosawa, Tianxiang Huang, Mizu Kuwahara, Reza Dana, and Akira Murakami. 2022. DryEyeRhythm: A reliable and valid smartphone application for the diagnosis assistance of dry eye. *The Ocular Surface* 25 (2022), 19–25. <https://doi.org/10.1016/j.jtos.2022.04.005>
- [47] Sudi Patel and Larysa Tutchenko. 2019. The refractive index of the human cornea: A review. *Cont Lens Anterior Eye* 42, 5 (May 2019), 575–580.
- [48] Stephen Pflugfelder, Lauren Nakhleh, Yasushi Kikukawa, Shin Tanaka, and Takuya Kosugi. 2022. Non-invasive tear break-up detection with the Kowa DR-1 α and its relationship to dry eye clinical severity. *Int. J. Mol. Sci.* 23, 23 (Nov. 2022), 14774.
- [49] Ruth Miyuki Santo, Felipe Ribeiro-Ferreira, Milton Ruiz Alves, Jonathan Epstein, and Priscila Novaes. 2015. Enhancing the cross-cultural adaptation and validation process: linguistic and psychometric testing of the Brazilian–Portuguese version of a self-report measure for dry eye. *Journal of Clinical Epidemiology* 68, 4 (2015), 370–378. <https://doi.org/10.1016/j.jclinepi.2014.07.009>
- [50] Rhett M. Schiffman, Murray Dale Christianson, Gordon Jacobsen, Jan D. Hirsch, and Brenda L. Reis. 2000. Reliability and Validity of the Ocular Surface Disease Index. *Archives of Ophthalmology* 118, 5 (05 2000), 615–621. <https://doi.org/10.1001/archophht.118.5.615> arXiv:<https://jamanetwork.com/journals/jamaophthalmology/articlepdf/413145/ecs90139.pdf>
- [51] Chika Shigeyasu, Masakazu Yamada, Norihiko Yokoi, Motoko Kawashima, Kazuhisa Suwaki, Miki Uchino, Yoshimune Hiratsuka, Kazuo Tsubota, and On Behalf Of The Decs-J Study Group. 2020. Characteristics and utility of fluorescein breakup patterns among dry eyes in clinic-based settings. *Diagnostics (Basel)* 10, 9 (Sept. 2020).
- [52] Swati Singh, Saumya Srivastav, Zarin Modiwala, Mohammed Hasnat Ali, and Sayan Basu. 2023. Repeatability, reproducibility and agreement between three different diagnostic imaging platforms for tear film evaluation of normal and dry eye disease. *Eye* 37, 10 (01 Jul 2023), 2042–2047. <https://doi.org/10.1038/s41433-022-02281-2>
- [53] John E. Sutphin, Gui-shuang Ying, Vatinee Y. Bunya, Yinxin Yu, Meng C. Lin, Kathleen McWilliams, Elizabeth Schmucker, Eric J. Kuklinski, Penny A. Asbell, Maureen G. Maguire, for the Dry Eye Assessment, and Management (DREAM) Study Research Group. 2022. Correlation of Measures From the OCULUS Keratograph and Clinical Assessments of Dry Eye Disease in the Dry Eye Assessment and Management Study. *Cornea* 41, 7 (2022). https://journals.lww.com/corneajnl/fulltext/2022/07000/correlation_of_measures_from_the_oculus.7.aspx

- [54] Deborah F. Sweeney, Thomas J. Millar, and Shiwani R. Raju. 2013. Tear film stability: A review. *Experimental Eye Research* 117 (2013), 28–38. <https://doi.org/10.1016/j.exer.2013.08.010> Tears: A Unique Mucosal Surface Secretion.
- [55] Maho Taniai, Kaito Okazaki, and Makoto Hasegawa. 2023. Improvement of Tear Meniscus Measurement for Dry Eye Detection Using Smartphone and Ring Light. In *2023 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*. 1–4. <https://doi.org/10.1109/ICCE-Asia59966.2023.10326431>
- [56] Sania Vidas Pauk, Igor Petriček, Tomislav Jukić, Smiljka Popović-Suić, Martina Tomić, Miro Kalauz, Sonja Jandroković, and Sanja Masnec. 2019. Noninvasive tear film break-up time assessment using handheld lipid layer examination instrument. *Acta Clin. Croat.* 58, 1 (March 2019), 63–71.
- [57] Sania Vidas Pauk, Igor Petriček, Tomislav Jukić, Smiljka Popović-Suić, Martina Tomić, Miro Kalauz, Sonja Jandroković, and Sanja Masnec. 2019. Noninvasive tear film break-up time assessment using handheld lipid layer examination instrument. *Acta Clin. Croat.* 58, 1 (March 2019), 63–71.
- [58] Sania Vidas Pauk, Igor Petriček, Martina Tomić, Tomislav Bulum, Sonja Jandroković, Maja Pauk Gulić, Miro Kalauz, and Dina Lešin Gaćina. 2023. Diagnostic accuracy of non-invasive tear film break-up time assessed by the simple manual interferometric device. *Contact Lens and Anterior Eye* 46, 2 (2023), 101776. <https://doi.org/10.1016/j.clae.2022.101776>
- [59] Luca Vigo, Marco Pellegrini, Federico Bernabei, Francesco Carones, Vincenzo Scorcia, and Giuseppe Giannaccare. 2020. Diagnostic performance of a novel noninvasive workup in the setting of dry eye disease. *J. Ophthalmol.* 2020 (Dec. 2020), 5804123.
- [60] Edward Jay Wang, William Li, Doug Hawkins, Terry Gernsheimer, Colette Norby-Slycord, and Shwetak N. Patel. 2016. HemaApp: noninvasive blood screening of hemoglobin using smartphone cameras. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (Heidelberg, Germany) (UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 593–604. <https://doi.org/10.1145/2971648.2971653>
- [61] Edward Jay Wang, Junyi Zhu, Mohit Jain, Tien-Jui Lee, Elliot Saba, Lama Nachman, and Shwetak N. Patel. 2018. Seismo: Blood Pressure Monitoring using Built-in Smartphone Accelerometer and Camera. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal QC, Canada) (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3173574.3173999>
- [62] Mark D.P. Willcox, Pablo Argüeso, Georgi A. Georgiev, Juha M. Holopainen, Gordon W. Laurie, Tom J. Millar, Eric B. Papas, Jannick P. Rolland, Tannin A. Schmidt, Ulrike Stahl, Tatiana Suarez, Lakshman N. Subbaraman, Omur Ö. Uçakhan, and Lyndon Jones. 2017. TFOS DEWS II Tear Film Report. *The Ocular Surface* 15, 3 (2017), 366–403. <https://doi.org/10.1016/j.jtos.2017.03.006> TFOS International Dry Eye WorkShop (DEWS II).
- [63] James S Wolffsohn, Reiko Arita, Robin Chalmers, Ali Djalilian, Murat Dogru, Kathy Dumbleton, Preety K Gupta, Paul Karpecki, Sihem Lazreg, Heiko Pult, Benjamin D Sullivan, Alan Tomlinson, Louis Tong, Edoardo Villani, Kyung Chul Yoon, Lyndon Jones, and Jennifer P Craig. 2017. TFOS DEWS II Diagnostic Methodology report. *Ocul. Surf.* 15, 3 (July 2017), 539–574.
- [64] Meng Xue, Yuyang Zeng, Shengkang Gu, Qian Zhang, Bowei Tian, and Changzheng Chen. 2024. SDE: Early Screening for Dry Eye Disease with Wireless Signals. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 4, Article 194 (jan 2024), 23 pages. <https://doi.org/10.1145/3631438>
- [65] Sharon Yeo and Louis Tong. 2018. Coping with dry eyes: a qualitative approach. *BMC ophthalmology* 18 (2018), 1–9.
- [66] Junhua Yu, Carl V Asche, and Carol J Fairchild. 2011. The economic burden of dry eye disease in the United States: a decision tree analysis. *Cornea* 30, 4 (April 2011), 379–387.
- [67] Fabrizio Zeri, Giulia Carlotta Rizzo, Erika Ponzini, and Silvia Tavazzi. 2024. Comparing automated and manual assessments of tear break-up time using different non-invasive devices and a fluorescein procedure. *Scientific Reports* 14, 1 (30 Jan 2024), 2516. <https://doi.org/10.1038/s41598-024-52686-0>
- [68] Sydney Zhang and Julio Echegoyen. 2023. Design and Usability Study of a Point of Care mHealth App for Early Dry Eye Screening and Detection. *Journal of Clinical Medicine* 12, 20 (2023). <https://doi.org/10.3390/jcm12206479>

A Appendix

A.1 DED Types

Every time we blink, a film of tear spreads over the eye, which protects and lubricates the eye, and keeps the surface of the eye smooth and clear. This tear film comprises of three layers: Mucin layer, Aqueous layer and Lipid layer [8]. *Mucin layer*, also known as mucus layer, is the innermost layer residing at the surface of cornea, and is secreted by conjunctival goblet cells. It lubricates and protects cornea, and also helps spread the aqueous layer over the eye's surface after blink. It is 1 µm thick. *Aqueous layer*, also known as watery layer, is the middle layer, secreted by lacrimal glands above the eyelids. It fights against outside impurity, and is 4 µm thick. *Lipid layer*, also known as oily layer, is the outermost layer of the tear film, secreted by the meibomian glands lining

the margins of the eyelids. It is $0.1\text{ }\mu\text{m}$ thick. It prevents the aqueous layer from evaporating; if the lipid layer is compromised, the aqueous layer gets exposed.

There are two types of dry eyes disease: Evaporative Dry Eye and Aqueous Tear Deficient Dry Eye. *Evaporative dry eye* is caused by deficient lipid layer leading to increased tear evaporation, while *aqueous tear deficient dry eye* is caused by deficient aqueous layer causing reduced tear secretion. Both evaporative and aqueous deficient types exist on a spectrum, and they can frequently co-occur [7].

A.2 Causes and Symptoms

One of the most common causes of DED is reduced blink rate, which usually happens when an individual is concentrating, e.g., while reading, driving, and especially working on computer. Computer vision syndrome (CVS) is a condition resulting from prolonged computer, tablet, e-reader and cell phone use. One of the main symptoms of CVS is dry eyes. Normal blink rate of an individual is 12-16 blinks/minute, however if the blink rate falls below 6 blinks/min, it can lead to dry eye. Hence DED is fairly common among the IT (Information Technology) workforce. With the deep penetration of smartphones, even children who are using smartphone for more than 3 hours a day are reporting DED [43]. Moreover, individuals are found to have significantly higher number of incomplete blinks when reading from the computer (7.02% of blinks) versus reading from a printed hard copy (4.33%) [6]. Incomplete or partial blinks also increase dry eye symptoms, as the tear film does not get formed. Apart from that, the common causes of eye dryness are blockage or dysfunction of meibomian glands, loosening and hanging of lower eye lids during old age (especially women after menopause), smoke or dry air hitting the eye, lens usage post eye-surgeries, and side effect of blood pressure and anti-histamine medicines.

The main symptom of dry eyes is a severe burning or scratchy sensation in the eyes. Other symptoms are pain in the eye, high sensitivity to light, mild redness in the eye, difficulty in wearing contact lenses, and blurred vision. Also individuals with dry eye usually complains of watery eyes, which seems counter-intuitive. It is called Reflex Tear production, which is the body's response to the irritation of dry eyes.

A.3 Diagnostics

DED is diagnosed using both subjective and clinical tests. Subjective tests for DED are carried through questionnaires measuring the frequency and severity of symptoms, environmental triggers and general vision-related quality of life. The Ocular Surface Disease Index (OSDI) questionnaire [50] is the most widely used questionnaire for DED clinical trials [63]. It consists of 12 questions aimed to gauge visual disturbances (blurred vision, gritty eyes, etc.), problems in performing visual actions (reading, driving or watching a TV/computer screen), and problems in various environmental conditions (humid, windy and air-conditioned). Based on the responses to these questions, the OSDI score can be calculated. A high OSDI score is indicative of DED. The commonly used clinical tests to diagnose DED are described below.

A.3.1 Fluorescein Break-Up Time (FBUT) test. The fluorescein break-up time test evaluates the stability of tear film. To measure FBUT, a sodium fluorescein (NaFl) impregnated paper strip is used, which is orange in color. The doctor places a drop of fluorescein dye on the tip of the strip. Using the hanging-drop method, that drop is placed onto the ocular surface. The patient is asked to blink a few times to spread the fluorescein dye and coats the tear film to cover the corneal surface. The patient is asked to blink twice and then to hold their eyes open for as long as possible. The doctor then examines the patient's eye using a broad-beam of slit lamp with a cobalt blue filter illumination. At the end of the second blink, the doctor starts a stopwatch and stop it after noticing the first tear break-up. Due to the staining by the dye, the tear break-up appears green under the blue light. This time taken from blink to the appearance of the first dry spot in the tear film is called FBUT. For each eye, three consecutive measurements are taken, and the time is averaged to obtain the FBUT. If FBUT is above 10s, the eyes are normal;

and if the FBUT is below 10s, the patient has dry eye. Basically, when FBUT is less than the blink rate threshold of 6 blinks/min, the ocular surface is left unprotected, thus exacerbating the signs and symptoms of dry eye.

A.3.2 Corneal Staining. Corneal staining involves instillation of dyes (such as sodium fluorescein, rose bengal, and lissamine green) on the ocular surface in an attempt to identify corneal epithelial cell defects. Spot-like staining of the ocular surface is an indicator of DED. To perform the corneal staining test, fluorescein-impregnated paper strips are used. (Other dyes such as lissamine green can also be used). The paper strip is touched gently on the inner lining of the lower eyelid. The patient is then asked to blink a few times to uniformly distribute dye. Damaged/abnormal cells retain the dyes and are stained in this procedure. Similar to the FBUT test, observation under a cobalt blue filter makes the stained cells appear green in color. Corneal staining is indicative of severe DED, but it is not well-correlated with disease severity in cases of mild/moderate DED [63].

A.3.3 Schirmer's test. Despite DEWS II [63] identifying hyperosmolarity as the third diagnostic test, our discussions with seven ophthalmologists working across three eye hospitals in India revealed a lack of available medical devices to measure hyperosmolarity. Instead, they predominantly rely on Schirmer's test for diagnosing dry eye. Schirmer's test measure the volume of tear production. To perform Schirmer's test, a Whatman# 41 filter paper is used. It is 5mm wide and 35mm long. The filter paper is folded at the notch (clearly indicated on the paper). The folded tip is inserted into the lower lid of the patient's eye by the doctor, ensuring that it is not touching the cornea or lashes. The patient is asked to keep their eyes closed for 5 minutes. After that, the filter paper is removed and the amount of wetting from the fold is measured. Sometimes a topical anesthetic is put in the eyes before the filter paper to prevent reflex tears, i.e., tears due to the irritation from the paper. Based on the wetting of the paper, dryness in the eye is determined: Normal ($\geq 15\text{mm}$), Mild (10-14mm), Moderate (5-9mm), and Severe ($< 5\text{mm}$). As per the DEWS II Diagnostic Methodology report [63], Schirmer's test should be used to diagnose severe Aqueous Tear Deficient Dry Eyes.

All three tests—FBUT, Corneal Staining, and Schirmer's—are cheap as they only require filter papers and dyes. However, it requires a doctor to administer these tests, and these tests are invasive and uncomfortable for the patient. Also, maintaining digital records of these tests is challenging, often resulting in doctors manually recording the results on patient's assessment sheets.

A.4 Treatment

One of the best and cheapest way to treat dry eye is by increasing the blink rate and relaxing the eye using the 20-20-20 rule. The rule says that for every 20 minutes spent looking at a digital display, an individual should look at something 20 feet away for 20 seconds. Doctors also usually prescribe lubricating eye drops, called artificial tears. These eye drops simply moisten the eyes, and needs to be applied several times a day. In severe DED, instead of eye drops, ointment or thick gel needs to be used. To clear blocked meibomian gland, hot fermentation is performed. Even there are devices like LipiFlow thermal pulsation which helps in unblocking the gland. If none of these method prove effective, the doctor may partially or completely plug the tear ducts. This procedure ensures that tears remain in the eyes for longer periods, as there is no outlet for drainage.