Advancements in ThermOcular image processing algorithms for precise ocular surface temperature analysis

by

Navid Shahsavari

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Science
in
Vision Science and Systems Design Engineering

Waterloo, Ontario, Canada, 2024

© Navid Shahsavari 2024

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The following publications have resulted from the work presented in this thesis:

- Shahsavari, Navid, et al. "Towards Precision in Ocular Surface Temperature Measurement: Improvements in the ThermOcular System's Image Segmentation." Journal of Computational Vision and Imaging Systems 9.1 (2023): 15-18.

Abstract

In recent advancements in the field of Ocular Surface Temperature (OST) measurement, this thesis presents significant enhancements to the ThermOcular system, primarily focusing on its application in real-world clinical and diagnostic settings. The ThermOcular system, initially developed for precise OST measurement using infrared thermography, serves as a pivotal tool in measuring and tracking the OST in different components of the ocular surface. Recognizing the potential and limitations of the existing system, this work aims to introduce improvements that significantly elevate its diagnostic accuracy, user-independence, and overall applicability in clinical environments.

In this thesis, a comprehensive strategy to refine the ThermOcular system is introduced, highlighting the simplification of the control point selection for clinicians using an innovative eye tag integration. This development is aimed at enhancing the registration process's efficiency by reducing the need for manual input and the associated error margin prevalent in the prior methodology. Alongside, the thesis describes an enhancement of image segmentation accuracy, using state-of-the-art machine learning models trained on a comprehensive dataset prepared for this purpose. These models enable more precise classification of ocular components, crucial for accurate OST measurement.

The thesis also addresses the challenge of artifacts due to the presence of eyelashes in thermal images and the effect of blinks on tracked OST, which previously compromised measurement accuracy. By developing and implementing algorithms for artifact detection and elimination, the thesis ensures that these common issues no longer compromise the reliability of OST assessments. This not only enhances measurement precision, but also contributes to the system's robustness against varied conditions.

In conclusion, the thesis encapsulates a significant advance in the domain of ocular

health diagnostics using the ThermOcular system. By focusing on automation, accuracy, and artifact mitigation, this work contributes to the development of a more reliable, efficient, and clinically applicable system for OST measurement. The advancements highlighted in this thesis not only underscore the potential of infrared thermography in ocular diagnostics, but also pave the way for future research in this evolving field.

Acknowledgements

This journey would not have been possible without the support and guidance of many individuals who have been instrumental in both my personal and professional growth during the course of my master's program. I would like to extend my heartfelt gratitude to each one of them.

First and foremost, I am immensely thankful to my supervisors, Dr. Paul Murphy and Dr. Alex Wong, for their unwavering support, guidance, and invaluable insights throughout my research. Their expertise and mentorship have been pivotal in shaping this work, and I am deeply grateful for their patience and encouragement.

I would also like to express my sincere appreciation to Dr. Ehsan Zare Bidaki for his guidance and assistance. His contributions and advice have been immensely helpful, and his willingness to share his knowledge and experience has been a great source of inspiration.

To my wife, Yasaman, words cannot adequately express my gratitude. You are my best friend, my confidant, and my greatest supporter. Embarking on the journey of immigration together, you have been my only friend, my family, and my homeland, all embodied in one incredible person. Your belief in me has been a guiding light, turning the challenges we faced into stepping stones. For every moment of doubt that you've dispelled, for every hardship that you've turned into a lesson of strength, and for every joy we've shared, I am profoundly thankful. Your presence in my life is a blessing and I am incredibly fortunate to have you by my side.

My family—mom, dad, brother, sister—has been my backbone throughout this journey. Being away from you has been challenging, but your endless love, encouragement, and sacrifices have not gone unnoticed. Your belief in my dreams has been a constant source

of comfort and motivation, and I am thankful for your unwavering support despite the distance.

Last but certainly not least, I want to thank Mana Monfared, my labmate and friend. Your help and support in my research have been invaluable, but it is your friendship and companionship that I cherish the most. Thank you for being there through the highs and lows, for the countless hours of discussion, and for making this journey a memorable one.

This thesis is not just a reflection of my work, but a testament to the collective effort and support of all those mentioned above and those I haven't named. To each and every one of you who contributed to this journey, whether named or unnamed, I am deeply grateful.

Dedication

To my wife, Yasaman, and my beloved family.

Table of Contents

Author's Declaration ii

Statement of Contributions iii

Abstract iv

Acknowledgements vi

Dedication viii

List of Figures xi

List of Tables xiv

1 Introduction 1

2 Background 4

3 ThermOcular 14

4 Segmentation 22

5 Registration 32

6 Artifact Mitigation Techniques 42

7 Software Design 48

8 Conclusion and future research 56

References 59

List of Figures

1.1 Sample images captured by the ThermOcular system. The left shows an IR thermography image highlighting the temperature gradients across the eye with the eyelashes appearing prominently due to their lower temperature. The right panel displays a high-resolution visual spectrum image, capturing detailed features such as iris, sclera, and pupil. 2

3.1 Figure showing ThermOcular dual IR-Visual camera system mounted on slit lamp. 15

3.2 Schematic representation of the ThermOcular system's operational workflow. 18

3.3 PSPNet Architecture: From the processing of an input image, through feature extraction and multi-level pooling, to the final pixel-level semantic segmentation output[60]. 20

4.1 Figure showing sample images from the 'Apricot' dataset. 25

4.2 Figure showing sample images from the 'TEyeD' dataset. Red, green, and blue colors represent sclera, iris, and pupil, respectively. 26

4.3 The architecture of the OCR-Net. [58] 28

4.4 Illustration of the Intersection over Union (IoU) Metric. This figure demonstrates the calculation of IoU, which is the ratio of the area of overlap to the area of the union between the predicted segmentation and the ground truth. 30

5.1 A visual comparison of eye tags as seen in IR (left) and visual (right) imaging, demonstrating their clear visibility and distinct pattern, which facilitate accurate registration in the ThermOcular system 36

6.1 Sample IR frames of the cornea with eyelashes effectively filtered out, depicted in the same black color as the background for clarity. 45

6.2 Diagram of the Eye Aspect Ratio (EAR) calculation. The EAR is computed using the distances between key landmarks: P2 to P6 and P3 to P5 for the vertical component, and P1 to P4 for the horizontal component, as indicated by the red lines. 46

7.1 An overview of the ThermOcular software's MVC architecture, illustrating the interconnectivity between user interfaces (Views), data processing units (Models), and the central coordinating components (Controllers) for each functional module. 49

7.2 Entity-Relationship Diagram illustrating the database schema for patient visits. 50

7.3 Figure showing the "Patient Management" tab Environment. 51

7.4 Figure showing the "Recording" tab Environment. 52

7.5 Figure showing the "Player" tab Environment. 53

7.6 Figure showing the "Processing" tab Environment. 54

7.7 Figure showing the "Analysis" tab Environment. 55

List of Tables

4.1 Benchmarking Results for Semantic Segmentation Models 31

5.1 Benchmarking results for point following methods with and without EyeTags in thermal and visual frames. 39

5.2 Calculated fiducial registration error (FRE) for each point following method and strategy. 40

Chapter 1

Introduction

The ThermOcular system [6] represents an advancement in ocular health diagnostics, particularly emphasizing the precise assessment of Ocular Surface Temperature (OST), a biomarker for identifying various ocular conditions. The mechanism of the ThermOcular system is rooted in the integration of infrared (IR) and visual imaging technologies. This system, mounted on a slit lamp bio-microscope, employs a dual-camera setup to capture synchronized video streams from both IR and visual spectrums (Figure 3.1). The initial phase involves the identification of control points within these captured frames, which are critical for the subsequent image registration algorithm (Figure 1.1). This initial step ensures the spatial alignment between the IR and visual datasets, a prerequisite for the precise temperature mapping across the ocular surface.

Following the registration phase, the ThermOcular system engages a segmentation algorithm, primarily focused on the cornea, to accurately locate the Region of Interest (ROI). The precision in segmentation is crucial, allowing for a targeted analysis of corneal temperature data, for comprehensive OST assessments. The system subsequently derives the

corneal temperature by overlaying IR and visual frames. Following this, the system employs an optical flow algorithm to identify control points in subsequent frames, allowing for the automated extraction of corneal temperature in the next images (Figure 3.2).

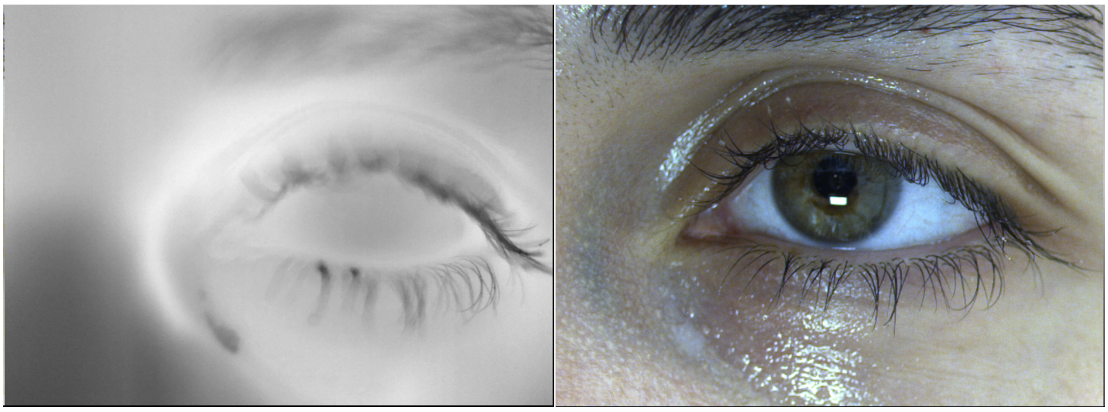


Figure 1.1: Sample images captured by the ThermOcular system. The left shows an IR thermography image highlighting the temperature gradients across the eye with the eyelashes appearing prominently due to their lower temperature. The right panel displays a high-resolution visual spectrum image, capturing detailed features such as iris, sclera, and pupil.

This thesis aims to enhance the precision and functionality of the ThermOcular system. The core objectives outlined in this thesis are as follows:

1. **Improvement of Segmentation:** Refinement of the segmentation process is necessary to accurately delineate not only the cornea but also the pupil, iris, and sclera. This enhancement directly influences the ability to generate precise temperature readings across various sections of the ocular surface. The goal is to surpass the current system's focus on corneal segmentation, thereby enabling a more comprehensive analysis of ocular surface temperature.

2. **Artifact Removal:** The presence of artifacts such as eyelashes and blinks introduces measurement errors, compromising diagnostic reliability. The objective is to develop robust methods to effectively eliminate these artifacts from thermal frames, ensuring that temperature readings accurately reflect the ocular surface condition.

3. **Registration Method Improvements:** The registration process is crucial for accurately mapping temperature readings across the ocular surface. This process will be refined to eliminate the need for extensive training on control point selection, enhancing the user experience and reducing potential errors. Additionally, the overall performance of the registration process will be improved to accurately track temperature changes, even with minor patient movements.

4. **Development of a Comprehensive Software System:** Designing and implementing a software system that manages patient records and facilitates the recording, processing, and analysis of videos captured from patients is a final objective. This system will integrate seamlessly with the hardware components to provide a user-friendly interface for clinicians, automating many processes that currently require manual intervention.

Chapter 2

Background

Ocular Surface Temperature

The measurement of Ocular Surface Temperature (OST) is a pivotal aspect in the field of ocular diagnostics, encompassing a range of invasive and non-invasive methods. Given the inherent sensitivity of the eye, non-invasive approaches, particularly infrared thermography, have gained precedence due to their contactless nature, offering a safe and effective means to assess the ocular surface's thermal state. The utility of thermography in ocular studies is highlighted by numerous investigations leveraging OST measurements for diagnosing various conditions. For example, Rosenstock et al. [46] employed OST measurements to identify inflammation in the lacrimal drainage system, illustrating the method's broad diagnostic applicability. Further, OST measurements have proven instrumental in research on dry eye disease [38], glaucoma [36], unilateral proptosis [7], ophthalmic post-herpetic neuralgia [8], vascular neuritis of the optic nerve [18], and retinoblastoma [26].

Beyond disease diagnosis, ocular thermography provides deep insights into the dynamics

of the tear film. This multi-layered fluid, essential for ocular health, undergoes continuous cycles of formation, destabilization, and reformation. The tear film's stability, crucial for preventing excessive evaporation, is primarily maintained by its lipid layer. Disturbances in this delicate balance can lead to increased evaporation rates [44, 51], a common precursor to dry eye syndrome [25]. In patients with dry eye, OST exhibits more rapid changes, with a significant magnitude of change and broader variations [38, 11, 39, 50, 41], detectable through infrared thermography [51, 12, 33]. The technology's ability to track these changes offers a unique perspective on tear film stability and the blink cycle's role in tear film reformation. Tear-Film Break-Up Time (TBUT) [14], the interval between a blink and the tear film's breakup, serves as a key indicator of tear film quality.

Despite its capabilities, the current methodology for OST measurement within the tear film breakup area is constrained, relying predominantly on manual selection and assignment. This underscores the need for advancements in OST measurement techniques to enhance diagnostic accuracy, reduce manual intervention, and improve overall efficiency in ocular health assessments.

OST Measurement Methods

The pursuit of Ocular Surface Temperature (OST) measurement has historically encompassed a range of methodologies, initially gravitating towards invasive techniques. Early efforts, as outlined by Tanj *et al.* [50], predominantly utilized thermistors either directly placed in probes or interposed between two contact lenses. These approaches, however, faced challenges in delivering consistent normal corneal temperature readings. The inconsistencies arose from a multitude of variables including the sensor type, probe placement on the ocular surface, specific measurement locales, the influence of administered anesthe-

sia, reliance on single-point data collection, probe temperature itself, and the prevailing ambient conditions.

The paradigm shifted with Mapstone's pioneering development of a non-contact OST measurement method [32], utilizing a bolometer to gauge the eye's surface radiation. This innovation marked a significant departure from contact-based methods, setting the stage for further technological advancements in the field.

The 1970s witnessed the emergence of a new generation of cameras, characterized by enhanced sensitivity and magnification capabilities. These advancements enabled the detailed capture of temperature data from the ocular surface. Initial iterations of these cameras depended on liquid nitrogen to cool the detectors, a requirement that underscored the technological limitations of the era. The subsequent advent of semiconductor Charge-Coupled Devices (CCDs) revolutionized detector design, significantly improving sensitivity and operational speed. Unlike their predecessors, these semiconductor-based detectors could function effectively at room temperature, obviating the need for external cooling agents.

The evolution of camera technology has established thermography's role as an invaluable non-invasive tool in OST measurement. Infrared thermography, in particular, has been the focal point of numerous research studies, exploring its potential across various applications. Within this domain, literature delineates multiple methods for OST assessment, underscored by the utilization of infrared thermography to analyze ocular surface images. Notably, the methodologies diverge into single and dual camera systems, each offering distinct approaches to OST estimation, thereby broadening the spectrum of non-invasive diagnostic tools available to researchers and clinicians.

Single Camera

Manual Methods

Initially, thermal cameras were slow to respond and could only capture a few frames in succession. The user could analyze each image simply by selecting a single point or multiple points in the area of interest. A majority of OST studies were conducted by comparing several points or areas of the corneal surface which were manually selected. In a study conducted by Efron et al.[15], 11 points were measured on the surface of the cornea in order to determine OST. A horizontal line was drawn across the geometric center of the cornea, and a distance of 0.5 millimeters was used to separate the points. After blinking, the cornea temperature was measured using this method.

According to another study[37], the authors used five 10x10 pixels boxes, each approximately $1mm^2$ in size, located on the horizontal axis of the cornea. A statistical analysis was conducted using the mean temperature of each area. According to their findings, dry eye patients have a higher OST than people with normal eyes of the same age. Furthermore, tears evaporate approximately three times faster in dry eyes than in normal eyes.

The study carried out by Galassi et al.[23] used five different points instead of five boxes. A horizontal line was drawn across the cornea, with one point located in the center, one on each of the internal and external canthi, and one half way between the internal canthus and the nasal limbus, and one halfway between the temporal limbus and external canthus. They conducted the study on subjects with primary open-angle glaucoma (POAG) and found that the OST for POAG patients was significantly lower than that for healthy eyes. The cornea's horizontal line was assigned five distinct points by Sodi et al.[47]. Each point was placed at an equal distance from the others, so that one of the points was located in the estimated corneal center. Temperatures were measured using IR thermography in

CRVO patients and compared to those in a control group.

Murphy et al.[40] used five boxes placed in the middle of the cornea on the horizontal line. Based on their findings, it was determined that corneal nerve stimulation was primarily caused by the rate of OST change during the use of the non-contact corneal aesthesiometer (NCCA) air pulse.

A box of 20x20 pixels or $3.3mm^2$ located at the center of the cornea was used by Mori et al.[39] to measure the OST. A considerable difference was found between dry eye patients and normal subjects with regard to the rate of decrease in OST after blinking.

An encircled area was placed in the center of the cornea by Chiang et al.[11]. This data was used to diagnose dry eye disease. Similarly, Ng et al.[41] placed a circular area at the center of the cornea. A relationship was found between OST and age based on the data collected.

Using 23 points across the cornea, Purslow et al.[44] divided the corneal surface into five regions: central region, higher and lower regions, nasal region, and temporal region. To measure the effect of wearing different types of contact lenses on the body's temperature, thermography was used.

Using thermography, Matteoli et al.[35] examined the temperature changes in the eyes of individuals who had age-related macular degeneration (AMD). To account for tear evaporation, they took only the first frame of the opened eye for analysis in order to establish a baseline. Initially, the images were acquired in RGB and converted to greyscale in order to be processed in an image processing algorithm. They then used MATLAB software to analyze selected frames of a recorded movie in five areas of the eye's surface. Their study found that the OST of AMD patients was significantly lower than that of a control group.

An investigation of the repeatability of thermal measurements in dry eye patients was conducted by Tan et al.[52]. The area of interest (ROI) was defined using the diamond method.

Data acquisition was carried out at 0, 5, and 10 seconds. The coefficients of repeatability (COR) of infrared ocular thermography were calculated using inter-image, inter-occasion, and inter-examiner comparisons. A majority of the coefficients used in their experiment were highly reproducible, according to the study.

In addition, Sudarshan et al.[49] studied dry eye disease using thermal imaging and manually extracted the OST ROI from the original image. For feature extraction, they used discrete wavelet transforms and Gabor transforms. As a result of their analysis of the features, they determined that the OST changes in the lower part of the cornea were more accurate for the diagnosis of dry eye disease.

Semi-Automatic Methods

Despite their simplicity, manual methods have shortcomings that result either from the limitations of the camera technology itself – low thermal and temporal accuracy, varying degrees of invasiveness – or from the limitations of data analysis – they do not operate automatically and require considerable post-processing. Researchers have improved the thermal camera capabilities and quality in order to improve their results from measurements taken from the eye surface. Due to this, a number of image processing algorithms have been implemented on thermal images to reduce the amount of manual input by the user.

A semi-automated method for acquiring OST was proposed by Acharya et al.[41]. In order to process the images, the original image of the eye was converted to grayscale. After manually cropping the image, the ROI of the image was located and the image was resized to 400x200 pixels. By assuming that the radius of the cornea represents one-fourth of the entire anterior surface of the eye, their algorithm detected the circular center of the

cornea. Following that, a profile line was drawn in the middle of the cornea for the purpose of measuring the temperature. OST and temperature distribution decreased with age, as was the main finding of the study.

A further study by Acharya et al.[2] used thermography and feature extraction to determine the effectiveness of treatment in dry eye patients. After a blink, thermographs (thermal images) were captured. Afterwards, the ocular region was segmented from the image. Using feature extraction, some texture information was extracted from the thermal images. Their analysis of the extracted information enabled them to determine whether a patient had responded to treatment.

An image processing technique was proposed in [34] for analyzing thermal images. An ideal eye would have been composed of a fixed shape, bound by two arches, symmetrical both horizontally and vertically, and narrower in the area representing both canthi. This fixed eye shape was then adapted by reversible geometric transformations to the thermal images of the eyes. Using their method, they were able to compare OST between left and right eyes, or even between different groups of people, without having to take into account ocular geometry.

Fully-Automatic Methods

In spite of advances in both hardware and software processing, semi-automatic methods still required user input. Some groups developed a fully automated method in response. Locating the corneal boundary in an ocular thermogram is the key technical challenge here. The researchers in [50] developed an automated method for measuring OST. Their method, which involved the use of a snake algorithm and a target tracing function, allowed them to localize the eye on the image. In addition, they were able to remove the effects of eyelashes.

In order to determine corneal temperatures, a circle was drawn in the center of the eye. In spite of this, the algorithm was not able to completely eliminate the influence of the eyelash, and the segmentation accuracy was not satisfactory.

An active contour model was used by Shuang et al.[62] to detect the eye from the rest of the image. The algorithm could handle low resolution and blurry images. A comparison was made between the algorithm and the Tan and Li algorithm. Although it took more time to process the data, the accuracy of the method was improved over the other methods despite the longer processing time.

Dual Camera Systems

Thus far, manual, semi-automatic, and fully automated methods have been described based on a single camera system. Due to their manual nature or low accuracy in locating the cornea ROI, they are not suitable for use in OST measurements. The algorithms were all designed to work only with thermal images. Nevertheless, since thermal images are generated based on temperature distribution, there is no clear distinction between the cornea and the surrounding region. As a result, it is not possible to segment the corneal area from the surface of the eye in a thermal image. There have been attempts to improve segmentation accuracy by using a dual camera system in order to solve this problem. As a result, they were able to identify the corneal boundary in the thermal image by using the visible image.

The researchers in [28] employed a semi-automatic technique to measure OST in dry eye patients using a visible (radiation) camera embedded in a thermal camera system. The pupil of the subject's eye was initially located in the center of the visible camera image by

a computer system during the pre-processing stage. A thermal camera was also aligned on the pupil of the subject's eye through camera alignment. They used two mirrors in their system, one fixed and one moving, to reflect visible light from the subject's eye into the visible camera detector in order to provide semi-simultaneous viewing of both images. In order to reflect visible light to the visible camera, the moving mirror rotated 45 degrees every 0.25 seconds. Infrared radiation was prevented from entering the thermal camera during the transmission of visible light to the camera by the rotating mirror. For further processing of the data, customised software was used after the data had been recorded. After manually extracting the ROI from the images, the extracted data was analyzed. Visual comparisons of the coincident images were conducted to determine tear film break-up times in dry eye patients.

Introducing this method was a big step forward in 2011. Compared to the past, an experienced user could identify the corneal boundary with significantly greater accuracy, and the location of tear-film break-up on the surface of the eye could be estimated more precisely by identifying the corneal boundary. In spite of this, their method was not synchronous and still required considerable user input in order to extract the data. In another study, Su et al.[48], analysed tear film break-up patterns using a dual camera system, incorporating a thermal and visual camera. A Germanium beam splitter was employed in their system to facilitate simultaneous imaging of thermal and visible cameras. By means of the beam-splitter, IR light was transmitted and visible light was reflected. The thermal camera and beam splitter were placed in front of the patient's eye, while the visible camera was placed beside it. In terms of synchronization, the new system was an improvement over the previous one. Nevertheless, user input was required to analyze tear-film patterns. An experiment conducted by Li et al.[30] in 2015 used a different camera installation to record simultaneous video files from the surface of the eye. A thermal camera and a visible

camera were mounted in front of the cornea approximately -15° and +15° off-axis from the geometric center of the cornea. In the next step, two movies of the eye were recorded, one from each camera, for a specific period of time. Afterward, the videos were synchronized using FINAL CUT PRO X software (Apple Inc.). An experienced user viewed both the thermal and visible videos concurrently to estimate the tear film break-up area.

In 2017, Kricancic et al.[29] described another method for installing cameras on the surface of the eye in order to monitor the change in temperature. A germanium filter was placed in a fixed position in front of the patient's eye. Contact lens wearers were monitored for changes in corneal temperature using the system. The user, however, was still required to enter information manually.

Eventually, Zare Bidaki et al[6] in 2021 proposed a novel semi-automatic dual camera system to measure OST. Their method involves the use of a visual camera to detect and segment the cornea and a thermographic camera to take an image of the eye. In the next step of their algorithm, an operator selects at least three control points in both thermal and visual frames. Then, an image registration algorithm is applied to find the cornea segment of the eye in the thermal image. They also employed an optical flow algorithm to follow the operator's chosen control points in the sequence of the images. Only the first frames of the thermal and visual cameras require the operator to select the control points.

Chapter 3

ThermOcular

The ThermOcular system represents a significant advancement in the field of ocular diagnostics, employing an innovative approach to measure Ocular Surface Temperature (OST), a biomarker in assessing ocular health. Traditionally, the system integrates infrared (IR) thermography with visual imaging technologies to provide a comprehensive analysis of the eye's surface temperature. Mounted on a slit lamp biomicroscope 3.1, the ThermOcular device captures synchronized video streams from both IR and visual spectrums. This dual-camera setup facilitates the precise location and tracking of temperature variations across different regions of interest, including Pupil, Iris, and Sclera, laying the groundwork for accurate and non-invasive tracking of the ocular surface temperature.



Figure 3.1: Figure showing ThermOcular dual IR-Visual camera system mounted on slit lamp.

The frames are captured synchronized from both IR and visual cameras. The synchronization process has been done using hardware-triggering. When the IR camera captures a frame, it triggers an opto-input pin of the visual camera to capture a frame. The IR

camera captures 25 frames per second (fps) which means that the whole system works with 25 fps.

Following the initial setup and the capture of synchronized video streams, as shown in the Figure 3.2, the ThermOcular system's workflow is meticulously organized into a series of critical steps to ensure accurate OST measurement. The process begins with the selection of control points within both IR and visual frames, a fundamental step for the subsequent image registration algorithm. The system uses the control points to align both IR and visual frames horziontally and vertically. This crucial phase guarantees the spatial congruence between the visual and infrared data, enabling precise temperature mapping across the ocular surface.

The registration process in the ThermOcular system utilizes affine transformation, a method that is prevalent due to its ability to preserve the collinearity and ratios of distances between points. This means that points initially lying on a line still lie on a line after transformation, and all parallel lines remain parallel, although the angles between lines are not necessarily preserved. The affine transformation is composed of translation, rotation, scaling, and shearing components, where shearing can either produce or correct distortion in the image by moving different parts of the image in various directions.

An affine transformation requires at least three pairs of corresponding points between the images to be calculated accurately. The importance of using this method in the ThermOcular system lies in its capability to adjust the thermal image to align with the visible image, despite the differences in image acquisition parameters, such as lens characteristics and viewpoint angles. By applying the affine transformation, the system ensures that the thermal and visible images are properly overlaid, which is critical for accurate temperature mapping.

The calculation of the affine transformation is based on a matrix operation that applies scaling, rotation, and translation to the points in the thermal image to match them with the corresponding points in the visible image. This mathematical operation is essential for aligning the two sets of images, allowing for the precise mapping of temperature data from the thermal image to the corresponding anatomical features in the visible image.

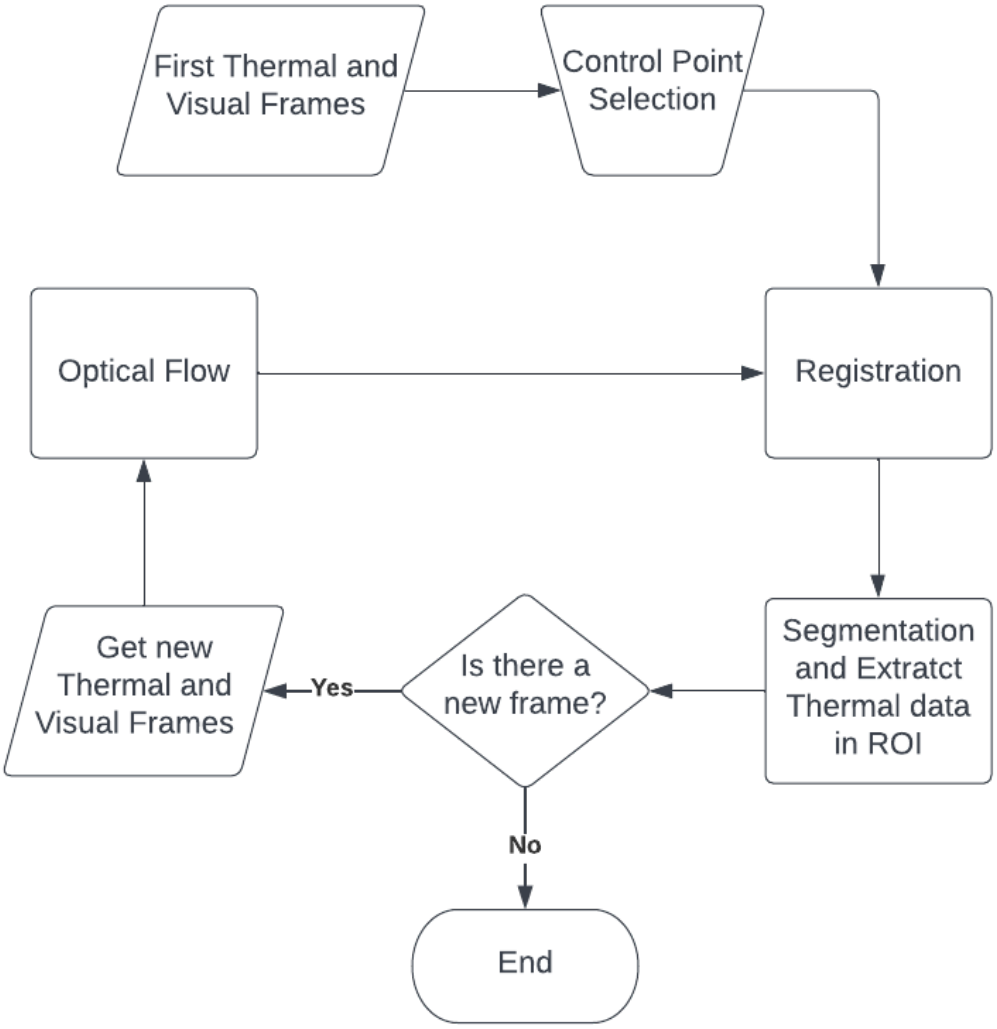


Figure 3.2: Schematic representation of the ThermOcular system's operational workflow.

Post-registration, the system invokes a semantic segmentation algorithm focused on the cornea to delineate the Region of Interest (ROI). This segmentation is paramount, as it isolates the cornea for targeted temperature analysis, thereby enhancing the accuracy of

OST assessment. The Pyramid Scene Parsing Network (PSPNet) [60] (Figure 3.3) was selected due to its proficiency in learning global features at different scales, making it suited for the segmentation stage of the project. The chosen model, Resnet50-PSPNet, was pre-trained on the ImageNet dataset [13], leveraging the extensive variety of images available in ImageNet to gain a broad understanding of visual features before being fine-tuned specifically for the task of cornea segmentation. Despite a slightly longer processing time compared to other networks, PSPNet stood out for its superior accuracy, achieving a mean Intersection over Union (IoU) of 94.6% on the project dataset, surpassing other models like FCN32[31], SegNet[4], and U-Net[45] in comparative analysis.

The project dataset consisted of 160 images captured from the eyes of 10 subjects using the Teledyne FLIR BFS visible camera, under varied gaze directions and eyelid positions to ensure a comprehensive representation of different ocular conditions. A total of 250 images were initially taken for each subject, from which 10 images per subject were meticulously chosen for network training and testing. This dataset was captured in a controlled environment with consistent illumination and room temperature, and subjects were instructed to move their eyes and blink naturally. For the training phase, 120 images were utilized, while 40 images were reserved for validating the network's performance, ensuring that the PSPNet model was both robust and reliable for accurate segmentation in the ThermOcular system.

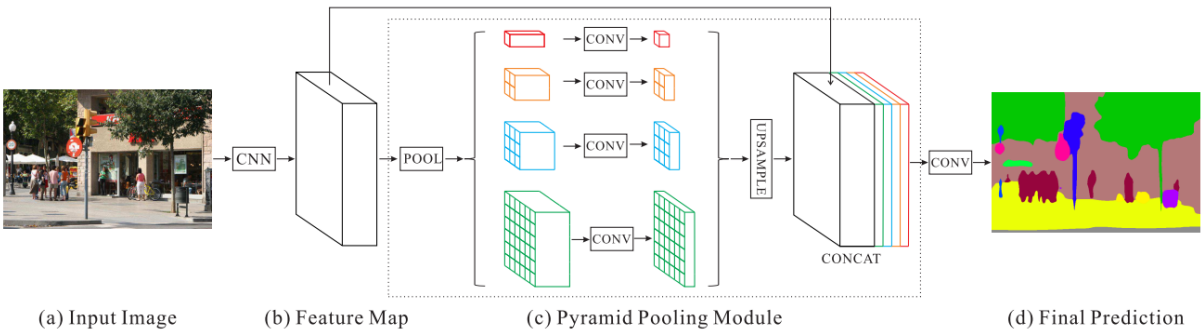


Figure 3.3: PSPNet Architecture: From the processing of an input image, through feature extraction and multi-level pooling, to the final pixel-level semantic segmentation output[60].

.

Following segmentation, an optical flow algorithm facilitates the automatic identification of control points in subsequent frames. This automation is key to the system's efficiency, allowing for the continuous extraction of temperature data without manual intervention.

The optical flow algorithm is a critical component in the ThermOcular system, particularly for tracking the movement of the eye between frames. When selecting control points for image registration, distinctive edges such as eyelashes and eyebrows are integral because they provide clear features that can be easily identified across IR and visual frames. However, for an effective optical flow, which estimates the motion of these control points, the algorithm requires a rich texture around these points to track their movement accurately.

Edges alone, like those provided by eyelashes and eyebrows, often lack the necessary textural information for reliable optical flow calculations. They present a high-contrast

boundary, but do not offer the varied intensity patterns that help in estimating motion over time. Optical flow algorithms depend on the assumption that the patterns of pixel intensities are consistent across consecutive frames, and this is where highly textured regions excel compared to edges.

This disparity between the requirements for control point selection and those for optical flow presents a challenge. While edges are excellent for the static identification of features, the dynamic tracking of these features through optical flow requires more texture than edges typically provide. This can lead to inaccuracies in motion estimation, making the process of maintaining robust and precise tracking throughout the video sequence quite tricky.

Chapter 4

Segmentation

Image Segmentation

Image segmentation is a process within the field of computer vision that involves the segmentation of a digital image into numerous distinct sections or regions. These segments are formed by grouping pixels together based on shared characteristics, allowing for a detailed partitioning of the image. This procedure is designed to simplify the complexity of an image by breaking it down into its constituent parts, thereby transforming the image into a format that is significantly easier to analyze and interpret.

The two primary methodologies within image segmentation, semantic and instance segmentation, offer unique approaches to how pixels are classified and grouped. Semantic segmentation focuses on the classification of each pixel into predefined categories, enabling the identification of various objects or features within an image without distinguishing between different occurrences of the same object. Conversely, instance segmentation takes this process a step further by not only classifying pixels into categories but also iden-

tifying individual entities of objects within those categories, thereby providing a more detailed analysis. This level of detailed segmentation is indispensable in fields such as medical imaging, where it enables the precise identification, analysis, and differentiation of anatomical structures and pathological conditions. By accurately segmenting medical images into relevant regions, such as differentiating between healthy tissue and tumors, or identifying specific organs, medical professionals can gain insights crucial for diagnosis, treatment planning, and patient care.

In the application of the ThermOcular system, which focuses specifically on the eye and its components, the segmentation task primarily involves semantic segmentation. This approach is utilized to meticulously classify each pixel of the eye into relevant categories,including the cornea, iris, and pupil, thereby facilitating a detailed analysis of the ocular surface. The segmentation is essential for accurately dividing the eye into its components to be able to extract and track the temperature within the regions of interest.

Dataset Preparation

For a comprehensive understanding of the dataset preparation for the segmentation model used in the ThermOcular system, it is crucial to focus deeper into the methodologies employed in collecting and preparing the data. The initial step in this process was the execution of a study known as 'Apricot', conducted at the Murphy Laboratory of Experimental Optometry (MLEO) at the School of Optometry, University of Waterloo. This study aimed to capture a wide array of ocular conditions through machine vision videos, which were recorded for a duration of 15 seconds each for 20 different subjects. These subjects were instructed to blink naturally, thereby ensuring the capture of a diverse range of eye states including fully open, half-open, and fully closed eyes, among others.

From these video recordings, a careful selection process has been used, wherein 180 images were meticulously chosen. This selection was not random; instead, it was strategically made to encompass a broad spectrum of conditions . This approach was fundamental in ensuring that the dataset accurately represented the variability found in real-world ocular states. Subsequently, these images were annotated with the Supervisely toolbox (Supervisely OU, Tallinn, Estonia), a robust platform designed specifically for different applications in image annotations, including semantic segmentation.

Recognizing the necessity of a robust and diverse dataset for the training of effective machine learning models, additional steps were taken to augment the dataset. This included the incorporation of 2,000 images from the TEyeD[20] dataset (Figure 4.2), a publicly available data set known for its extensive collection of eye images annotated with ocular segments. The integration of the TEyeD dataset significantly broadened the scope and diversity of the training data.

However, a notable imbalance existed between the newly annotated 'Apricot' images and the pre-existing images from the TEyeD dataset. To address this imbalance and ensure a comprehensive representation of various ocular conditions, the 'Apricot' images were replicated tenfold. This replication was a critical step in maintaining a balanced dataset, which is vital for the effective training of machine learning models and preventing the model to be over-fitted.

Beyond mere replication, a series of data augmentation techniques were employed to further enhance the dataset's diversity. These techniques included rotation, flipping, and cropping of images, among others. The purpose of these augmentations was twofold: first, to simulate a wider range of ocular conditions and variations than what was originally captured, and second, to ensure the robustness of the segmentation models against variations in real-world ocular images. By artificially creating variations in the dataset, the mod-

els could learn to generalize better, making them more effective in accurately segmenting ocular components across a wide array of eye images.



Figure 4.1: Figure showing sample images from the 'Apricot' dataset.

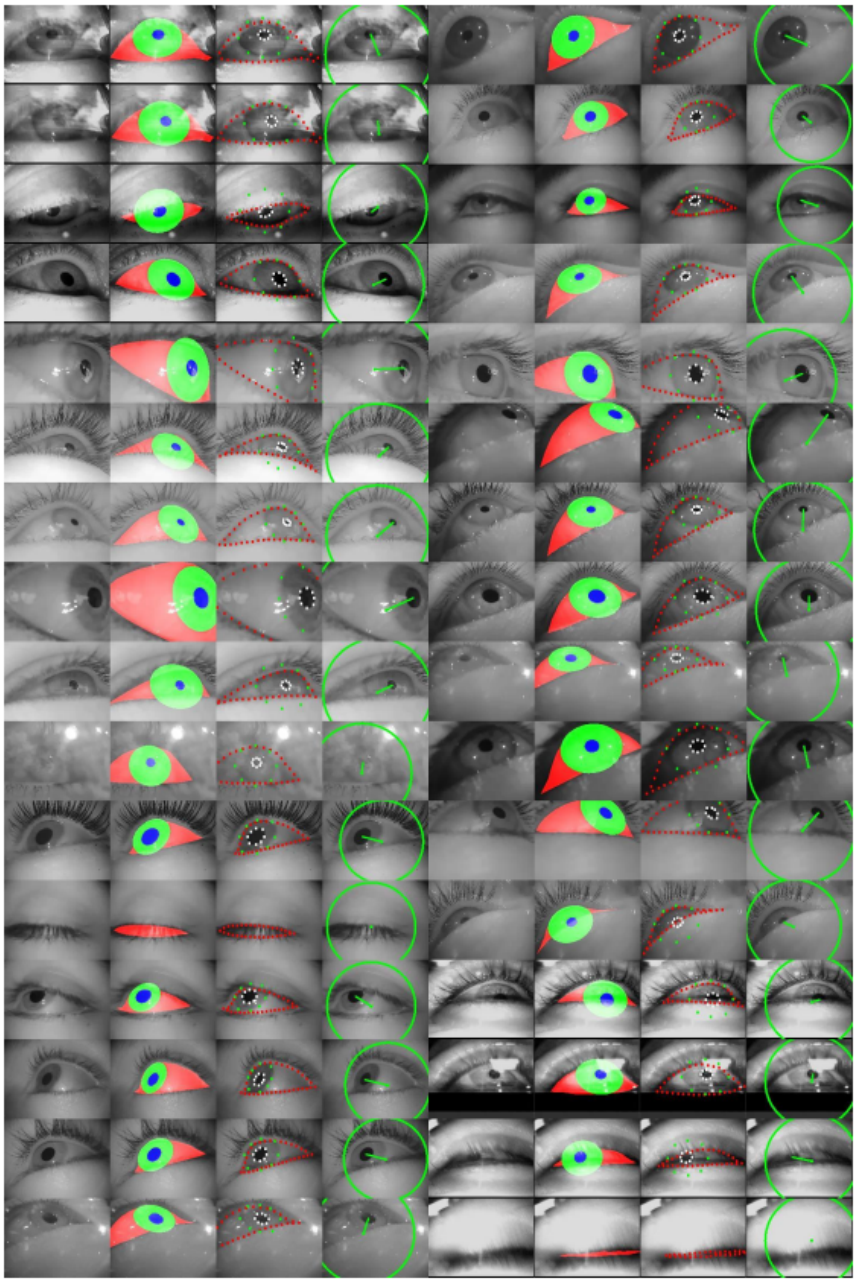


Figure 4.2: Figure showing sample images from the 'TEyeD' dataset. Red, green, and blue colors represent sclera, iris, and pupil, respectively.

Model Selection

In order to find the most accurate and efficient segmentation model for the ThermOcular system, an array of state-of-the-art semantic segmentation models was evaluated. This evaluation encompassed a diverse set of models renowned for their performance in semantic segmentation tasks, particularly in the domain of medical image segmentation. The selected models for comparison included PSPNet, Deeplabv3[10], PP-LiteSeg[43], Unet++[63], SegFormer[55], and OCRNet[58]. Each of these models has demonstrated exceptional performance and effectiveness across various medical image segmentation challenges, making them prime candidates for this comparative analysis.

OCR-Net

The OCR (Object-Contextual Representations) Net introduces a novel approach for semantic segmentation, focusing on enhancing pixel representation through context aggregation at the object level. This method is grounded in the understanding that the label of a pixel is determined by the object class it belongs to. The OCR Net framework operates in three primary stages (Also represented in Figure 4.3):

1. **Learning Object Regions:** Initially, the model learns to identify soft object regions within an image, guided by ground-truth segmentation. These regions correlate with the different classes present in the image, effectively creating a coarse segmentation that serves as the foundation for further processing.

2. **Object Region Representation:** Once the object regions are identified, the model aggregates the pixel representations within each region to form a comprehensive representation of that region. This step is crucial for capturing the essence of each

object class within the image, providing a rich contextual background against which individual pixel classifications are made.

3. **Pixel-Object Relation and Augmentation:** The final step involves computing the relation between each pixel and the identified object regions, followed by augmenting the pixel's representation with this contextual information. This process employs a weighted aggregation of all object region representations, with the weights derived from the pixel-object relations. The augmentation with object-contextual representations significantly improves the model's ability to accurately segment complex images by providing additional contextual cues that inform the classification of each pixel.

The OCR Net's methodology differs significantly from traditional multi-scale context schemes by explicitly differentiating between pixels belonging to the same object class and those belonging to different classes. This distinction allows for more nuanced context aggregation, leading to improved segmentation accuracy. Moreover, the model's efficiency in processing and its competitive performance across various benchmarks highlight its potential as a robust solution for semantic segmentation tasks.

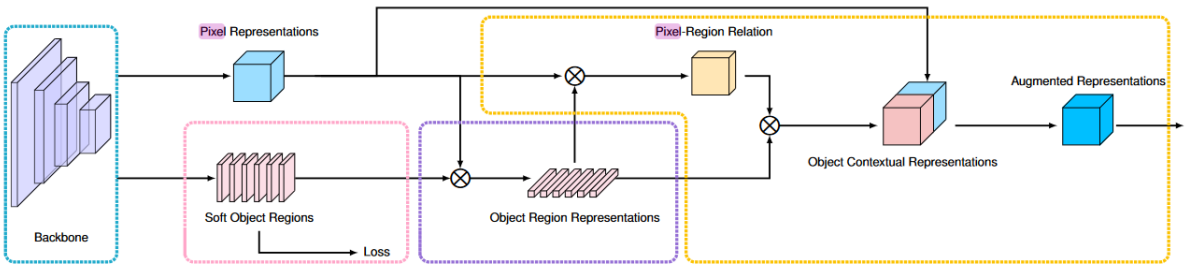


Figure 4.3: The architecture of the OCR-Net. [58]

Training Details

In the training process, a batch size of 5 images was chosen for a good balance between computational efficiency and the model's ability to generalize. The model underwent 15,000 iterations, ensuring thorough training. A learning rate of 0.01 was selected to balance the speed and stability of learning. Stochastic Gradient Descent (SGD) optimizer was used alongside the cross-entropy loss function. The HRNet-w18 architecture served as the backbone, pre-trained on the VOC12[19] dataset and fine-tuned on a proprietary dataset, from 'Apricot' study, enhancing the model's accuracy in segmentation.

Benchmarking

In the benchmarking process, the mean Intersection Over Union (mIOU) metric (Figure 4.4) was utilized to evaluate the segmentation effectiveness. This metric offers a comprehensive measure of a model's precision in distinguishing between different ocular regions, essentially indicating how well the predicted segmentation aligns with the ground truth. mIOU is crucial for assessing the model's capability to accurately segment ocular components.

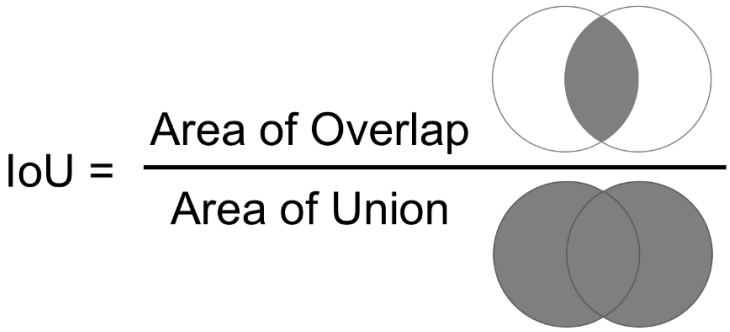


Figure 4.4: Illustration of the Intersection over Union (IoU) Metric. This figure demonstrates the calculation of IoU, which is the ratio of the area of overlap to the area of the union between the predicted segmentation and the ground truth.

The choice of mIOU as the primary evaluation metric is due to its ability to provide a balanced assessment across various classes within an image. It calculates the average overlap between the predicted and actual labels for each class, thus giving insight into the model's performance across the board. This is particularly important in medical imaging, where accurate delineation of each component can be vital for diagnosis and treatment planning.

To establish a fair and meaningful comparison between the newly proposed segmentation model and the original model employed by the ThermOcular system, a focused analysis on specific classes was essential. Given the ThermOcular system's primary function of accurately segmenting the cornea, the comparison concentrated on the mIOU scores for the cornea (represented by the pupil and iris in the segmentation task) and the background, which includes all other parts of the image, including the sclera. This approach ensures a direct and equitable evaluation of each model's ability to identify and isolate the cornea from the rest of the ocular surface.

Table 5.1 summarizes the performance of each semantic segmentation model explored in the study, highlighting their mIOU scores for both overall segmentation and specifically for the cornea and background.

Table 4.1: Benchmarking Results for Semantic Segmentation Models

Model-Backbone	**Pre-train DS**	**MIOU(total) (%)**	**MIOU Cornea&Background**
PSPNet-ResNet50	ImageNet	91.68	94.6
Deeplabv3-ResNet50	VOC12	95.28	97.11
PP-LiteSeg-STDC2	Cityscapes	92.13	93.54
Unet	Cityscapes	90.21	93.91
Unet+++	Cityscapes	96.34	96.88
SegFormer	Cityscapes	96.34	97.36
OCRNet-HRNet-w18	VOC12	**97.61**	**98.24**

Chapter 5

Registration

In the realm of ocular temperature measurement, precise assessment and tracking of the ocular surface temperature (OST) in different components of the eye are paramount for diagnosing and monitoring various eye conditions. The ThermOcular system represents a significant advancement towards achieving unparalleled accuracy in OST evaluations. The performance of this system's efficacy lies in its ability to accurately register thermal images with their corresponding visual counterparts. This process, known as registration, is critical for aligning images from different modalities, ensuring that temperature measurements are accurately mapped onto the correct anatomical locations of the eye.

Registration, in the context of the ThermOcular system, involves the precise alignment of infrared (IR) and visual images captured concurrently. This step facilitates the exact overlay of temperature data onto the visual representation of the ocular surface, allowing for extracting the temperature of the region of interest. The complexity of ocular anatomy, combined with the dynamic nature of eye movements, requires a robust registration process to maintain the accuracy of temperature measurements across consecutive frames.

Historically, the registration process within the ThermOcular system relied heavily on manual interventions, where clinicians would identify control points within the imagery to assist in the alignment. While effective, this method introduced potential for human error and inconsistencies, underscoring the need for improvements in the system's registration capabilities.

In the process of image registration within the ThermOcular system, the identification of control points and the application of optical flow are pivotal for ensuring precision and accuracy. Control points, essential for aligning IR and visual images, are predominantly determined by identifying edges within the images. Edges, which delineate boundaries between different regions of an image, serve as critical markers for registration, providing the spatial references needed for accurate overlay of temperature data onto the visual representation of the ocular surface. However, while edges are instrumental for locating control points, they present limitations when applied to optical flow algorithms.

Optical flow, a method used to estimate motion between two image frames, relies heavily on texture within the images to track movement across sequences. Unlike edges, which mark the transition between different visual features, texture offers a rich pattern of variations in intensity, which is crucial for optical flow algorithms to accurately track movement. Unfortunately, areas with prominent edges, but minimal texture can be challenging for optical flow, as the lack of detailed patterns makes it difficult to ascertain motion accurately. This discrepancy between the need for edges to identify control points and the preference for texture in optical flow underscores a fundamental challenge: optical flow can introduce additional error in the registration process. The reliance on texture for motion estimation means that in images where edges dominate, but texture is sparse, optical flow algorithms are at a disadvantage, potentially compromising the registration accuracy and, by extension, the performance of the ThermOcular system's temperature mapping.

In response to these challenges, this thesis aims to reduce dependency on manual input and elevate the system's efficiency and diagnostic accuracy. By comparing the current optical flow-based, static control-points-based, and EyeTag-based methodologies, the registration process has been improved.

In subsequent sections, the focus shifts to the specifics of the registration improvements implemented in the ThermOcular system. This discussion will cover the introduction of static control point retention and the innovative use of EyeTags.

Static Control Point Retention

The traditional approach to image registration within the system relied on dynamic updates of control points across sequences. This method, while effective in accommodating minor patient movements, often introduced errors due to the optical flow's dependence on texture variations within the images—a feature not always consistently present, especially in regions delineated by sharp edges without significant texture.

Static control point retention is a strategic response to these challenges. By maintaining a consistent set of control points throughout the imaging sequence, the system uses the initial frame's accurately identified markers across subsequent frames without requiring continuous recalibration. This approach not only streamlines the registration process, but also significantly reduces the potential for error introduced by optical flow algorithms, which can struggle with the precise tracking of movements in texture-sparse regions.

The efficacy of static control point retention lies in its simplicity and precision. Initial control points, selected for their clear visibility and definable edges, serve as stable references that anchor the registration process. This method assumes that the ocular surface

and the patient's head position remain relatively stable throughout the imaging period, a reasonable assumption for the controlled environment in which ThermOcular assessments are conducted.

Importantly, the assumption of minimal patient movement is practically maintained by the use of the slit lamp's chin rest. Patients undergoing assessment with the ThermOcular system place their heads on the chin rest, providing a stable base that significantly minimizes head movements. This stability is critical, as static control point retention relies on the premise that the patient's head will not shift significantly during the imaging process. The chin rest helps keep movements very small, so they stay within what the system can handle without affecting how accurately it maps temperature.

This technique addresses the limitations associated with optical flow by avoiding the need for detailed texture analysis in every frame. Instead, by relying on the robustness of initial control points, the system ensures a consistent basis for image alignment, enhancing the accuracy of temperature mapping across the ocular surface.

Static control point retention thus represents an enhancement in the ThermOcular system's registration methodology. It not only improves the system's operational efficiency by reducing the steps required for accurate registration but also increases the accuracy of ocular surface temperature assessments.

Integration of EyeTags

An advancement of the ThermOcular system is introducing innovative component "EyeTags" to enhance the registration accuracy and streamline the registration process. EyeTags are specially designed markers that are easily identifiable in both IR and visual

images.These markers are made of a stick-on aluminium material, which stands out in both visual and IR images due to the contrast in temperature between the aluminium and the skin. These markers serve as consistent, identifiable points across images, assisting in the precise alignment of IR data with the corresponding visual representation of the ocular surface (Figure 5.1).

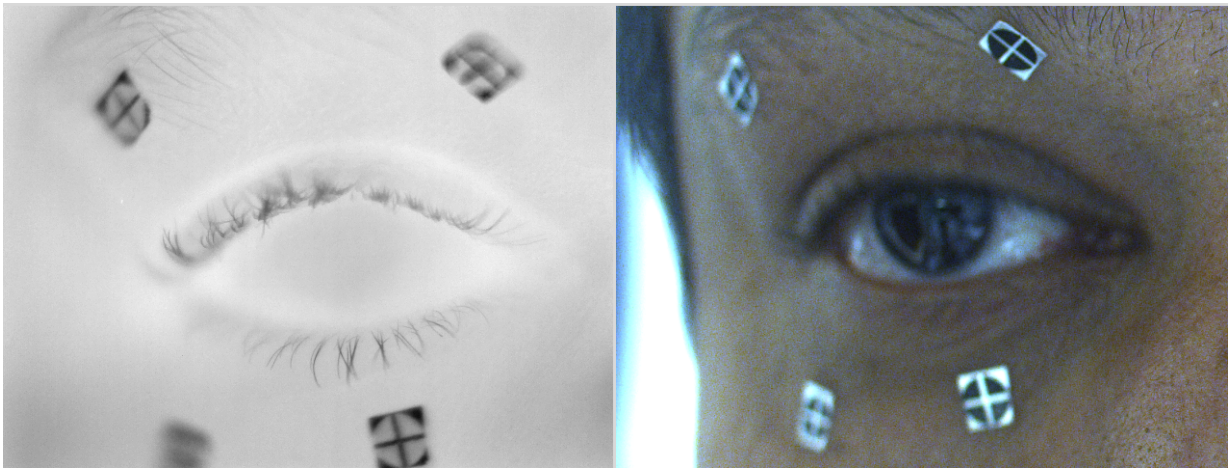


Figure 5.1: A visual comparison of eye tags as seen in IR (left) and visual (right) imaging, demonstrating their clear visibility and distinct pattern, which facilitate accurate registration in the ThermOcular system

EyeTags greatly aid clinicians in locating control points. Without reference points, identifying control points can be challenging due to the difference between IR and visual frames. EyeTags provide clear markers, eliminating the guesswork and simplifying the selection process. This convenience means clinicians no longer require training for control point selection, as the EyeTags serve as intuitive guides.

EyeTags are engineered to be distinctly recognizable within the imagery captured by the ThermOcular system. This is particularly beneficial in situations where the texture of

the ocular surface might not provide sufficient detail for accurate optical flow calculations.

The addition of EyeTags significantly improves the performance of the ThermOcular system by:

- **Enhancing Precision:** With EyeTags as fixed reference points, the system can achieve more precise temperature mapping across the ocular surface, improving the registration and, hence, the whole system accuracy.
- **Reducing Manual Intervention:** The need for clinicians to manually identify and select control points is greatly diminished, making the process faster and less prone to human error.

Benchmarking

Point Tracking Evaluation

In order to compare all the registration strategies, a video encompassing both thermal and visual frames was captured, containing over 500 frames. A set of 8 control points were selected on every other frame, selected twice: directly on the eyelashes or eyebrows, and subsequently on the EyeTags. This established four experimental scenarios to assess the efficacy of static versus optical flow methods of control point retention and the utility of EyeTags:

- **Scenario 1: Points on Eye with Static Control Point Retention**
  - Control points were manually selected directly on the eye.

  - The static method was used, assuming minimal eye movement and retaining the initial control points for alignment.

- **Scenario 2: Points on Eye with Optical Flow Control Point Following**
  - Control points were manually selected directly on the eye.
  - The optical flow method dynamically tracked the movement of the eye to adjust control points across frames.

- **Scenario 3: Points on EyeTags with Static Control Point Retention**
  - Control points were manually selected on the EyeTags surrounding the eye.
  - The static method was used with EyeTags serving as fixed reference points.

- **Scenario 4: Points on EyeTags with Optical Flow Control Point Following**
  - Control points were manually selected on the EyeTags surrounding the eye.
  - The optical flow method was employed, allowing for the dynamic tracking of EyeTags to maintain registration accuracy.

To calculate the registration error, the Euclidean distance was calculated between the established control points in both thermal and visual frames, and the ground truth control point selected manually. By averaging these distances, the overall RMSE was computed for each strategy over the full video sequence as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2} \tag{5.1}$$

where P_i denotes the predicted control point location in either thermal or visual frame, O_i is the observed location in the ground truth frame, and n represents the total number of control points. This calculation was iterated across all corresponding pairs of thermal and visual frames within the dataset, and the average RMSE is calculated for each thermal and visual videos.

Strategy	**Point Following Method**	**Frame Type**	**RMSE (in Pixels)**
Without EyeTags	Static	Thermal	10.36
		Visual	9.27
	Optical Flow	Thermal	13.09
		Visual	11.81
With EyeTags	Static	Thermal	9.54
		Visual	9.33
	Optical Flow	Thermal	6.28
		Visual	4.39

Table 5.1: Benchmarking results for point following methods with and without EyeTags in thermal and visual frames.

Fiducial Registration Error Calculation

To assess the precision of the registration process in the ThermOcular system, the fiducial registration error (FRE) was calculated. FRE is a common metric used to quantify the misplacement of control points during registration, representing the overall misalignment of points. The Euclidean distance between the reference points in the fixed frame (visible) and the corresponding points in the transformed image (thermal) was determined. The

FRE was computed as the average of these Euclidean distances.

In practice, 8 control points were identified in both fixed and moving frames. A transformation matrix was then calculated using three pairs of corresponding points. Subsequently, the remaining points were transformed using this matrix, and the distances between the transformed points and their original counterparts were computed. The mean of these distances for each frame was taken as the FRE for that frame.

This process was systematically applied to all frames in the sequence, and the resulting FREs were averaged to yield the mean FRE for each registration scenario. The results are presented in the following table:

Strategy	**Point Following Method**	**Mean FRE ± SD (in Pixels)**
Without EyeTags	Static	11.04
	Optical Flow	13.36
With EyeTags	Static	9.84
	Optical Flow	5.63

Table 5.2: Calculated fiducial registration error (FRE) for each point following method and strategy.

Discussion

The benchmarking results depicted in Tables 5.1 and 5.2 offer insightful revelations about the performance of the ThermOcular system's registration strategies. The analysis of point tracking results indicates a trend where the inclusion of EyeTags consistently improved point tracking accuracy. Scenarios utilizing EyeTags demonstrated lower RMSE values in

comparison to their counterparts without EyeTags, suggesting that the additional reference points provided by EyeTags enhance the system's ability to maintain alignment despite eye movements.

When comparing the static and optical flow methods, the static approach yielded lower RMSE values for scenarios without EyeTags. This outcome could be attributed to the controlled setting assumption that dictates minimal eye movement during imaging, which generally holds true when the subject's head movements are restrained. Additionally, the improved performance of the static method validates the assumption regarding errors introduced by the optical flow method. Since control points are typically selected on prominent edges of the eye for registration purposes, these points may not be ideal for optical flow tracking. Edges can induce inaccuracies in optical flow computation due to the lack of detailed textural information, leading to an increase in registration error. Therefore, the static approach, by bypassing the optical flow's susceptibility to edge-related inaccuracies, presents a more reliable method in controlled imaging environments. However, in scenarios with EyeTags, optical flow yielded lower RMSE compared to static approach. This is due to the fact that textural information around the control points on EyeTags help the optical flow algorithm to predict the control points more accurately.

The calculation of FRE adds another layer of analysis, revealing the transformation accuracy of control points from the visual to thermal frames. Scenarios with static control point retention again showed a trend of lower FRE values, underscoring their potential for achieving precise registration in static or slowly changing environments. However, the optical flow method showcased a significant improvement in performance when EyeTags were implemented, as evidenced by the reduced FRE values. This suggests that EyeTags could serve as a stabilizing factor, enhancing the optical flow method's adaptability and accuracy in a dynamic tracking context.

Chapter 6

Artifact Mitigation Techniques

In the attempt of refining the ThermOcular system's capabilities for Ocular Surface Temperature (OST) assessment, two pivotal challenges were identified and addressed: the presence of eyelashes in IR images, and the detection of blinks. These challenges, if unaddressed, compromise the accuracy and reliability of OST measurements, thus affecting the system's diagnostic precision.

The current configuration of the ThermOcular system, while robust in its design and application, does not specifically address the elimination of eyelashes or the detection of blinks. Eyelashes, as external artifacts, can significantly influence thermal readings by blocking the thermal data corresponding to the ocular surface. This introduces errors in temperature measurements.

Furthermore, the system's existing approach to blink detection relies primarily on the segmentation model's ability to identify the cornea. In instances where the model fails to locate the cornea, the frame is marked as a blink and subsequently discarded. This method, while straightforward, does not account for partial blinks or situations where the

eye is slightly open. In such cases, the cornea may still be detectable, but the collected IR data may not be of diagnostic value due to the partial obstruction or interference from eyelids and eyelashes.

To address these limitations, artifact mitigation techniques are proposed. These techniques aim to enhance the system's ability to accurately identify and eliminate the influence of eyelashes and to detect blinks with greater precision. By doing so, the overall accuracy of the OST measurements will be improved and, by extension, the reliability of the ThermOcular system's diagnostic outputs.

The following sections will delve into the specific methodologies developed for eyelash elimination and blink detection, underscoring their significance in advancing the ThermOcular system towards a more accurate and efficient tool for ocular surface temperature measurement.

Eyelash Elimination

In order to enhance the ThermOcular system's precision in ocular surface temperature (OST) measurement, a significant barrier identified is the system's current segmentation algorithm's incapacity to isolate and disregard eyelashes in thermal images. Eyelashes present a unique challenge due to their differential thermal data compared to the ocular surface. This distinct temperature variance can markedly distort the corneal temperature readings, resulting in outcomes that may not accurately reflect the true thermal state of the ocular surface. Such variations are pivotal, as they can potentially affect the diagnostic accuracy of the system..

To address this problem, the present study introduces a statistical approach that is

designed to eliminate the impact of eyelashes on thermal images. Recognizing the distinct and significantly different temperature of eyelashes compared to the ocular surface, the eyelashes emerge as statistical outliers within the designated Region of Interest (ROI)—which includes the pupil, iris, and sclera. This distinction is central to the approach used, which is based on the premise that the temperature distribution of the ocular surface within the ROI follows a normal distribution pattern.

The core of the proposed method involves utilizing a statistical algorithm designed for the detection of outliers. By calculating the mean (μ) and standard deviation (σ) of the temperature distribution within the ROI, thresholds are established that specify the expected bounds of normal thermal variation:

$$\mu = \text{mean}(\text{ROI_temperature}) \tag{6.1}$$

$$\sigma = \text{std_dev}(\text{ROI_temperature}) \tag{6.2}$$

$$\text{Lower_Limit} = \mu - 2\sigma \tag{6.3}$$

$$\text{Upper_Limit} = \mu + 2\sigma \tag{6.4}$$

Pixels with temperatures that exceed these bounds—whether warmer or cooler than the established 2-standard deviation range—are classified as outliers.

The results of this methodology underscores its efficacy, revealing thermal images from which eyelash artifacts have been eliminated (Figure 6.1). Such images afford more accurate reflection of the corneal thermal data, significantly improving the ThermOcular system's reliability by ensuring that temperature readings are not affected by extraneous thermal influences.

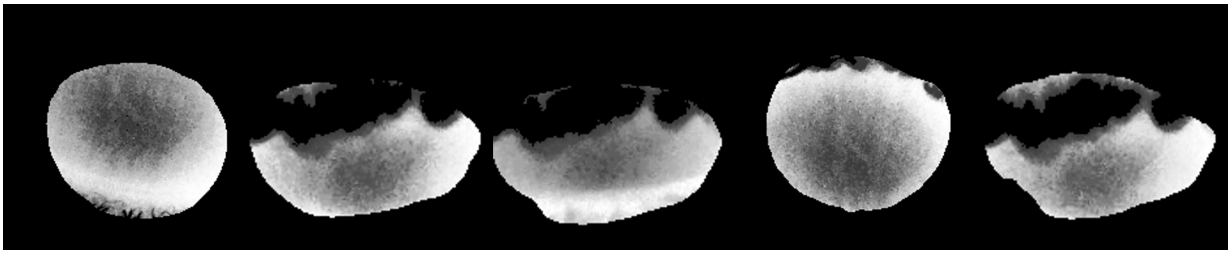


Figure 6.1: Sample IR frames of the cornea with eyelashes effectively filtered out, depicted in the same black color as the background for clarity.

Blink Detection

In the attempt to enhance the ThermOcular system's operational efficacy, particularly in accurately assessing ocular surface temperature, blinks present a noteworthy analytical challenge. Blinks, or the transient closure of the eyelid over the ocular surface, can disrupt the continuity of thermal data collection, requiring a robust mechanism for their detection and exclusion from the analysis. The existing strategy within the ThermOcular system is the exclusion of frames in which the cornea is not detected. This method, while straightforward, is prone to inaccuracies, especially in scenarios where the eyelid is not fully closed. In such instances, the captured thermal data may be unreliable, potentially skewing the OST measurement and, by extension, the diagnostic interpretation.

To address this challenge and refine the precision of blink detection, the current study proposes the adoption of the Eye Aspect Ratio (EAR) as a metric, which previously introduced in [9]. The EAR methodology helps assessing eye openness by quantitatively measuring the spatial relationship between specific anatomical landmarks around the eye. This metric is calculated as follows:

$$EAR = \frac{\|P2 - P6\| + \|P3 - P5\|}{2 \times \|P1 - P4\|} \tag{6.5}$$

Here, $P1$ through $P6$ represent the six critical landmarks identified around the eye, with $P2 - P6$ and $P3 - P5$ measuring vertical distances, and $P1 - P4$ the horizontal distance across the eye. This formula yields a scalar value indicative of the eye's openness level.

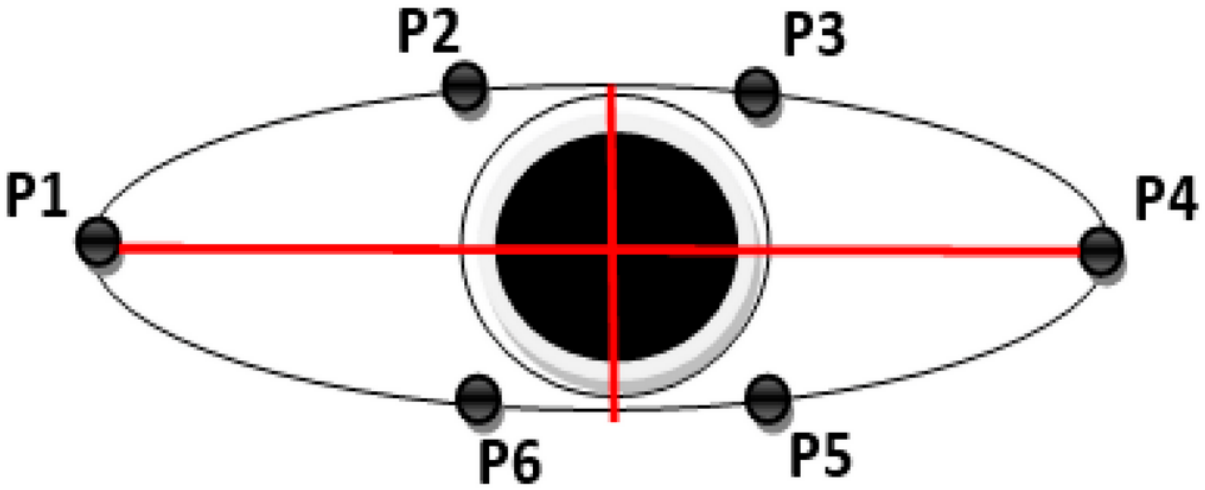


Figure 6.2: Diagram of the Eye Aspect Ratio (EAR) calculation. The EAR is computed using the distances between key landmarks: P2 to P6 and P3 to P5 for the vertical component, and P1 to P4 for the horizontal component, as indicated by the red lines.

A distinctive feature of the EAR-based approach is its ability to detect fully completed blinks and partial blinks or instances where the eye remains slightly open. The threshold is calculated for every patient individually. The median of the Eye Aspect Ratio (EAR) across all frames for a particular patient is determined, and half of that median value is

set as the threshold for blinking. Frames that fall below this threshold are flagged as blink instances and consequently omitted from the subsequent analysis.

The introduction of the EAR metric into the blink detection algorithm represents a significant methodological advancement. It allows for a more granular and accurate identification of blink events across the sequence of frames. By doing so, it ensures the integrity of the thermal data being analyzed, significantly enhancing the ThermOcular system's reliability and diagnostic accuracy.

Chapter 7

Software Design

The development of the ThermOcular system software was specifically driven by clinicians' needs for a user-friendly software to work with the system. The objective of this software was to conduct patient visits, record diagnostic videos, efficiently manage patient information, and analyze Ocular Surface Temperature (OST) data. This software is purpose-built to address these specific requirements by automating and optimizing the processes of video recording, patient data management, and OST data analysis. Its design significantly reduce the manual workload on healthcare professionals, enhances the precision of diagnostics, and streamlines patient management workflows.

Architecture Design

The core structure of the software developed for the ThermOcular system is rooted in the Model-View-Controller (MVC) design pattern (Figure 7.1), in order to make the system's software modular and scalable. This pattern facilitates a clean separation of different parts,

allowing for independent development, testing, and maintenance of each component. The software is developed in Python programming language. Tkinter library is used for the graphical user interface (GUI) to ensure cross-platform compatibility and provide a familiar and user-friendly environment for clinicians and researchers.

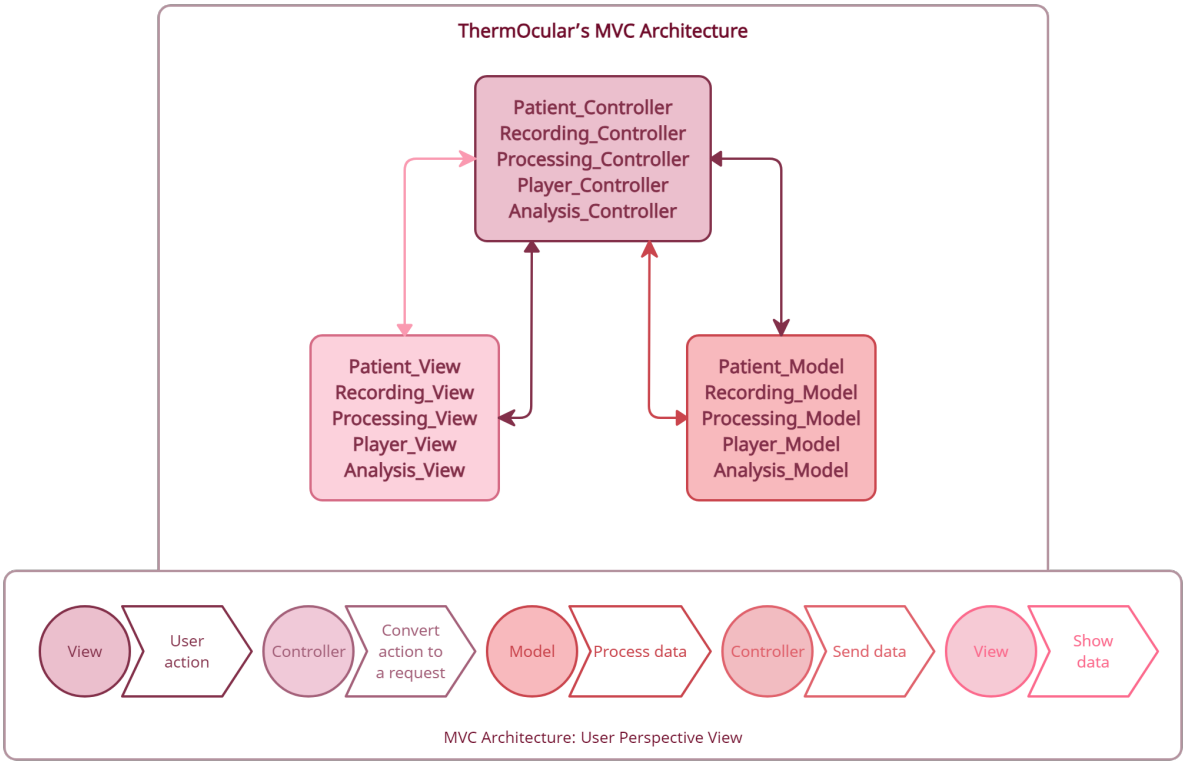


Figure 7.1: An overview of the ThermOcular software's MVC architecture, illustrating the interconnectivity between user interfaces (Views), data processing units (Models), and the central coordinating components (Controllers) for each functional module.

Database and File Management

At the heart of the system's data management is a PostgreSQL database, renowned for its robustness and reliability. The entity-relationship (ER) diagram of the database (Figure 7.2) meticulously outlines the relationships between patients, visits, video recordings, and analysis results, ensuring a structured and efficient approach to data storage and retrieval. This database foundation supports the comprehensive patient management capabilities of the software, allowing for seamless tracking of patient visits, video recordings, and analytical insights.

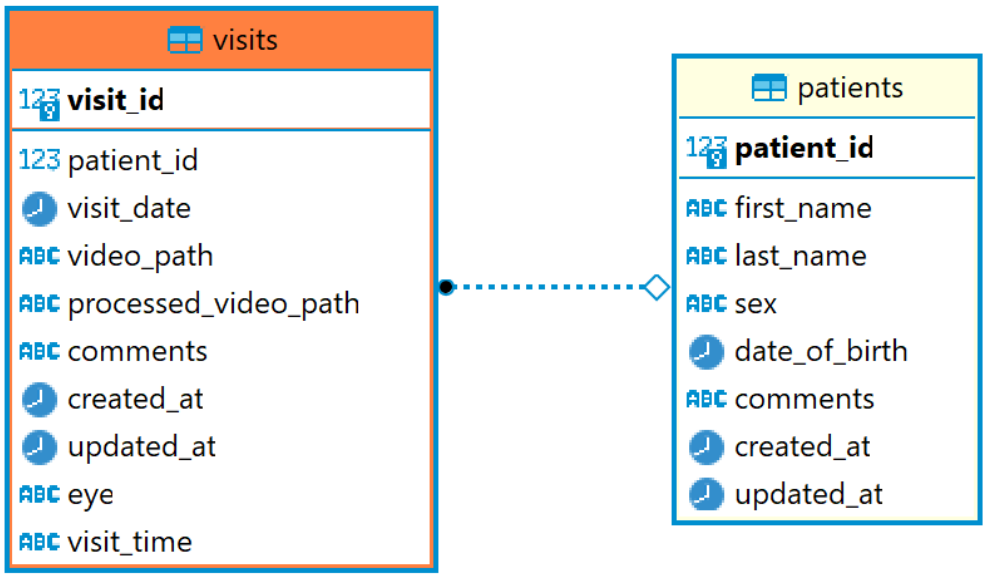


Figure 7.2: Entity-Relationship Diagram illustrating the database schema for patient visits.

The file storage strategy is designed to accommodate storing of the IR and visual images. For each patient, two distinct files are stored: one for infrared (IR) frames and another

for visual frames. This dual-file approach ensures the integrity and accessibility of the raw data for future processing and analysis and other possible projects on ThermOcular. Metadata of the patient and the path to their respective files are stored in the database, facilitating easy access and management of historical data. Clinicians are able to select specific visits and work with the data through the processing, player, and analysis modules. The software is flexible to store files either in pickle or DICOM formats, in order to have the ability to be viewed in the DICOM viewer applications.

Modular Design and Functionality

The software's interface is organized into five main tabs, each serving a distinct function within the process:

1. **Patient Management**: This tab serves as the cornerstone for managing patient information. It enables clinicians to add, update, and retrieve patient details, thereby streamlining the administrative aspect of patient care (Figure 7.3).

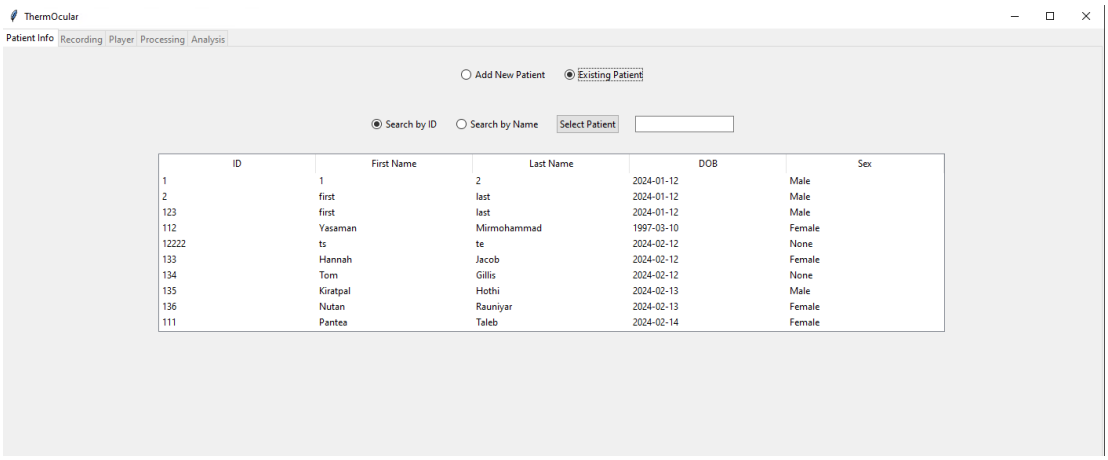


Figure 7.3: Figure showing the "Patient Management" tab Environment.

2. **Recording Videos**: The purpose of this tab is capturing and storing video recordings of patients' eyes, this module simplifies the data collection process, ensuring that high-quality IR and visual images are available for subsequent analysis. The software stores the files without compression so that high quality images are available for each patient (Figure 7.4).

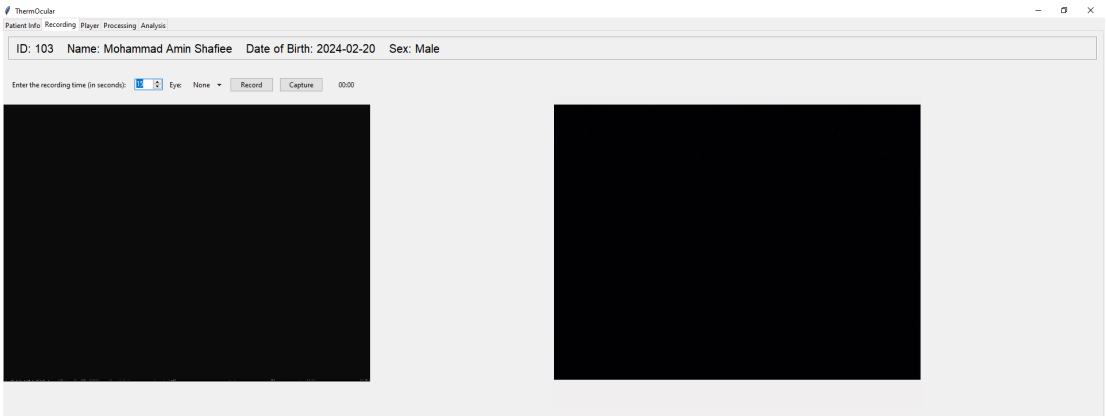


Figure 7.4: Figure showing the "Recording" tab Environment.

3. **Video Playback**: This tab provides a comprehensive overview of a patient's visits, allowing clinicians to review and select, view, or delete specific video recordings of patients(Figure 7.3).

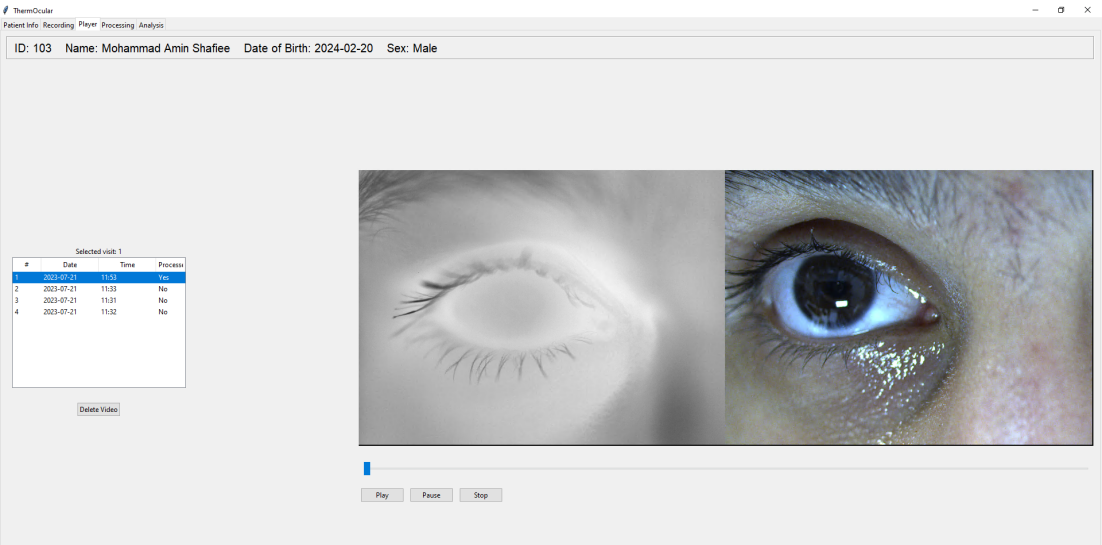


Figure 7.5: Figure showing the "Player" tab Environment.

4. **Video Processing**: At the core of the system's analytical capabilities, this module includes the implemented algorithms to process the thermal and visual videos. Clinicians can interactively select control points and process each visit of the selected patient. The software stores the processed results of the visits and updates the database with the processed video path so that for analyzing each visit clinicians need to process the videos only once (Figure 7.6).

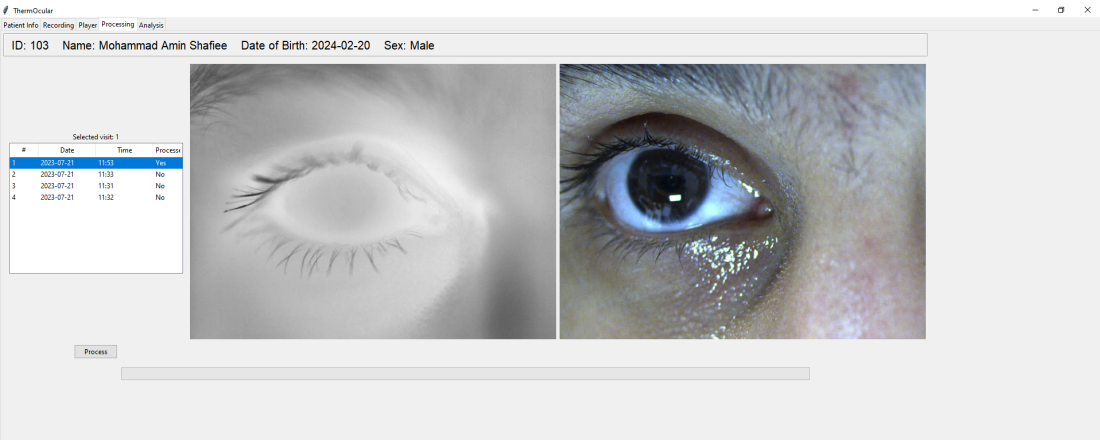


Figure 7.6: Figure showing the "Processing" tab Environment.

5. **Analysis Results**: Designed for in-depth examination of temperature variations across different regions of the eye, this tab offers clinicians a powerful tool for tracking temperature changes and extracting valuable insights into OST of the patients (Figure 7.7).

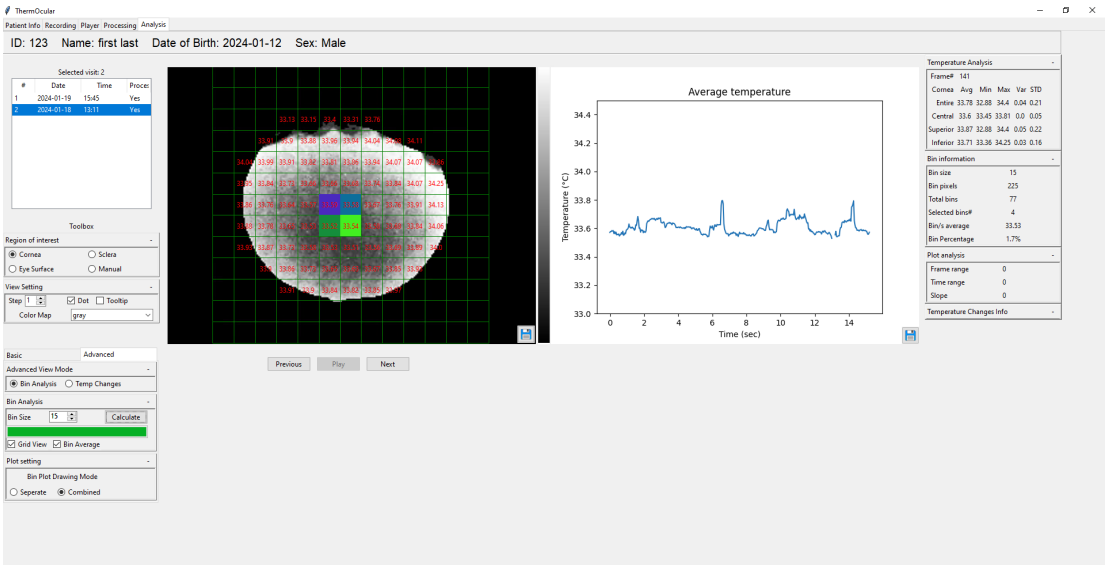


Figure 7.7: Figure showing the "Analysis" tab Environment.

Each tab is constructed using the MVC design pattern, ensuring scalability and modular design of the software. The **View** component is responsible for rendering the GUI, providing a straightforward and efficient interface for users to interact with the software. The **Model** component handles data processing, including interactions with the database, file management, and the implementation of specialized algorithms for the ThermOcular system. Lastly, the **Controller** merges the Model and View, managing the functionality of each tab and facilitating seamless data flow throughout the system.

The segmentation model, a pivotal element in the processing of ocular images, is implemented using PyTorch[42]. PyTorch's flexibility and powerful computational capabilities enable the software to perform complex image segmentation tasks, essential for accurate temperature analysis and diagnostics.

Chapter 8

Conclusion and future research

Conclusion

This thesis has introduced significant enhancements to the ThermOcular system, advancing its capabilities in ocular surface temperature (OST) assessment. The primary improvements include:

- **Enhanced Segmentation:** The segmentation process was refined to accurately identify multiple ocular components, including the pupil, iris, sclera, and especially the cornea. This refinement allows for precise temperature readings across the ocular surface, providing a comprehensive analysis beneficial for diagnosing a variety of ocular conditions.
- **Artifact Removal:** The development of robust methods for the removal of artifacts such as eyelashes and blinks from thermal images significantly improves the diagnostic reliability of the system. By ensuring that temperature readings accurately reflect

the ocular surface's condition, the system can offer more reliable information about ocular surface temperature.

- **Improved Registration Method:** The registration process was enhanced to facilitate ease of use for clinicians, eliminating the need for extensive training on control point selection. This improvement not only enhances the system's user experience but also increases its accuracy by reducing potential errors associated with manual input and the traditional registration strategy of the ThermOcular system.

- **Comprehensive Software System:** A software system was developed to manage patient records and automate the recording, processing, and analysis of videos. This system significantly streamlines operations, making the diagnostic process more efficient and user-friendly.

Future Works

While the current enhancements have substantially improved the ThermOcular system, there remain areas for future development to further augment its capabilities:

- **Automated EyeTag Detection:** Future work could focus on perfecting the automation of eye tag detection, reducing, or completely eliminating the need for manual selection of control points. This would streamline the system's operation, making it fully autonomous and even more user-friendly.

- **Enhanced Computational Infrastructure:** Future developments could introduce an embedded GPU-based system within the ThermOcular device. This enhancement aims to mitigate the computational limitations of the existing clinical workstations,

which are often outdated and not equipped for deep learning tasks. Such an embedded system would:

  - Facilitate the local processing of deep learning models directly on the device, ensuring rapid analysis and real-time application in clinical diagnostics.
  - Comply with strict healthcare privacy policies, allowing all data to be processed in situ, thereby maintaining the integrity and confidentiality of patient information.

- **Disease Diagnosis through OST Pattern Analysis:** Future enhancements could leverage the ThermOcular system to study the patterns of Ocular Surface Temperature (OST) changes in patients afflicted with various diseases that show ocular temperature fluctuations. Machine learning models could be developed to assist clinicians in diagnosing diseases based on OST changes. This approach not only opens new avenues for non-invasive diagnostics but also capitalizes on the advanced capabilities of the ThermOcular system to provide a deeper understanding of ocular health and its relation to diseases.

References

[1] Rahib Hidayat Abiyev and Koray Altunkaya. Neural network based biometric personal identification with fast iris segmentation. *International Journal of Control, Automation and Systems*, 7:17–23, 2009.

[2] U Rajendra Acharya, Jen Hong Tan, S Vidya, Sharon Yeo, Cheah Loon Too, Wei Jie Eugene Lim, Kuang Chua Chua, and Louis Tong. Diagnosis of response and non-response to dry eye treatment using infrared thermography images. *Infrared Physics & Technology*, 67:497–503, 2014.

[3] Muhammad Arsalan, Rizwan Ali Naqvi, Dong Seop Kim, Phong Ha Nguyen, Muhammad Owais, and Kang Ryoung Park. Irisdensenet: Robust iris segmentation using densely connected fully convolutional networks in the images by visible light and near-infrared light camera sensors. *Sensors*, 18(5):1501, 2018.

[4] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.

[5] Cides S Bezerra, Rayson Laroca, Diego R Lucio, Evair Severo, Lucas F Oliveira, Alceu S Britto, and David Menotti. Robust iris segmentation based on fully convolutional

networks and generative adversarial networks. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 281–288. IEEE, 2018.

[6] Ehsan Zare Bidaki, Alexander Wong, and Paul Murphy. A novel computational thermal-visual imaging system for automatic cornea temperature measurement and tracking. *Journal of Computational Vision and Imaging Systems*, 8(1):20–23, 2022.

[7] P Bourjat and M Gautherie. Unilateral exophthalmos investigated by infrared thermography. *Modern problems in ophthalmology*, 14:278–285, 1975.

[8] Genis Cardona, Philip B Morgan, Nathan Efron, and Andrew B Tullo. Ocular and skin temperature in ophthalmic postherpetic neuralgia. *Pain Clinic*, 9(2):145–150, 1996.

[9] Jan Cech and Tereza Soukupova. Real-time eye blink detection using facial landmarks. *Cent. Mach. Perception, Dep. Cybern. Fac. Electr. Eng. Czech Tech. Univ. Prague*, pages 1–8, 2016.

[10] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017.

[11] Huihua Kenny Chiang, Chih Yen Chen, Hung You Cheng, Ko-Hua Chen, and David O Chang. Development of infrared thermal imager for dry eye diagnosis. In *Infrared and Photoelectronic Imagers and Detector Devices II*, volume 6294, pages 36–43. SPIE, 2006.

[12] Jennifer P Craig, Iqbal Singh, Alan Tomlinson, Philip B Morgan, and Nathan Efron. The role of tear physiology in ocular surface temperature. *Eye*, 14(4):635–641, 2000.

[13] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.

[14] Thomas J Dursch, Wing Li, Baseem Taraz, Meng C Lin, and Clayton J Radke. Tear-film evaporation rate from simultaneous ocular-surface temperature and tear-breakup area. *Optometry and Vision Science*, 95(1):5–12, 2018.

[15] Brennan NA Efron N, Young G. Ocular surface temperature. *Curr Eye Res*, 1989.

[16] Hazem M El-Bakry. Human iris detection using fast cooperative modular neural nets and image decomposition. *Machine Graphics & Vision International Journal*, 11(4):499–512, 2002.

[17] HM El-Bakry. Fast iris detection using neural nets. In *Canadian Conference on Electrical and Computer Engineering 2001. Conference Proceedings (Cat. No. 01TH8555)*, volume 2, pages 1409–1414. IEEE, 2001.

[18] AI Eremenko. Thermography in the diagnosis of vascular neuritis of the optic nerve. *Oftalmologicheskii zhurnal*, (4):235–239, 1990.

[19] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results. http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html.

[20] Wolfgang Fuhl, Gjergji Kasneci, and Enkelejda Kasneci. Teyed: Over 20 million real-world eye images with pupil, eyelid, and iris 2d and 3d segmentations, 2d and 3d landmarks, 3d eyeball, gaze vector, and eye movement types. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 367–375. IEEE, 2021.

[21] Wolfgang Fuhl, Thiago Santini, Gjergji Kasneci, and Enkelejda Kasneci. Pupilnet: Convolutional neural networks for robust pupil detection. *arXiv preprint arXiv:1601.04902*, 2016.

[22] Wolfgang Fuhl, Thiago Santini, Gjergji Kasneci, Wolfgang Rosenstiel, and Enkelejda Kasneci. Pupilnet v2. 0: Convolutional neural networks for cpu based real time robust pupil detection. *arXiv preprint arXiv:1711.00112*, 2017.

[23] Fernando Galassi, Barbara Giambene, Andrea Corvi, and Giacomo Falaschi. Evaluation of ocular surface temperature and retrobulbar haemodynamics by infrared thermography and colour doppler imaging in patients with glaucoma. *British Journal of Ophthalmology*, 91(7):878–881, 2007.

[24] Abhishek Gangwar, Akanksha Joshi, Ashutosh Singh, Fernando Alonso-Fernandez, and Josef Bigun. Irisseg: A fast and robust iris segmentation framework for non-ideal iris images. In *2016 international conference on biometrics (ICB)*, pages 1–8. IEEE, 2016.

[25] Eri Hosaka, Takushi Kawamorita, Yuko Ogasawara, Nanami Nakayama, Hiroshi Uozato, Kimiya Shimizu, Murat Dogru, Kazuo Tsubota, and Eiki Goto. Interferometry in the evaluation of precorneal tear film thickness in dry eye. *American journal of ophthalmology*, 151(1):18–23, 2011.

[26] Bogdasarov IuB, OP Lenskaia, and BM Belkina. Thermography in the diagnosis of retinoblastoma in children. *Meditsinskaia radiologiia*, 30(12):19–21, 1985.

[27] Ranjeet Ranjan Jha, Gaurav Jaswal, Divij Gupta, Shreshth Saini, and Aditya Nigam. Pixisegnet: pixel-level iris segmentation network using convolutional encoder–decoder with stacked hourglass bottleneck. *IET biometrics*, 9(1):11–24, 2020.

[28] Tomoyuki Kamao, Masahiko Yamaguchi, Shiro Kawasaki, Shiro Mizoue, Atsushi Shiraishi, and Yuichi Ohashi. Screening for dry eye with newly developed ocular surface thermographer. *American journal of ophthalmology*, 151(5):782–791, 2011.

[29] Henry Kricancic, Hamish McNeill, Michael Titze, David Alonso-Caneiro, and Michael J Collins. Instrument for simultaneous assessment of fluorescein and thermal dynamics of the tear film. In *2017 ARVO ASIA Meeting*, 2017.

[30] Wing Li, Andrew D Graham, Steve Selvin, and Meng C Lin. Ocular surface cooling corresponds to tear film thinning and breakup. *Optometry and Vision Science*, 92(9):e248–e256, 2015.

[31] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[32] R Mapstone. Measurement of corneal temperature. *Experimental Eye Research*, 7(2):237–IN29, 1968.

[33] William Mathers. Evaporation from the ocular surface. *Experimental eye research*, 78(3):389–394, 2004.

[34] Sara Matteoli, Davide Coppini, and Andrea Corvi. A novel image processing procedure for thermographic image analysis. *Medical & Biological Engineering & Computing*, 56:1747–1756, 2018.

[35] Sara Matteoli, Lucia Finocchio, Ilaria Biagini, Giovanni Giacomelli, Andrea Sodi, Andrea Corvi, Gianni Virgili, and Stanislao Rizzo. A thermographic study on eyes affected by age-related macular degeneration: Comparison among various forms of

the pathology and analysis of risk factors. *Infrared Physics & Technology*, 76:402–407, 2016.

[36] Philip B Morgan, J Vincent Smyth, Andrew B Tullo, and Nathan Efron. Ocular temperature in carotid artery stenosis. *Optometry and vision science*, 76(12):850–854, 1999.

[37] Philip B Morgan, Andrew B Tullo, and Nathan Efron. Infrared thermography of the tear film in dry eye. *Eye*, 9(5):615–618, 1995.

[38] Philip B Morgan, Andrew B Tullo, and Nathan Efron. Ocular surface cooling in dry eye—a pilot study. *Journal of the British Contact Lens Association*, 19(1):7–10, 1996.

[39] Asako Mori, Yoshihisa Oguchi, Yukio Okusawa, Masafumi Ono, Hiroshi Fujishima, and Kazuo Tsubota. Use of high-speed, high-resolution thermography to evaluate the tear film layer. *American journal of ophthalmology*, 124(6):729–735, 1997.

[40] Paul J Murphy, Philip B Morgan, Sudi Patel, and John Marshall. Corneal surface temperature change as the mode of stimulation of the non-contact corneal aesthesiometer. *Cornea*, 18(3):333, 1999.

[41] EYK Ng, Gerk Chang Yee, Tan Jian Hua, Manjunath Kagathi, et al. Analysis of normal human eye with different age groups using infrared images. *Journal of medical systems*, 33(3):207–213, 2009.

[42] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

[43] Juncai Peng, Yi Liu, Shiyu Tang, Yuying Hao, Lutao Chu, Guowei Chen, Zewu Wu, Zeyu Chen, Zhiliang Yu, Yuning Du, et al. Pp-liteseg: A superior real-time semantic segmentation model. *arXiv preprint arXiv:2204.02681*, 2022.

[44] Christine Purslow and James S Wolffsohn. Ocular surface temperature: a review. *Eye & contact lens*, 31(3):117–123, 2005.

[45] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015.

[46] Ted Rosenstock, P Chart, and JJ Hurwitz. Inflammation of the lacrimal drainage system-assessment by thermography, 1983.

[47] A Sodi, B Giambene, G Falaschi, R Caputo, Bernardo Innocenti, Andrea Corvi, and Ugo Menchini. Ocular surface temperature in central retinal vein occlusion: preliminary data. *European journal of ophthalmology*, 17(5):755–759, 2007.

[48] Tai-Yuan Su, Shu-Wen Chang, Chiao-Ju Yang, and Huihua Kenny Chiang. Direct observation and validation of fluorescein tear film break-up patterns by using a dual thermal-fluorescent imaging system. *Biomedical optics express*, 5(8):2614–2619, 2014.

[49] Vidya K Sudarshan, Joel EW Koh, U Rajendra Acharya, Jen Hong Tan, Muthu Rama Krishnan Mookiah, Chua Kuang Chua, and Louis Tong. Evaluation of evaporative dry eye disease using thermal images of ocular surface regions with dwt and gabor transform. *Application of infrared to biomedical sciences*, pages 359–375, 2017.

[50] Jen-Hong Tan, Eddie YK Ng, and Rajendra Acharya U. Automated detection of eye and cornea on infrared thermogram using snake and target tracing function coupled

with genetic algorithm. *Quantitative InfraRed Thermography Journal*, 6(1):21–36, 2009.

[51] Jen Hong Tan, EYK Ng, and Rajendra Acharya U. An efficient automated algorithm to detect ocular surface temperature on sequence of thermograms using snake and target tracing function. *Journal of medical systems*, 35:949–958, 2011.

[52] Li Li Tan, Srinivasan Sanjay, and Philip B Morgan. Repeatability of infrared ocular thermography in assessing healthy and dry eyes. *Contact Lens and Anterior Eye*, 39(4):284–292, 2016.

[53] Francisco Javier Vera-Olmos and Norberto Malpica. Deconvolutional neural network for pupil detection in real-world environments. In *Biomedical Applications Based on Natural and Artificial Computing: International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2017, Corunna, Spain, June 19-23, 2017, Proceedings, Part II*, pages 223–231. Springer, 2017.

[54] Richard P Wildes. Iris recognition: an emerging biometric technology. *Proceedings of the IEEE*, 85(9):1348–1363, 1997.

[55] Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M Alvarez, and Ping Luo. Segformer: Simple and efficient design for semantic segmentation with transformers. *Advances in Neural Information Processing Systems*, 34:12077–12090, 2021.

[56] Yuting Yang, Peisong Shen, and Chi Chen. A robust iris segmentation using fully convolutional network with dilated convolutions. In *2018 IEEE International Symposium on Multimedia (ISM)*, pages 9–16. IEEE, 2018.

[57] Yuk-Hoi Yiu, Moustafa Aboulatta, Theresa Raiser, Leoni Ophey, Virginia L Flanagin, Peter Zu Eulenburg, and Seyed-Ahmad Ahmadi. Deepvog: Open-source pupil segmentation and gaze estimation in neuroscience using deep learning. *Journal of neuroscience methods*, 324:108307, 2019.

[58] Yuhui Yuan, Xilin Chen, and Jingdong Wang. Object-contextual representations for semantic segmentation. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VI 16*, pages 173–190. Springer, 2020.

[59] Zhong Bo Zhang, Si Liang Ma, Ping Zuo, and Jie Ma. Fast iris detection and localization algorithm based on adaboost algorithm and neural networks. In *2005 International Conference on Neural Networks and Brain*, volume 2, pages 1085–1088. IEEE, 2005.

[60] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2881–2890, 2017.

[61] Tianming Zhao, Yuanning Liu, Guang Huo, and Xiaodong Zhu. A deep learning iris recognition method based on capsule network architecture. *IEEE Access*, 7:49691–49701, 2019.

[62] Shuang Zheng, Dongmei Fu, Tao Yang, Ling Luo, and Xinshuai Nan. A novel method for eye contour extraction from blurred infrared images. In *2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, volume 1, pages 103–106. IEEE, 2016.

[63] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*, pages 3–11. Springer, 2018.