

Machine Learning Models for the Diagnosis of Dry Eyes Using Real-world Clinical Data

Tamer N Jarada,^{1,2} Karl Stonecipher,^{3,4} Olivia Perez³ and Ahmed R Al-Ghoul^{1,2}

1. CSI Dry Eye Inc., Calgary, Alberta, Canada; 2. Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; 3. Department of Ophthalmology, University of North Carolina, Chapel Hill, NC, USA; 4. Department of Ophthalmology, Tulane University, New Orleans, LA, USA

DOI: <https://doi.org/10.17925/USOR.2024.18.1.9>

Introduction: Dry eye disease (DED) is a common ocular condition marked by discomfort and tear film instability. Machine learning (ML) techniques are increasingly recognized for their potential to transform healthcare, particularly in ophthalmology. **Methods:** This study aims to develop ML models for predicting the severity and type of DED using demographic and clinical data. Real-world clinical data collected over 18 months served as the basis for creating two diagnostic models. The data set comprised 1,313 well-structured samples, each annotated by domain experts, showcasing multiple demographic and clinical features. A correlation feature selection technique was used to eliminate redundant and irrelevant features. Through stratified 10-fold cross-validation, two support vector machine models – dry eye severity model (SM) and dry eye type model (TM) – were developed to predict the severity and type of DED. **Results:** The SM achieved a moderate performance in predicting the severity of dry eye cases, with an area under the receiver operating characteristic (AUC-ROC) of 0.79 and an area under the precision-recall (AUC-PR) of 0.61. Furthermore, the TM demonstrated effectiveness in predicting different dry eye types, yielding an AUC-ROC of 0.91 and an AUC-PR of 0.83. We also verified the robustness of both SM and TM by comparing their performance with nine baseline ML methods. Both SM and TM consistently outperformed the baseline methods in terms of AUC-ROC and AUC-PR. **Conclusion:** The potential application of these models lies in improving health outcomes and offering early alerts to potentially prevent the progression of DED.

Keywords

Clinical diagnostics, dry eye, dry eye disease, dry eye type, dry eye severity, machine learning, real-world data

Disclosures: Tamer N Jarada is the Data Science Lead at CSI Dry Eye Inc. Karl Stonecipher does research, speaks and consults for CSI Dry Eye. Ahmed R Al-Ghoul is the Founder and CEO of CSI Dry Eye Inc. Olivia Perez has no financial or non-financial relationships or activities to declare in relation to this article.

Review process: Double-blind peer review.

Compliance with ethics: This article is a retrospective study and is not subject to Health Research Ethics Board of Alberta (HREBA) approval. It was conducted in accordance with the principles outlined in the Helsinki Declaration of 1964 and its later amendments. Written informed consent to participate in the study and for the publication of this article was obtained from all study participants, or their parent/guardian or next of kin if the participant was deceased or unable to provide consent, and the article does not contain any identifying information about the participants.

Data availability: The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Authorship: The named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship of this manuscript, take responsibility for the integrity of the work as a whole, and have given final approval for the version to be published.

Access: This article is freely accessible at touchOPHTHALMOLOGY.com © Touch Medical Media 2024.

Received: 4 September 2023

Accepted: 2 February 2024

Citation: *touchREVIEWS in Ophthalmology*. 2024;18(1):46-48

Corresponding author: Dr Karl Stonecipher, 1002 N. Church Street, Suite 101, Greensboro, NC 27401, USA. E: stonenc@gmail.com

Support: No funding was received in the publication of this article.

Dry eye disease (DED) is a prevalent ocular condition characterized by discomfort, visual disturbances and tear film instability due to insufficient tear production or rapid tear evaporation.¹ With an increasing global incidence (i.e. up to 50%) and its significant impact on the quality of life, there is a growing need for accurate and efficient diagnostic methods to aid in its timely identification and management.² Traditional diagnostic approaches rely on subjective assessments and clinical evaluations, often leading to variability and suboptimal accuracy.

In recent years, machine learning (ML) techniques have emerged as potential tools in revolutionizing healthcare, including in ophthalmology.³ ML models offer the possibility of improving accuracy, objectivity and speed in medical diagnosis by leveraging extensive real-world clinical data.⁴ These algorithms analyze complex patterns and relationships within data sets, extracting valuable insights not readily apparent to human clinicians. This potential could significantly benefit the diagnosis and management of DED.

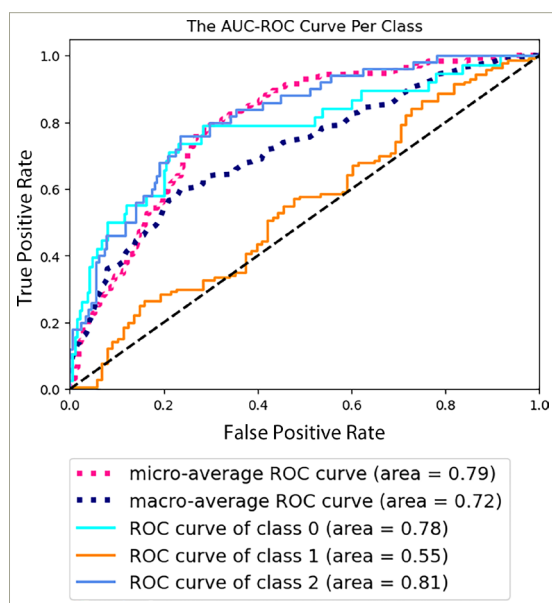
Recent studies have showcased the potential of ML models in ophthalmology. For instance, Ting et al. developed a deep learning system for the diagnosis of diabetic retinopathy.³ While not specific to dry eye, this study illustrated how ML models can effectively use real-world clinical data to classify and diagnose complex eye conditions. Leveraging a large data set of retinal images, the authors achieved an AUC-ROC of 0.936, underscoring the power of ML techniques in medical image analysis.

In this study, we explore the application of ML models for the diagnosis of DED using real-world clinical data. We examine various approaches and methodologies used by researchers to harness the potential of ML in identifying and classifying DED. The insights gained from this study have the potential to contribute to the evolution of the diagnosis of dry eyes, providing more accurate and efficient methods that may lead to improved patient outcomes and an enhanced quality of life.

Background

The study aims to enhance diagnostics in DED by specifically examining the type and severity of the disease using ML in a typical clinical setting. The objectives include the use of real-world clinical data to create ML models for assessing the type and severity of DED, focussing on facilitating the implementation by clinicians in a clinical setting.

Figure 1: The area under the receiver operating characteristic curve per class of the dry eye severity model



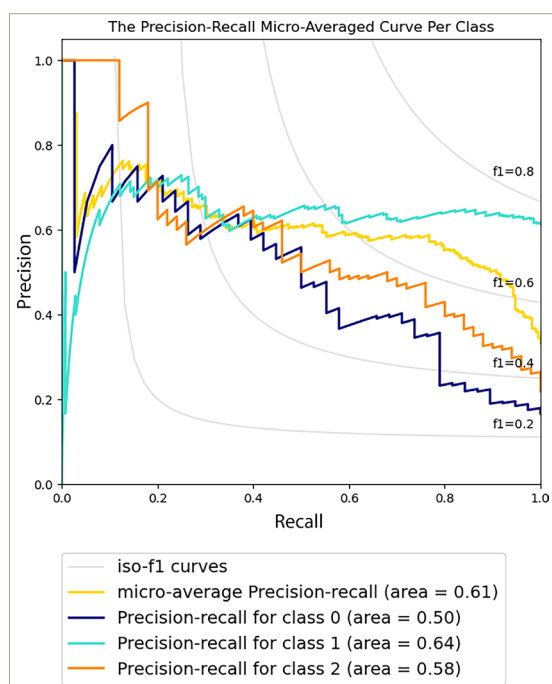
AUC-ROC = area under the receiver operating characteristic.

Real-world clinical data collected over 18 months were used to develop two ML diagnostic models for the severity and type of dry eyes. The data set from a Canadian clinic comprised 1,313 well-structured samples, each annotated by domain experts, showcasing numerous features.

Methods

Real-world data was collected in Canadian clinics for over 28 months from December 1, 2018 to March 31, 2021 and then anonymized. These data were used to develop two ML diagnostic models for assessing the severity and type of DED. The data set consisted of 1,313 patient samples, each carefully annotated by domain specialists, showcasing various features. The use of such extensive real-world clinical data

Figure 2: The precision-recall micro-averaged curve per class of the dry eye severity model



aims to enhance the generalizability and relevance of the models in practical settings.

To optimize the performance of the ML models, we applied a correlation feature selection technique.⁵ This method systematically identified and eliminated redundant and irrelevant features, refining the input variables to focus on the most pertinent aspects for accurate diagnosis. This step streamlined the models and improved their interpretability.

To assess the performance and robustness of the models comprehensively, we used a stratified 10-fold cross-validation methodology.⁶ This technique divides the data set into subsets, maintaining the distribution of the original data to minimize bias and generate a more reliable assessment of the predictive capabilities of the models.

Among the various ML techniques, support vector machines (SVMs) demonstrated proficiency in handling the intricacies of the data set.⁷ SVMs excel in categorizing the data into distinct groups, making them well-suited for discerning different levels of severity and types of dry eyes.

We developed and fine-tuned two distinct SVM models, the dry eye severity model (SM) and the dry eye type model (TM). We used the p-value as a statistical measure to validate our hypothesis against the observed data and ensure that our proposed models achieve sufficient statistical power. The SM quantifies the severity of DED, offering insights into its progression and intensity. The TM classifies dry eye types, aiding clinicians in making more precise treatment decisions. Importantly, both models use the same set of features for predictions, ensuring consistency and comparability between severity and type assessments. This uniformity in feature selection enhances the applicability of the models and simplifies their integration into clinical practice.

This study was performed in accordance with the Helsinki Declaration of 1964 and its later amendments. Written informed consent was received from the patients to participate in this study and for the publication of this article. This article does not contain any identifying information about the participants.

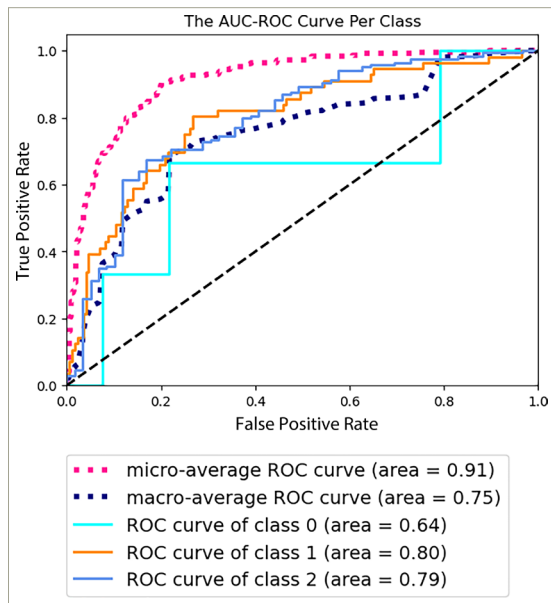
Results

The SM demonstrated satisfactory performance in predicting the severity of dry eyes in various cases, with an area under the receiver operating characteristic (AUC-ROC) curve of 0.79 ($p\text{-value} = 1.41e^{-05}$) (Figure 1) and an area under the precision-recall (AUC-PR) curve of 0.61 ($p\text{-value} = 1.34e^{-03}$) (Figure 2). These metrics offer a comprehensive evaluation of the capability of the models to differentiate between different severity levels and effectively identify true-positive cases while managing false positives.

On the other hand, the TM showed reliable predictive capabilities in distinguishing various dry eye types, as indicated by an AUC-ROC value of 0.91 ($p\text{-value} = 1.28e^{-04}$) (Figure 3) and an AUC-PR value of 0.83 ($p\text{-value} = 1.37e^{-04}$) (Figure 4). These values suggest the effectiveness of the model in accurately identifying different dry eye classifications. The notable metrics highlight the strength and capacity of the models to capture true-positive cases while minimizing false positives.

The robustness and performance of both the SM and TM were confirmed through a thorough comparison with nine baseline ML methods. These baseline methods included various traditional algorithms commonly used in classification tasks. The results of the comparative analysis

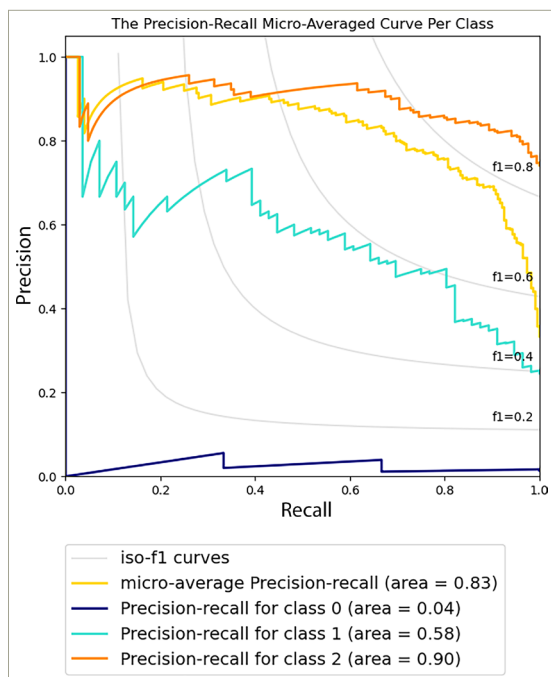
Figure 3: The area under the receiver operating characteristic curve per class of the dry eye type model



AUC-ROC = area under the receiver operating characteristic.

emphasized the consistent performance of the SM and TM. Across both AUC-ROC and AUC-PR metrics, both models consistently outperformed all baseline methods, indicating their effectiveness in terms of predictive accuracy, sensitivity and specificity.

Figure 4: The precision-recall micro-averaged curve per class of the dry eye type model



Discussion

The AUC-ROC and AUC-PR values achieved by the SM and TM, coupled with their consistent outperformance over baseline methods, emphasize their usefulness in accurately diagnosing the severity and type of dry eyes. These findings suggest the potential of ML techniques to contribute to ophthalmology and refine diagnostic procedures.

In the realm of the diagnosis of dry eyes, the ML models introduced in this study serve as a validated prediction framework. Their ability to predict both severity and type of dry eye reflects their robustness and potential impact. By providing clinicians with a tool for reliably anticipating dry eye conditions, these models have the potential to improve health outcomes. Their predictive capabilities may play a role in initiating timely interventions, potentially preventing the progression of DED.

Beyond diagnostic accuracy, these models could proactively contribute to patient care by offering early alerts. This could enable healthcare professionals to implement preventive measures or tailor treatment strategies promptly, potentially mitigating the discomfort and impairment associated with dry eyes and enhancing the quality of life of patients.

While already demonstrating predictive capabilities, the trajectory of these ML models is set for further refinement and augmentation. The integration of future data, encompassing a broader array of clinical scenarios and patient profiles, has the potential to enhance these models further. As these models continue to learn from diverse data sets, their precision and reliability are expected to improve, allowing them to adapt to the nuanced complexities inherent in DED.

To implement the findings of this study in clinical practice, we have taken a proactive step by seamlessly integrating the SM and TM with the widely adopted CSI Dry Eye Software (CSI Dry Eye Inc., Calgary, Alberta, Canada). This strategic integration enables a streamlined incorporation of our ML models into the existing workflows of clinicians globally who rely on the CSI Dry Eye Software. By leveraging this established platform, we aim to facilitate the accessibility and adoption of our predictive models, ensuring that clinicians worldwide can readily incorporate these models into their routine diagnostic practices. This collaborative approach not only enhances the practical applicability of our models but also fosters a more efficient and widespread integration of advanced diagnostic capabilities into real-world clinical settings. As a result, clinicians using the CSI Dry Eye Software can now benefit from the enhanced predictive accuracy and tailored insights provided by our SM and TM, contributing to more precise and personalized management of DED for patients worldwide.

Conclusion

By using ML and incorporating real-world clinical data, the models presented in this study contribute to the field of dry eye diagnosis and highlight the potential of ML to support medical decision-making processes. The validation of these ML models as the initial framework for predicting the severity and type of dry eyes marks a significant advancement. With an ongoing focus on improvement through the integration of future data, these models have the potential to bring changes to the landscape of dry eye diagnosis and treatment. □

- Wolffsohn JS, Arita R, Chalmers R, et al. TFOS DEWS II diagnostic methodology report. *Ocul Surf.* 2017;15:539–74. DOI: 10.1016/j.jtos.2017.05.001.
- Storås AM, Strömke I, Riegler MA, et al. Artificial intelligence in dry eye disease. *Ocul Surf.* 2022;23:74–86. DOI: 10.1016/j.jtos.2021.11.004.
- Ting DSW, Cheung CY-L, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA.* 2017;318:2211–23. DOI: 10.1001/jama.2017.18152.
- Burlina PM, Joshi N, Pekala M, et al. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA Ophthalmol.* 2017;135:1170–6. DOI: 10.1001/jamaophthalmol.2017.3782.
- Guyon I, Elisseeff A. An introduction to variable and feature selection. *J Mach Learn Res.* 2003;3:1157–82.
- Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artificial Intelligence.* 1995;14:1137–45.
- Cortes C, Vapnik V. Support-vector networks. *Mach Learn.* 1995;20:273–97. DOI: 10.1007/BF00994018.