Predicting Meibomian Gland Dropout and Feature Importance Analysis with Explainable Artificial Intelligence

Fredrik A. Fineide*†‡, Andrea M. Storås†§, Michael A. Riegler§¶, Tor Paaske Utheim*†‡‖,
*Department of Medical Biochemistry, Oslo University Hospital, Oslo, Norway
†Department of Computer Science, Oslo Metropolitan University, Oslo, Norway
‡The Norwegian Dry Eye Clinic, Oslo, Norway
§Department of Holistic Systems, SimulaMet, Oslo, Norway
¶Department of Computer Science, UiT The Arctic University of Norway, Tromsø, Norway
‖Department of Ophthalmology, Oslo University Hospital, Norway

***Abstract*—Dry eye disease is a common and potentially debilitating medical condition. Meibum secreted from the meibomian glands is the largest contributor to the outermost, protective lipid layer of the tear film. Dysfunction of the meibomian glands is the most common cause of dry eye disease. As meibomian gland dysfunction progresses, gradual atrophy of the glands is observed. The meibomian glands are commonly visualized through meibography, a technique requiring specialist equipment and knowledge that might not be available to the physician. In the present project we use machine learning on clinical tabular data to predict the degree of meibomian gland dropout. Moreover, we employ explainable artificial intelligence on the best performing algorithms for feature importance evaluation. The best performing algorithms were *AdaBoost*, *multilayer perceptron* and *LightGBM* which outperformed the majority vote baseline classifier in every included evaluation metric for both multioutput and binary classification. Through explainable artificial intelligence known associations are validated and novel connections identified and discussed.**

***Index Terms*—Machine Learning, Explainable Artificial Intelligence, Dry Eye Disease, Meibomian Gland Dysfunction, Meibography**

I. Introduction

The meibomian glands (MGs) are modified sebaceous glands in the eyelids. These glands are responsible for producing and secreting meibum, a lipid substance forming the majority of the lipid layer of the ocular tear film [1].

A schematic illustration of the tear film is included in Figure 1. The lipid layer is the outermost layer of the tear film, decreasing evaporation of the underlying mucoaqueous component, thus preventing dehydration of the ocular surface and providing protection from the external environment [1]. Meibomian gland dysfunction (MGD) is the most common cause of dry eye disease (DED), which affects 5-50% of the population [2], [3]. The hallmark of DED is a disruption of tear film homeostasis, while MGD is defined as "a chronic, diffuse abnormality of the meibomian glands, commonly characterized by terminal duct obstruction and/or qualitative/quantitative changes in the glandular secretion. This may result in alteration of the tear film, symptoms of eye irritation, clinically apparent inflammation, and ocular surface disease" [3], [4]. The loss of MGs, often referred to as meibomian gland (MG) dropout, is associated with several underlying local and systemic disease states, their treatment, senescence, and MGD [5].

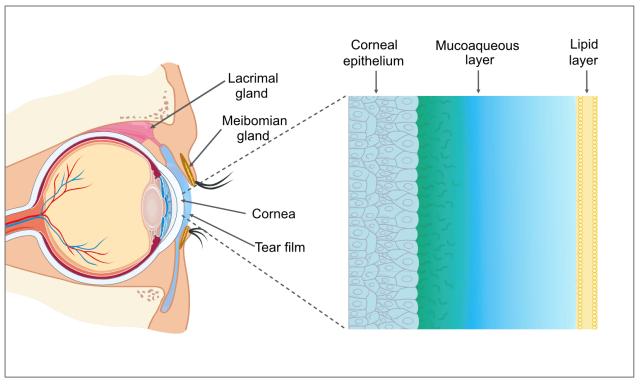


Figure 1: The ocular tear film, its layers and contributors. Illustration by Sara Nøland.

The degree of MG dropout is commonly determined through contact, or more recently, non-contact meibography, depicting the inside of the eyelids and the MGs with infrared photography. Following imaging, the amount of MG dropout is either calculated with software or visually graded by the physician, commonly placed in groups of thirds or quartiles representing the percentage of glandular loss. Determining the degree of MG dropout is of importance regarding disease staging, rate of progression and treatment alternatives. Some treatment options, such as intense pulsed light (IPL) treatment, aim to improve tear film stability by increasing the function of the MGs and quality of the secreted meibum. For such modalities to be effective, a minimum of residual glands are required. Moreover, for proper calibration and dosage the correct amount of MG dropout is often a necessary input.

The equipment necessary to perform meibography can be

expensive and unavailable to physicians. Moreover, the procedure may be difficult and unpleasant, and the interpretation of the images requires expertise. As a result, the current study examined whether machine learning (ML) algorithms can predict the degree of MG dropout based on questionnaires and routine dry eye clinical tests. Furthermore, we utilise explainable artificial intelligence (XAI) to investigate which clinical features are the most important predictors concerning MG dropout.

The main contributions of this work are the following:

- Evaluate the ability of ML algorithms to predict the degree of MG dropout based on tabular clinical data.
- Employ XAI to delineate the most important predictors.
- Compare results from different XAI techniques used on different classifiers of both tree-type architecture and neural networks.
- Discuss the clinical implications of the features recognized as the most important, both known associations and novel connections.

II. Related Work

In ophthalmology and DED, ML has mostly been used for image analysis [6]–[8]. However, our group recently used principal components analysis comparing salivary and ocular lipids in DED patients and healthy controls [9]. Also, we employed various ML algorithms on clinical tabular data in DED patients to predict instability of the tear film [10].

When using ML, especially in medicine, there is a need for explainability [11]. Lack of explainability in a black box diagnostic tool leads to several legal, ethical and medical challenges [12]. For example, how can a patient really give informed consent to a diagnostic procedure that can not be explained? Who is legally liable in the case of errors? From a medical perspective it is of utmost importance for the physician to understand why a given model makes the recommendation it does. If the the model output differs from the conclusion reached by the clinician, the physician should be able to verify the clinical validity of the features weighted by the model. From a medical research standpoint, XAI will help discover novel associations and potentially contribute to establishing causative factors. Approaches such as visualization through knowledge graphs, combining deep learning (DL) and ontology, integrating human-in-the-loop, transfer learning and multi-task learning have been pointed out as possibilities to increase human understanding of opaque ML algorithms [13].

Although neural networks are represented herein with the multilayer perceptron (MLP), our main focus is on tree-based models as they are computationally less expensive and still tend to outperform DL architectures on tabular data [14], [15].

As far as we know, this is the first time predictions on MG dropout through ML based on tabular data has been attempted. Moreover, this is the first time XAI has been used to explore the causes of MG atrophy.

Table I: Distribution of MG dropout.

Meiboscale	Number of observations (right eye / left eye)
0	2/2
1	44/33
2	215/196
3	183/173
4	129/169

III. Methodology

A. Participants

The dataset included clinical data from 582 patients examined in the Norwegian Dry Eye Clinic between September 2021 and December 2022. The inclusion criteria were age over 18 years, a diagnosis of DED and the ability to provide informed consent. No specific exclusion criteria were specified for the inclusion of clinical data in the present project.

All subjects signed written, informed consent. The study was approved by the Regional Medical Ethics Committee of South-East Norway (reference id 6892). The study adhered to the Declaration of Helsinki. The dataset is not publicly available since it contains patient-sensitive data but all code produced to conduct the experiments is made public available.

B. Ophthalmological examinations

All clinical examinations were performed by a single dry eye expert physician (FAF). Prior to the clinical examination, patients answered the Ocular Surface Disease Index (OSDI), 5-Item Dry Eye Questionnaire (DEQ-5) as well as questionnaires developed in our clinic regarding presence and duration of various symptoms, medications, diet and more. Following completion of the questionnaires, dry eye specific tests were performed in the following order: Schirmer test without anesthesia, tear film break-up time (TBUT) measured with fluorescein under slit lamp microscopy (average of three measurements used), ocular surface staining (OSS) (assessed according to the Oxford grading scheme [16]), number of expressible glands (counted among the central eight on the lower eyelids), meibum expressibility (ME) (grading the number of expressible glands among the central five glands of the lower eyelids), meibum quality (MQ), intraocular pressure (IOP), and finally meibography performed with the Keratograph 5M (Oculus Optikgeräte GmbH, Wetzlar, Germany). The degree of MG dropout was graded between 0 and 4 according to the meiboscale as presented by Pult and Riede-Pult in which grade 0 equals 0% dropout and grade 1 to 4 represent increasing quartiles [17].

C. Data preprocessing

For both the classification task and XAI the same dataset including tabular data from 582 DED patients examined at the Norwegian Dry Eye Clinic was analysed. The degree of MG dropout was missing for nine patients which were excluded. The final dataset included 573 patients. Details regarding distribution of MG dropout is presented in Table I.

Only two patients had MG dropout grade 0 and were included among patients with grade 1 since the models are unlikely to learn from such a low number of instances. For both problems, models predicting the degree of MG dropout on both lower eyelids were developed.

D. Classification

In total $9,150$ out of $53,544$ values were blank, which equals 17.1%. Missing data were handled with the *KNNImputer* with the number of neighbors set to three and assigned equal weighting. Because the dataset was comprised of both binary and numerical variables, *StandardScaler* was used to normalise the data by removing the mean and scale to unit variance. This way of scaling avoids inappropriately weighted importance of a given feature based on high numerical values alone. In this study, our primary focus was on the explainability aspect of the ML models rather than their predictive performance. We aimed to provide insights into the decision-making process of the models and understand how the input features contribute to the predictions. Our main objective was to enhance the interpretability and transparency of the models, which is essential for trust and adoption in real-world applications. Given this focus, 80% of the dataset was used for training and validation while 20% was held out as an independent test set, rather than employment of a non-overlapping dataset or cross-validation. As a baseline, a classifier that places all instances in the most frequent category was used. Also, a linear *support vector classifier (SVC)* was included as linear regression is commonly used for classification tasks in medicine.

Due to the imbalanced data and multioutput classification, model performance was evaluated using balanced accuracy, Matthews correlation coefficient (MCC) and the F1 score.

In addition to multioutput classification, binary classifiers for each grade of MG dropout were trained. That is, models were trained to predict whether a patient had MG dropout grade 1 or not, grade 2 or not, grade 3 or not, and finally, grade 4 or not.

E. Explainable artificial intelligence

Feature importance was examined for all the best performing classifiers, which included the *MLP*, *AdaBoostClassifier*, *HistGradientBoostingClassifier*, *XGBoost Classifier* and the *LightGBM Classifer*. For *MLP*, *AdaBoost* and *HistGradientBoosting* classifiers permutation importance was used. For the *XGBoost* and *LightGBM* classifiers the built-in feature importance evaluators were employed. Permutation feature importance disrupts the relationship between the feature and the true outcome by permuting, or rearranging, the values of a given feature [18]. Thus the importance of a feature is measured by calculating the change in the models prediction error following permutation of a feature. The split count importance feature ranking used by the *LightGBM* classifier bases a feature´s importance value on the number of times the feature is used to split the data. All experiments regarding classification and XAI were performed using *scikit-learn* in *Python* 3.9.15 on an ASUS ROG Zephyrus M16 with Intel Core i9-12900H CPU, NVIDIA GeForce RTX 3080 Ti GPU and 32 GB memory. Experiments for XAI were performed on the training data from the classification task. The source code is publicly available, including results from classification and feature importance analysis [1].

Table II: Demographics and clinical data.

Sex	Frequency			
Men	145			
Women	428			
Parameter	Min	Max	Mean	SD
Age	18	91	53.53	16.98
Schirmer test (mm/5min)				
Right eye	1	35	17.59	10.75
Left eye	1	35	15.93	9.99
TBUT (seconds)				
Right eye	1	11	4.40	3.02
Left eye	1	11	4.38	2.95
OSS				
Right eye	0	12	2.28	1.93
Left eye	0	12	2.35	1.86
Number of expressible glands				
Right eye	0	8	5.16	2.30
Left eye	0	8	4.68	2.31
ME				
Right eye	0	3	1.15	1.03
Left eye	0	3	1.38	1.06
MQ				
Right eye	0	20	4.88	3.57
Left eye	0	21	5.33	3.93
Osmolarity (mOsm/L)				
Right eye	275	350	304.39	19.74
Left eye	275	347	300.63	17.41
IOP				
Right eye	5	26	12.72	3.82
Left eye	4	35	12.52	3.98
OSDI	0	100	39.79	22.36
DEQ-5	0	22	13.02	4.30

DEQ-5: dry eye questionnaire-5; IOP: intraocular pressure; L: liter; Max: maximum; ME: meibum expressibility; min: minimum; mm: millimeter; mOsm: milliosmolar: MQ: meibum quality; OSDI: ocular surface disease index; OSS: ocular surface staining; SD: standard deviation; TBUT: tear film break-up time

IV. Results

The patients were from 18 to 91 years old, and the mean age was 53.53 years. There were 428 females and 145 males. Demographic and clinical data is presented in Table II.

A. Classifier Experiments

Classification with the baseline classifier on the right eye gave a balanced accuracy score of .25, F1 score of .26 and a

[1] https://github.com/freafin/MeibomianGlandClassification

Table III: Results for the right eye.

Classifier	Balanced accuracy	MCC	F1 score
Baseline	.25	.0	.26
Linear SVC	.36	.10	.40
XGBoost	.43	**.20**	**.47**
AdaBoost	.45	.15	.41
MLP	.46	.16	.44
HistGradientBoost	.49	.18	.45
LightGBM	**.50**	**.20**	**.47**

MCC: Matthews correlation coefficient; MLP: multilayer perceptron; SVC: support vector classifier. The best results are in bold.

Table IV: Results for the left eye.

Classifier	Balanced accuracy	MCC	F1 score
Baseline	.25	.0	.21
Linear SVC	.27	.0	.26
XGBoost	.33	.11	.37
AdaBoost	.34	**.13**	**.39**
MLP	.34	.10	.37
HistGradientBoost	.34	.09	.37
LightGBM	**.37**	.10	**.39**

MCC: Matthews correlation coefficient; MLP: multilayer perceptron; SVC: support vector classifier. The best results are in bold.

Table V: Balanced accuracy for binary classification of dropout grades in the right eye.

MG dropout grade	SVC	AdaBoost	LightGBM	MLP
1	.63	.73	.75	**.79**
2	.53	.51	.51	**.56**
3	.51	.52	.56	**.57**
4	.59	.62	.59	**.64**

MG: meibomian gland; MLP: multilayer perceptron; SVC: support vector classifier. The best results are in bold.

MCC of .0. With the linear *SVC*, balanced accuracy increased to .36, while the best performing classifiers, *XGBoost*, *AdaBoost*, *MLP*, *HistGradientBoost* and *LightGBM* reached .44, .45, .46, .49 and .50, respectively. Results for the right eye are presented in Table III and results for the left eye in Table IV.

For binary classification the baseline classifier had a balanced accuracy of .5 for all classes due to assigning all subjects to the majority class and there only being two possibilities. Results from the binary classifiers are presented in Table V.

B. Feature Importance Experiments

For the sake of brevity, only feature importances from the *MLP*, *AdaBoost* and *LightGBM* classifiers are included. The results from the remaining experiments are available online. First, feature importances according to the *MLP* and *AdaBoost* classifiers using permutation importance are presented followed by inspection of the in-built split count method used by the *LightGBM* classifier.

1) Multilayer Perceptron Permutation Importance: The 20 most important features for predicting MG dropout with the *MLP* are presented in Figure 2a. For the right eye, age was the most important predictor. Features from previously undergone treatments such as laser eye surgery, autoimmune diseases like rheumatoid arthritis, restrictive diets and clinical signs are all represented. Their clinical implications are discussed below. Some of the specific features differed between the eyes, however, the overall clinical motif remained similar.

2) AdaBoost Permutation Importance: The 20 features regarded as most important by the *AdaBoost* classifier are listed in Figure 2b. For the right eye, features such as age, number of expressible glands, osmolarity and tear film stability were of importance. Also, clinical signs such as telangiectasias and cicatricial disease are high ranking. Moreover, self-reported presence of hyperemia, itching, epiphora as well as symptoms in the morning and in the summer were seen as important.

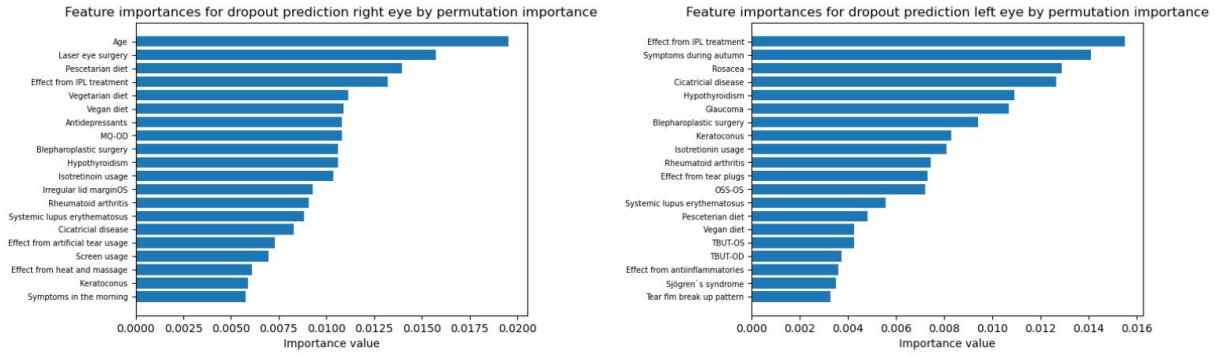
3) LightGBM Classifier Split Count Importance: The 20 features ranked as most important by *LightGBM* in predicting MG dropout are listed in Figure 2c. Unlike the features seen as important by the *MLP*, no previous treatments, systemic autoimmune diseases or dietary factors are included. Rather, clinical features including age, osmolarity, tear film stability, the number of expressible MGs, MQ, and OSDI were considered important. Clinical implications are discussed below.

V. Discussion

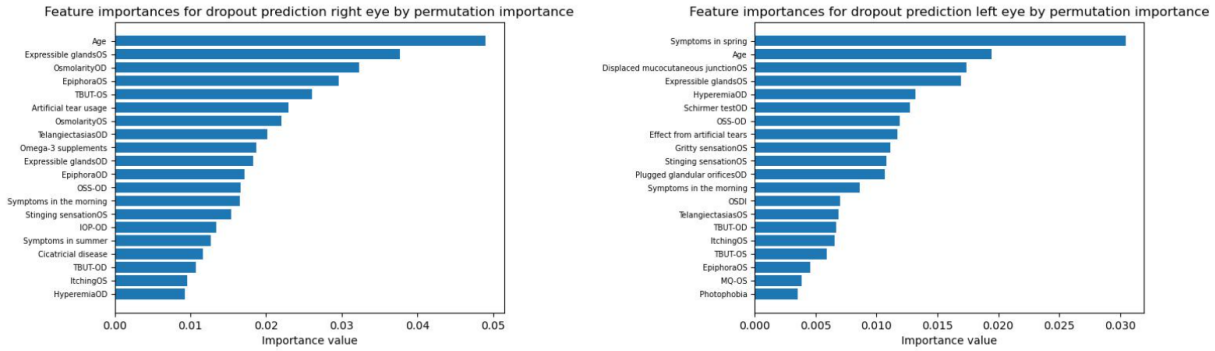
The main aim of the present project was two-fold. First, we wanted to determine whether ML algorithms could make accurate predictions of MG dropout based on clinical data and questionnaires. Second, we set out to illuminate which features were the most important predictors contributing to the loss of MGs.

Concerning multiclass classification, our models outperformed the baseline and linear *SVC* classifiers for both eyes. The F1 score increased from .26 at baseline to .47 with *LightGBM*, indicating superior model performance. However, the MCC remained relatively low, increasing from .0 to only .20 for the best performing algorithm. Since the MCC include both true and false positives and negatives, globally good results are necessary for a value closer to 1 which represents perfect predictions [19]. In addition to outperforming the baseline classifiers, the models presented an estimated probability that a given subject belongs to a given class. These probabilities were typically centred around the class considered most probable by the model in a decreasing fashion. Thus, although the models might not be absolutely certain, the clinician is presented with a probability overview, indicating to which class a patient belong. As a result, the physician will have an estimated probability of whether the patient has more than 75% MG dropout, which may be of specific importance when deciding treatment alternatives.

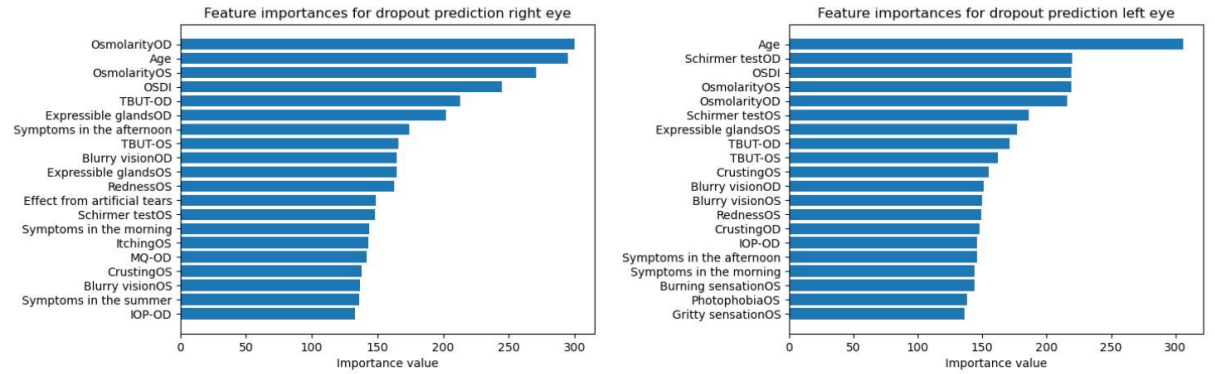
The binary classifiers outperformed baseline classification metrics with the *MLP* achieving the overall greatest results.



(a) Multilayer Perceptron



(b) AdaBoost Classifier



(c) LightGBM Classifier

Figure 2: Feature importance from the Multilayer perceptron, AdaBoost and LghtGBM Classifiers. Left: right eye. Right: left eye. Feature importance is arranged top to bottom.

When predicting in a binary fashion the data becomes highly skewed with proportionally few patients belonging to the positive class. A balanced accuracy of .5 represents predicting all subjects in the majority class. As can be seen from Table V and the confusion matrices presented in the source code, the classifiers managed to overcome the skewness of the data in varying degree depending on classifier and class to be predicted.

When looking at feature importance, age was one of the most important features of MG dropout according to all included algorithms, validating previous observations [5], [20].

For both *AdaBoost* and *LightGBM*, the number of expressible glands was of particular importance in predicting the amount of MG dropout. This is understandable because totally atrophied glands will not be expressible. However, this lack of expressibility might also be due to reversible obstruction of only partially atrophied glands by thickened, stagnated meibum. As such, the quality of the secreted meibum was considered among the most important predictors by both the *MLP* and *LightGBM*. These findings are very interesting as previous reports have questioned whether MG dropout might be preceded by increased intraluminal pressure resulting from meibum stagnation [21].

The fact that TBUT was one of the most important features for all three classifiers substantiates the role of the MGs and meibum in stabilising the tear film. This is in accordance with our previous findings that the degree of MG dropout is an important predictor of an unstable tear film [10]. Interestingly, lid margin abnormalities such as telangiectasias, cicatricial disease, an irregular lid margin and displacement of the mucocutaneous junction stand out as important features. The latter of these has previously been proposed as one of the diagnostic criteria for MGD [22]. The same article indicated the ocular symptom score as the most important diagnostic tool regarding MGD. In the present work both the OSDI and several elements from our in-house questionnaires were among the most important features. Self-reported epiphora, stinging and gritty sensations, itching, blurry vision, symptoms at various times of day and year are all included among the 20 most important features. Since these results stem from our in-house questionnaires they are all previously unconfirmed. Data gathering in the clinic is still ongoing and we aim to reassess the results on larger datasets when possible. Also, we want to evaluate whether any of these self-reported elements might predict effect (or lack thereof) to various treatments used in the clinic.

As the only classifier among the three presented here, the *MLP* included risk factors, such as blepharoplasty and laser eye surgery, medications, such as antidepressants and isotretinoin, as well as underlying autoinflammatory diseases like rheumatoid arthritis (RA), systemic lupus erythematosus (SLE) and hypothyroidism. An increased degree of MG dropout has been described among patients with SLE [23], [24], and both RA and SLE are associated with secondary Sjögren‘s syndrome [25]. Refractive surgery and cataract surgery might pose deleterious effects on the MGs [26], [27]. However, the impact of surgery and the remaining risk factors accentuated by the *MLP* on the MGs are all in need of further investigation.

Interestingly, several restrictive diets such as veganism, vegetarianism and pescetarianism were among the most important features for the *MLP* in predicting MG dropout. Studies have indicated an ameliorating effect from omega-3 and omega-6 fatty acids on DED [28]–[30], as well as from supplementation of fish oil [31], and a mediterranean diet [32], [33]. Moreover, a recent study indicated a possible protective effect from animal fats in regards to DED [34]. However, we do not claim that there is any causal relationship between these diets and the atrophy of MGs. These results may be because of selection bias as all included subjects are DED patients. Also, since patients with a greater degree of dropout may experience more symptoms, there is a possibility that they are more prone to try different diets for alleviation. The possible link between diet and MG dropout is of potential great consequence and in need of further research.

The model performance varied between the eyes with better performance on the right eye. Also, the feature ranking differed between the eyes. This can be explained by the higher number of eyes with dropout grade 4 among the left eyes. Also, for both models, measurements from the contralateral eye was important. This is due to both eyes being affected in a large number of patients.

VI. Conclusion

This study uses ML methods to classify DED patients according to the degree of MG atrophy. Moreover, we employ two XAI methodologies, namely permutation importance and split count importance. The predictive capabilities of the models presented herein cannot replace meibography in the clinic yet. However, we believe that the models are accurate enough to prove valuable for clinicians lacking the necessary equipment when considering treatment alternatives. Based on our findings with XAI in this study, age, the number of expressible MGs, tear film stability and symptoms were the most important features when predicting the degree of MG dropout. Furthermore, several interesting novel associations were presented by the *MLP*. This shows that XAI can be used to discover interesting and potentially novel medical knowledge.

For future work we plan to compare the predictive capabilities of the model in clinical practice. In addition we are also working on collecting additional meibographic images and combine these in a multimodal analysis.

Conflicting Interests: Fredrik A. Fineide is co-owner of the Norwegian Dry Eye Clinic. Tor P. Utheim is the founder and co-owner of the Norwegian Dry Eye Clinic.

References

[1] M. D. Willcox, P. Argüeso, G. A. Georgiev, J. M. Holopainen, G. W. Laurie, T. J. Millar, E. B. Papas, J. P. Rolland, T. A. Schmidt, and U. Stahl, “Tfos dews ii tear film report,” *The ocular surface*, vol. 15, no. 3, pp. 366–403, 2017.

[2] F. Stapleton, M. Alves, V. Y. Bunya, I. Jalbert, K. Lekhanont, F. Malet, K. S. Na, D. Schaumberg, M. Uchino, J. Vehof, E. Viso, S. Vitale, and L. Jones, "Tfos dews ii epidemiology report," *Ocul Surf*, vol. 15, no. 3, pp. 334–365, 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/28736337

[3] J. P. Craig, K. K. Nichols, E. K. Akpek, B. Caffery, H. S. Dua, C. K. Joo, Z. Liu, J. D. Nelson, J. J. Nichols, K. Tsubota, and F. Stapleton, "Tfos dews ii definition and classification report," *Ocul Surf*, vol. 15, no. 3, pp. 276–283, 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/28736335

[4] J. D. Nelson, J. Shimazaki, J. M. Benitez-del Castillo, J. P. Craig, J. P. McCulley, S. Den, and G. N. Foulks, "The international workshop on meibomian gland dysfunction: report of the definition and classification subcommittee," *Invest Ophthalmol Vis Sci*, vol. 52, no. 4, pp. 1930–7, 2011. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/21450914

[5] F. Fineide, R. Arita, and T. P. Utheim, "The role of meibography in ocular surface diagnostics: A review," *Ocul Surf*, 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/32416235

[6] A. M. Storas, I. Strumke, M. A. Riegler, J. Grauslund, H. L. Hammer, A. Yazidi, P. Halvorsen, K. G. Gundersen, T. P. Utheim, and C. J. Jackson, "Artificial intelligence in dry eye disease," *Ocul Surf*, vol. 23, pp. 74–86, 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/34843999

[7] R. K. Saha, A. M. Chowdhury, K.-S. Na, G. D. Hwang, Y. Eom, J. Kim, H.-G. Jeon, H. S. Hwang, and E. Chung, "Automated quantification of meibomian gland dropout in infrared meibography using deep learning," *The Ocular Surface*, vol. 26, pp. 283–294, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1542012422000519

[8] Z. Zhang, X. Lin, X. Yu, Y. Fu, X. Chen, W. Yang, and Q. Dai, "Meibomian gland density: An effective evaluation index of meibomian gland dysfunction based on deep learning and transfer learning," *Journal of Clinical Medicine*, vol. 11, no. 9, 2022. [Online]. Available: https://www.mdpi.com/2077-0383/11/9/2396

[9] F. Fineide, X. Chen, T. Bjellaas, V. Vitelli, T. P. Utheim, J. L. Jensen, and H. K. Galtung, "Characterization of lipids in saliva, tears and minor salivary glands of sjogren's syndrome patients using an hplc/ms-based approach," *Int J Mol Sci*, vol. 22, no. 16, 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/34445702

[10] F. Fineide, A. M. Storas, X. Chen, M. S. Magno, A. Yazidi, M. A. Riegler, and T. P. Utheim, "Predicting an unstable tear film through artificial intelligence," *Sci Rep*, vol. 12, no. 1, p. 21416, 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/36496510

[11] S. Reddy, "Explainability and artificial intelligence in medicine," *The Lancet Digital Health*, vol. 4, pp. e214–e215, 04 2022.

[12] J. Amann, A. Blasimme, E. Vayena, D. Frey, and V. Madai, "Explainability for artificial intelligence in healthcare: a multidisciplinary perspective," *BMC Medical Informatics and Decision Making*, vol. 20, 11 2020.

[13] A. Holzinger, "From machine learning to explainable ai," in *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA)*, 2018, pp. 55–66.

[14] L. Grinsztajn, E. Oyallon, and G. Varoquaux, "Why do tree-based models still outperform deep learning on typical tabular data?" in *Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2022. [Online]. Available: https://openreview.net/forum?id=Fp7__phQszn

[15] Y. Gorishniy, I. Rubachev, V. Khrulkov, and A. Babenko, "Revisiting deep learning models for tabular data," *CoRR*, vol. abs/2106.11959, 2021. [Online]. Available: https://arxiv.org/abs/2106.11959

[16] A. J. Bron, V. E. Evans, and J. A. Smith, "Grading of corneal and conjunctival staining in the context of other dry eye tests," *Cornea*, vol. 22, no. 7, pp. 640–50, 2003. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/14508260

[17] H. Pult and B. Riede-Pult, "Comparison of subjective grading and objective assessment in meibography," *Cont Lens Anterior Eye*, vol. 36, no. 1, pp. 22–7, 2013. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/23108007

[18] C. Molnar, *Interpretable Machine Learning*, 2nd ed., 2022. [Online]. Available: https://christophm.github.io/interpretable-ml-book

[19] S. A. Hicks, I. Strumke, V. Thambawita, M. Hammou, M. A. Riegler, P. Halvorsen, and S. Parasa, "On evaluation metrics for medical applications of artificial intelligence," *Sci Rep*, vol. 12, no. 1, p. 5979, 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/35395867

[20] R. Arita, K. Itoh, K. Inoue, and S. Amano, "Noncontact infrared meibography to document age-related changes of the meibomian glands in a normal population," *Ophthalmology*, vol. 115, no. 5, pp. 911–5, 2008. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/18452765

[21] A. Tomlinson, A. J. Bron, D. R. Korb, S. Amano, J. R. Paugh, E. I. Pearce, R. Yee, N. Yokoi, R. Arita, and M. Dogru, "The international workshop on meibomian gland dysfunction: Report of the diagnosis subcommittee," *Investigative Ophthalmology and Visual Science*, vol. 52, no. 4, pp. 2006–2049, 03 2011.

[22] R. Arita, K. Itoh, S. Maeda, K. Maeda, A. Furuta, S. Fukuoka, A. Tomidokoro, and S. Amano, "Proposed diagnostic criteria for obstructive meibomian gland dysfunction," *Ophthalmology*, vol. 116, no. 11, pp. 2058–63 e1, 2009. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/19744718

[23] Z. Gu, Q. Lu, A. Zhang, Z. W. Shuai, and R. Liao, "Analysis of ocular surface characteristics and incidence of dry eye disease in systemic lupus erythematosus patients without secondary sjögren's syndrome," *Frontiers in Medicine*, vol. 9, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fmed.2022.833995

[24] Y. Li, S. Ou, S. Lin, H. Qian, Z. Zhao, M. Zhang, S. Li, Y. Liu, and G. Shi, "Meibomian gland alteration in patients with systemic lupus erythematosus," *Lupus*, vol. 31, no. 4, pp. 407–414, 2022, pMID: 35246003. [Online]. Available: https://doi.org/10.1177/09612033221079760

[25] A. J. Bron, C. S. de Paiva, S. K. Chauhan, S. Bonini, E. E. Gabison, S. Jain, E. Knop, M. Markoulli, Y. Ogawa, V. Perez, Y. Uchino, N. Yokoi, D. Zoukhri, and D. A. Sullivan, "Tfos dews ii pathophysiology report," *The Ocular Surface*, vol. 15, no. 3, pp. 438–510, 2017, tFOS International Dry Eye WorkShop (DEWS II). [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1542012417301349

[26] J. W. Jung, J. Y. Kim, H. S. Chin, Y. J. Suh, T.-i. Kim, and K. Y. Seo, "Assessment of meibomian glands and tear film in post-refractive surgery patients," *Clinical & Experimental Ophthalmology*, vol. 45, no. 9, pp. 857–866, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/ceo.12993

[27] A. El Ameen, S. Majzoub, G. Vandermeer, and P.-J. Pisella, "Influence of cataract surgery on meibomian gland dysfunction," *Journal Français d'Ophtalmologie*, vol. 41, no. 5, pp. e173–e180, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0181551218301220

[28] A. Ng, J. Woods, T. Jahn, L. W. Jones, and J. Sullivan Ritter, "Effect of a novel omega-3 and omega-6 fatty acid supplement on dry eye disease: A 3-month randomized controlled trial," *Optom Vis Sci*, vol. 99, no. 1, pp. 67–75, 2022.

[29] I. Molina-Leyva, A. Molina-Leyva, and A. Bueno-Cavanillas, "Efficacy of nutritional supplementation with omega-3 and omega-6 fatty acids in dry eye syndrome: a systematic review of randomized clinical trials," *Acta Ophthalmologica*, vol. 95, no. 8, pp. e677–e685, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/aos.13428

[30] M. Pellegrini, C. Senni, F. Bernabei, A. F. G. Cicero, A. Vagge, A. Maestri, V. Scorcia, and G. Giannaccare, "The role of nutrition and nutritional supplements in ocular surface diseases," *Nutrients*, vol. 12, no. 4, 2020. [Online]. Available: https://www.mdpi.com/2072-6643/12/4/952

[31] T. KAWAKITA, F. KAWABATA, T. TSUJI, M. KAWASHIMA, S. SHIMMURA, and K. TSUBOTA, "Effects of dietary supplementation with fish oil on dry eye syndrome subjects: randomized controlled trial," *Biomedical Research*, vol. 34, no. 5, pp. 215–220, 2013.

[32] I. Molina-Leyva, A. Molina-Leyva, B. Riquelme-Gallego, N. Cano-Ibáñez, L. García-Molina, and A. Bueno-Cavanillas, "Effectiveness of mediterranean diet implementation in dry eye parameters: A study of predimed-plus trial," *Nutrients*, vol. 12, no. 5, 2020. [Online]. Available: https://www.mdpi.com/2072-6643/12/5/1289

[33] A. Galor, H. Gardener, B. Pouyeh, W. Feuer, and H. Florez, "Effect of a mediterranean dietary pattern and vitamin d levels on dry eye syndrome," *Cornea*, vol. 33, no. 5, pp. 437–441, 2014. [Online]. Available: https://journals.lww.com/corneajrnl/Fulltext/2014/05000/Effect_of_a_Mediterranean_Dietary_Pattern_and.1.aspx

[34] S. Fukuoka, R. Arita, T. Mizoguchi, M. Kawashima, S. Koh, R. Shirakawa, T. Suzuki, S. Sasaki, and N. Morishige, "Relation of dietary fatty acids and vitamin d to the prevalence of meibomian

gland dysfunction in japanese adults: The hirado–takushima study," *Journal of Clinical Medicine*, vol. 10, no. 2, 2021. [Online]. Available: https://www.mdpi.com/2077-0383/10/2/350