

Glaucoma Precognition: Recognizing Preclinical Visual Functional Signs of Glaucoma

Krati Gupta^{1,3}, Anshul Thakur¹, Michael Goldbaum², and Siamak Yousefi^{3,4}

¹School of Computing & Electrical Engineering, Indian Institute of Technology Mandi, Mandi, India

²Department of Ophthalmology, University of California, San Diego, USA

³Department of Ophthalmology, University of Tennessee Health Science Center, Memphis, USA

⁴Department of Genetics, Genomics, and Informatics, University of Tennessee Health Science Center, Memphis, USA

*krati.gupta89@gmail.com, anshul.iitmandi@gmail.com, mgoldbaum@ucsd.edu,
siamak.yousefi@uthsc.edu*

Abstract

Deep archetypal analysis (DAA) has recently been proposed as an unsupervised approach for discovering latent structures in data. However, while a few approaches have used classical archetypal analysis (AA), DAA has not been incorporated in medical image analysis as yet. The purpose of this study is to develop a precognition framework to identify preclinical signs of glaucomatous vision loss using convex representations derived from DAA. We first develop an AA structure and a novel DAA framework to recognize hidden patterns of visual functional loss, and then project visual field data over the identified patterns to obtain a representation for glaucoma precognition several years prior to disease onset. We then develop a glaucoma classification framework using class-balanced bagging with neural networks to address the class imbalance problem. In contrast to other classification approaches, DAA, applied to a unique prospective longitudinal dataset with approximately eight years of visual field tests from normal eyes that developed glaucoma, has allowed visualization of the early signs of glaucoma and development of a construct for glaucoma precognition. Our findings suggest that our proposed glaucoma precognition approach could significantly advance state-of-the-art glaucoma prediction.

1. Introduction

Dictionary learning approaches can discover latent structures in data in an unsupervised manner. Archetypal analysis (AA) and its several variations have been long

used in different applications including dictionary learning [2, 3, 10, 25]. Recently, deep archetypal analysis (DAA) has been proposed to address several limitations of the AA [16, 26, 27]. Deep AA could learn relevant transformations and incorporate appropriate information into the learning process thus providing an effective representation in most applications. Such models may aid uncovering hidden visual functional patterns of vision loss that may lead to glaucoma.

Glaucoma is the second leading causes of blindness worldwide [21]. Risk factors for glaucoma include advanced age, African American ethnicity, elevated intraocular pressure (IOP), and thinner central corneal thickness [13, 22]. A key issue is that glaucoma is typically asymptomatic, particularly at the early stages of the disease, thus most subjects with glaucoma are often unaware of the disease until visual functional loss becomes significant [29]. These issues underlie the challenges inherent in attempts to forecast glaucoma and likewise, highlight the potential clinical importance and economic impact of developing new methods for early prediction of glaucoma.

Currently, glaucoma-induced visual field loss is mainly assessed using well-established standard automated perimetry (SAP) [15]. The Humphrey 30-2 (30° field with 2 dB resolution in brightness) testing system generates a map of 74 (after excluding two test locations corresponding to blind spot) local retinal sensitivities to the light. A visual field map is typically used by clinicians to subjectively determine the severity of glaucoma-induced functional loss and thus, is accepted as an important component of glaucoma assessment. Nonetheless, this process is highly subjective and prone to inter- and intra-observer variability. Another chal-

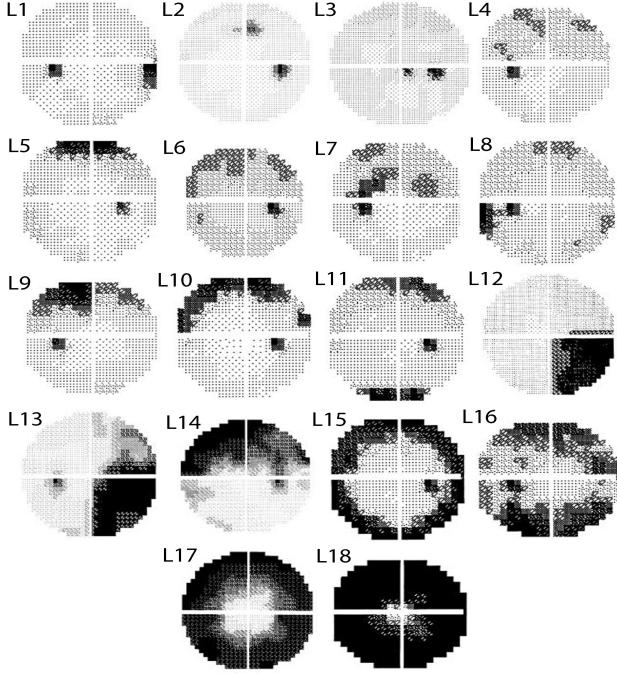


Figure 1. Three certified visual field readers identified 18 patterns of visual field loss (in total deviation format) from the OHTS participants who developed glaucoma in 2003 [17].

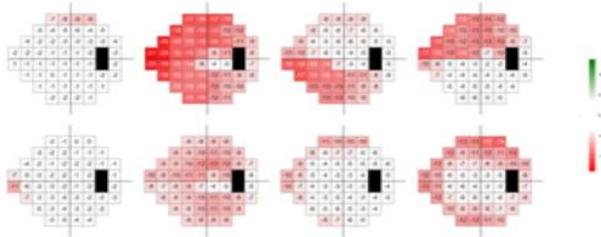


Figure 2. Sample patterns of visual field loss identified by GMM-EM [30] (visualized in total deviation format).

lence is that visual fields are highly variable (noisy) particularly as the visual field of the patient deteriorates [14, 20].

To the best of our knowledge, there has been no published study that has attempted to visualize preclinical signs of visual functional loss in patients with glaucoma. For our study, we first developed a DAA framework to generate a convex representation of visual fields and then identified hidden patterns of visual field loss. We visualized DAA-identified patterns of visual field loss. Our results suggest that the identified patterns of visual field loss we identified, represent preclinical signs of glaucoma that may be unknown to glaucoma specialists. We then developed a glaucoma precognition construct using DAA framework.

Several studies have attempted to identify patterns of visual field loss in patients with existing glaucoma. For instance, in 2003, three visual field readers and glaucoma ex-

perts of a very well-known glaucoma clinical trial, called the ocular hypertension treatment study (OHTS), manually identified 18 prevalent patterns of visual field loss in patients who had developed glaucoma [17] (Fig. 1). Others have used Gaussian Mixture Model Expectation Maximization (GMM-EM) to automatically identify patterns of visual field loss in patients with existing glaucoma and have used these patterns for further glaucoma monitoring [30] (Fig. 2). Other studies have used classical AA to identify patterns of visual field loss of patients with glaucoma [11] and then used those patterns to detect glaucoma progression[28]. However, the OHTS study was manual and subjective thus prone to human observer selection bias. There is no report of pattern assessment methods that rely on GMM-EM or classical AA to be applied on visual fields of suspected glaucoma (elevated IOP but normal visual field based on current clinical guidelines). To date, pattern assessment has only been applied to visual fields of normal or confirmed glaucoma patients. While some previous studies have used patterns of visual field loss for detecting progression in glaucoma [17, 23, 30], to our knowledge, our studies are the first to propose a glaucoma precognition construct based on unsupervised DAA and supervised neural network.

The main contributions of this paper are: (1) We recognize preclinical signs of glaucoma using DAA and visualize those patterns. (2) We develop a class-balanced neural network to address imbalanced samples in positive and negative groups. (3) We propose incorporating DAA in class-balanced neural network to develop a construct for glaucoma precognition.

2. Medical Background

Several deep learning approaches have been successfully applied in ophthalmology including glaucoma [1, 4, 8, 18]. However, most of deep learning models in glaucoma have been centered on diagnosis. Since glaucoma diagnosis requires only cross-sectional data, it is easier to access. Moreover, models typically perform better for diagnosis because disease signs are already present and identifiable by human experts; thus, making the signs easier to identify by machine. However, precognition of the disease from baseline parameters is more challenging. Access to prospective and longitudinal data prior and after disease onset is not trivial and identifying preclinical signs of the disease, which are hidden to human expert, is significantly more involved. Hence, there are no reports of deep learning models that can forecast glaucoma.

However, a few studies have attempted to predict glaucoma prior to disease onset using conventional statistical or classical machine learning approaches [5, 22, 24]. Such studies generally rely on conventional statistical analysis, and typically make strong assumptions on conditions, in or-

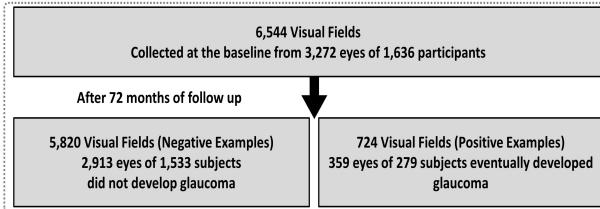


Figure 3. Visual fields and their labels. Out of 6,544 visual fields that were collected from the baseline visit of the ocular hypertension treatment study (OHTS) participants, 5,820 corresponded to eyes that eventually developed glaucoma after approximately six years and 724 visual fields corresponded to eyes that remained normal. The hypothesis was that the proposed precognition construct can identify hidden patterns of visual functional loss in eyes that eventually developed glaucoma.

der to determine risk factors that may lead to disease onset. For instance, a conventional cox hazard model was applied to structural features such as optic nerve head topography and retinal nerve fiber layer thickness to predict glaucoma [24]. Another study used multivariate cox hazard models to identify structural and functional risk factors for glaucoma [10]. However, the identified risk factors through statistical analysis were found to be too imprecise for prediction of glaucoma in advance of disease onset. To address this challenge, a classical machine learning model using relevance vector machines (RVM) was developed to predict glaucoma from a set of structural and functional features [5]. However, in this study, only raw visual fields were used as input features to the RVM classifier, while more clinically relevant representations could have improved recognizing subtle (preclinical) signs of the disease. Moreover, the sample sizes of all the aforementioned studies are relatively small, limiting the ability to generalize findings. In this study we use a large-scale prospective dataset and apply DAA to obtain clinically relevant patterns of visual field loss. We show that these patterns can serve as preclinical signs of glaucoma, which are specific and sensitive in forecasting glaucoma onset.

3. Dataset

The Ocular Hypertension Treatment Study (OHTS) was conducted across 22 centers in the US. The study investigated the role of lowering IOP in preventing or delaying the onset of visual field loss in patients at moderate risk of developing glaucoma in the future [13]. The OHTS was a retrospective study in which all risk factors and data were collected at the baseline (when all subjects were normal, based on clinical guidelines) and afterwards for approximately eight years routinely. Therefore, the OHTS dataset allows testing hypotheses on factors, and hidden patterns of visual field loss, that may lead to glaucoma. Our study was conducted according to the tenets of Helsinki Declaration

and we received relevant institutional review board (IRB) and appropriately signed data use agreements.

For each subject, two or three visual field tests (collected by Humphrey; Carl Zeiss Meditec, Dublin, California) full threshold (SITA Standard; 30-2 procedure). A total of 7,248 visual fields collected from the baseline visit of 3,272 eyes (1,636 subjects) with elevated IOP but normal appearing optic disc and normal visual field at the baseline (when participants entered the study). Visual field and optic disc evaluations are typically performed by clinicians for glaucoma assessment. Visual field and clinical parameters were collected twice annually for over six years. Eventually, 359 eyes from 279 participants developed glaucoma based on either visual field or optic disc assessments (Fig. 3). More specifically, two OHTS certified readers carefully had examined follow up visual fields and when they had identified obvious visual field abnormality, they had recalled subject for re-testing to confirm abnormality, which was further confirmed by an independent endpoint committee [13].

We labeled 5,820 visual fields that corresponded to eyes that did not develop glaucoma as negative examples and 724 visual fields that corresponded to eyes that eventually developed glaucoma as positive examples (Fig. 3). We then hypothesized that there may be hidden visual functional defect patterns in the visual fields of those eyes that eventually developed glaucoma that either were missed by, or unknown to clinicians. Our aim was thus 1) to identify and visualize those subtle visual field defect patterns and 2) develop a construct to forecast glaucoma from visual fields several years prior to disease onset.

4. Deep Archetypal Analysis (DAA)

Archetypal analysis and DAA were introduced for discovering latent factors from high-dimensional data by performing matrix factorization. Archetypal analysis [10] is a matrix factorization where a matrix, $\mathbf{X}(\mathbb{R}^{d \times n})$, whose columns represent d -dimensional data points, is decomposed as $\mathbf{X} = \mathbf{DA}$. Matrix $\mathbf{D}(\mathbb{R}^{d \times k})$ contains k archetypes lying on convex hull (external points) of the data, and $\mathbf{A}(\mathbb{R}^{k \times n})$ is a convex representation matrix, which implies that data points can be represented as a convex combination of archetypes, and archetypes can also be represented as a convex combination of the individual data points, that is $\mathbf{D} = \mathbf{XB}$, where $\mathbf{B}(\mathbb{R}^{n \times k})$ is a convex representation matrix. Archetypes present a convenient way for capturing extremal properties of the input data points. Incorporating appropriate optimization frameworks, one could identify archetypes \mathbf{D} from the input data points \mathbf{X} [9]:

$$\begin{aligned} \underset{\mathbf{B}, \mathbf{A}}{\operatorname{argmin}} \|\mathbf{X} - \mathbf{DA}\|_F^2 &= \underset{\mathbf{B}, \mathbf{A}}{\operatorname{argmin}} \|\mathbf{X} - \mathbf{XBA}\|_F^2 \\ \mathbf{b}_j \in \Delta_n, \mathbf{a}_i \in \Delta_k &\quad b_j \in \Delta_n, \mathbf{a}_i \in \Delta_k \\ \Delta_n \triangleq [\mathbf{b}_j \geq 0, \|\mathbf{b}_j\|_1 = 1], \quad \Delta_k \triangleq [\mathbf{a}_i \geq 0, \|\mathbf{a}_i\|_1 = 1] \end{aligned} \quad (1)$$

where \mathbf{a}_i and \mathbf{b}_j represent columns of \mathbf{A} ($\mathbb{R}^{k \times n}$) and \mathbf{B} ($\mathbb{R}^{n \times k}$), respectively. The block-coordinate descent method [9] can be employed to solve this non-convex optimization problem. Although AA may effectively model the convex hull of data, AA is prone to outlier samples (as an external point). It also limited in modeling either the average or local characteristics of data. The motivation of deep archetypal analysis was to address these limitations [16, 26, 27].

Deep AA basically performs multiple AA-based factorizations on the input matrix and its subsequent factors. At the first layer of DAA, the input matrix \mathbf{X} is decomposed into an archetypal dictionary \mathbf{D}_1 and convex-sparse representation matrix \mathbf{A}_1 similar to AA process using equation 1. \mathbf{A}_1 is then serve as input to the second layer and is again factorized using AA leading to dictionary \mathbf{D}_2 and convex-sparse representations \mathbf{A}_2 . Thus, $\mathbf{X} \approx \mathbf{D}_1 \mathbf{A}_1 \approx \mathbf{D}_1 \mathbf{D}_2 \mathbf{A}_2 = \mathbf{D}_{L2} \mathbf{A}_2$, where \mathbf{D}_{L2} represents DAA dictionary obtained at the second layer of DAA framework. A user-defined depth of factorization can stop this process. Therefore, DAA decomposes \mathbf{X} into $L + 1$ factors with L representing the number of layers: $X \approx \mathbf{D}_1 \mathbf{D}_2 \mathbf{D}_3 \dots \mathbf{D}_L \mathbf{A}_L$. The factorization at different layers of DAA framework can be represented as:

$$\begin{aligned}\mathbf{X} &\approx \mathbf{D}_1 \mathbf{A}_1 = \mathbf{X} \mathbf{B}_1 \mathbf{A}_1 \\ \mathbf{A}_1 &\approx= \mathbf{A}_1 \mathbf{B}_2 \mathbf{A}_2 \\ \mathbf{A}_2 &\approx= \mathbf{A}_2 \mathbf{B}_3 \mathbf{A}_3 \\ &\vdots \\ \mathbf{A}_{L-1} &\approx= \mathbf{A}_{L-1} \mathbf{B}_L \mathbf{A}_L \\ \mathbf{A}_L &\approx= \mathbf{A}_L \mathbf{B}_{L+1} \mathbf{A}_{L+1}\end{aligned}$$

Deep archetypes model both local (by archetypes that lie on extremal points) and global (by archetypes lie on data average) characteristics of the data. Dictionaries obtained at deeper layers ($L > 1$) are convex combinations of the archetypes obtained at the first layer [16]. Therefore, deeper dictionaries atoms can lie on the boundary as well as inside the convex hull. Therefore, DAA systematically captures both local and global characteristics of data.

DAA is appropriate for visual field data analysis because of two major reasons: 1) most of the clinically known glaucomatous patterns of visual field loss are local and thus lie on or near the boundary of the visual field data in the initial 76-d space. The convex hull modelling properties of DAA can identify these local patterns, and hence, provides a convex representation that is consistent with glaucoma clinical knowledge, and 2) unlike many other dictionary learning models, such as principal or independent component analysis or singular value decomposition, DAA does not project the data to any latent space and most archetypes are, in fact, data points. Therefore, convex representations obtained by

DAA are interpretable and clinically explainable. We use this property to visualize patterns of visual field loss.

5. Proposed Framework

In this section, we will discuss the proposed glaucoma preognition framework. The framework has two components: feature extraction using DAA and classification using class-balanced neural network as follows:

5.1. Feature Extraction Using DAA

We will compute the DAA dictionary (\mathbf{D}_{Li}), where L_i represents the dictionary obtained at i th layer of the DAA framework. Each visual field example is represented as \mathbf{x} . Essentially, each atom of this dictionary represents a vertex of a high-dimensional simplex. We will then project visual fields on this simplex to obtain convex representations as:

$$\underset{y \in \Delta_k}{\operatorname{argmin}} \|\mathbf{x} - \mathbf{D}_{Li} y\|_F^2 \quad (2)$$

such that $\Delta_k \triangleq [a_i \geq 0, \|a_i\|_1 = 1]$. Here y represents corresponding convex representation of each visual field \mathbf{x} and k indicates the number of atoms in \mathbf{D}_{Li} . These convex representations are sparse, and we will show that how convex representations highlight the preclinical signs of visual functional loss that may not be captured by clinincals or clinical instruments. This is an unsupervised procedure since no class-specific information was used to obtain DAA dictionary through simplex projection.

5.2. Class-balanced Bagging for Classification

Similar to the general scenario of datasets in the real-world healthcare settings, the dataset used in this study was also imbalanced, due to the less number of eyes, that eventually developed glaucoma than the number of eyes that did not develop the disease throughout the course of OHTS. It is reported that class imbalance could generate bias towards the class with greater number of samples in strong classifiers such as support vector machines (SVM) and neural networks. We thus propose to use a bagging-based approach where as each individual classifier, a feed-forward neural network, was fed with class-balanced training examples, as illustrated in Fig. 4.

In the training step, this approach divides the samples of the negative class (those eyes that did not develop glaucoma) into smaller non-overlapping subsets such that the number of samples in each subset was almost equal to the number of positive examples (eyes that eventually developed glaucoma). The samples of each subset are then used as input to a neural network (multi-layer perceptron; all with similar parameters) to learn the classification rule. This framework is different from traditional bagging approach where each sample has the same likelihood of being selected for training in any of the classifiers. During testing,

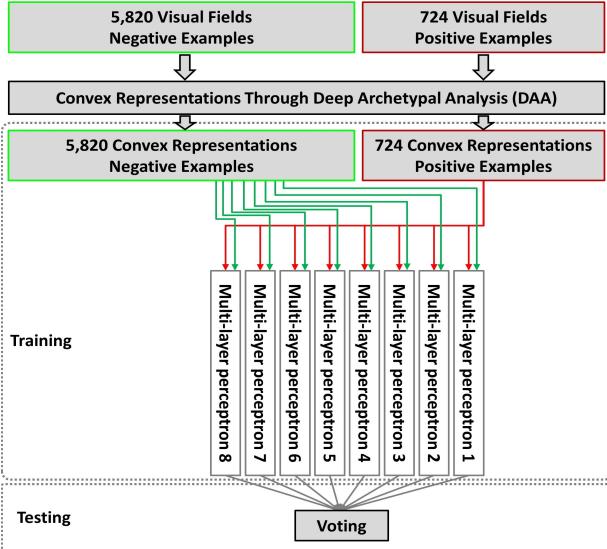


Figure 4. Class-balanced bagging approach. Examples in negative group are randomly divided to eight subsets and then each subset along with positive examples are used to train eight classifiers.

each neural network is considered as an independent classifier, and final prediction is acquired via a majority voting rule, applied over individual predictions.

5.3. Training and Testing Classifiers and Comparison

We use 10-fold stratified cross-validation and area under the receiver operating characteristics (AUC) for comparing neural networks applied to raw visual fields, convex representation through AA, convex representation through DAA, relevance vector machine (RVM) applied on raw visual fields, and major glaucoma risk factors including IOP, CCT, and age. As discussed in Section 2, the only machine learning based method for glaucoma prediction, known to the authors, was RVM [5], which was compared against the proposed frameworks.

5.4. Parameter Setting and Performance Metric

All the parameters such as number of dictionary atoms (archetypes), the number of layers in DAA, the number of nodes and layers in neural network were selected such that the model provides an optimal performance on the cross-validation data. More specifically, we selected these parameters based on an extensive grid search to provide maximum AUC and least missed detection rate. Each neural network classifier included a single hidden layer composed of 200 neurons. We used the Adam optimizer with a fixed learning rate of 0.0001 for training each neural network. For class-balanced bagging, the negative class was divided into eight subsets, and hence, the proposed framework was an ensemble of eight different neural networks. Gaussian kernel with

a width of 0.9 was used in the baseline method for training the RVM. Similar to DAA framework, we selected the RVM parameters using a grid-search on the cross-validation data to maximize AUC. We selected the same configuration for neural networks in all experiments. The DAA, AA, multi-layer perceptron, and AUC performance metrics were implemented in Python using scikit-learn library, while RVM was implemented in Matlab. All statistical analyses were performed in R. We used the implementation of Chen et al. for AA [9] and used the implementations in [16, 26, 27] for DAA analysis.

5.5. Visualizing Preclinical Signs of Glaucoma

To visualize patterns of visual field loss obtained through convex representation of DAA, we apply the identified DAA coefficients on visual fields and subjectively evaluated the identified patterns and excluded archetypes that have a significant correlation with other archetypes. This process is performed under the supervision of a glaucoma expert. The selected patterns are visualized as preclinical signs of glaucoma.

6. Results and Discussion

A total of 6,544 visual fields at the baseline visit of each eye were reliable and normal (according to clinical guidelines), of which 724 visual fields labeled as positive and 5,820 visual fields labeled as negative (Fig. 3). The mean age (standard deviation; SD) of subjects in the negative and positive groups were 55.7 (9.6) and 58.8 (9.0) years, respectively (P value < 0.001; based on generalized estimating equation; GEE). Approximately 42% of subjects in the negative group were male while 56% of the subjects in the positive group were male (P value < 0.001) indicating more males compare to females developed glaucoma. Mean IOP of eyes in the negative and positive groups were 24.8 mmHg (2.9) and 26.1 mmHg (3.3), respectively (P value < 0.001). Mean CCT of eyes in the negative and positive groups were 574.7 mmHg (38.3) and 558.7 mmHg (39.0), respectively (P value < 0.001). Older age, elevated IOP, and thinner CCT are glaucoma risk factors. We identified that subjects corresponding to positive samples (those that eventually developed glaucoma) had older age, higher IOP, and thinner CCT, all statistically significantly different from negative group. However, none of these risk factors performed better than 0.53 (in terms of AUC) in predicting glaucoma, which is analyzed by evaluating the performance of few significant clinical parameters of positive and negative groups.

Visual field testing through standard automated perimetry (SAP) remains a gold standard for glaucoma assessment. Patterns of visual field loss play a major role in glaucoma diagnosis, severity identification, and therapy adjustments based on the type of visual functional defect [7]. However, manual classification of glaucoma through visual fields is

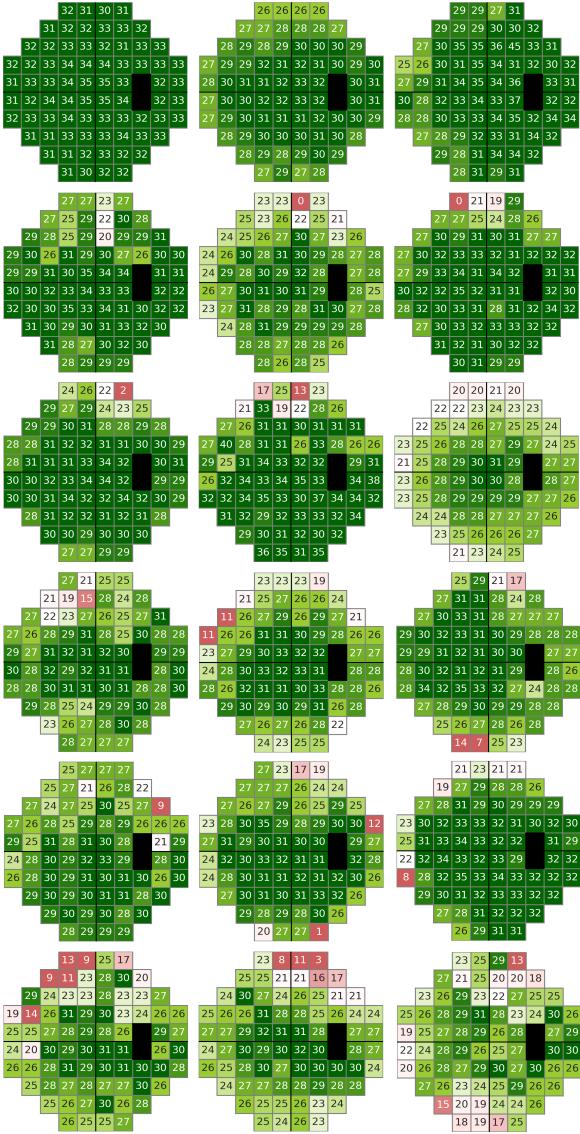


Figure 5. Preclinical visual functional signs of glaucoma identified by archetypal analysis (AA) of visual fields that were collected from the Ocular Hypertension Treatment Study (OHTS) participants at the baseline visit.

labor intensive and requires significant clinical training yet is highly subjective with limited agreement even among glaucoma specialists [12, 19]. Thus, automatically identifying (early) patterns of visual field loss can impact glaucoma management. Several researchers, us included, have used unsupervised learning to discover (glaucomatous) patterns of visual field loss [6, 11, 28, 30–32]. We have extensively used Gaussian mixture modeling (GMM) to discover patterns of visual field loss and to identify glaucoma progression along those GMM-identified patterns [6, 30–32]. Other teams have used classical AA for such goals [11, 28].

Fig. 5 shows 18 patterns of visual field loss identified by

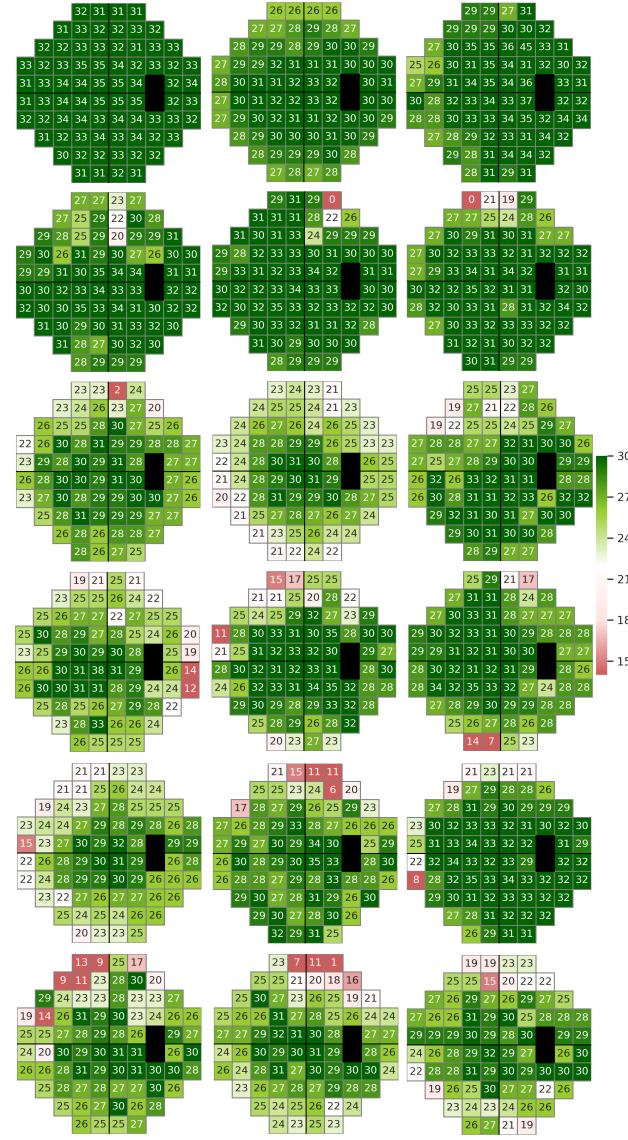


Figure 6. Preclinical visual functional signs of glaucoma identified by deep archetypal analysis (DAA) of visual fields that were collected from the Ocular Hypertension Treatment Study (OHTS) participants at the baseline visit.

applying classical archetypal analysis (AA) on OHTS visual fields. These patterns were evaluated objectively (using correlation) and subjectively by a glaucoma expert to identify the smallest subset of patterns that are clinically relevant patterns. We identified 18 patterns; the top-left pattern was identified as normal while other patterns were preclinical signs of glaucoma. Fig. 6 represents 18 patterns of visual field loss identified by deep archetypal analysis (DAA). These patterns were also evaluated subjectively by a glaucoma expert. It is worth mentioning that manual assessment of visual fields by three glaucoma experts also identified 18 prevalent yet mutually exclusive patterns of visual fields

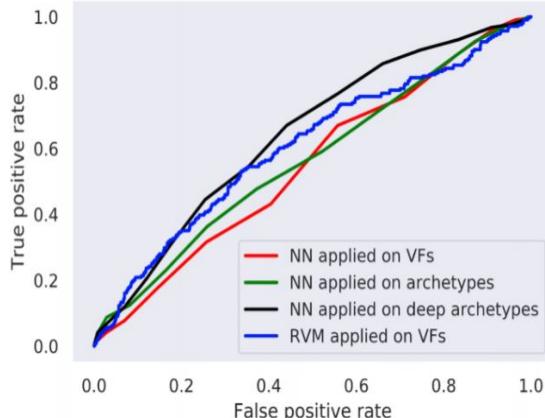


Figure 7. ROC curves of neural networks (NN) applied on original visual fields (VFs), convex representations of visual fields obtained by AA and DAA, and relevance vector machine (RVM).

loss in patients with glaucoma [17]. Since glaucoma experts had not identified any suspicious glaucomatous patterns of visual field loss in their subjective evaluation of the baseline visual fields (recall that all 6,544 visual fields at the baseline were normal according to clinical guidelines; Fig. 3), we suggest these DAA patterns of visual field loss are subtle defect patterns that serve as novel signatures of developing glaucoma in the future (Fig. 6). However, objective (based on correlation) and subjective evaluations were performed to select 18 mutually exclusive patterns shown in Fig. 6.

It is challenging to compare DAA- and AA-derived patterns of visual field loss. Therefore, to further assess the effectiveness of DAA objectively, we developed machine learning classifiers to identify positive and negative samples based on deep archetypes, classical archetypes, raw visual fields, and clinical parameters. We used 128 DAA patterns as input features to the classifier for predicting glaucoma after performing an extensive grid search to identify maximum AUC. To provide a fair comparison, we used 128 classical AA patterns as was used in DAA assessment. Machine learning analyses showed that DAA patterns were significantly more effective in predicting glaucoma compared to other approaches (Fig. 7, black curve with AUC of 0.71).

To avoid any bias due to multiple visual field tests from same eyes of subjects, we accounted for correlation between tests and eyes of same subjects using a nested structure in generalized estimating equation (GEE) [33]. To account for multiple VFs from same eyes in training and testing of machine learning models, we selected the training and testing examples based on subjects rather than eyes or visual fields.

We compared the proposed framework compared against RVM [5], classical AA approach, and raw VFs. The AUC of model on DAA representation of visual fields was 0.71 while the AUC of model on classical AA representation of visual fields and raw visual fields were 0.61 and 0.55, re-

spectively. The AUC of RVM [5] was 0.64. In fact, AUC of DAA was significantly higher on both cross-validation and held-out datasets (P value < 0.001). This highlights that the deep convex representation, obtained by simplicial projection, is more discriminative than the input raw visual fields as well AA and classical RVM. The AUC of age, CCT, and IOP in predicting glaucoma was 0.56, 0.52, and 0.51, respectively. These findings suggest that none of the well-known glaucoma risk factors could predict glaucoma well ahead of time. We also investigated the role of two major clinical instrument parameters including visual field mean deviation (MD) and pattern standard deviation (PSD) in predicting glaucoma. The AUCs of MD and PSD were 0.50, and 0.51, respectively (Fig. 8). Thus, our proposed glaucoma precognition outperformed glaucoma risk factors and visual field instrument parameters in predicting glaucoma as well (Fig. 8).

The AUC value of 0.71 seems to be low compared to several approaches for identifying glaucoma with higher accuracy. While from a statistical perspective this may seem a valid argument, from clinical perspective, the story is different. Glaucoma precognition from baseline visual fields approximately five years prior to disease onset is a challenging task. In fact, relatively, glaucoma diagnosis is the easiest task since clinicians already have observed clinical signs of the disease, however, in prediction there is no clinical sign and one would need to identify hidden pre-clinical patterns of the disease.

This study was conducted on visual field tests with Humphrey 30-2 pattern. Other studies using visual fields with Humphrey 24-2 or central 10-2 patterns may shed light on the effectiveness of DAA in predicting glaucoma using other visual field test patterns. Nevertheless, visual field testing is subjective, time-consuming and contains a significant degree of variability. Therefore, future studies could investigate the role of structural data such as fundus photographs or optical coherence tomography (OCT) data in predicting glaucoma prior to disease onset.

7. Conclusion

In this work, a framework was developed and implemented using deep archetypal analysis, to effectively predict glaucoma, several years prior to disease onset. The approach obtains unsupervised convex representations of visual fields, using simplex projections. It is shown that these convex representations are clinically meaningful and more discriminative than raw visual fields or other classical approach for visual field analysis. To overcome the class-imbalance issue, an effective class-imbalance bagging approach has been applied. As a proof of concept, the OHTS glaucoma clinical trial dataset was used to assess the effectiveness of approach for early glaucoma prediction. Experimental results indicate that a system of deep archetypal

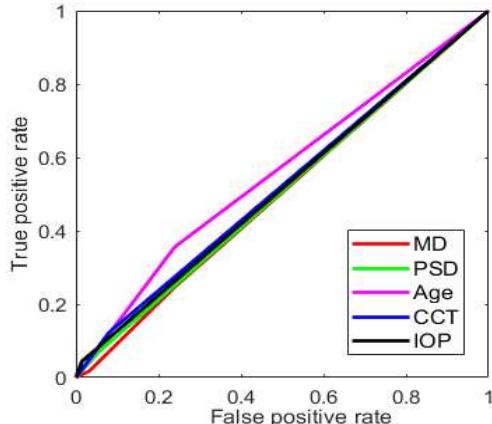


Figure 8. ROC curves of various clinical parameters (MD, PSD, Age, CCT, and IOP).

representation, integrated with class-balanced bagging provides improved predictions of glaucoma development from baseline measurements several years prior to disease development. Future work with independent datasets may be required to verify the findings of this study.

8. Acknowledgment

This work was supported by the National Institute of Health (NIH) grant EY030142.

References

- [1] Jin Mo Ahn, Sangsoo Kim, Kwang-Sung Ahn, Sunghoon Cho, Kwan Bok Lee, and Ungsoo Samuel Kim. A deep learning model for the detection of both advanced and early glaucoma using fundus photography. In *PloS one*, volume 13, 2018. [2](#)
- [2] C. Bauckhage and K. Manshaei. Kernel archetypal analysis for clustering web search frequency time series. In *2014 22nd International Conference on Pattern Recognition*, pages 1544–1549, Aug 2014. [1](#)
- [3] Christian Bauckhage and Christian Thurau. Making archetypal analysis practical. In *Joint Pattern Recognition Symposium*, pages 272–281, 2009. [1](#)
- [4] Karine D. Bojikian, Cecilia S. Lee, and Aaron Y. Lee. Finding Glaucoma in Color Fundus Photographs Using Deep Learning. *JAMA Ophthalmology*, 137(12):1361–1362, 2019. [2](#)
- [5] Christopher Bowd, Intae Lee, Michael H. Goldbaum, Madhusudhanan Balasubramanian, Felipe A. Medeiros, Linda M. Zangwill, Christopher A. Girkin, Jeffrey M. Liebmann, and Robert N. Weinreb. Predicting glaucomatous progression in glaucoma suspect eyes using relevance vector machine classifiers for combined structural and functional measurements. *Investigative ophthalmology visual science*, 53(4):2382–9, 2012. [2, 3, 5, 7](#)
- [6] Christopher Bowd, Robert N. Weinreb, Madhusudhanan Balasubramanian, Intae Lee, Gil-Jin Jang, Siamak Yousefi, Linda M. Zangwill, Felipe A. Medeiros, Christopher A. Girkin, Jeffrey M. Liebmann, and Michael H. Goldbaum. Glaucomatous patterns in frequency doubling technology (fdt) perimetry data identified by unsupervised machine learning classifiers. In *PloS one*, 2014. [6](#)
- [7] Paolo Brusini and Chris A Johnson. Staging functional damage in glaucoma: review of different classification methods. *Survey of ophthalmology*, 52(2):156–179, 2007. [5](#)
- [8] X. Chen, Y. Xu, D. W. Kee Wong, T. Y. Wong, and J. Liu. Glaucoma detection based on deep convolutional neural network. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 715–718, 2015. [2](#)
- [9] Yuansi Chen, Julien Mairal, and Zaid Harchaoui. Fast and robust archetypal analysis for representation learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1478–1485, 2014. [3, 4, 5](#)
- [10] Adele Cutler and Leo Breiman. Archetypal analysis. *Technometrics*, 36(4):338–347, 1994. [1, 3](#)
- [11] Tobias Elze, Louis R. Pasquale, Lucy Q. Shen, Teresa C. Chen, Janey L. Wiggs, and Peter J. Bex. Patterns of functional vision loss in glaucoma determined with archetypal analysis. *Journal of The Royal Society Interface*, 12(103):20141118, 2015. [2, 6](#)
- [12] Jampel HD et al. Agreement among glaucoma specialists in assessing progressive disc changes from photographs in open-angle glaucoma patients. *Am J Ophthalmol*, 147:39–44, 2009. [6](#)
- [13] Mae O. Gordon, Julia A. Beiser, James D. Brandt, Dale K. Heuer, Eve J. Higginbotham, Chris A. Johnson, John L. Keltner, J. Philip Miller, II Parrish, Richard K., M. Roy Wilson, Michael A. Kass, and for the Ocular Hypertension Treatment Study Group. The Ocular Hypertension Treatment Study: Baseline Factors That Predict the Onset of Primary Open-Angle Glaucoma. *Archives of Ophthalmology*, 120(6):714–720, 06 2002. [1, 3](#)
- [14] David B. Henson, Shaila Chaudry, Paul H. Artes, E. Brian Faragher, and Alec Ansons. Response Variability in the Visual Field: Comparison of Optic Neuritis, Glaucoma, Ocular Hypertension, and Normal Eyes. *Investigative Ophthalmology Visual Science*, 41(2):417–421, 02 2000. [2](#)
- [15] Chris A Johnson, Pamela A Sample, George A Cioffi, Jeffrey R Liebmann, and Robert N Weinreb. Structure and function evaluation (SAFE): I. criteria for glaucomatous visual field loss using standard automated perimetry (SAP) and short wavelength automated perimetry (swap). *American Journal of Ophthalmology*, 134:177–185, 2002. [1](#)
- [16] Sebastian Mathias Keller, Maxim Samarin, Mario Wieser, and Volker Roth. Deep archetypal analysis. In Gernot A. Fink, Simone Frintrop, and Xiaoyi Jiang, editors, *Pattern Recognition, DAGM-GCPR*, pages 171–185, 2019. [1, 4, 5](#)
- [17] John L. Keltner, Chris A. Johnson, Kimberly E. Cello, Mary A. Edwards, Shannan E. Bandermann, Michael A. Kass, Mae O. Gordon, and for the Ocular Hypertension Treatment Study Group. Classification of Visual Field Abnormalities in the Ocular Hypertension Treatment Study. *Archives of Ophthalmology*, 121(5):643–650, 05 2003. [2, 7](#)

- [18] Wangmin Liao, Beiji Zou, Rongchang Zhao, Yuanqiong Chen, Zhiyou He, and Mengjie Zhou. Clinical interpretable deep learning model for glaucoma diagnosis. *IEEE journal of biomedical and health informatics*, 2019. 2
- [19] Paul R. Lichter. Variability of expert observers in evaluating the optic disc. *Transactions of the American Ophthalmological Society*, 74:532–72, 1976. 6
- [20] Frederick S. Mikelberg, C M Parfitt, Nicholas V. Swindale, Stuart L. Graham, Stephen Michael Drance, and Ray G. Gosine. Ability of the heidelberg retina tomograph to detect early glaucomatous visual field loss. *Journal of glaucoma*, 4(4):242–7, 1995. 2
- [21] Harry A. Quiagley. Glaucoma. *Lancet*, 377(9774):1367–77, 2011. 1
- [22] Maria Salvat, Marco Zeppieri, C Tosoni, and Paolo Brusini. Baseline factors predicting the risk of conversion from ocular hypertension to primary open-angle glaucoma during a 10-year follow-up. *Eye (London, England)*, 30:784–795, 05 2016. 1, 2
- [23] Pamela A. Sample, Catherine Boden, Zuohua Zhang, John Pascual, Te-Won Lee, Linda M. Zangwill, Robert N. Weinreb, Jonathan G. Crowston, Esther M. Hoffmann, Felipe A. Medeiros, Terrence Sejnowski, and Michael Goldbaum. Unsupervised Machine Learning with Independent Component Analysis to Identify Areas of Progression in Glaucomatous Visual Fields. *Investigative Ophthalmology Visual Science*, 46(10):3684–3692, 2005. 2
- [24] Mitra Sehi, Namita Bhardwaj, Yun Suk Chung, and David S. Greenfield. Evaluation of baseline structural factors for predicting glaucomatous visual-field progression using optical coherence tomography, scanning laser polarimetry and confocal scanning laser ophthalmoscopy. *Eye*, 26:1527–1535, 2012. 2, 3
- [25] Sohan Seth and Manuel JA Eugster. Archetypal analysis for nominal observations. *IEEE transactions on pattern analysis and machine intelligence*, 38(5):849–861, 2015. 1
- [26] Anshul Thakur, Vinayak Abrol, Pulkit Sharma, and Padmanabhan Rajan. Deep convex representations: Feature representations for bioacoustics classification. In *Interspeech 2018*, pages 2127–2131, 2018. 1, 4, 5
- [27] Anshul Thakur and Padmanabhan Rajan. Deep archetypal analysis based intermediate matching kernel for bioacoustic classification. *J. Sel. Topics Signal Processing*, 13(2):298–309, 2019. 1, 4, 5
- [28] Mengyu Wang, Lucy Q. Shen, Louis R. Pasquale, Paul Petrakos, Sydney Formica, Michael V. Boland, Sarah R. Wellik, Carlos Gustavo De Moraes, Jonathan S. Myers, Osamah Saeedi, Hui Wang, Neda Baniasadi, Dian Li, Jorriyt Tichelaar, Peter J. Bex, and Tobias Elze. An Artificial Intelligence Approach to Detect Visual Field Progression in Glaucoma Based on Spatial Pattern Analysis. *Investigative Ophthalmology Visual Science*, 60(1):365–375, 2019. 2, 6
- [29] Robert N Weinreb, Tin Aung, and Felipe A Medeiros. The pathophysiology and treatment of glaucoma: a review. *JAMA*, 311(18):1901–1911, 2014. 1
- [30] Siamak Yousefi, Michael H. Goldbaum, Madhusudhanan Balasubramanian, Felipe A. Medeiros, Linda M. Zangwill, Jeffrey M. Liebmann, Christopher A. Girkin, Robert N. Weinreb, and Christopher Bowd. Learning from data: Recognizing glaucomatous defect patterns and detecting progression from visual field measurements. *IEEE Transactions on Biomedical Engineering*, 61:2112–2124, 2014. 2, 6
- [31] Siamak Yousefi, Michael H Goldbaum, Linda M Zangwill, Felipe A Medeiros, and Christopher Bowd. Recognizing patterns of visual field loss using unsupervised machine learning. In *Medical Imaging 2014: Image Processing*, volume 9034, page 9034M. International Society for Optics and Photonics, 2014. 6
- [32] Siamak Yousefi et al. Unsupervised gaussian mixture-model with expectation maximization for detecting glaucomatous progression in standard automated perimetry visual fields. *Translational vision science technology*, 5, 2016. 6
- [33] Scott L Zeger, Kung-Yee Liang, and Paul S Albert. Models for longitudinal data: a generalized estimating equation approach. *Biometrics*, pages 1049–1060, 1988. 7