Rule-based classification of visual field defects

Enkelejda Kasneci[1], Gjergji Kasneci[2], Ulrich Schiefer[3], and Wolfgang Rosenstiel[1]
[1]*Department of Computer Engineering, University of Tübingen, Germany*
[2]*Hasso-Plattner-Institute, Germany*
[3]*Centre for Ophthalmology, Institute for Ophthalmic Research, University of Tübingen, Germany*
enkelejda.kasneci@uni-tuebingen.de

Keywords:
scotoma, rule-based, classification, visual field, defect

Abstract:
The automated recognition of the visual field defect type from results of visual field testing is crucial for the adequate diagnosis and treatment of the underlying disease of the visual system. This paper presents a reliable rule-based classifier that emulates the decision strategies of expert ophthalmologists based on a two-level approach that combines methods of unsupervised learning.

1 INTRODUCTION

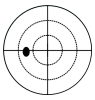
Diseases affecting the optic nerve (e.g., glaucoma), or the brain (e.g., stroke, trauma, brain injury) may lead to blind areas or areas of reduced visual perception in the visual field. Such areas, also known as visual field defects (or scotoma), are identified through the measurement of the visual field, i.e., perimetry. During a perimetric examination, light stimuli of different luminance levels are projected onto a uniform background at predefined locations in the visual field. The subject confirms stimulus perception by pressing a button; no response to a stimulus projection is interpreted as failure to see the stimulus. Visual field defects are areas at which stimuli are not perceived or only perceived when the stimuli have high light intensity.

Figure 2(a) depicts a perimetric result showing a central visual field defect. Locations of perceived stimuli are marked by black dots. Rectangles represent stimuli locations with reduced visual perception. The darker the rectangle, the lower the sensitivity at that location. The detected location, shape and size of defects in the visual field are hints for the underlying disease of the visual system. In the clinical routine, the classification of the visual field defect type from perimetric measurements is performed manually based on the Tübingen Scotoma Classification Scheme (TSCS), which combines expert know-how and long-term experience. This scheme distinguishes between eight classes of visual field defects as depicted in Figure 1.

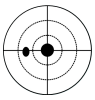
The high variability in the manifestation of visual field defect types makes the automated classification of visual field defects from perimetric results highly challenging. Moreover, the measurement data may be sparse or noisy.

Most prior techniques [Bizios et al., 2007, Boden et al., 2007, Brigatti et al., 1997, Goldbaum et al., 1994, Goldbaum et al., 2002, Goldbaum, 2005, Goldbaum et al., 2009, Henson et al., 1997] have focused on the automated detection of glaucoma from perimetric results, as it is a progressive disease that is becoming more present due to demographic aging. Other methods have considered a subset of the visual field defects in terms of the TSCS [Keating et al., 1993, Mutlukan and Keating, 1994]. The extension and applicability of these methods to the recognition of other defect types has not been investigated. Classification techniques for the detection of all visual field defect types based on neural networks have been presented in [Fink, 2004, Jürgens et al.,

Normal visual field: Less than 7 relative defects may occur anywhere in the visual field, but not in the edges of the grid. No defect clusters besides the physiological blind spot are visible.

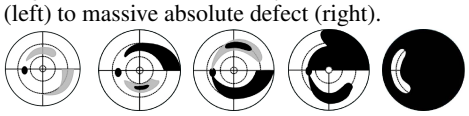


Central scotoma: Defect that occurs within 5° eccentricity without respecting the vertical or the horizontal meridian. Includes also the *Paracentral scotoma* (normal blind spot position) and the *Centrocecal scotoma* (extending from the blind spot towards the central area symmetrically above and below the midline).

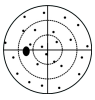


Concentric constriction: Visual field loss sparing the central visual field.

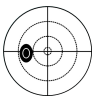
Glaucoma: Progressive defect that occurs in five stages [Aulhorn and Karmeyer, 1977], from relative (left) to massive absolute defect (right).



Diffuse visual field defect: More than 7 relative defects that are disseminated across the visual field.



Blind spot enlargement: Defect involving at least two points that are contiguous with the blind spot.



Sector- oder wedge-shaped defect: Wedge-shaped defect respecting neither the vertical nor the horizontal meridian.

Hemianopic: Defect respecting, at least locally, the vertical meridian.

Others: Defects that cannot be attributed to any of the above mentioned classes.

Figure 1: Tübingen Scotoma Classification Scheme (TSCS) [Nevalainen et al., 2007]

2001]. These approaches, however, do not report how the performance of the algorithms varies with decreasing quality of the data from perimetric results. The main drawback of methods that are based on neural networks is their dependence on the training procedure, as the overall classification performance can be negatively influenced by missing correlations and noise in the input data [Bengtsson et al., 2005, Henson et al., 1997].

In this work we consider all the TSCS types of visual field defects and provide a highly accurate rule-based classification technique that integrates expert knowledge with parameters that were statistically established from a large number of real-world visual field examinations. Thus, our method is not necessarily dependent on correlations in the training or input data. In contrast to related methods, our approach has been evaluated on perimetric results of different quality levels.

2 INPUT DATA

Perimetric results consisting of 192 concentrically arranged stimuli locations (e.g., Figure 2(a)) are input data to our classification method. The result concerning the detection of a stimulus location by the test subject is represented by a stimulus vector $s_i = (x_i, y_i, dd_i)$, $i \in \{0, \ldots, 191\}$. (x_i, y_i) represents the position of the stimulus with $x_i, y_i \in \{-30, \cdots, 30\}$. dd_i represents the defect depth that is related to the measured differential luminance sensitivity at location (x_i, y_i). The range of defect values from 0 dB to over 30 dB is subdivided into 7 intervals, i.e., $dd_i \in \{0, \cdots, 6\}$ classes. A perimetric result can be represented as a matrix of stimuli vectors:

$$\mathbf{M} = \left[\begin{pmatrix} x_0 \\ y_0 \\ dd_0 \end{pmatrix} \cdots \begin{pmatrix} x_{191} \\ y_{191} \\ dd_{191} \end{pmatrix} \right] \qquad (1)$$

The classification method was evaluated on 8,868 anonymized visual field examination results provided by the University Eye Hospital Tübingen. The data was hand-labeled by expert ophthalmologists. In addition, a classification quality, $Q \in \{2, 3, 4\}$, representing the physician's certainty for the identification of the visual field defect, was provided. $Q = 4$ is the highest quality level, i.e., the defect type could clearly be identified. The distribution of perimetric results among the defect classes is presented in Table 1.

Q4	Q3	Q2	Visual field defect class
690	1,622	123	C_1: Normal visual field
16	634	430	C_2: Central scotoma
13	111	51	C_3: Concentric constriction
125	2,956	999	C_4: Glaucoma
1	16	2	C_5: Diffuse visual field defects
9	268	153	C_6: Blind spot enlargement
0	19	8	C_7: Sector- or wedge-shaped defect
105	315	202	C_8: Hemianopic defect
959	**5,941**	**1,968**	

Table 1: Distribution of defect types in the evaluation dataset of perimetric results with quality level (Q4-Q2).

3 CLASSIFICATION METHOD

Our method for the recognition of the visual field defect type from a perimetric result consists of three steps:

1. In a first step, structures in perimetric examinations, i.e., clusters of stimuli locations with impaired visual perception, are found. This is achieved by two methods of unsupervised learning, namely *Hierarchical Agglomerative Clustering* (HAC) [Duda et al., 2000] and *Self-Organizing Maps* (SOMs) [Kohonen, 1990]. In a two-level approach, SOMs are used to compute prototype clusters, which are then combined to final clusters using HAC. The most decisive benefit of the SOM-based pre-clustering is noise reduction [Mangiameli et al., 1996, Henson et al., 1997, Tafaj et al., 2011a].

2. In the next step, for the original perimetric result new features are derived from the clusters found in Step 1, e.g., centroid position, cluster size, average defect depth of stimuli locations in the cluster, etc.

3. Finally, first-order logic rules check the class membership of the enriched perimetric result (i.e., with the above features). The rules have been designed in close collaboration with expert ophthalmologists. Uncertain decision parameters (e.g., decision thresholds) have been optimized by means of correlation analysis.

Step 1: Clustering for the feature enrichment of perimetric results

Pre-clustering through Self-Organizing Maps A Self-Organizing Map (SOM) is an unsupervised learning method introduced by Kohonen [Kohonen, 1990]. In a SOM network, the N-dimensional input data is mapped to a lower dimensional arrangement (grid) of neurons, such that the topological order is maintained [Kohonen, 1990]. Thus, adjacent units map to similar data points. The particularity of a SOM is that it can be used at the same time for both the reduction of data complexity (by clustering) and the nonlinear projection onto a lower-dimensional space [Kaski, 1997]. Each neuron, also called *unit*, is assigned a reference vector m_i containing the synaptic weights. These reference vectors are initialized prior to the SOM training, usually randomly. During training, the reference vectors of the map units are iteratively adapted to fit the input data best. Given the input data point x, for each unit, the distance between its reference vector and x is computed. The Kohonen unit c with the minimum distance to x (usually the Euclidean metric is used) is the winner neuron, also called *Best Matching Unit, BMU*, Equation 2:

$$c = c(x) = \arg min_i \{\|x - m_i\|^2\} \qquad (2)$$

Thus, the BMU determines the spatial location of a topological neighborhood of excited neurons, thereby providing the basis for cooperation among neighboring neurons. In each new iteration step $t+1$, the reference vectors are adapted to the input from the previous iteration step, $x(t)$, according to the gradient descent rule in Equation 3:

$$m_i(t+1) = m_i(t) + \alpha(t) h_{ci} [x(t) - m_i(t)] \qquad (3)$$

$\alpha(t)$ represents the learning rate over time (which typically decreases with learning progress), whereas h_{ci} represents the neighborhood kernel around the BMU c.

Hierarchical Agglomerative Clustering In HAC clusters of points are constructed in a bottom-up manner; starting with each point as a singleton cluster, clusters that are close to each other, according to a predefined similarity or distance measure, e.g., the Euclidean distance, are merged iteratively until a single,

all-encompassing cluster remains [Jain et al., 1999, Tan et al., 2005]. We use HAC with *centroid-linkage* on top of the SOM-based preclustering of stimuli locations to derive coherent defect clusters. Centroid-linkage clustering defines the cluster distance as the distance between their centroids [Flach, 2012]. The goal of every merge step is to find and merge a pair of clusters c_i, c_j that minimize the linkage function $L_{centroid}(c_i, c_j)$:

$$L_{centroid}(c_i, c_j) = D(\mu_i, \mu_j) \tag{4}$$

The SOM-based clustering approach has several advantages over other popular alternatives, such as k-means or expectation-maximization-based methods. Such methods require that the number of clusters is known in advance. In our scenario, we have no prior knowledge on the number of clusters. In the SOM approach, the only parameter that is decisive for the number of clusters is an empirically established threshold on the distance of neighboring neurons. That is, neighboring neurons with a distance greater than the threshold belong to different clusters. While several flat-clustering algorithms can be adapted to take such a threshold into account, they often face the problem of unstable results in the presence of noise. In fact, noise is the most difficult aspect to deal with in the scenario of clustering stimuli locations.

Figure 2 illustrates how this two-level approach was used for the analysis of visual field defects in perimetric examination results. Figure 2(a) depicts a perimetric examination result showing a central visual field defect. A SOM grid (12×6 units) was trained on the normalized perimetric data from Figure 2(a) based on the Matlab SOM Toolbox[1]. Figure 2(b) shows the mapping of the SOM units (represented by black circles) on the normalized perimetric data. Note that a 3-dimensional representation has been used for the sake of better visualization. In this example, the visual field defect and the intact area of the visual field form well-separated clusters. Figure 2(c) shows a so-called U-Matrix for the SOM trained in Figure 2(b) visualizing the Euclidean distances between reference vectors of neighboring units (black represents small Euclidean distance). As is can be seen in Figure 2(c), a bright set of cells separates two dark areas that correspond to the clusters in the perimetric result. Note that separating the clusters of the SOM from Figure 2(b) according to the corresponding U-Matrix is straightforward for the presented example, which was chosen for the sake of a comprehensible demonstration of the method. However, note that in general, the U-Matrix may suggest multiple clusters with non-obvious boundaries between them. For such cases further analysis is needed to separate the clusters. Therefore, we apply HAC on top of the SOM results to separate clusters based on their hierarchical dependencies derived from the centroid-linkage strategy. This was also done to detect the clusters of units on the SOM from Figure 2(c). The corresponding dendrogram and the clustering result are presented in Figure 2(d).

[1]Matlab SOM Toolbox, `http://www.cis.hut.fi/projects/somtoolbox/`

Step 2: Feature enrichment

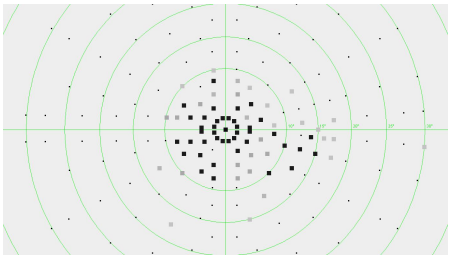
The features of the clusters derived in Step 1 (e.g., cluster centroid, cluster size, average defect depth of stimuli locations in a cluster, etc.) are used to enrich the original perimetric result. The enriched feature vector of a perimetric result is $\mathbf{x} \in \mathscr{X}$ contains the following features:

- the set $\mathbf{M}$ of 192 three-dimensional vectors representing the stimuli (see Equation 1)
- the number n_d of defects in the perimetric examination result
- the number n_c of clusters derived by the SOM-based clustering
- the set of cluster means $\{\mu_1, ..., \mu_{n_c}\}$
- the set of cluster sizes $\{S_1, ..., S_{n_c}\}$
- the set of the numbers of stimuli locations in each cluster $\{sc_1, ..., sc_{n_c}\}$
- the set containing the average defect depth in each cluster $\{\overline{dd}_1, ..., \overline{dd}_{n_c}\}$

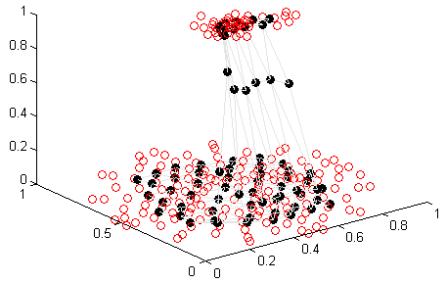
Such an enriched perimetric result becomes the input of the classifier in Step 3, by which it is mapped to one of the eight visual field defect classes depicted in Figure 1.

Step 3: Recognition of visual field defect types

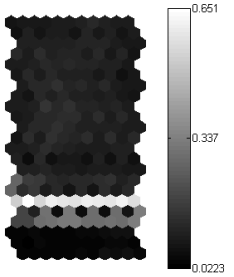
The classifier has been designed as a function $f : \mathscr{X} \to \mathscr{C}$, $\mathscr{C} = \{C_1, C_2, \ldots, C_8\}$, that maps an



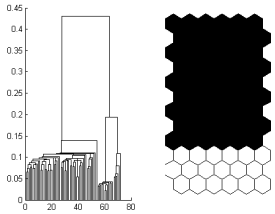
(a) Perimetric result showing a central visual field defect



(b) The normalized perimetric data from (a) (red circles) and a mapped SOM with a lattice of 12x6 units (black circles)



(c) U-Matrix visualizing the distances between neighboring units of the SOM from (b)



(d) Dendrogram and clusters resulting from HAC with centroid-linkage on the SOM from (b)

Figure 2: An example of SOM usage for clustering perimetric results.

input $\mathbf{x} \in \mathscr{X}$ to one of the 8 classes in $\mathscr{C}$ in a way that is similar to how an expert ophthalmologist would decide, namely based on rules derived from frequently occurring observations. Hence, the classification method consists of a collection of rule-based binary classifiers that operate based on a one-versus-all scheme. For each of the 8 classes, a first-order logic rule has been manually assembled by considering decision rules used by ophthalmologists and also by empirically analyzing a hand-labeled data set of perimetric results of quality levels 3 and 4, see Table 1. This data set was recommended by an expert ophthalmologist and included 20 representatives of each scotoma class, except for the classes C_5 and C_7, which were represented by 10 perimetric results each. In order to identify appropriate values for some of the input features (i.e., for values that are prone to uncertainty), a correlation analysis was run on this data set with the goal of increasing the accuracy of the scotoma classifier. To this end, the *Matthews Correlation Coefficient, MCC*, was computed on the number of *true positives (TP)*, *true negatives (TN)*, *false positives (FP)*, and *false negatives (FN)* for a given class: $MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$ [Baldi et al., 2000]. The characterization of each class is based on the TSCS from Figure 1.

Normal visual field (C_1) The normal visual field is characterized by only one defect cluster that corresponds to the physiological blind spot and few sporadic defects that occur due to the patient's inattention. The correlation analysis on the data revealed a maximum number of 10 defects. Note that this threshold is in accordance with typical thresholds used by expert ophthalmologists, which range between 8 and 11. This number includes the failed stimuli locations within the blind spot cluster $c_{blindspot}$ with centroid $\mu_{blindspot}$ and size $S_{blindspot}$, and few sporadic defects. The values $\mu_{blindspot}$ and $S_{blindspot}$ stem from the TSCS.

$$(n_d \leq 10 \ \wedge \ n_c = 1 \ \wedge \ \mu \approx \mu_{blindspot} \ \wedge \ S \approx S_{blindspot}) \rightarrow C_1 \quad (5)$$

Central scotoma (C_2) A central scotoma is characterized by absolute or relative defects that do not respect the vertical or horizontal meridian in the central visual field. A perimetric result is assigned to this class if it fulfills the rule $(RC \vee RCC) \rightarrow C_2$, where *RC* and *RCC* are defined in the Formulas 6 and 7, respectively.

The *RC* part of the rule stems entirely from expert knowledge: All 13 stimuli locations within 2° eccentricity are checked; if more than 50% of these locations, i.e., more than 6, are defects, the presence of a central scotoma is assumed.

$$[RC] : \sum_{s_i \in M} (\|(x_{s_i}, y_{s_i})\|_2 \leq 2)(\overline{dd}_{s_i} \geq 1)) \geq 6 \quad (6)$$

RCC stems also to a large extent from expert knowledge: A *paracentral scotoma* is assumed whenever there exists a central cluster of defects with centroid within 10° eccentricity, containing more than 5 stimuli. The minimum number of stimuli locations within the central cluster was derived from correlation analysis (i.e., MCC analysis) on the data. The value 5 was revealed as a reliable threshold.

$$[RCC] : \exists c_i : \|\mu_i\|_2 \leq 10 \wedge sc_i \geq 5 \wedge \overline{dd}_i \geq 3 \quad (7)$$

Concentric constriction (C_3) This defect type is manifested by an intact central visual field and a peripheral field constriction. Thus, we consider a central (within 15° eccentricity) and a peripheral defect cluster and check their positions. In addition, the average defect depths in the clusters are compared. The defect depth in the peripheral defect cluster is expected to be at least 1.3 times larger than the defect depth in the central cluster. The values 15 and 1.3 were established by means of MCC analysis on the data.

$$(\exists c_i, c_j : \mu_i \approx (0,0) \wedge \forall s_k \in c_i : \left\|(x_{s_k}, y_{s_k})\right\|_2 \leq 15 \wedge \mu_j \approx (0,0) \wedge \overline{dd}_j \geq \overline{dd}_i * 1.3) \rightarrow C_3 \quad (8)$$

Glaucoma (C_4) This class is characterized by arcuate-shaped defects. However, the perimetric results for this class only rarely showed well-separated defect clusters as depicted in the schematic view of Figure 1. In reality, the defect clusters were often distorted. Furthermore, the five stages of glaucoma reveal different cluster shapes. A perimetric examination is classified as glaucomatous if there exists a cluster with at least 15 defects spanning over the right and left hemifield according to the rule in Equation 9. While a minimum average defect depth of $\overline{dd}_i = 3$ is a typical value observed by experts, the remaining thresholds were established by means of MCC analysis.

$$(\exists c_i : sc_i \geq 15 \wedge \overline{dd}_i \geq 3 \wedge \exists s_j, s_k \in c_i : x_{s_j} > 10 \wedge x_{s_k} < -10) \rightarrow C_4 \quad (9)$$

Diffuse visual field defects (C_5) These are defects that are spread across the visual field, see Figure 1. Clustering typically yields one large cluster encompassing more than half of the visual field (i.e., ≥0.5*VF). Experts also expect to observe an average defect depth of at least 2 in this cluster.

$$(\exists c_i : S_i \geq 0.5 * VF \wedge \overline{dd}_i \geq 2) \rightarrow C_5 \quad (10)$$

Blind spot enlargement (C_6) These are defects that occur due to the enlargement of the blind spot. Such cases are typically assumed when there is a defect cluster close to the blind spot area, for which the centroid or the size does not correspond to the centroid or the size of the normal physiological blind spot cluster. The following is a decision rule used by expert ophthalmologists for this class of defects.

$$\left(\exists c_i : \mu_i \neq \mu_{blindspot} \vee S > S_{blindspot}\right) \rightarrow C_6 \quad (11)$$

Sector- or wedge-shaped defects (C_7) These are defect areas that are wedge-shaped. A perimetric result is assigned to this class, if there is a defect cluster with the following features: (i) it contains more than 8 defect stimuli, (ii) it is not the blind spot, and (iii) the angle between the sides of the wedge-shaped cluster (l_l and l_r) is between 30° and 90°. The values for the minimum number of defect stimuli in (i) and the minimum average defect depth in (ii) were established by means of MCC analysis. The angle between the cluster sides stems from expert knowledge.

$$(\exists c_i : \mu_i \neq \mu_{blindspot} \wedge S_i \neq S_{blindspot} \wedge sc_i \geq 8 \wedge \overline{dd}_i \geq 2 \wedge 30 \leq \angle(l_l(c_i), l_r(c_i)) \leq 90) \rightarrow C_7 \quad (12)$$

Hemianopic defect (C_8) These defects expand over one hemifield and respect the vertical meridian of the visual field. The corresponding perimetric results are typically characterized by an impaired left or right visual field side, where most of the stimuli were not recognized. The other half of the visual field is intact, except for the blind spot area and few relative or absolute defects, e.g., due to patient's inattention. To detect this defect type, the algorithm counts the number of defects in each half of the visual field. Correlation analysis on the data, by means of MCC, revealed that a left-sided hemianopic defect is found, if at least 60% of the stimuli locations of the left side (i.e., with x coordinate $x \leq 0$) of the visual field are absolute defects (i.e., stimuli locations with defect depth $dd \geq 3$). The threshold 3 for the defect

depth of a stimulus is based on expert knowledge. The right side of the visual field is considered intact, if there are less than 10% of relative defects (i.e., stimuli locations with defect depth $dd \leq 2$). The rule for the detection of the left-sided hemianopic defect is defined by *RHL* in Formula 13. The detection of a right-sided hemianopic defect is defined by *RHR* in the Formula 14 and is symmetric to *RHL*. In summary, a perimetric examination result is assigned to this defect class if it fulfills the rule $(RHL \vee RHR) \rightarrow C_8$.

$$[RHL]: \frac{\sum_{s_i \in M}((x_{s_i} \leq 0)(dd_{s_i} \geq 3))}{\sum_{s_i \in M}(x_{s_i} \leq 0)} \geq 0.6 \ \wedge \ \frac{\sum_{s_j \in M}((x_{s_j} > 0)(dd_{s_j} \leq 2))}{\sum_{s_j \in M}(x_{s_j} > 0)} \leq 0.1 \quad (13)$$

$$[RHR]: \frac{\sum_{s_i \in M}((x_{s_i} \geq 0)(dd_{s_i} \geq 3))}{\sum_{s_i \in M}(x_{s_i} \geq 0)} \geq 0.6 \ \wedge \ \frac{\sum_{s_j \in M}((x_{s_j} < 0)(dd_{s_j} \leq 2))}{\sum_{s_j \in M}(x_{s_j} < 0)} \leq 0.1 \quad (14)$$

Note that in general more than one of the presented decision rules may fire for the same input. Under the assumption that each perimetric examination result belongs to one and only one defect class, we iterate through the rules in the order they were introduced and the class corresponding to the first rule that fires.

4 EXPERIMENTAL RESULTS

The presented algorithm was evaluated on all 8,848 perimetric examination results of quality levels 4, 3, and 2 from Table 1, in a one-against-all scheme. The classification performance was measured by means of *accuracy* (ACC) and *specificity* (SP). The evaluation results of the classification method are presented in Table 2. For each scotoma class and the corresponding quality level, the ACC and SP values are reported. Furthermore, weighted average values for ACC and SP are shown per defect type, as well as for the entire data set.

The overall classification ACC and SP of the presented rule-based classification method were 83% and 85%, respectively. For examination results of quality level 4, the average (weighted) ACC and SP values were 92% and 98%, respectively. Considering the three defect classes with the largest number of instances, namely *normal visual field*, *glaucoma*, and *hemianopic defect*, the algorithm shows very good performance, i.e., ACC and SP values of above 90%.

With decreasing data quality (from examination results of quality level 4 to examination results of quality level 2), which corresponds to increasing uncertainty in classification by the ophthalmologists, we found that the performance of our algorithm also decreases, see Table 2. For examination results of quality levels 3 and 2, there was a decrease in the average ACC and SP values to 87% and 85% (for quality level 3) and 79% and 78% (for quality level 2), respectively. This effect could be observed for each scotoma class. The main reason for this effect is the high diversity of manifestations of a scotoma type, which may lead to uncertainty in the classification result by the physician or a miss-classification by the algorithm.

The largest loss in ACC (i.e., from 93% for data of quality level 4 to 74% for data of quality level 2) is found in the recognition of glaucoma. Indeed, the recognition of glaucoma is considered particularly challenging [Nouri-Mahdavi et al., 2011], mainly because of its manifestation in 5 Aulhorn-stages, as depicted schematically in Figure 1. In the clinical routine, the classification of glaucoma is not only based on the perimetric result, but also under consideration of patient-related features, such as the intraocular pressure. Such features, however, are often difficult to quantify and were thus not considered in our rule-based classification method.

In some cases, especially for data of low quality, (e.g., quality level 2) more than one of the presented decision rules might fire. Under the assumption that each perimetric examination result belongs to one and only one defect class, the presented rule iteration order interferes with the classification performance.

Comparison to related work

The accuracy achieved by our method regarding the detection of glaucoma varied from 74% (for data of quality level 2) to 93% (for data of quality level 4), see Table 2. The overall weighted average accuracy for glaucoma recognition was 76%,

	Weighted avg.		Quality 4		Quality 3		Quality 2	
Visual field defect class	**ACC (%)**	**SP (%)**	**ACC (%)**	**SP (%)**	**ACC (%)**	**SP (%)**	**ACC (%)**	**SP (%)**
C_1 : Normal visual field	91	97	91	98	91	97	90	92
C_2 : Central scotoma	83	82	91	91	86	85	80	78
C_3 : Concentric constriction	88	89	97	98	89	90	84	85
C_4 : Glaucoma	76	78	93	95	77	79	74	73
C_5 : Diffuse visual field defects	88	88	95	95	88	88	86	86
C_6 : Blind spot enlargement	83	84	94	95	85	85	81	82
C_7 : Sector- or wedge-shaped defect	88	88	97	97	88	89	87	87
C_8 : Hemianopic defect	91	92	97	98	90	91	89	90
Overall	**83**	**85**	**92**	**98**	**87**	**85**	**79**	**78**

Table 2: Evaluation of the accuracy (ACC) and specificity (SP) of the rule-based classifier on 8,848 hand-labeled perimetric examination results.

which is well in line with other approaches [Boden et al., 2007, Brigatti et al., 1997, Goldbaum, 2005, Goldbaum et al., 2009, Goldbaum et al., 2002, Goldbaum et al., 1994, Henson et al., 1996, Mutlukan and Keating, 1994]. However, note that these approaches were evaluated on relatively small data sets consisting of a few hundred perimetric examination results, whereas our algorithm was evaluated on 4,080 perimetric results of glaucomatous visual field defect cases.

An approach that has considered a larger subset of visual field defect classes was presented in [Keating et al., 1993]. The achieved accuracy reported there varied from 74% to 96% on simulated perimetric results of the following defect types: (1) central visual field defects, (2) concentric constriction, and (3) hemianopic defects. Our algorithm outperforms this approach for all the mentioned defect types and all data quality levels.

In comparison to the classification accuracy of 80% that was achieved by [Fink, 2004], where a Hopfield-attractor network was used to recognize the visual field defect classes, our algorithm performs better when considering the overall accuracy on data of quality levels 4 and 3. For data of quality level 2, our approach yields similar accuracy results as the approach of [Fink, 2004].

In comparison to the approach presented in [Jürgens et al., 2001] that achieves a classification accuracy of 95%, our algorithm shows similar performance on data of quality level 4, where about 1,000 perimetric results were classified. It is important to note that it is unclear how the above approaches would have performed on large data sets of varying quality levels.

In summary, the main advantage of our rule-based scotoma classifier is the integration of expert knowledge into the automated classification method. Furthermore, the rules can be easily adapted and new features can be included.

5 CONCLUSION

We presented a rule-based classification method for the automated recognition of the visual field defect types from visual field measurements. This approach was evaluated on a large set of visual field measurements of different quality levels and showed a high classification accuracy. The presented method can also be used beyond the classification purpose. More specifically, the two-level approach can be employed to automatically extract areas of reduced perception in the visual field assessed with methods other than perimetry, e.g., with EFOV [Tafaj et al., 2013], which measures the visual exploration capability of a subject based on the online analysis of eye movements [Tafaj et al., 2012]. Besides its usage in local diagnostic processes, e.g., assisting the clinical routine, the method could also be used in tele-medicine. Further improvements of the presented algorithm include: (1) the refinement of the decision rules and the investigation of further features, such as features related to the patient's general health condition, with the focus on the glaucoma defect type, (2) the development of a user-friendly interface for individual threshold adaptation, and (3) integration with other software tools for vision research, such as Vishnoo [Tafaj et al., 2011b].

REFERENCES

Aulhorn, E. and Karmeyer, H. (1977). Frequency distribution in early glaucomatous visual field defects. *Documenta Ophthalmologica Proceedings Series*, 14:75–83.

Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A. F., and Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics*, 16(5):412–424.

Bengtsson, B., Bizios, D., and Heijl, A. (2005). Effects of input data on the performance of a neural network in distinguishing normal and glaucomatous visual fields. *Investigative Ophthalmology & Visual Science*, 46(10):3730–3736.

Bizios, D., Heijl, A., and Bengtsson, B. (2007). Trained artificial neural network for glaucoma diagnosis using visual field data: a comparison with conventional algorithms. *Journal of Glaucoma*, 16(1):20–28.

Boden, C., Chan, K., Sample, P. A., Hao, J., Lee, T., Zangwill, L. M., Weinreb, R. N., and Goldbaum, M. H. (2007). Assessing Visual Field Clustering Schemes Using Machine Learning Classifiers in Standard Perimetry. *Investigative Ophthalmology & Visual Science*, 48(12):5582–5590.

Brigatti, L., Nouri-Mahdavi, K., Weitzmann, M., and Caprioli, J. (1997). Automatic Detection of Glaucomatous Visual Field Progression with Neural Networks. *Archives of Ophthalmology*, 115(6):725–728.

Duda, R. O., Hart, P. E., and Stork, D. G. (2000). *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2 edition.

Fink, W. (2004). Neural attractor network for application in visual field data classification. *Physics in Medicine and Biology*, 49(13):2799–2809.

Flach, P. (2012). *Machine Learning: The art and science of algorithms that make sense of data*. Cambridge University Press.

Goldbaum, M. H. (2005). Unsupervised learning with independent component analysis can identify patterns of glaucomatous visual field defects. *Transactions of the Americal Ophthalmological Society*, 103:270–280.

Goldbaum, M. H., Jang, G. J., Bowd, C., Hao, J., Zangwill, L. M., Liebmann, J., Girkin, C., Jung, T. P., Weinreb, R. N., and Sample, P. A. (2009). Patterns of glaucomatous visual field loss in SITA fields automatically identified using independent component analysis. *Transactions of the Americal Ophthalmological Society*, 107:136–144.

Goldbaum, M. H., Sample, P. A., Chan, K., Williams, J., Lee, T. W., Blumenthal, E., Girkin, C. A., Zangwill, L. M., Bowd, C., Sejnowski, T., and Weinreb, R. N. (2002). Comparing machine learning classifiers for diagnosing glaucoma from standard automated perimetry. *Investigative Ophthalmology & Visual Science*, 43(1):162–169.

Goldbaum, M. H., Sample, P. A., White, H., Colt, B., Raphaelian, P., Fechtner, R. D., and Weinreb, R. N. (1994). Interpretation of automated perimetry for glaucoma by neural network. *Investigative Ophthalmololgy & Visual Science*, 35(9):3362–3373.

Henson, D. B., Spenceley, S. E., and Bull, D. R. (1996). Spatial classification of glaucomatous visual field loss. *British Journal of Ophthalmology*, 80(6):526–531.

Henson, D. B., Spenceley, S. E., and Bull, D. R. (1997). Artificial neural network analysis of noisy visual field data in glaucoma. *Artificial Intelligence in Medicine*, 10(2):99–113.

Jain, A. K., Murty, M. N., and Flynn, P. J. (1999). Data clustering: a review. *ACM Computing Surveys*, 31(3):264–323.

Jürgens, C., Koch, T., Burth, R., Schiefer, U., and Zell, A. (2001). Classification of perimetric results and reduction of number of test locations using artificial neural networks [arvo abstract nr. 4539]. *Investigative Ophthalmology & Visual Science*, 42(4):846.

Kaski, S. (1997). Data exploration using self-organizing maps. In *Acta Polytechnica Scandinavica: Mathematics, Computing and Management in Engineering Series No 82*.

Keating, D., Mutlukan, E., Evans, A., McGarvie, J., and Damato, B. (1993). A back propagation neural network for the classification of visual field data. *Physics in Medicine and Biology*, 38(9):1263.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480.

Mangiameli, P., Chen, S. K., and West, D. (1996). A comparison of som neural network and hierarchical clustering methods. *European Journal of Operational Research*, 93(2):402–417.

Mutlukan, E. and Keating, D. (1994). Visual field interpretation with a personal computer based neural network. *Eye*, 8(3):321–323.

Nevalainen, J., Krapp, E., Paetzold, J., Mildenberger, I., Besch, D., Vonthein, R., Keltner, J. L., Johnson, C. A., and Schiefer, U. (2007). Visual field defects in acute optic neuritis - distribution of different types of defect pattern, assessed with threshold-related supraliminal perimetry, ensuring high spatial resolution. *Graefe's Arch Clin Exp Ophthalmol*, 246(4):599–607.

Nouri-Mahdavi, K., Nassiri, N., Giangiacomo, A., and Caprioli, J. (2011). Detection of visual field progression in glaucoma with standard achromatic perimetry: A review and practical implications. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 249(11):1593–1616.

Tafaj, E., Dietzsch, J., Bogdan, M., Schiefer, U., and Rosenstiel, W. (2011a). Reliable classificatin of visual field defects in automated perimetry using clustering. In *Proceedings of the* 8*th IASTED International Conference on Biomedical Engineering*, BIOMED '11.

Tafaj, E., Hempel, S., Heister, M., Aehling, K., Schaeffel, F., Dietzsch, J., Rosenstiel, W., and Schiefer, U. (2013). A New Method for Assessing the Exploratory Field of View (EFOV). In Stacey, D., Sol-Casals, J., Fred, A. L. N., and Gamboa, H., editors, *HEALTHINF 2012*, pages 5–11. SciTePress.

Tafaj, E., Kasneci, G., Rosenstiel, W., and Bogdan, M. (2012). Bayesian online clustering of eye movement data. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, pages 285–288, New York, NY, USA. ACM.

Tafaj, E., Kübler, T., Peter, J., Schiefer, U., Bogdan, M., and Rosenstiel, W. (2011b). Vishnoo - an open-source software for vision research. In *Proceedings of the* 24^{th} *IEEE International Symposium on Computer-Based Medical Systems*, CBMS' 11, pages 1–6. IEEE.

Tan, P. N., Steinbach, M., and Kumar, V. (2005). *Introduction to Data Mining*. Addison Wesley.