Article

Characterization of Pupillary Light Response Features for the Classification of Patients with Optic Neuritis

Stefano Polizzi [1,†], Nico Curti [1,†], Lorenzo Dall'Olio [1], Laura Cercenelli [2,*], Luigi Fontana [3,4], Nicola Valsecchi [3,4], Emanuela Marcelli [2], Gastone Castellani [5,‡] and Piera Versura [3,4,‡]

1   Department of Physics and Astronomy, University of Bologna, 40127 Bologna, Italy
2   eDIMES Lab, Department of Medical and Surgical Sciences, University of Bologna, 40138 Bologna, Italy
3   Ophthalmology Unit, Department of Medical and Surgical Sciences, University of Bologna, 40138 Bologna, Italy
4   IRCCS Azienda Ospedaliero-Universitaria di Bologna, 40138 Bologna, Italy
5   Department of Medical and Surgical Sciences, University of Bologna, 40138 Bologna, Italy
*   Correspondence: laura.cercenelli@unibo.it
†   These authors contributed equally to this work.
‡   These authors contributed equally to this work.

Abstract: Pupillometry is a promising technique for the potential diagnosis of several neurological pathologies. However, its potential is not fully explored yet, especially for prediction purposes and results interpretation. In this work, we analyzed 100 pupillometric curves obtained by 12 subjects, applying both advanced signal processing techniques and physics methods to extract typically collected features and newly proposed ones. We used machine learning techniques for the classification of Optic Neuritis (ON) vs. Healthy subjects, controlling for overfitting and ranking the features by random permutation, following their importance in prediction. All the extracted features, except one, turned out to have significant importance for prediction, with an average accuracy of 76%, showing the complexity of the processes involved in the pupillary light response. Furthermore, we provided a possible neurological interpretation of this new set of pupillometry features in relation to ON vs. Healthy classification.

Keywords: machine learning; pupillometry; neurodegeneration; optic neuritis; multiple sclerosis

1. Introduction

Recently, pupillometric response to different stimuli has been used as a mirror for several brain conditions, pathological or not [1]. The eye, which is embryologically, anatomically, and physiologically an extension of the brain [2], has been an early explored target for the identification of neurodegeneration biomarkers [3].

Optic Neuritis (ON) refers to the inflammation of the optic nerve, and it is a frequent cause of optic nerve injury that can occur in children and adults. Several diseases can determine ON, although it is mainly idiopathic in nature. Infections, autoimmunity, paraneoplastic disorders, granulomatous diseases, and demyelination are some of the causes that can determine ON [4]. Moreover, ON is the earliest clinical symptom in about 25% of cases of Multiple Sclerosis (MS) and occurs in the course of the disease in about 70% [5]. Pupil size is controlled by the parasympathetic constriction pathway and the sympathetic dilation pathway. When light falls on the retina, nerve stimulus is sent along the optic nerve to the optic chiasma, terminating at the pretectal nucleus in the midbrain and not at the lateral geniculate nucleus of the thalamus. From the pretectal nucleus, the stimulus is delivered to the Edinger-Westphal nucleus (EWN) for the efferent pathway that determines the activation of the iris sphincter muscle in both eyes. On the other hand, the dilation pathway starts at the hypothalamus and the locus coeruleus (LC) and connects to the iris dilator muscle. However, the dilatation pathway is less understood than that leading to

constriction because it is regulated by cortical and sub-cortical areas, also implicated in many aspects of cognition. The two pathways are interconnected, as the LC inhibits the EWN. In eyes with ON, a relative afferent pupillary defect (RAPD), or Marcus-Gunn pupil, presents itself with pupillary asymmetry during the swinging flashlight test, in which the pupils constrict less when a flashlight is swung from the unaffected eye to the affected eye, secondary to a defect in the afferent pathway. Thus, the pupillary reflex is a useful indicator of both the afferent and autonomic nervous system function, reflecting both afferent, sympathetic, and parasympathetic activities. Sympathetic innervation determines the initial pupil diameter, whereas parasympathetic activity determines the amplitude and velocity of pupil contraction.

Thus, pupillometry, i.e., the measurement of the pupil size and reactivity to light, is a cheap, easy to apply, and non-invasive technique to study many neurodegenerative diseases and even many characteristic brain states [6], such as concentration [7] and performance in doing complex tasks [8]. In our study, we focused on acute ON.

It is not clear, however, which parameters of the pupil dynamics are useful to consider to assess the brain state, and there is no common agreement between researchers neither on a standard protocol to apply in terms of type and duration of the stimulus, nor on the characteristics that a pupillometric response should have to be considered as pathological [9,10]. Indeed, pupil response is affected by significant inter- and intra-observer variability due to differences in ambient lighting, the intensity of the light stimulus, the method used to direct that stimulus, and the background characteristics of the examined subjects [10], such as eye disorders, age [11], smoking [12], drugs assumption, mental concentration [6], or even the emotional state [13]. This may be the main reason why a general consensus on the procedure is difficult and for neurodegenerative diseases, contrasting results exist [9]. Another debated point is the possibility that abnormalities in the pupillary response can be a premonitory sign that can help in an early diagnosis of the pathology [10,14]; however, one first needs to identify the most important and stable features from pupillometric data before addressing this question and relating the response to the stage of the pathology.

A first possible solution would be, as suggested for instance by Ishikawa [12] in 2021, to adjust the statistical analysis on all the possible confounders. The problem in doing this is that many possible confounders are known to date in the literature and just as many are probably still unknown. Thus, we would lose all the simplicity and convenience, both practical and economical, of pupillometric experiments. For instance, every pupillometric measurement should probably be preceded by a complete ophthalmologic visit, a full questionnaire on the past and present pathologies, a full list of the drugs assumed, and probably even psychological and cognitive tests. Finally, even if all these were possible, other possible confounders, such as the daytime of the experiment or the wakeful state of the subject are simply impossible to control. In our work we applied a different approach, we tried to exploit the maximum of the information contained in our pupillometric data, introducing some new parameters coming from methods used in physics (cinematic and random walk theory) or deeper signal processing (signal transformations, Hilbert transform, peaks detection) of the data, in order to increase the accuracy of the prediction of the pathology, without any additional test.

From a methodological point of view, most of the recent works applying experimental conditions like ours on several neurologic diseases used simple descriptive statistical tests or analysis of variance or covariance [15–18], which were based on standard computerized methods of feature extraction [19]. Some of them used generalized linear regression methods [20], sometimes considering random effects [12], to take into account the intra-subject variability. Very recently, an interesting chromatic pupillometry-based machine learning model has been used on Alzheimer's disease patients, with the goal to study the effect of the disease on several pupillometric parameters at different localizations of the light stimulus, central and peripheral retina [21]. Their work is remarkable because of the use of more advanced techniques, but their experimental data were quite different from ours.

Only a few studies used pupillometry specifically to assess ON patients [18,22–25]. Given the strict connection between ON and MS, sometimes they were studied in concomitance [18,23,26,27]. Overall, it turns out a general slowing down of pupillary response for ON patients, but parameters considered in most studies are not complete and the result is sometimes contradicting, especially on whether parasympathetic (constriction and rest) or sympathetic (dilation) or both pupillary functions are impacted in ON or ON/MS. Even concerning the effect of age and sex. there was no general consensus always associated with the pupillary light response, for instance in [28].

All the works addressing ON or ON/MS used simple comparison tests, ANOVA, ANCOVA, or linear regression to compare patients with controls exclusively from an association viewpoint and the pupillometric parameters the most often considered were the latency time to constriction and the amplitude of constriction. To the best of the authors' knowledge, there are no works, except the seminal one reported previously on Alzheimer's disease [21], using more advanced machine learning methods to exploit the entire information contained in pupillometric data, for instance, computing a more thorough list of parameters.

Here, we compared ON patients with Healthy controls to test a multi-dimension method, with the goal to give a more thorough picture of the pupillary response, possibly mitigating the problem of large variability, and better classification accuracy. This can also improve our understanding of the effect of the pathology on specific pupillary functions. As we will show, some of the features that we introduced significantly improve the prediction of the pathology, mitigating the inter- and intra-subject variability of the data.

We considered as features characterizing the pupillometric sequence both, the main features considered in the literature as affected by common neurodegenerative diseases (Parkinson's disease, Alzheimer's disease, and Multiple Sclerosis) [29], and other new features coming from a deeper analysis of the signal. Since our patients were diagnosed with unilateral ON, we did not address the question of whether pupillometry can help in an early diagnosis of the pathology, nonetheless, our work helps in identifying the most important parameters that must be collected for the prediction of the pathology. The conclusion is that pupillometric data must be considered in its completeness and is very rich in information.

2. Materials and Methods

2.1. Patients Selection

We collected data from 12 subjects, equally divided between ON (50%) and Healthy subjects (50%). The images were retrieved from subjects who gave their voluntary consent to research and signed an informed consent to use data from analysis. All the data were obtained during routine ophthalmological examinations in the Ophthalmic Unit at IRCCS Azienda Ospedaliero-Universitaria di Bologna. The pupillometric signals acquired from the patients were totally anonymized, discarding any metadata about sex, age, and clinical covariates, except by the presence of ON. The lack of these metadata represents one of the main limits of our work, avoiding the possibility of testing covariates' importance in patients' classification. On the other hand, it is also a strength of this study, in that we need to rely exclusively on characteristics of the pupillometric curves and exploit all the information contained therein. All the subjects were examined with the same experimental conditions, using two different protocols of enlightening.

For each patient we collected the pupillometric exams of both eyes, repeating the acquisition for each protocol procedure twice for a more robust estimation of the pupillary light response (PLR) signals.

2.2. Image Acquisition Protocols

We collected the PLR signals via the Corneal Analyser CA-800 corneal topography from Topcon Europe Medical B.V. and the related software, equipped with white LEDs for dynamic and static pupillometry examination. We applied two acquisition protocols

for each patient, varying the exposition time to the lights from 0.5 s to 1 s. The entire acquisition protocol could be summarized as follows: 3 s in scotopic conditions (dark), 0.5 s/1 s of photopic vision (white LEDs source), and finally 12.5 s/12 s of scotopic vision again. In the following, we refer to Protocol 1 as the acquisition related to 0.5 s–12.5 s of photopic/scotopic visions and to Protocol 2 as the acquisition with 1 s–12 s ones. Between each patient and each different measure, the eye was at rest for at least 1 min in a dark room. A representation of the PLR signals analyzed in this work is shown in Figure 1a.
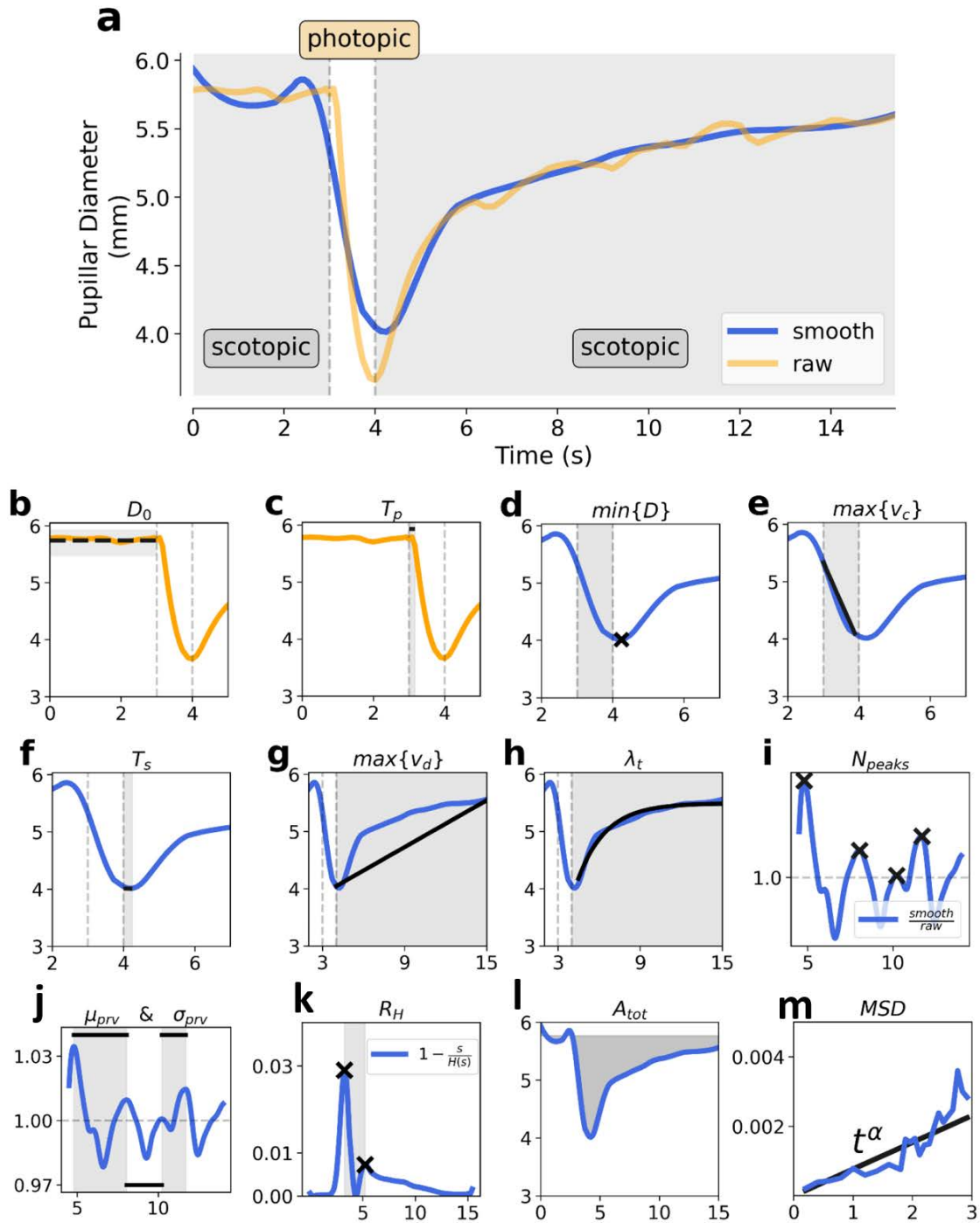

Figure 1. (**a**) Representation of the PLR signal acquired and analyzed in this work, with the highlight of the scotopic and photopic regions. In orange, we show the raw signal obtained by the pupillometer, and in blue the signal obtained by the application of a smoothing algorithm. (**b**–**m**) PLR features extracted and analyzed in this work. In detail, each subplot refers to: the baseline of the raw pupil

diameter (D_0) (**b**); the latency time of constriction (T_P) (**c**) the minimal diameter ($min\{D\}$) (**d**); the maximal contraction velocity ($max\{v_c\}$) (**e**); the latency time (T_s) between light off and minimum diameter (**f**); the maximum re-dilation velocity ($max\{v_d\}$) (**g**); the exponential re-dilation time constant (λ_t) (**h**); the number of oscillations (crosses, $N_{peaks,}$) (**i**); and the average (μ_{prv}) and standard deviation (σ_{prv}) of the time distances between peaks (**j**); the ratio of the two highest peaks (the two crosses) detected from the proportion between the signal and the envelope, (R_H) (**k**); the total area between the signal and the baseline (A_{tot}) (**l**); and the exponent (α) of mean squared displacement (MSD) as a function of time of the pupil center (**m**).

2.3. PLR Features Extraction

The acquisition protocol could be split into two conditions: the scotopic vision, related to the starting and ending points of the acquisition, in which the eye is in a low-light condition, and the photopic one, in which the eye is exposed to a light source. From both these conditions we can extract quantitative features able to describe the response of the pupil in terms of time and velocity. Analogous analyses could be performed considering the transition ranges from these two conditions. Many authors have already proposed several features for the quantitative description of the PLR signals [9,11,12,15,28,29]. In this work we aimed to provide a global overview of the features obtainable by a PLR signal, considering the ones already used in the literature (some of which are in a limited number of works) together with a new set of features, given by the application of signal processing algorithms. All of these new proposed features rely on the quantification of biological interpretable aspects of the PLR signals. A global overview of the features already considered in the literature is shown in Table 1.

Table 1. Summary of the PLR signal features already proposed in the literature. For each study, we reported the number of patients involved, the analyzed disease, and the set of features proposed for the quantitative analysis. All the features cited in this table were considered in the current work, providing a mathematical equation for their evaluation.

Study (*n* = Sample Size)	Case Study	Proposed Feature(s)
Chougule et al. (review)	Alzheimer's disease	• Latency time of pupil constriction
Straub et al. (*n* = 222)	Diabetes	• Latency time of pupil constriction
Ishikawa (*n* = 1406)	Aging, Sex, Smoking	• Minimum pupil diameter • Maximum constriction velocity • Latency time scotopic phase
Fotiou et al. (*n* = 50)	Alzheimer's disease	• Signal enclosing area
Surakka et al. (*n* = 176)	Multiple sclerosis	• Average pupil diameter • Latency time of pupil constriction • Minimum pupil diameter • Maximum constriction velocity • Latency time scotopic phase
Bitirgen et al. (*n* = 25)	Multiple sclerosis	• Minimum pupil diameter • Maximum constriction velocity • Latency time scotopic phase
Yoo (*n* = 30)	Optic Neuritis	• Maximal constriction velocity • Constriction ratio • Latency time of pupil constriction

The most direct features obtainable from the PLR signal (s(t)) are given by the diameter of the pupil at the initial scotopic phase of the acquisition. This measure represents the baseline of the entire signal and could be used to standardize the inter- and intrapatient acquisitions. For its quantification, we used the average value of the PLR signal D_0 obtained in the first scotopic phase, i.e., from the beginning of the acquisition to the point in which the derivative of the signal changes its sign [29] (ref. Figure 1b)

$$D_0 = \frac{1}{N} \sum_1^N s(t) \qquad where\ N \in 1^{st}\ scotopic\ region \tag{1}$$

Starting from the initial scotopic phase, a second feature is given by the latency time of constriction (T_p), at the transition between the scotopic region and the photopic region. It consists of the time difference between the starting time of the photopic region (fixed to 3 s in our acquisitions) and the first response of the pupil ($t_{photopic}$), detected as the time at which the largest second derivative of the signal occurs [9,11,29] (ref. Figure 1c). Identifying the starting point of the photopic phase as described above, the latency time to constriction is given

$$T_p = t_{photopic} - 3s \tag{2}$$

Moving to the photopic region, the s(t) signal will drastically drop-down, achieving its minimum value in terms of pupil diameter [12,29,30] (ref. Figure 1d). The minimum value achieved in this window ($min\{D\}$) was measured as a percentage of the signal centered on the baseline (ref. Equation (1)), rescaling the s(t) with respect to the baseline, often occurring immediately after the switch off of the light:

$$min\ \{D\} = min\{/\ \}_t\ where\ t \in [3s,\ 5s] \tag{3}$$

Another meaningful measure of the pupillary response is given by the maximum absolute velocity ($max\{v_c\}$) with which the minimum value of the pupil diameter is achieved, i.e., the velocity of the rescaled PLR signal estimated in the transition time between the initial scotopic phase and the scotopic one [12,29,30] (ref. Figure 1e).

$$max\ \{v_{contraction}\} = \frac{1}{D_0} max\left\{\left|\frac{ds(t)}{dt}\right|\right\}_t \qquad where\ t \in \left[t_{photopic},\ t_{min\{PLR\}}\right] \tag{4}$$

Moving to the second scotopic phase, we obtained the portion of the PLR signal related to the return to the rest state of the pupil diameter. In this phase, two informative features for the characterization of the PLR signal are given by the latency time (T_s) between the real starting point of the scotopic phase (fixed to 3.5/4 s) and minimum diameter $min\{s(t)\}$, and the maximum velocity to achieve the re-dilation of the pupil ($max\{v_d\}$) [12,29,30] (ref. Figure 1f,g).

$$T_s = [3.5s,\ 4s] - t_{min\{s(t)\}} \tag{5}$$

$$max\ \{v_{dilation}\} = max\ \left\{\frac{ds(t)}{dt}\right\}_t \qquad where\ t \in \left[t_{min\{s(t)\}},\ t_N\right],\ N\ end\ point \tag{6}$$

This second scotopic phase represents the most informative section of the PLR signal and the longest one in terms of time points. The returning phase to the rest conditions $\rho(x)$ could be modeled by an exponential recovery, providing information about the characteristic relaxation time of the process. This exponential relaxation is typical of many living processes involving ion channels in cells, such as when an electric signal is transmitted to the pupil [31,32]. In our work, we tested several fitting functions for the best representation of the process, among which the Michaelis-Menten function, which is used to model several biological processes and the exponential obtained the best results in terms of root-mean-square error. Starting from this hypothesis we computed as PLR feature the characteristic

time (λ_t) of the exponential recovery obtained by fitting the range of points between the $t_{min\{s(t)\}}$ and the endpoint of the acquisition (ref. Figure 1h).

$$\rho(x) = -c_0 e^{-\lambda_t x} + c_1 \tag{7}$$

where c_0 and c_1 are the parameters of the exponential equations.

A deeper analysis of the scotopic phase showed the presence of oscillatory behavior in the recovery section of the signals. This effect can be monitored by applying classical signal processing algorithms to that portion of the signal. We applied a strong smoothing algorithm to the scotopic section, evaluating the ratio between the raw signal and the smoothed one. In this way, we could emphasize the oscillatory behavior, obtaining a signal centered in 1, on which we applied a peak detection algorithm (ref. Figure 1i). We monitored the number of peaks (N_{peaks}) identified on the processed signal as a feature for our study, i.e.,

$$N_{peaks} = \# \, of \, peaks \left\{ \frac{\Lambda(s(t))}{s(t)} \right\}_{scotopic} \quad \text{where } \Lambda \text{ smoothing algorithm} \tag{8}$$

For a better description of this oscillatory behavior, we also added as features the peaks rate variability expressed in terms of the average value (μ_{prv}) of the time distances among peaks (p_i) and corresponding standard deviation (σ_{prv}) (ref. Figure 1j), i.e.,

$$\mu_{prv}(p) = \frac{1}{N_{peaks}} \left(\sum_{i=2}^{N_{peaks}} p_i - p_{i-1} \right) \tag{9}$$

$$\sigma_{prv}(p) = \sqrt{\mu_{prv}(p^2) - \mu_{prv}(p)^2} \tag{10}$$

where p_i ranges in $[1, N_{peaks}]$.

Further analysis of the smoothed signal obtained on the whole PLR signal highlighted the presence of further important features. Computing the absolute value of the Hilbert transform of the smoothed signal $|H(\Lambda(s(t))|$, we obtained the global envelope of the PLR signal. The ratio between the filtered signal $\Lambda(s(t))$ and the envelope represents the demodulated signal, i.e., where all the peaks have the same height. This part of the analysis can be used to better identify peaks and compare different signals and was already used in the literature on other biological signals [33]. During our analyses, we noticed an important difference in the amplitude of these two peaks (namely in position H_0 and H_1) in relation to the two classes under analysis (ON vs. Healthy) (ref. Figure 1k). The quantification of this feature (R_H) is given by

$$R_H = \frac{peak \, \{H_0\}}{peak \, \{H_1\}} \quad \text{where } H = 1 - \frac{\Lambda(s(t))}{|H(\Lambda(s(t)))|} \tag{11}$$

where $peak \, \{H_i\}$ indicates the value of the signal in position H_i.

Considering the total PLR signal, a further obtainable feature is given by the area enclosed between the signal and its baseline (A_{tot}) [15]. This kind of feature takes care of the entire trend of the PLR signal and could provide information that aggregates the entire set of features described up to now. In detail, we estimated the area under the signal using the trapezoid method for approximating the definite integral (ref. Figure 1l), i.e.,

$$A_{tot} = \left| \sum_{k=1}^{N-1} f(t_k) + \frac{f(t_N) + f(t_0)}{2} (t_k - t_{k-1}) \right| \tag{12}$$

where $f(t)$ is the PLR centered on the baseline, i.e., $f(t) = s(t) - D_0$ (ref. Equation (1)).

A final feature obtained from the whole pupillary response, but from an independent type of data collected at the same time as the pupil diameter, is the exponent of mean squared displacement (MSD) as a function of the time of the pupil center (ref. Figure 1m).

Indeed, from the acquired images the net displacement of the center of the pupil can be tracked with time. Therefore, the *MSD* can be compared with the one of standard Brownian motion and it gives some information on the erratic movement of the pupil during the experiment [34,35]. *MSD* is computed from:

$$MSD(\delta t) = \frac{\sum_{\delta t} \vec{r}_c(\delta t)}{N_{\delta t}} \tag{13}$$

where $\vec{r}_c(\delta t)$ is the distance traveled from the pupil center after time δt, averaged over the corresponding number of time intervals ($N_{\delta t}$) present in the time series. From the MSD, then the exponent of its function of time, α, is fitted (ref Figure 1m):

$$MSD(\delta t) = c_0 \delta t^{\alpha} \tag{14}$$

We stress that this feature is independent of pupil diameter.

2.4. Machine Learning Pipeline

We analyzed the entire set of features using a machine-learning pipeline. We split the available samples of PLR curves into a series of train/test sets via stratified 5-fold cross-validation: in this way, we ensured that in each fold an equal number of samples of each class was considered by the model for both the training and validation.

We rescaled the set of features using their median values, normalizing according to the 1st and 3rd quartiles, i.e., a robust scaling algorithm: in this way we minimized the dependency from possible outliers. The parameters for the features standardization were estimated from the training set and propagated to the related test set, avoiding cross contaminations and biases to the model.

The processed set of features was used as input for a penalized classification model, trained for the automated prediction of the two classes (ON vs. Healthy). We used a Ridge classifier for this task, performing a further 10-fold cross-validation on the training set for the estimation and tuning of the best hyperparameters.

The pipeline performances were estimated by monitoring the classification accuracy score (*ACC*) and the Matthews Correlation Coefficient (MCC), given by:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{16}$$

where *TP*, *TN*, *FP*, and *FN* represent the number of True Positives, True Negatives, False Positives, and False Negatives, respectively. The use of the *MCC* score is recommended for binary classification tasks for the quantification of the correct predictions (accuracy) in relation to the numerosity of the two classes [36].

We repeated the entire proposed pipeline for 200 different cross-validation procedures, monitoring the score metrics along them. In this way we could test the robustness of the proposed pipeline in relation to different train/test splits, quantifying its efficiency according to the distribution of the obtained scores.

Feature ranking was assessed by fitting the pipeline on the whole dataset and computing the accuracy drop after a random permutation of each variable. For each variable, 500 different random permutations were simulated to get the average and the corresponding error bars. Since the importance of each variable is contained in the interval [0, 1], to avoid unrealistic confidence intervals outside that range and be more precise, we determined the 95% confidence interval by bootstrapping. For each variable importance, we resampled with replacement 5000 times the original data. Variables with higher importance are then the ones having the most impact on the model prediction and the error bars give information about the statistical significance of each variable, with a *p*-value of 5%.

To assess the performance of the classification at the patient level, we classified a subject as ON if it had at least 2 curves classified as pathologic. We checked the robustness of this condition by using a different number of curves for the classification of the patients. For instance, we considered as well a subject as a case with at least 4 curves classified as pathologic, which is equivalent to a majority vote. The rationale behind our first choice is that one subject was classified as a pathological case if they have at least two PLR curves classified as pathological because pupil response issues may not arise simultaneously on both eyes and may not be detected at the same time by both protocols.

The entire pipeline was written in Python v3.10 using the scikit-learn library [37] for the feature computation as well as for the implementation of the pre-processing steps and the machine-learning model.

3. Results

We processed 100 PLR signals, extracting the set of 13 features from each of them. Due to the low number of samples, the repeated measurements of each patient have been considered independent signals. We tested if the PLR curves obtained from the 2 different protocols were leading to significantly different values of the features, using a Wilcoxon test. Observing that this was not the case, i.e., differences were never significant, we considered the features extracted from both protocols as two sets of values indexed by the same label (ON or Healthy).

A preliminary statistical analysis was performed on the ON signals, verifying the starting assumption that unilateral ON should affect only one of the eyes of the patients, halving the number of available samples for classification. However, in this analysis we monitored the agreement of the features extracted from the ON patients, testing the presence of statistically significant evidence of discrepancies between the behavior of the two eyes. We applied the proposed classification pipeline aiming to classify the pathological eye of the ON population, obtaining not significantly different performances from random classification (*p*-value 0.63). The impossibility of correctly classifying as different the two eyes of ON patients, statistically confirmed the choice of using the PLR obtained by both eyes of ON patients altogether in our main classification task.

The distribution of the metric scores obtained by the ridge classifier along the 200 simulations is shown in Figure 2. The set of features extracted from the PLR signals could correctly classify the PLR signals into ON vs. Healthy, with an accuracy of 0.76 ± 0.02 and a corresponding MCC score of 0.52 ± 0.04. Features ranking, with corresponding 95% bootstrapped confidence intervals are plotted in Figure 2c, showing that all the features except for R_H, are significantly important for the prediction. The patient classification, done as described in Section 2.4, had an average accuracy of 83%, therefore outperforming the classification of the individual PLR curves. Despite the improvement in accuracy, we will not discuss in detail the patient classification, due to the low number of subjects. Indeed, our main analysis was done on the classification of the PLR signals, which gives more reliable results.
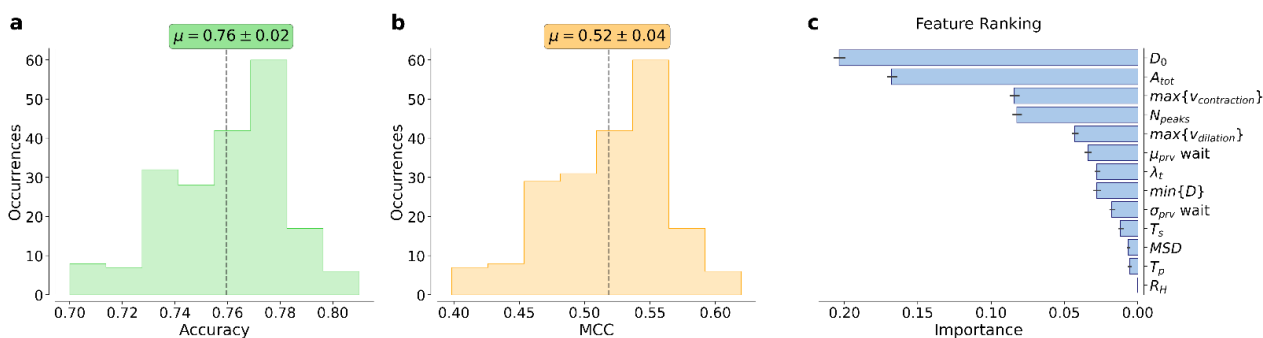
Figure 2. Results obtained by the proposed machine learning pipeline for the automated classification of ON vs. Healthy patients analyzing the set of features extracted from the PLR signals. (**a**) Distribution of the accuracy scores and mean $\pm$ standard deviation, achieved by the proposed pipeline on 200 different 5-fold cross-validations procedures. (**b**) Distribution of the Matthews Correlation Coefficient scores and mean $\pm$ standard deviation, achieved by the proposed pipeline on 200 different 5-fold cross-validations procedures. (**c**) Ranking of the top-performing features identified by the Ridge classifier. The feature importance was assessed by computing the loss of accuracy when permutating each variable. Bars are the 95% confidence intervals of 5000 bootstrap resamples from 500 different random permutations.

4. Discussion

The Ridge classifier used a weighted combination of the extracted features, allowing a ranking of them in relation to their importance in the prediction. The re-iteration of the proposed pipeline for 200 different cross-validations confirms the robustness of the model and the extracted features for the prediction.

The robustness of the extracted features was confirmed also by the possibility to use the entire set of PLR signals belonging to the ON population. Thus, our results support the hypothesis of a bilateral alteration of the pupillary response in the ON population compared to healthy subjects [27]. This result could be explained by different theories. One possible explanation could be the alteration of the pupillary response in both eyes is secondary to a previous or concomitant sub-clinical episode of ON in the contralateral eye, supporting the role of pupillometry as a diagnostic tool in detecting inapparent signs of optic nerve inflammation. However, another explanation could be that patients with unilateral ON might present an alteration of the pupillary response in the contralateral eye, in the absence of an inflammation of the optic nerve, secondary to a reduced afferent stimulus, and a subsequently reduced activation of the efferent pathway. Moreover, because of the strict connection between ON and MS, some of the patients could be in an early stage of the MS pathology, undergoing neurological degeneration in both eyes.

The most informative feature for the classification (20% of mean importance) is the baseline diameter in the first scotopic region. The D_0 feature is simple and easy to acquire since it does not necessarily need a light stimulus. It was already considered in a previous study on Multiple Sclerosis, where ON is the first inflammatory event in 15–20% of patients (ref. Surakka et al. [29]). Multiple Sclerosis patients tend to have smaller baseline values, and then smaller pupil diameter in rest dark conditions, which confirmed, even if with different techniques, the essence of the result of Surakka et al. [29]. This feature measures a balance between the parasympathetic and the sympathetic function [19].

The second most important feature (mean of 17% importance) is the global area under the PLR signal. The A_{tot} feature was designed to capture the overall behavior of the PLR signal, regardless of the scotopic/photopic subdivisions. This result highlights the coexistence of multiple factors in the PLR signal capable of providing information on the classification of ON vs. Healthy patients and the complexity of their interconnections. Indeed, it suggests that the afferent signal is affected in ON, with a subsequent alteration in both the parasympathetic and sympathetic function in a non-trivial way, not easily detectable with previously measured features, such as maximum contraction velocity

($max\{v_{contraction}\}$), minimum relative amplitude ($min\{D\}$) for the parasympathetic ones, and maximal re-dilation velocity ($max\{v_{dilation}\}$) and a generic re-dilation time (λ_t) for the sympathetic ones. To our knowledge, we are the first to introduce this parameter for the study of ON, although it was suggested as a possible parameter by Fotiou et al. [15], but then never used in the following pupillometric studies, probably because it was not therein considered as reliable. It is interesting to note that A_{tot} is instead a very good predictor of ON.

Below the first two features, there is a big drop in the importance of around 9% and we can find, with essentially the same ranking, $max\{v_{contraction}\}$, and N_{peaks}. The first one, as we said previously, is a variable typically collected in pupillometric studies, even on Multiple Sclerosis [12,29,30] and the second one has been introduced in this work as an automatically detected variable. This is again a variable that includes both the sympathetic and parasympathetic nervous system, reflecting the competition between them, because noradrenaline and acetylcholine are both released in the neuromuscular junction of the dilator muscle, when a contraction occurs, and this causes oscillations in absence of a net light stimulus [38,39]. We noticed a prevalence of higher values of N_{peaks} in the Healthy class, highlighting the importance of this oscillatory trend for healthy subjects. Smaller values of the N_{peaks} are a direct consequence of the lack of these oscillations or (at least) of the less pronounced condition of this behavior. The comparable role played by the velocity of pupil contraction, which tends to be smaller in ON patients, is strictly related to the role of N_{peaks}, and therefore is not surprising to obtain an equivalent (within the error bars) contribution to the classification. Indeed, the absence of the oscillatory trend and at the same time reduced contraction velocity could be a sign of the reduced release or sensitivity of/to acetylcholine neurotransmitter in the neuromuscular junction of the dilator muscle, causing a decrease of the inhibition of the dilation necessary for the contraction. However, it is interesting how the classification model used for the forecast tends to give priority to both these features' contributions, confirming their complementarity. This means that the characterization of the observed oscillations is not yet perfectly captured by the proposed features and a combination of them is mandatory for the correct classification.

Then, there is again a drop of around 4% in importance and we find 4 variables with importance between 3% and 5%. Two of them, $max\{v_{dilation}\}$, and $min\{D\}$, are features typically collected in pupillometric studies even on Multiple Sclerosis [12,29,30], and the other two were introduced in this work: $\mu_{prv}(p)$ and λ_t. The first one can be seen as the average waiting time between oscillations, and therefore is a signature again of both parasympathetic and sympathetic functions. The second one can be interpreted as the typical time decay of acetylcholine, causing the restoration of the sympathetic function.

On the bottom of the obtained ranked list of features, we find the last four significant features for the prediction, even though with an importance of less than 2%. Two of them are features introduced in this work (MSD and σ_{prv}), while the others are commonly collected features (T_p and T_s). It is worth noticing however that, even if T_p was considered one of the most associated features related to the pathology, in Multiple Sclerosis [40], Alzheimer's disease [9,11], and diabetes [9,11], it ends up being not very important for the prediction task. The feature obtained from the mean square displacement although not very important for prediction, can be interesting since it is independent of the pupil diameter. Its physical interpretation is that for ON patients the erratic movement of the pupil is slowed down, and with an exponent smaller than the one of healthy subjects, indicating a sub-diffusive movement (the mean α exponent is smaller than one), comparable to diffusion in crowded media. Finally, R_H was the only feature not significant for the prediction of ON among the considered ones.

We would like to stress that the proposed model and consequent results were obtained considering a combination of almost the totality of the features. We did not find equivalent results manually excluding a subset of them, confirming the importance of the whole set of extracted features.

It is important to reiterate that in our study, for the first time, we focused on prediction and not association, and it is possible that features important for prediction are not the most associated with the outcome [41].

This study is not without limitations: the first one being the small number of samples, which represents a limit for results generalization, despite the use of the appropriate statistical methods to control it. Moreover, we could not consider the information of private data of the subjects, but the whole accuracy of the classification was due to the information extracted solely from the pupillometric data, with no need for further, often expensive, and time-consuming examinations.

5. Conclusions

In this work, we proposed a fully automated pipeline for the analysis of PLR signals and classification of ON vs. Healthy patients. We introduced a new set of features able to better characterize the PLR signal, proving their effectiveness in the classification task. We characterized the features extracted according to their significant importance for the classification of ON vs. Healthy (average accuracy of 76% and MCC of 0.52). The most important features for such a task ended up being the baseline diameter in the first scotopic region (namely D_0, 20% of mean importance), and the global area under the PLR signal (namely A_{tot}, 17% of mean importance), after which a considerable decreasing gap in mean importance was observed. Furthermore, a possible neurological interpretation of them was provided in relation to the optic neuritis inflammation. These results suggested that the afferent signal is affected in ON, with an interaction between both the parasympathetic and sympathetic functions, in a non-trivial way. In this work, we provided a guideline for future analysis of PLR signals, showing how ad hoc algorithms could capture interesting characteristics strictly related to neurological responses.

References

1. La Morgia, C.; Carelli, V.; Carbonelli, M. Melanopsin Retinal Ganglion Cells and Pupil: Clinical Implications for Neuro-Ophthalmology. *Front. Neurol.* **2018**, *9*, 1047. [CrossRef] [PubMed]
2. Chow, R.; Lang, R. Early Eye Development in Vertebrates. *Annu. Rev. Cell Dev. Biol.* **2001**, *17*, 255–296. [CrossRef] [PubMed]
3. Javaid, F.Z.; Brenton, J.; Guo, L.; Cordeiro, M.F. Visual and Ocular Manifestations of Alzheimer's Disease and Their Use as Biomarkers for Diagnosis and Progression. *Front. Neurol.* **2016**, *7*, 55. [CrossRef] [PubMed]
4. Bennett, J.L. Optic Neuritis. *Continuum* **2019**, *25*, 1236–1264. [CrossRef]
5. Toosy, A.T.; Mason, D.F.; Miller, D.H. Optic neuritis. *Lancet Neurol.* **2014**, *13*, 83–99. [CrossRef]
6. Hartmann, M.; Fischer, M.H. Pupillometry: The Eyes Shed Fresh Light on the Mind. *Curr. Biol.* **2014**, *24*, R281–R282. [CrossRef]

7. Naber, M.; Alvarez, G.; Nakayama, K. Tracking the allocation of attention using human pupillary oscillations. *Front. Psychol.* **2013**, *4*, 919. [CrossRef]

8. De Rodez Benavent, S.A.; Nygaard, G.O.; Harbo, H.F.; Tønnesen, S.; Sowa, P.; Landrø, N.I.; Wendel-Haga, M.; Etholm, L.; Nilsen, K.B.; Drolsum, L.; et al. Fatigue and cognition: Pupillary responses to problem-solving in early multiple sclerosis patients. *Brain Behav.* **2017**, *7*, e00717. [CrossRef] [PubMed]

9. Chougule, P.S.; Najjar, R.P.; Finkelstein, M.T.; Kandiah, N.; Milea, D. Light-Induced Pupillary Responses in Alzheimer's Disease. *Front. Neurol.* **2019**, *10*, 360. [CrossRef] [PubMed]

10. Martínez-Ricarte, F.; Castro, A.; Poca, M.A.; Sahuquillo, J.; Expósito, L.; Arribas, M.; Aparicio, J. Infrared pupillometry. Basic principles and their application in the non-invasive monitoring of neurocritical patients. *Neurol. Engl. Ed.* **2013**, *28*, 41–51. [CrossRef]

11. Straub, R.H.; Thies, U.; Jeron, A.; Palitzsch, K.D.; Schölmerich, J. Valid parameters for investigation of the pupillary light reflex in normal and diabetic subjects shown by factor analysis and partial correlation. *Diabetologia* **1994**, *37*, 414–419. [CrossRef] [PubMed]

12. Ishikawa, M. Clinical factors affecting pupillary light reflex parameters: A single-centre, cross-sectional study. *Ophthalmic Physiol. Opt.* **2021**, *41*, 952–960. [CrossRef] [PubMed]

13. Fountoulakis, K.; Fotiou, F.; Iacovides, A.; Tsiptsios, J.; Goulas, A.; Tsolaki, M.; Ierodiakonou, C. Changes in pupil reaction to light in melancholic patients. *Int. J. Psychophysiol.* **1999**, *31*, 121–128. [CrossRef]

14. Mary Ann Liebert Inc.; The Brain Trauma Foundation; The American Association of Neurological Surgeons; The Joint Section on Neurotrauma and Critical Care. Pupillary diameter and light reflex. *J. Neurotrauma* **2000**, *17*, 583–590. [CrossRef]

15. Fotiou, F.; Fountoulakis, K.N.; Goulas, A.; Alexopoulos, L.; Palikaras, A. Automated standardized pupillometry with optical method for purposes of clinical practice and research. *Clin. Physiol.* **2000**, *20*, 336–347. [CrossRef]

16. Muppidi, S.; Adams-Huet, B.; Tajzoy, E.; Scribner, M.; Blazek, P.; Spaeth, E.B.; Frohman, E.; Davis, S.; Vernino, S. Dynamic pupillometry as an autonomic testing tool. *Clin. Auton. Res.* **2013**, *23*, 297–303. [CrossRef]

17. Tekin, K.; Sekeroglu, M.A.; Kiziltoprak, H.; Doguizi, S.; Inanc, M.; Yilmazbas, P. Static and dynamic pupillometry data of healthy individuals. *Clin. Exp. Optom.* **2018**, *101*, 659–665. [CrossRef] [PubMed]

18. Yoo, Y.J.; Hwang, J.M.; Yang, H.K. Differences in pupillary light reflex between optic neuritis and ischemic optic neuropathy. *PLoS ONE* **2017**, *12*, e0186741. [CrossRef]

19. Venkata Sivakumar, A.; Kalburgi-Narayana, M.; Kuppusamy, M.; Ramaswamy, P.; Bachali, S. Computerized dynamic pupillometry as a screening tool for evaluation of autonomic activity. *Neurophysiol. Clin.* **2020**, *50*, 321–329. [CrossRef]

20. Frost, S.; Robinson, L.; Rowe, C.C.; Ames, D.; Masters, C.L.; Taddei, K.; Rainey-Smith, S.R.; Martins, R.N.; Kanagasingam, Y. Evaluation of Cholinergic Deficiency in Preclinical Alzheimer's Disease Using Pupillometry. *J. Ophthalmol.* **2017**, *2017*, 7935406. [CrossRef]

21. Lustig-Barzelay, Y.; Sher, I.; Sharvit-Ginon, I.; Feldman, Y.; Mrejen, M.; Dallasheh, S.; Livny, A.; Schnaider Beeri, M.; Weller, A.; Ravona-Springer, R.; et al. Machine learning for comprehensive prediction of high risk for Alzheimer's disease based on chromatic pupilloperimetry. *Sci. Rep.* **2022**, *12*, 9945. [CrossRef]

22. Meethal, N.S.K.; Mazumdar, D.; Morshchavka, S.; Robben, J.; van der Steen, J.; George, R.; Pel, J.J. A haploscope based binocular pupillometer system to quantify the dynamics of direct and consensual Pupillary Light Reflex. *Sci. Rep.* **2021**, *11*, 21090. [CrossRef]

23. Ellis, C.J. The afferent pupillary defect in acute optic neuritis. *J. Neurol. Neurosurg. Psychiatry* **1979**, *42*, 1008–1017. [CrossRef] [PubMed]

24. Pintér, A.; Cseh, D.; Sárközi, A.; Illigens, B.M.; Siepmann, T. Autonomic Dysregulation in Multiple Sclerosis. *Int. J. Mol. Sci.* **2015**, *16*, 16920–16952. [CrossRef]

25. Zheng, D.; Huang, Z.; Chen, W.; Zhang, Q.; Shi, Y.; Chen, J.; Cen, L.; Li, T. Repeatability and clinical use of pupillary light reflex measurement using RAPDx® pupillometer. *Int. Ophthalmol.* **2022**, *42*, 2227–2234. [CrossRef] [PubMed]

26. Bos, J.E. Clinical quantitative pupillometry: The latency of pupillary constriction as measured with the IRIS system (preliminary results). *Neuro-Ophthalmology* **1988**, *8*, 299–305. [CrossRef]

27. Bos, J.E.; Fock, J.M.; Reulen, J.P.H.; Lith, G.H.M.V. Pupil constriction and visual evoked potential latency in optic neuritis. *Neuro-Ophthalmology* **1990**, *10*, 187–196. [CrossRef]

28. De Seze, J.; Arndt, C.; Stojkovic, T.; Ayachi, M.; Gauvrit, J.Y.; Bughin, M.; Saint Michel, T.; Pruvo, J.P.; Hache, J.C.; Vermersch, P. Pupillary disturbances in multiple sclerosis: Correlation with MRI findings. *J. Neurol. Sci.* **2001**, *188*, 37–41. [CrossRef]

29. Surakka, J.; Ruutiainen, J.; Romberg, A.; Puukka, P.; Kronholm, E.; Karanko, H. Pupillary function in early multiple sclerosis. *Clin. Auton. Res.* **2008**, *18*, 150–154. [CrossRef]

30. Bitirgen, G.; Akpinar, Z.; Turk, H.B.; Malik, R.A. Abnormal Dynamic Pupillometry Relates to Neurologic Disability and Retinal Axonal Loss in Patients with Multiple Sclerosis. *Transl. Vis. Sci. Technol.* **2021**, *10*, 30. [CrossRef]

31. Nekouzadeh, A.; Rudy, Y. Statistical properties of ion channel records. Part I: Relationship to the macroscopic current. *Math. Biosci.* **2007**, *210*, 291–314. [CrossRef]

32. Soudry, D.; Meir, R. History-Dependent Dynamics in a Generic Model of Ion Channels—An Analytic Study. *Front. Comput. Neurosci.* **2010**, *4*, 3. [CrossRef] [PubMed]

33. Dall'Olio, L.; Curti, N.; Remondini, D.; Safi Harb, Y.; Asselbergs, F.W.; Castellani, G.; Uh, H.W. Prediction of vascular aging based on smartphone acquired PPG signals. *Sci. Rep.* **2020**, *10*, 19756. [CrossRef]

34. Einstein, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys.* **1905**, *322*, 549–560. [CrossRef]

35. Soula, H.; Caré, B.; Beslon, G.; Berry, H. Anomalous versus Slowed-Down Brownian Diffusion in the Ligand-Binding Equilibrium. *Biophys. J.* **2013**, *105*, 2064–2073. [CrossRef] [PubMed]

36. Chicco, D.; Tötsch, N.; Jurman, G. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Min.* **2021**, *14*, 13. [CrossRef] [PubMed]

37. Buitinck, L.; Louppe, G.; Blondel, M.; Pedregosa, F.; Mueller, A.; Grisel, O.; Niculae, V.; Prettenhofer, P.; Gramfort, A.; Grobler, J.; et al. API design for machine learning software: Experiences from the scikit-learn project. *arXiv* **2013**, arXiv:1309.0238.

38. Lowenstein, O.; Loewenfeld, I.E. Role of sympathetic and parasympathetic systems in reflex dilatation of the pupil: Pupillographic Studies. *Arch. Neurol. Psychiatry* **1950**, *64*, 313–340. [CrossRef]

39. Davson, H. The Pupil. In *Physiology of the Eye*; Macmillan Education: London, UK, 1990; pp. 754–766. [CrossRef]

40. Jakobsen, J. Pupillary function in multiple sclerosis. *Acta Neurol. Scand.* **2009**, *82*, 392–395. [CrossRef]

41. Varga, T.V.; Niss, K.; Estampador, A.C.; Collin, C.B.; Moseley, P.L. Association is not prediction: A landscape of confused reporting in diabetes—A systematic review. *Diabetes Res. Clin. Pract.* **2020**, *170*, 108497. [CrossRef] [PubMed]