Article

Optimal Combination of Mother Wavelet and AI Model for Precise Classification of Pediatric Electroretinogram Signals

Mikhail Kulyabin [*,‡] , Aleksei Zhdanov [‡] , Anton Dolganov , Andreas Maier

from any ideas, methods, instructions, or products referred to in the content.

Article

Optimal Combination of Mother Wavelet and AI Model for Precise Classification of Pediatric Electroretinogram Signals

Mikhail Kulyabin [1,‡,*] , **Aleksei Zhdanov** [2,‡] , **Anton Dolganov** [2] and **Andreas Maier** [1]

1   Pattern Recognition Lab, University of Erlangen-Nuremberg, Erlangen 91058, Germany
2   Engineering School of Information Technologies, Telecommunications and Control Systems, Ural Federal
      University Named after the First President of Russia B. N. Yeltsin, Yekaterinburg 620002, Russia
*   Correspondence: mikhail.kulyabin@fau.de
‡   These authors contributed equally to this work.

Abstract: The continuous advancements in healthcare technology have empowered the discovery, diagnosis, and prediction of diseases, revolutionizing the field. Artificial intelligence (AI) is expected to play a pivotal role in achieving the goals of precision medicine, particularly in disease prevention, detection, and personalized treatment. This study aims to determine the optimal combination of the mother wavelet and AI model for the analysis of pediatric electroretinogram (ERG) signals. The dataset, consisting of signals and corresponding diagnoses, undergoes Continuous Wavelet Transform (CWT) using commonly used wavelets to obtain a time-frequency representation. Wavelet images were used for the training of five widely used deep learning models: VGG-11, ResNet-50, DensNet-121, ResNext-50, and Vision Transformer, to evaluate their accuracy in classifying healthy and unhealthy patients. The findings demonstrate that the combination of Ricker Wavelet and Vision Transformer consistently yields the highest median accuracy values for ERG analysis, as evidenced by the upper and lower quartile values. The median balanced accuracy of the obtained combination of the three considered types of ERG signals in the article are 0.83, 0.85, and 0.88. However, other wavelet types also achieved high accuracy levels, indicating the importance of carefully selecting the mother wavelet for accurate classification. The study provides valuable insights into the effectiveness of different combinations of wavelets and models in classifying ERG wavelet scalograms.

Keywords: biomedical research; electroretinography; electroretinogram; ERG; classification; deep learning; cnn; transformer; wavelet; scalogram

---

1. Introduction

The pediatric electroretinogram (ERG) is a measure of the electrical activity of the retina in response to light stimulation, typically performed on infants and children. The ERG signal consists of a series of positive and negative waveforms, labeled as a-wave, b-wave, and c-wave, which reflect the activity of different retinal cells [1]. The a-wave represents the photoreceptor response, while the b-wave reflects the activity of the bipolar cells and Müller cells [2]. The amplitude and latency of the a-wave and b-wave are commonly used as parameters to evaluate the function of the retina in pediatric patients. Abnormalities in the pediatric ERG signal can be indicative of a range of retinal diseases or disorders, including congenital stationary night blindness, retinitis pigmentosa, and Leber's congenital amaurosis.

The c-wave of the ERG is often considered less prominent and less studied compared to the a-wave and b-wave. The a-wave represents the initial negative deflection, reflecting the hyperpolarization of photoreceptors, while the b-wave represents the subsequent positive deflection, primarily originating from bipolar cells [3]. These waves are well-established and have been extensively researched due to their direct relevance to the visual pathway.

In contrast, the c-wave represents a slower, positive deflection following the b-wave, originating from the retinal pigment epithelium (RPE) and Müller cells [4]. Its amplitude is smaller and its

clinical significance is not as clearly understood [5]. Consequently, it receives less attention in scientific literature and clinical practice.

While the a-wave and b-wave are directly related to photoreceptor and bipolar cell function [6], the c-wave is thought to reflect RPE and Müller cell activity, which are involved in the retinal pigment epithelium-photoreceptor complex. The complex nature of these cells and their functions may contribute to the relatively limited emphasis on the c-wave in the text compared to the a-wave and b-wave. However, further research is needed to fully understand the c-wave's role and its potential clinical applications.

ERG may be of various types, depending on the specific electrophysiological protocol used. One type of ERG signal is the Scotopic 2.0 ERG response, which is obtained under conditions of low light intensity. This response is mainly generated by the rod photoreceptor cells in the retina and is characterized by a relatively slow and low-amplitude waveform. Another type of ERG signal is the Maximum 2.0 ERG response, which represents the maximum electrical response that can be elicited from the retina. This response is obtained under conditions of high light intensity and is mainly generated by the cone photoreceptor cells in the retina. The Maximum 2.0 ERG response is characterized by a faster and higher-amplitude waveform compared to the Scotopic 2.0 ERG response. The Photopic 2.0 ERG response is a third type of ERG signal that is obtained under conditions of moderate light intensity. This response is also mainly generated by the cone photoreceptor cells in the retina and is characterized by a waveform that is intermediate in both amplitude and latency between the Scotopic 2.0 ERG response and the Maximum 2.0 ERG response. It is important to note that the specific electrophysiological protocol used to obtain these different types of ERG signals can vary depending on the research question and the clinical application. Detailed information about the parameters of the electrophysiological study, including the light intensity, wavelength, and duration of the light stimulus, as well as the recording electrodes and amplification settings are shown in a previous study [7].

The Figure 1 depicts pediatric ERG signals of both healthy and unhealthy subjects, along with the designation of the parameters that clinicians analyze. By analyzing the parameters of the ERG waveform, such as the amplitude (a, b) and latency of the a-wave and b-wave (la, lb), clinicians can identify abnormalities and diagnose a range of retinal disorders [8].

In a healthy subject in Figure 1a, the temporal representation of the ERG signal typically exhibits distinct and recognizable waveforms. The signal begins with a negative deflection called the a-wave, which represents the hyperpolarization of photoreceptors in response to light stimulation [1]. Following the a-wave, there is a positive deflection known as the b-wave, which primarily reflects the activity of bipolar cells in the retina. The b-wave is usually larger in amplitude compared to the a-wave [9].

In an unhealthy subject in Figure 1b, the shape of the ERG signal in temporal representation can vary depending on the underlying pathology [10]. In some cases, there may be a significant reduction or absence of both the a-wave and b-wave, indicating a severe dysfunction or loss of photoreceptor and bipolar cell activity [11]. This can be observed in conditions such as advanced retinitis pigmentosa or severe macular degeneration.

Alternatively, certain diseases may selectively affect specific components of the ERG waveform. For example, in some cases of congenital stationary night blindness, the b-wave may be reduced or absent while the a-wave remains relatively normal, indicating a specific defect in bipolar cell function [12,13].

Thus, the shape of the ERG signal in temporal representation provides valuable insights into the integrity and function of retinal cells and can aid in the diagnosis and understanding of various retinal diseases and disorders [14].

In this study, we implemented trained deep-learning models using images of wavelet scalograms to determine the optimal mother wavelet. This approach differs from previous works, where the authors independently searched for features from the time-frequency representation of the signal. It should be noted that deep learning is optimal for image classification because it allows for the

extraction of high-level features from raw data, which can result in higher accuracy. While adult ERGs can be standardized by establishing norms for various parameters, pediatric ERGs are less specific as the amplitude and latency of such ERGs can vary considerably. Consequently, diagnosis often necessitates the use of supplementary diagnostic methods. The study's scientific novelty resides in utilizing deep learning techniques to identify optimal mother wavelet for pediatric ERGs, an approach that can also be applied to other types of ERG signals.
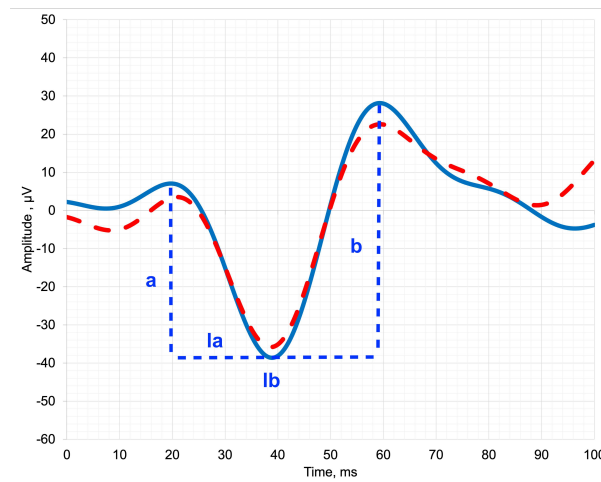
(**a**) Healthy subject.
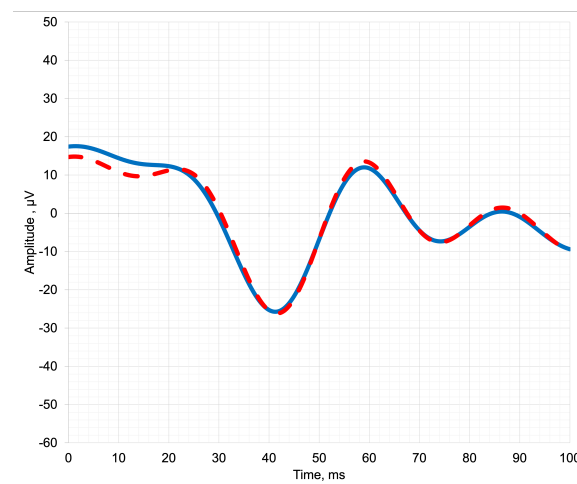
(**b**) Unhealthy subject.

Figure 1. Dark-adapted full-field electroretinogram: (**a**) healthy subject; (**b**) unhealthy subject.

The Related Works section explores prior studies using diverse wavelets to analyze adult ERG data and a recent study utilizing deep learning to identify the optimal mother wavelet for pediatric ERGs. In the Materials and Methods section, we describe an approach to address ERGs class imbalance through under-sampling the majority class, applying wavelet transformation, and training five deep learning models for ERGs classification. The Results section presents findings from the experiment, demonstrating the highest median accuracy with the Ricker Wavelet and Vision Transformer combination for ERG wavelet scalogram classification. The Discussion section addresses limitations and emphasizes the significance of expanding the feature space via continuous wavelet transform for effective classification. Finally, the Conclusions highlight the efficacy of the combination of Ricker Wavelet and Vision Transformer in achieving high accuracy for ERG wavelet scalogram classification.

2. Related Works

Wavelet analysis has been widely used to study ERG in the field of ophthalmology. In recent publications shown in Table 1, the selection of the mother wavelet has been motivated by various factors. In the study presented in [15], mother wavelets were optimized to analyze normal adults' ERG waveforms by minimizing scatter in the results. This approach led to improved accuracy and allowed for a more precise analysis of the data.

Different wavelets emphasize various features of a signal, making it crucial to choose the most appropriate mother wavelet. In a previous study [16], researchers conducted a preliminary analysis and concluded that the Ricker wavelet was the best fit for their waveforms due to its conformity to the shape of the adult ERG data. Similarly, another study [17] suggested that the Morlet wavelet was appropriate for adult ERG analysis, although there is still no consensus on the optimal mother wavelet. The aforementioned articles successfully addressed the classification problem and provided frequency pattern estimates for ERG.

The use of the Morlet wavelet transform in ERG analysis has been shown to provide a more comprehensive analysis of the data. For example, in [18], the Morlet wavelet transform was used for the first time to quantify the frequency, peak time, and power spectrum of the OP components of the adults ERG, providing more information than other wavelet transforms.

In [19], the aim was to classify glaucomatous and healthy sectors based on differences in frequency content within adults ERG using the Morlet wavelet transform and potentially the CWT. This approach could improve discrimination between normal and abnormal waveform signals in optic nerve diseases, which is essential for accurate diagnosis and treatment.

Finally, in [20], the Gaussian wavelet was chosen for its convenience in pediatric and adult ERG semi-automatic parameter extraction and better time domain properties. However, challenges remain in achieving simultaneous localization in both the frequency and time domains, indicating a need for further improvement in wavelet analysis techniques.

In summary, the selection of the mother wavelet plays a crucial role in ERG analysis, and various factors should be taken into consideration to ensure an accurate and comprehensive analysis of the data.

Table 1. Comparative table of used mother wavelets for CWT and studied signals (subjects).

Year	First Author and Reference	Mother Wavelet	Number of Signals (Subjects)
2005	Penkala [16]	Morlet Wavelet, Ricker Wavelet	120 (N/A)
2007	Penkala [15]		102 (N/A)
2010	Barraco [21]	Ricker Wavelet	24 (N/A)
2011	Barraco [22]		N/A (10)
2011	Barraco [23]		N/A (10)
2014	Gauvin [19]	Morse Wavelet	N/A (40)
2014	Dimopoulos [18]	Morlet Wavelet	N/A (63)
2015	Miguel-Jiménez [24]		N/A (47)
2020	Ahmadieh [17]		N/A (36)
2022	Zhdanov [20]	Gaussian Wavelet	425 (N/A)

3. Materials and Methods

3.1. Dataset Balancing

In this study signals from the IEEEDataPort repository were used, which is a publicly accessible ophthalmic electrophysiological signals database [25]. The dataset encompasses three types of pediatric signals: Maximum 2.0 ERG Response, Scotopic 2.0 ERG Response, and Photopic 2.0 ERG Response.

Table 2 presents the Unbalanced Dataset column, which shows the number of signals in the dataset that belong to the healthy and unhealthy classes. The table reveals that the classes are imbalanced. To address this issue, the Imbalanced-learn package [26] was chosen as the solution, which has been utilized by researchers to solve such class imbalance issues. It is noteworthy that only pediatric signals were utilized, as they are the most representative and obviate the need for artificially generated signals.

Table 2. Comparative table of ERG signals before and after balancing.

Unbalanced Dataset		Balanced Dataset	
healthy	unhealthy	healthy	unhealthy
Maximum 2.0 ERG Response			
143	60	62	60
Scotopic 2.0 ERG Response			
52	48	52	48
Photopic 2.0 ERG Response			
171	68	68	63

An under-sampling technique was employed using the AllKNN function from the Imbalanced-learn package [26,27]. The AllKNN function uses the nearest neighbor algorithm to identify samples that contradict their neighborhood. The classical significant features of ERG signals were used as input to this function. To ensure the effectiveness of the nearest neighbor algorithm, the choice of the number of nearest neighbors to be considered is crucial. In our study, we use 13 as the number of nearest neighbors to achieve the desired class balance. The hyperparameter is chosen empirically: higher and lower values of nearest neighbors either remove too much data or do not remove it enough. The pairplot of the ERG signals distributions, presented in Figure 2, illustrates the results of this under-sampling technique, where orange and blue colors correspond to the healthy and unhealthy classes, respectively. It should be noted that the Scotopic signals were already balanced and did not require any under-sampling.

Thus, dataset balancing was implemented. Table 2 presents the distribution of healthy and unhealthy subjects within balanced dataset. In this work, we use balanced dataset for training experiments.
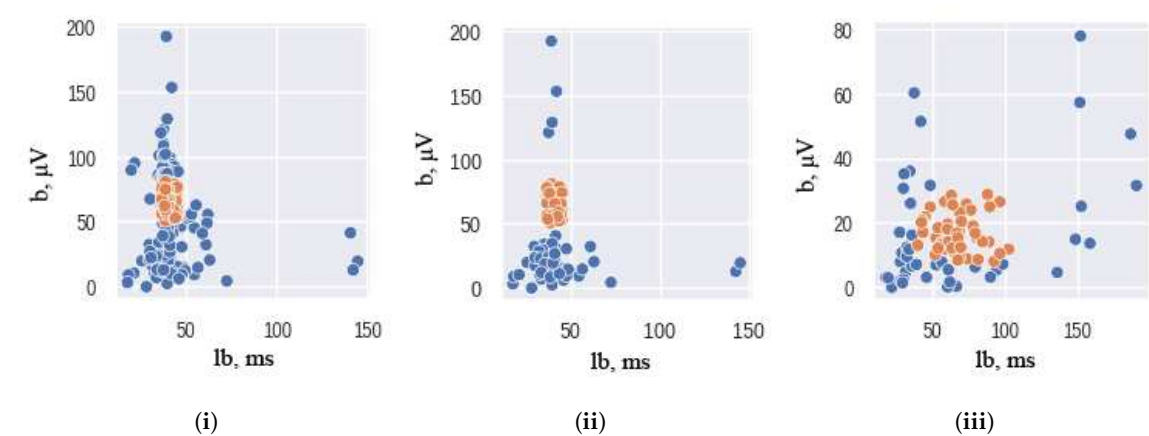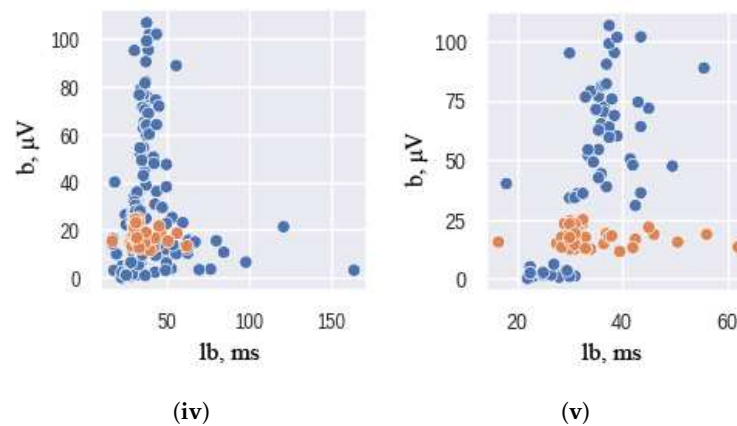

(i)　　　　　　　　**(ii)**　　　　　　　　**(iii)**

Figure 2. *Cont.*

Figure 2. Scatterplot Visualizations of the ERG signal classical features: Maximum 2.0 ERG Response - before (**i**) and after (**ii**) under-sampling; Scotopic 2.0 ERG Response - (**iii**); Photopic 2.0 ERG Response - before (**iv**) and after (**v**) under-sampling. Here b – is the amplitude of the b-wave (µV), lb is the latency of the b-wave (ms).

3.2. Training Pipeline

Figure 3 shows the training pipeline encompassing five distinct stages. During the Initial stage, the ERG signal dataset that is acquired in a time-domain representation is balanced. Subsequently, at the Transformation stage, the signal undergoes wavelet transformation, leading to a frequency-time representation, and then stored in image classification dataset format.

Further, we split the train and the test subsets: undersampling the test set can lead to a biased evaluation of a model's performance, which could be detrimental in real-world scenarios. Therefore, the test set represents the real-world distribution of healthy and unhealthy samples with an 85:15 ratio to the training subset.

At the following stages of Training and Cross-validation, the wavelet scalogram images are subjected to classification utilizing the training and validation datasets. Ultimately, the efficiency of the image classification process is assessed in the Evaluation stage using the balanced metrics.
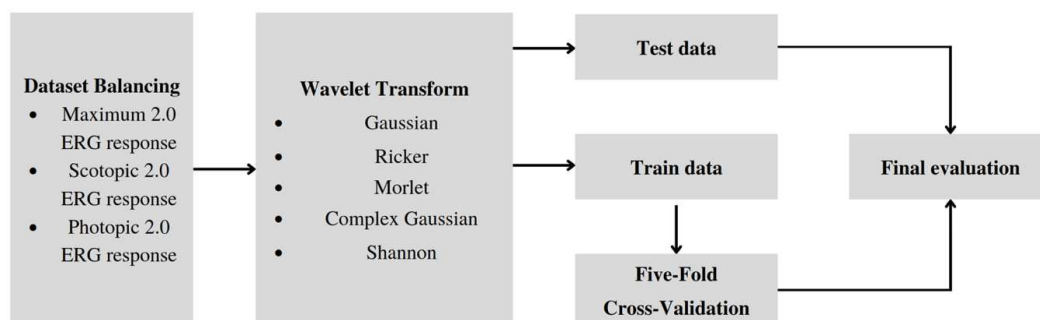
Figure 3. Training Pipeline: dataset balancing, wavelet transformation, splitting the data on test and train datasets, Cross-Validation and final evaluation of the model.

3.2.1. Data Preprocessing

The dataset under investigation contains signals comprising of 500 entries each, alongside their corresponding target (diagnosis). To perform an analysis, CWT was carried out on each signal using the PyWavelets library [28]. The base functions employed in this study were the commonly used ones, namely Ricker, Morlet, Gaussian Derivative, Complex Gaussian Derivative, and Shannon. The scaling parameters were adjusted to generate 512x512 gray-scale images.

3.2.2. Baseline

The training was independently performed on 5 widely used architectures in the field of deep learning: VGG-11, ResNet-50, DensNet-121, ResNext-50 and Vision Transformer. The choice of models above was based on their popularity and proven effectiveness in the field of image classification. These models have been widely used in various computer vision tasks, and their performance has been thoroughly evaluated on standard datasets such as ImageNet [29–31]. In particular, using Vision Transformer it is possible to investigate the effectiveness of this newer architecture compared to the more established models [32].

VGG-11 is one of the most popular pre-trained models for image classification. Introduced in ILSVRC 2014 Conference, it remains the model to beat even today [33]. VGG-11 contains seven convolutional layers, each followed by a ReLU activation function, and five max pooling operations.

Residual Network (ResNet) is a specific type of convolutional neural network introduced in the paper "Deep Residual Learning for Image Recognition" by He Kaiming, et al., 2016 [34]. ResNet-50 is a 50-layer CNN that consists of 48 convolutional layers, one MaxPool layer, and one average pool layer.

ResNext-50 is a simple, highly modularized network architecture for image classification. It was constructed by repeating a building block that aggregates a set of transformations with the same topology [35].

DenseNet name refers to Densely Connected Convolutional Networks developed by Gao Huang, et al. in 2017 [36]. In this work, we used DensNet-121 that consists of 120 Convolutions and 4 AvgPool layers.

Vision Transformer is a model for image classification that employs a Transformer-like architecture over patches of the image. An image is split into patches with fixed size, each of the patches then linearly embedded, position embeddings are added, and the resulting sequence of vectors is fed to Transformer encoder [37]. In current work, we used ViT_Small_r26_s32_224 model, pretrained on a large collection of images in a supervised fashion, namely ImageNet-21k, at a resolution of 224x224 pixels.

We used ADAM optimization with 0.001 initial learning rate. Each model was trained until convergence using the early stopping criteria on the validation loss with batch size of 16 on a single NVIDIA V100.

3.2.3. Loss function

The loss function plays a critical role in deep learning. In this work, we utilize the most commonly used Cross-entropy loss function for classification tasks [38], which represents negative log-likelihood of a Bernoulli distribution (1):

$$CE(\tilde{y}, \hat{y}) = -\frac{1}{N} \sum_{i=1}^{N} \tilde{y}_i \log(\hat{y}_i),$$

(1)

where

- $\tilde{y}$ - one-hot encoded ground truth distribution,
- $\hat{y}$ - predicted probability distribution,
- N - the size of the training set.

3.2.4. Data Augmentation

The incorporation of data augmentation techniques during the training process leads to an augmentation of the distributional variability of input images. This augmentation is known to enhance the resilience of models by increasing their capacity to perform well on a wider range of inputs. Given the characteristics of our dataset, we opted to apply exclusively geometric transformations such as random cropping, vertical flipping, and image translation to the images under consideration.

3.2.5. Cross Validation

Cross-validation is a resampling method that is employed to assess the effectiveness of deep learning models on a dataset with limited samples. The technique entails partitioning the dataset into k groups. To account for the limited nature of our dataset and facilitate a more objective evaluation of the trained models, we applied a five-fold cross-validation strategy in the present study. The test subset was first separated according to the real-world distribution of healthy and unhealthy clinical patients for each type of ERG response. The remaining shuffled training subset was then divided into five folds of which one is used for validation and four for training. The process is repeated for five experiments, using every fold once as the validation set.

3.2.6. Evaluation

For each experiment, a confusion matrix was constructed using the test dataset, and the evaluation metrics were subsequently computed [39]. This approach enabled the accurate assessment of the model's performance across different folds, thus ensuring a more comprehensive evaluation. Additionally, the confusion matrix provides a detailed overview of the model's performance, highlighting the number of correct and incorrect predictions made by the model.

For a complete understanding of the model performance, several metrics were computed: Precision, Recall, and F1 score [40]:

$$Precision = \frac{TP}{TP + FP}, \tag{2}$$

$$Recall = \frac{TP}{TP + FN}, \tag{3}$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall}, \tag{4}$$

where

- *TP = True Positive*,
- *FP = False Positive*,
- *FN = False Negative*.

Since the test subset reflects the real-world distribution and is not balanced, we should consider Balanced Accuracy [41]:

$$Balanced\ Accuracy = \frac{Sensitivity + Specificity}{2}, \tag{5}$$

where

$$Sensitivity = Recall = \frac{TP}{TP + FN}, \tag{6}$$

$$Specificity = \frac{TN}{TN + FP}. \tag{7}$$

4. Results

Figure 4 shows box-plot distributions of classification accuracy where a is Maximum 2.0 ERG Response, b - Scotopic 2.0 ERG Response, c - Photopic 2.0 ERG Response.

The findings from Figure 4a suggest that the Ricker Wavelet combined with the Vision Transformer produces the highest classification accuracy for ERG wavelet scalograms. Specifically, the median balanced accuracy value is 0.83, with the upper quartile being 0.85 and the lower quartile being 0.8. In contrast, when utilizing the Shannon Wavelet, the median accuracy value is 0.8, with the upper quartile at 0.82 and the lower quartile at 0.77. Additionally, for the Morlet Wavelet, the median balanced accuracy value is 0.78, with the upper quartile at 0.81 and the lower quartile at 0.77.
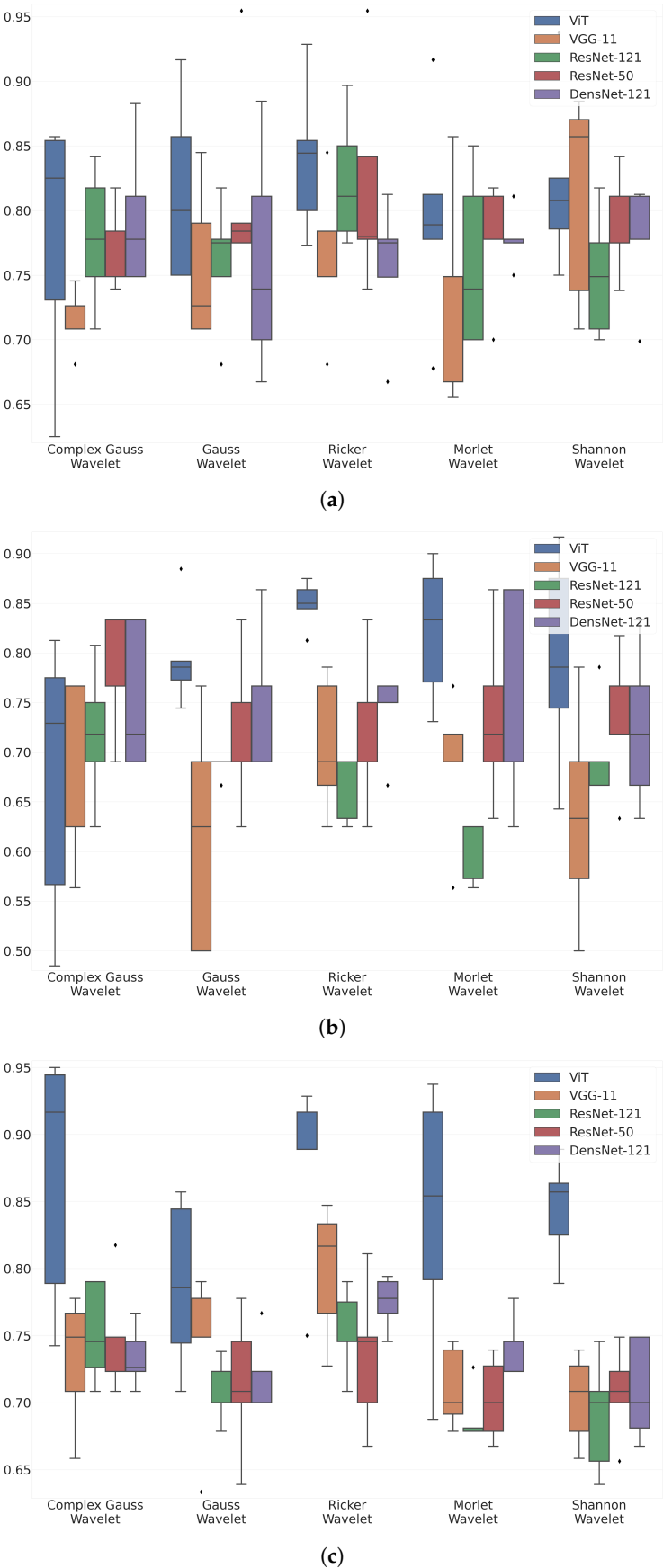
Figure 4. Box-plot distributions of classification accuracy: (a) Maximum 2.0 ERG Response; (b) Scotopic 2.0 ERG Response; (c) Photopic 2.0 ERG Response.

The findings from Figure 4b suggest that the Ricker Wavelet combined with the Vision Transformer produces the highest classification accuracy for ERG wavelet scalograms. Specifically, the median balanced accuracy value is 0.85, with the upper quartile being 0.87 and the lower quartile being 0.84. In contrast, when utilizing the Morlet Wavelet, the median accuracy value is 0.82, with the upper quartile at 0.86 and the lower quartile at 0.76. Additionally, for the Gaussian Wavelet, the median balanced accuracy value is 0.78, with the upper quartile at 0.79 and the lower quartile at 0.77.

The findings from Figure 4c suggest that the Ricker Wavelet combined with the Vision Transformer produces the highest classification accuracy for ERG wavelet scalograms. Specifically, the median balanced accuracy value is 0.88, with the upper quartile being 0.92 and the lower quartile being 0.88. In contrast, when utilizing the Shannon Wavelet, the median accuracy value is 0.86, with the upper quartile at 0.87 and the lower quartile at 0.83. Additionally, for the Morlet Wavelet, the median balanced accuracy value is 0.85, with the upper quartile at 0.93 and the lower quartile at 0.79.

More detailed values of the metrics are given in the appendix of the article in Tables A1–A3.

The results from Figure 4a–c provide valuable insights into the accuracy of ERG wavelet scalogram classification. In all three cases, the Ricker Wavelet combined with the Vision Transformer yielded the highest median accuracy values, demonstrating the effectiveness of this combination. Additionally, the upper and lower quartile values further support the superiority of this approach, showing consistently high accuracy levels. However, it is important to note that the performance of other wavelet types should not be overlooked. For example, when utilizing the Shannon Wavelet in Figure 4b, a median balanced accuracy value of 0.82 was achieved, which is still a relatively high level of accuracy. Similarly, in Figure 4c, the Morlet Wavelet produced a median balanced accuracy value of 0.85, which is also noteworthy. Overall, these findings suggest that the selection of the mother wavelet plays a critical role in determining the accuracy of the classification of ERG wavelet scalograms. By carefully selecting the most appropriate wavelet type and transformer architecture, it may be possible to achieve even higher accuracy levels, thereby advancing our understanding of ERG wavelet scalogram classification.

The selection of the mother wavelet in ERG analysis is of utmost importance as it directly influences the quality and interpretability of the results. Choosing an appropriate mother wavelet requires careful consideration of various factors to ensure accurate and comprehensive data analysis. Despite the abundance of literature on ERG analysis, there is a lack of a clearly formulated motivation for selecting a specific mother wavelet. Existing sources often fail to provide explicit reasoning or guidelines for choosing one wavelet over another in the context of ERG analysis. This gap in the literature hinders researchers from making informed decisions regarding the selection of the most suitable mother wavelet for their ERG data analysis, highlighting the need for further research and guidance in this area.

The dataset comprised signals with 500 entries each, and CWT was applied using the PyWavelets library to generate 512x512 gray-scale images. The CWT was performed using different base functions, including Ricker, Morlet, Gaussian Derivative, Complex Gaussian Derivative, and Shannon.

Deep learning architectures, namely VGG-11, ResNet-50, DenseNet-121, ResNext-50, and Vision Transformer, were independently trained on the dataset to establish baselines for performance comparison. These models have been extensively evaluated in computer vision tasks and have shown effectiveness in image classification. Training was carried out until convergence using the ADAM optimization with an initial learning rate of 0.001 and a batch size of 16 on a single NVIDIA V100.

The performance of the models was evaluated using cross-validation with a five-fold strategy to account for the limited sample size. The dataset was divided into training and test subsets, and the training subset was further partitioned into five folds for validation. Confusion matrices were constructed using the test dataset, enabling the computation of evaluation metrics such as precision, recall, F1 score, and balanced accuracy. This approach provided a comprehensive assessment of the models' performance across different folds and allowed for a detailed analysis of correct and incorrect predictions made by the models.

The results demonstrated the effectiveness of the deep learning models in classifying the signals and diagnosing the corresponding conditions. Overall, the Vision Transformer model exhibited the highest performance, achieving the highest precision, recall, F1 score, and balanced accuracy among the tested architectures. The VGG-11, ResNet-50, and DenseNet-121 models also displayed strong performance, while the ResNext-50 model achieved slightly lower metrics. These findings highlight the potential of deep learning models, particularly the Vision Transformer architecture, in analyzing signal datasets and facilitating accurate diagnoses. The study's results contribute to the understanding of the applicability of deep learning techniques in medical diagnostics and pave the way for future research in this domain.

5. Discussion

The choice of wavelet for ERG signal analysis depends on the waveform characteristics, with different wavelets having varying frequencies and temporal resolutions. An optimal wavelet should possess effective noise suppression capabilities [42], accurately capture transient and sustained components of the ERG signal, and provide interpretable coefficients for feature identification. Computational efficiency is important for handling large datasets and real-time applications. Additionally, the familiarity and expertise of the researcher or clinician in interpreting specific wavelets can enhance the accuracy and efficiency of the analysis. Careful wavelet selection is crucial to ensure reliable and meaningful results in clinical and research settings [43].

The Ricker Wavelet yielded the highest median accuracy values for ERG wavelet scalogram classification due to the following potential reasons:

1. Wavelet characteristics: the specific properties of the Ricker Wavelet, including its shape and frequency properties, align well with the features present in ERG wavelet scalograms, leading to improved accuracy in classification compared to other wavelet types.
2. Noise suppression capabilities: the Ricker Wavelet demonstrates superior noise suppression capabilities, effectively reducing unwanted noise in ERG wavelet scalograms while preserving important signal components, resulting in enhanced accuracy.
3. Time-frequency localization: the Ricker Wavelet excels in accurately localizing transient and sustained components of ERG waveforms across different time intervals, enabling better capture and representation of crucial temporal features, thereby increasing the discriminative power of the wavelet in classifying ERG responses.

It should be noted that the present study utilized a limited set of signals to identify the most appropriate mother wavelet for ERG analysis. Nevertheless, the sample was well-balanced, which lends confidence to the relatively stable classification outcomes.

In comparison to electrophysiological data, which typically contain numerous parameters describing motor function, ERG analysis necessitates the addition of significant parameters to ensure the efficient classification of specific states. As ERG analysis involves only four parameters [44], insufficient for precise diagnosis, expanding the feature space via continuous wavelet transform in the frequency-time domain is essential.

Selected neural networks may exhibit superior performance when trained on larger datasets. For instance, the accuracy distribution of the Transformer model displays a wide range [45], however, this variability would likely be reduced with an increase in the size of the training dataset. Moreover, it was essential to divide and keep the test data without modification based on the distribution observed in real-world scenarios, which affected the quantity of training data that is available.

As the ERG signals are equipment- and intensity-specific, it is important to exercise caution when combining different datasets. An area of research that holds promise involves the creation of synthetic signals, which can augment the available training data.

This research investigates the potential of AI algorithms in accurately classifying eye diseases and acknowledges their role as supportive tools to medical specialists. While the algorithms demonstrate

good accuracy, we emphasize the indispensability of specialist involvement. The complex nature of human health, the significance of empathetic care, and the unparalleled decision-making capabilities of doctors underscore their ongoing essential role in delivering comprehensive and holistic medical care. Classification algorithms can serve as clinical decision support systems, enhancing physicians' expertise and facilitating more efficient and effective healthcare delivery [46,47].

5.0.1. Limitations

Equipment limitations: The utilization of only the Tomey EP-1000 equipment for ERG registration introduces a potential limitation to the generalizability of the model's results. As different equipment may exhibit variations in signal acquisition and measurement precision, the model's performance and outcomes may differ when applied with alternative ERG registration devices. Furthermore, the use of a corneal electrode during ERG registration does not entirely eliminate the possibility of electrooculogram-induced noise stemming from the involuntary movement of the eye muscles [3]. Consequently, the presence of such noise may impact the accuracy and reliability of the obtained ERG signals, influencing the model's performance.

Study protocol considerations: The employment of specific ERG protocols, namely Maximum 2.0 ERG Response, Scotopic 2.0 ERG Response, and Photopic 2.0 ERG Response, within this investigation, ensures the acquisition of ERG recordings with optimal quality on the employed equipment [48]. However, it is crucial to acknowledge that altering the study protocol, such as modifying the brightness or timing of light stimuli, could yield varying results when utilizing the model. Changes in these protocol parameters may introduce variations in the recorded ERG signals, potentially impacting the model's predictive performance and its ability to generalize to different experimental conditions or protocols.

Dataset limitations: The dataset used in this study comprises data from both healthy subjects and subjects with retinal dystrophy [49]. While the inclusion of healthy subjects provides a baseline for comparison, the focus of the model's training and evaluation is on the detection and diagnosis of retinal dystrophy. Therefore, it is important to note that this specific model is designed and optimized for the diagnosis of retinal dystrophy and may not be applicable for the diagnosis of other diseases or conditions. The model's performance and generalizability to other diseases should be assessed separately using appropriate datasets and evaluation protocols.

Neural network feature limitations: While neural networks have shown improved performance with larger datasets, it is essential to consider the limitations associated with the available data. The accuracy of the chosen neural networks, especially the Transformer, increases with an increase in the training subset. However, the available datasets are limited, and the different nature of the origin of the signals makes it difficult to combine such datasets.

To overcome equipment limitations in ERG recordings, the use of Erg-Jet electrodes can be considered [50]. These electrodes help minimize noise caused by the movement of eye muscles, thus improving the quality of the recorded signals. Additionally, standardizing the study protocol across different equipment setups can help address the issue of using different equipment [44].

Regarding study protocol limitations, the adoption of standardized protocols ensures a uniform approach to ERG recordings. Furthermore, presenting detailed statistical data about the dataset used in the study helps address study protocol limitations by providing transparency and enabling researchers to evaluate the robustness and generalizability of the findings.

To mitigate dataset limitations, it is crucial to expand the ERG dataset and include a broader range of retinal diseases. By increasing the size and diversity of the dataset, researchers can enhance the representativeness of the data and improve the model's ability to generalize to various clinical scenarios. Incorporating additional retinal diseases beyond the scope of the current study would enable a more comprehensive understanding of the model's performance and its applicability to a wider range of clinical conditions. Expanding the dataset can also help identify potential subgroups or rare conditions that may have specific ERG characteristics, contributing to the advancement of

knowledge in the field. A promising development direction is the generation of synthetic ERG signals: combining the mathematical model of the signal and Generative Adversarial Networks [51]. This will increase the training subset with signals of a similar origin.

6. Conclusions

The results of this study indicate that the combination of Ricker Wavelet combined with Vision Transformer consistently achieves the highest median balanced accuracy values across all three ERG responses: 0.83, 0.85, and 0.88 consequently. The robust upper and lower quartile values provide compelling evidence for the superiority of this combination, consistently demonstrating high accuracy levels. However, it is important to acknowledge that other wavelet types also yield relatively high accuracy levels and should not be disregarded. These findings underscore the critical role of selecting an appropriate mother wavelet in determining the accuracy of ERG wavelet scalogram classification. Careful consideration of the wavelet type and transformer architecture holds significant potential for attaining even higher levels of accuracy. Overall, this study offers valuable insights into the effectiveness of different wavelet-model combinations, thereby contributing to the precise classification of pediatric ERG signals and advancing the field of healthcare.

The findings of this study will be utilized to develop an AI-based decision support system in ophthalmology, leveraging the insights gained from ERG wavelet scalogram classification. This system aims to enhance ophthalmology-related applications by incorporating accurate and efficient analysis of ERGs. Furthermore, the results of this study may hold value for manufacturers of electrophysiological stations used in ERGs. The understanding of which wavelet types, such as the Ricker Wavelet, yield superior classification accuracy may guide the development and optimization of electrophysiological stations, enabling them to provide more reliable and advanced diagnostic capabilities in the field of ophthalmology.

Author Contributions: Conceptualization, A.Z. and M.K.; methodology, A.Z. and M.K.; software, M.K. and A.D.; validation, M.K., A.Z., and A.M.; formal analysis, M.K.; investigation, A.D.; writing-original draft preparation, A.Z. and M.K.; writing-review and editing, A.M.; visualization, M.K. and A.D.; supervision, A.M.; project administration, A.Z.; funding acquisition, A.Z. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Zhdanov, A.E.; Dolganov, A.Y.; Borisov, V.I.; Lucian, E.; Bao, X.; Kazaijkin, V.N.; Ponomarev, V.O.; Lizunov, A.V.; Ivliev, S.A. 355 OculusGraphy: Pediatric and Adults Electroretinograms Database, 2020. https://dx.doi.org/10.21227/y0fh-5v04 (accessed on 29 November 2022).

Conflicts of Interest: The authors declare no conflict of interest.

Table A1. Maximum 2.0 ERG Response Metrics Table with Wavelet Function Variations.

VGG-11				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.714	0.885	0.745	0.65
Gaussian Wavelet	0.756	0.86	0.771	0.73
Ricker Wavelet	0.762	0.819	0.81	0.82
Morlet Wavelet	0.719	0.783	0.795	0.82
Shannon Wavelet	0.812	0.773	0.835	0.92

Table A1. *Cont.*

ResNet-50				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.779	0.858	0.841	0.83
Gaussian Wavelet	0.76	0.86	0.828	0.8
Ricker Wavelet	0.823	0.874	0.881	0.89
Morlet Wavelet	0.76	0.861	0.827	0.8
Shannon Wavelet	0.75	0.845	0.826	0.81
DensNet-121				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.779	0.858	0.841	0.83
Gaussian Wavelet	0.76	0.86	0.828	0.8
Ricker Wavelet	0.823	0.874	0.881	0.89
Morlet Wavelet	0.76	0.861	0.827	0.8
Shannon Wavelet	0.75	0.845	0.826	0.81
ResNext-50				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.768	0.856	0.835	0.82
Gaussian Wavelet	0.816	0.873	0.869	0.87
Ricker Wavelet	0.819	0.845	0.869	0.9
Morlet Wavelet	0.777	0.866	0.847	0.83
Shannon Wavelet	0.788	0.846	0.857	0.87
Vision Transformer				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.778	0.74	0.812	0.895
Gaussian Wavelet	0.815	0.658	0.77	0.891
Ricker Wavelet	0.84	0.727	0.802	0.867
Morlet Wavelet	0.795	0.738	0.789	0.833
Shannon Wavelet	0.821	0.72	0.782	0.84

Table A2. Scotopic 2.0 ERG Response Metrics Table with Wavelet Function Variations.

VGG-11				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.669	0.644	0.668	0.7
Gaussian Wavelet	0.616	0.598	0.574	0.575
Ricker Wavelet	0.707	0.629	0.707	0.805
Morlet Wavelet	0.691	0.675	0.674	0.7
Shannon Wavelet	0.636	0.59	0.613	0.655

Table A2. *Cont.*

ResNet-50				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.718	0.701	0.684	0.68
Gaussian Wavelet	0.686	0.688	0.673	0.675
Ricker Wavelet	0.655	0.677	0.62	0.575
Morlet Wavelet	0.602	0.606	0.6	0.625
Shannon Wavelet	0.7	0.657	0.66	0.705
DensNet-121				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.753	0.707	0.723	0.755
Gaussian Wavelet	0.756	0.79	0.714	0.65
Ricker Wavelet	0.743	0.8	0.679	0.6
Morlet Wavelet	0.747	0.761	0.707	0.65
Shannon Wavelet	0.724	0.79	0.614	0.5
ResNext-50				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.778	0.813	0.608	0.55
Gaussian Wavelet	0.718	0.742	0.691	0.675
Ricker Wavelet	0.718	0.685	0.71	0.73
Morlet Wavelet	0.734	0.783	0.667	0.575
Shannon Wavelet	0.731	0.796	0.668	0.6
Vision Transformer				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.674	0.739	0.712	0.718
Gaussian Wavelet	0.796	0.676	0.762	0.848
Ricker Wavelet	0.849	0.775	0.825	0.833
Morlet Wavelet	0.822	0.845	0.736	0.701
Shannon Wavelet	0.793	0.724	0.764	0.78

Table A3. Photopic 2.0 ERG Response Metrics Table with Wavelet Function Variations.

VGG-11				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.732	0.877	0.782	0.71
Gaussian Wavelet	0.74	0.878	0.797	0.73
Ricker Wavelet	0.798	0.964	0.798	0.7
Morlet Wavelet	0.711	0.881	0.736	0.64
Shannon Wavelet	0.702	0.876	0.711	0.62

Table A3. *Cont.*

ResNet-50				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.752	0.912	0.79	0.7
Gaussian Wavelet	0.713	0.895	0.725	0.62
Ricker Wavelet	0.753	0.896	0.8	0.73
Morlet Wavelet	0.689	0.857	0.731	0.64
Shannon Wavelet	0.69	0.842	0.74	0.67
DensNet-121				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.734	0.899	0.772	0.68
Gaussian Wavelet	0.718	0.868	0.773	0.7
Ricker Wavelet	0.775	0.938	0.806	0.71
Morlet Wavelet	0.743	0.899	0.785	0.7
Shannon Wavelet	0.709	0.849	0.778	0.72
ResNext-50				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.749	0.866	0.813	0.77
Gaussian Wavelet	0.714	0.859	0.775	0.71
Ricker Wavelet	0.735	0.872	0.792	0.73
Morlet Wavelet	0.703	0.858	0.733	0.71
Shannon Wavelet	0.707	0.854	0.769	0.7
Vision Transformer				
Mother Wavelet Function	Balanced Accuracy	Recall	F1	Precision
Complex Gaussian Wavelet	0.868	0.857	0.902	0.893
Gaussian Wavelet	0.788	0.785	0.787	0.791
Ricker Wavelet	0.875	0.758	0.852	0.895
Morlet Wavelet	0.838	0.863	0.851	0.807
Shannon Wavelet	0.845	0.778	0.83	0.856

References

1. Constable, P.; Marmolejo-Ramos, F.; Gauthier, M.; Lee, I.; Skuse, D.; Thompson, D. Discrete Wavelet Transform Analysis of the Electroretinogram in Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder. *Frontiers in Neuroscience* **2022**, *16*, 1–11. doi:10.3389/fnins.2022.890461.

2. Manjur, S.; Hossain, M.B.; Constable, P.; Thompson, D.; Marmolejo-Ramos, F.; Lee, I.; Skuse, D.; Posada-Quintero, H. Detecting Autism Spectrum Disorder Using Spectral Analysis of Electroretinogram and Machine Learning: Preliminary results **2022**. *2022*, 3435–3438. doi:10.1109/EMBC48229.2022.9871173.

3. Constable, P.; Bach, M.; Frishman, L.; Jeffrey, B.; Robson, A. ISCEV Standard for clinical electro-oculography (2017 update). *Documenta ophthalmologica. Advances in ophthalmology* **2017**, *134*. doi:10.1007/s10633-017-9573-2.

4. Arden, G.; Constable, P. The electro-oculogram. *Progress in retinal and eye research* **2006**, *25*, 207–48. doi:10.1016/j.preteyeres.2005.11.001.

5. Umeya, N.; Miyawaki, I.; Inada, H. Use of an alternating current amplifier when recording the ERG c-wave to evaluate the function of retinal pigment epithelial cells in rats. *Documenta Ophthalmologica* **2022**, *145*, 1–9. doi:10.1007/s10633-022-09890-y.

6. Zhdanov, A.; Constable, P.; Manjur, S.M.; Dolganov, A.; Posada-Quintero, H.F.; Lizunov, A. OculusGraphy: Signal Analysis of the Electroretinogram in a Rabbit Model of Endophthalmitis Using Discrete and Continuous Wavelet Transforms. *Bioengineering* **2023**, *10*. doi:10.3390/bioengineering10060708.

7. Zhdanov, A.; Evdochim, L.; Borisov, V.; Bao, X.; Dolganov, A.; Kazaijkin, V. OculusGraphy: Filtering of Electroretinography Response in Adults **2021**. doi:10.1109/EDM52169.2021.9507654.

8. Constable, P.; Gaigg, S.; Bowler, D.; Jägle, H.; Thompson, D. Full-field electroretinogram in autism spectrum disorder. *Documenta ophthalmologica. Advances in ophthalmology* **2016**, *132*. doi:10.1007/s10633-016-9529-y.

9. Constable, P.; Ritvo, R.A.; Ritvo, A.; Lee, I.; McNair, M.; Stahl, D.; Sowden, J.; Quinn, S.; Skuse, D.; Thompson, D.; McPartland, J. Light-Adapted Electroretinogram Differences in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders* **2020**, *50*. doi:10.1007/s10803-020-04396-5.

10. McAnany, J.J.; Persidina, O.; Park, J. Clinical electroretinography in diabetic retinopathy: A review. *Survey of Ophthalmology* **2021**, *67*. doi:10.1016/j.survophthal.2021.08.011.

11. Kim, T.H.; Wang, B.; Lu, Y.; Son, T.; Yao, X. Functional Optical Coherence Tomography Enables in vivo Optoretinography of Photoreceptor Dysfunction due to Retinal Degeneration. *Biomedical Optics Express* **2020**, *11*. doi:10.1364/BOE.399334.

12. Hayashi, T.; Hosono, K.; Kurata, K.; Katagiri, S.; Mizobuchi, K.; Ueno, S.; Kondo, M.; Nakano, T.; Hotta, Y. Coexistence of GNAT1 and ABCA4 variants associated with Nougaret-type congenital stationary night blindness and childhood-onset cone-rod dystrophy. *Documenta Ophthalmologica* **2020**, *140*. doi:10.1007/s10633-019-09727-1.

13. Kim, H.M.; Joo, K.; Han, J.; Woo, S.J. Clinical and Genetic Characteristics of Korean Congenital Stationary Night Blindness Patients. *Genes* **2021**, *12*.

14. Zhdanov, A. E., e.a. Evaluation of the effectiveness of the decision support algorithm for physicians in retinal dystrophy using machine learning methods. *Computer Optics* **2023**, *42*, 272–277. doi:10.18287/2412-6179-CO-1124.

15. Penkala, K.; Jaskuła, M.; Lubiński, W. [Improvement of the PERG parameters measurement accuracy in the continuous wavelet transform coefficients domain]. *Annales Academiae Medicae Stetinensis* **2007**, *53 Suppl 1*, 58–60; discussion 61.

16. Penkala, K. Analysis of bioelectrical signals of the human retina (PERG) and visual cortex (PVEP) evoked by pattern stimuli. *BULLETIN OF THE POLISH ACADEMY OF SCIENCES TECHNICAL SCIENCES* **2005**, *53*.

17. Ahmadieh, H.; Behbahani, S.; safi, S. Continuous wavelet transform analysis of ERG in patients with diabetic retinopathy. *Documenta Ophthalmologica* **2021**. doi:10.1007/s10633-020-09805-9.

18. Dimopoulos, I.; Freund, P.; Redel, T.; Dornstauder, B.; Gilmour, G.; Sauvé, Y. Changes in Rod and Cone-Driven Oscillatory Potentials in the Aging Human Retina. *Investigative ophthalmology & visual science* **2014**, *55*. doi:10.1167/iovs.14-14219.

19. Gauvin, M.; Lina, j.m.; Lachapelle, P. Advance in ERG Analysis: From Peak Time and Amplitude to Frequency, Power, and Energy. *BioMed research international* **2014**, *2014*, 246096. doi:10.1155/2014/246096.

20. Zhdanov, A.; Dolganov, A.; Zanca, D.; Borisov, V.; Ronkin, M. Advanced Analysis of Electroretinograms Based on Wavelet Scalogram Processing. *Applied Sciences* **2022**, *12*. doi:10.3390/app122312365.

21. Barraco, R.; Adorno, D.P.; Brai, M. Wavelet analysis of human photoreceptoral response. 2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2010), 2010, pp. 1–4. doi:10.1109/ISABEL.2010.5702846.

22. Barraco, R.; Persano Adorno, D.; Brai, M. An approach based on wavelet analysis for feature extraction in the a-wave of the electroretinogram. *Computer Methods and Programs in Biomedicine* **2011**, *104*, 316–324. doi:https://doi.org/10.1016/j.cmpb.2011.05.001.

23. Barraco, R.; Persano Adorno, D.; Brai, M. ERG signal analysis using wavelet transform. *Theory in biosciences = Theorie in den Biowissenschaften* **2011**, *130*, 155–63. doi:10.1007/s12064-011-0124-1.

24. Miguel-Jiménez, J.M.; Blanco, R.; De-Santiago, L.; Fernández, A.; Rodríguez-Ascariz, J.M.; Barea, R.; Martín-Sánchez, J.L.; Amo, C.; Sánchez-Morla, E.V.; Boquete, L. Continuous-wavelet-transform analysis of the multifocal ERG waveform in glaucoma diagnosis. *Medical & Biological Engineering & Computing* **2015**, *53*, 771 – 780.

25. Zhdanov, A.; Dolganov, A.; Borisov, V.; Ronkin, M.; Ponomarev, V.; Zanca, D. OculusGraphy: Ophthalmic Electrophysiological Signals Database **2022**. doi:10.21227/r1wb-pg25.

26. Lemaître, G.; Nogueira, F.; Aridas, C. Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning **2016**. *18*.

27. Fricke, M.; Bodendorf, F. Identifying Trendsetters in Online Social Networks – A Machine Learning Approach **2020**. pp. 3–9. doi:10.1007/978-3-030-51057-2_1.

28. Lee, G.; Gommers, R.; Waselewski, F.; Wohlfahrt, K.; Aaron. PyWavelets: A Python package for wavelet analysis. *Journal of Open Source Software* **2019**, *4*, 1237. doi:10.21105/joss.01237.

29. Xu, G.; Shen, X.; Chen, S.; Zong, Y.; Zhang, C.; Yue, H.; Liu, M.; Chen, F.; Che, W. A Deep Transfer Convolutional Neural Network Framework for EEG Signal Classification. *IEEE Access* **2019**, *7*, 112767–112776. doi:10.1109/ACCESS.2019.2930958.

30. Wu, Q.e.; Yu, Y.; Zhang, X. A Skin Cancer Classification Method Based on Discrete Wavelet Down-Sampling Feature Reconstruction. *Electronics* **2023**, *12*. doi:10.3390/electronics12092103.

31. Huang, G.H.; Fu, Q.J.; Gu, M.Z.; Lu, N.H.; Liu, K.Y.; Chen, T.B. Deep Transfer Learning for the Multilabel Classification of Chest X-ray Images. *Diagnostics* **2022**, *12*. doi:10.3390/diagnostics12061457.

32. Chen, C.F.R.; Fan, Q.; Panda, R. CrossViT: Cross-Attention Multi-Scale Vision Transformer for Image Classification. International Conference on Computer Vision (ICCV), 2021.

33. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv 1409.1556* **2014**.

34. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778. doi:10.1109/CVPR.2016.90.

35. Xie, S.; Girshick, R.; Dollar, P.; Tu, Z.; He, K. Aggregated Residual Transformations for Deep Neural Networks. 2017, pp. 5987–5995. doi:10.1109/CVPR.2017.634.

36. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.

37. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; Uszkoreit, J.; Houlsby, N. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *ICLR* **2021**.

38. Bishop, C.M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*; Springer-Verlag: Berlin, Heidelberg, 2006.

39. Mohammed, R.; Rawashdeh, J.; Abdullah, M. Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results. 2020 11th International Conference on Information and Communication Systems (ICICS), 2020, pp. 243–248. doi:10.1109/ICICS49469.2020.239556.

40. Goutte, C.; Gaussier, E. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. Advances in Information Retrieval; Losada, D.E.; Fernández-Luna, J.M., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2005; pp. 345–359.

41. García, V.; Mollineda, R.A.; Sánchez, J.S. Index of Balanced Accuracy: A Performance Measure for Skewed Class Distributions. Pattern Recognition and Image Analysis; Araujo, H.; Mendonça, A.M.; Pinho, A.J.; Torres, M.I., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2009; pp. 441–448.

42. Liao, C.C.; Yang, H.T.; Chang, H.H. Denoising Techniques With a Spatial Noise-Suppression Method for Wavelet-Based Power Quality Monitoring. *IEEE Transactions on Instrumentation and Measurement* **2011**, *60*, 1986–1996. doi:10.1109/TIM.2011.2115610.

43. Tzabazis, A.; Eisenried, A.; Yeomans, D.; Moore, H. Wavelet analysis of heart rate variability: Impact of wavelet. *Biomedical Signal Processing and Control* **2018**, *40*, 220–225. doi:10.1016/j.bspc.2017.09.027.

44. Robson, A.; Frishman, L.; Grigg, J.; Hamilton, R.; Jeffrey, B.; Kondo, M.; Li, S.; McCulloch, D. ISCEV Standard for full-field clinical electroretinography (2022 update). *Documenta Ophthalmologica* **2022**, *144*, 1–13. doi:10.1007/s10633-022-09872-0.

45. Guo, J.; Han, K.; Wu, H.; Tang, Y.; Chen, X.; Wang, Y.; Xu, C. CMT: Convolutional Neural Networks Meet Vision Transformers. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 12175–12185.

46. Lim, J.; Hong, M.; Lam, W.; Zhang, Z.; Teo, Z.; Liu, Y.; Ng, W.; Foo, L.; Ting, D. Novel technical and privacy-preserving technology for artificial intelligence in ophthalmology. *Current Opinion in Ophthalmology* **2022**, *Publish Ahead of Print*. doi:10.1097/ICU.0000000000000846.

47. Bouaziz, M.; Cheng, T.; Minuti, A.; Denisova, K.; Barmettler, A. Shared Decision Making in Ophthalmology: A Scoping Review. *American Journal of Ophthalmology* **2022**, *237*, 146–153. doi:https://doi.org/10.1016/j.ajo.2021.12.005.

48. Zhdanov, A.E.; Borisov, V.I.; Dolganov, A.Y.; Lucian, E.; Bao, X.; Kazaijkin, V.N. OculusGraphy: Norms for Electroretinogram Signals. 2021 IEEE 22nd International Conference of Young Professionals in Electron Devices and Materials (EDM), 2021, pp. 399–402. doi:10.1109/EDM52169.2021.9507597.

49. Zhdanov, A.E.; Borisov, V.I.; Lucian, E.; Kazaijkin, V.N.; Bao, X.; Ponomarev, V.O.; Dolganov, A.Y.; Lizunov, A.V. OculusGraphy: Description of Electroretinograms Database. 2021 Third International Conference Neurotechnologies and Neurointerfaces (CNN), 2021, pp. 132–135. doi:10.1109/CNN53494.2021.9580221.

50. Lu, Z.; Zhou, M.; Guo, T.; Liang, J.; Wu, W.; Gao, Q.; Li, L.; Li, H.; Chai, X. An in-silico analysis of retinal electric field distribution induced by different electrode design of trans-corneal electrical stimulation. *Journal of Neural Engineering* **2022**, *19*, 055004. doi:10.1088/1741-2552/ac8e32.

51. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. Advances in neural information processing systems, 2014, pp. 2672–2680.