IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2023.0322000

Attention to the Electroretinogram: Gated Multilayer Perceptron for ASD Classification

MIKHAIL KULYABIN[1], PAUL A. CONSTABLE[2], ALEKSEI ZHDANOV[3], IRENE O. LEE[4], DOROTHY A. THOMPSON[5], ANDREAS MAIER[1]

[1]Pattern Recognition Lab, University of Erlangen-Nuremberg, Germany (e-mail: mikhail.kulyabin@fau.de)
[2]College of Nursing and Health Sciences, Caring Futures Institute, Flinders University, Adelaide, Australia
[3]Siemens Healthineers, Erlangen, Germany
[4]Behavioural and Brain Sciences Unit, Population Policy and Practice Programme, UCL Great Ormond Street Institute of Child Health, University College London, London, UK
[5]The Tony Kriss Visual Electrophysiology Unit, Clinical and Academic Department of Ophthalmology, Great Ormond Street Hospital for Children NHS Trust, London, UK; UCL Great Ormond Street Institute of Child Health, University College London, London, UK

Corresponding author: Mikhail Kulyabin (e-mail: mikhail.kulyabin@fau.de).

ABSTRACT The electroretinogram (ERG) is a clinical test that records the retina's electrical response to a brief flash of light as a waveform signal. Analysis of the ERG signal offers a promising non-invasive method for studying different neurodevelopmental and neurodegenerative disorders. Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterized by poor communication, reduced reciprocal social interaction, and restricted and/or repetitive stereotyped behaviors that should be detected as early as possible to ensure timely and appropriate intervention to support the individual and their family. In this study, we applied gated Multilayer Perceptron (gMLP) for the light-adapted ERG waveform classification as an effective alternative to Transformers. In this first reported application of this model to ASD classification which consisted of basic multilayer perceptrons, with fewer parameters than Transformers. We compared the performance of different time-series models on an ASD-Control dataset and found that the superiority of gMLP in classification accuracy was the best at 89.7% compared to alternative models and supports the use of gMLP in classification models based on ERG recordings involving case-control comparisons.

INDEX TERMS ASD, Deep Learning, Electroretinogram, ERG, Gated MLP, Transformer, Waveform

I. INTRODUCTION

A. POTENTIAL FOR ERG DIAGNOSIS IN CNS DISORDERS

THE full-field electroretinogram (ERG) is the waveform recorded from the eye under dark- or light-adapted (DA or LA) conditions in response to a brief flash of light. Clinically, the ERG waveform can be used for the diagnosis of conditions affecting the retina, such as inherited or acquired diseases [1]. Because the retina is an extension of the central nervous system (CNS), and its function is readily accessible through the ERG, several studies have investigated changes in the ERG waveform in conditions affecting the CNS in human and animal studies [2]. For example, the analysis of the ERG waveform to identify potential biomarkers has also been proposed for the early detection of Attention Deficit Hyperactivity Disorder (ADHD) [3], bipolar disorder [4] and using a mouse model for Parkinson's disease [5].

The shape of the ERG waveform depends on the state of retinal adaptation with the DA- and LA-ERG responses dominated by rod and cone pathways, respectively [1], [6]. The main excitatory neurotransmitter of the retina is glutamate, which contributes to the main positive b-wave generated by the bipolar cells [7]. The preceding negative a-wave is formed by hyperpolarization of the photoreceptor outer segments [8] that reduces glutamate release into the post rececptoral synapse with bipolar and horizontal cells [9]. The horizontal cells provide inhibition to cone photoreceptors using gamma-aminobutyric acid (GABA) signaling that modulates the a-wave's amplitude [10]. Dopamine-driven responses from the amacrine cells also contribute to the high-frequency oscillatory potentials visible as small ripples on the ascending limb of the b-wave [11]. Given the contributions of these key neurotransmitters and their role in CNS disorders, changes in the ERG waveform have been associated with alterations in these neurotransmitters [2], [12]. For example, reduced dopamine in early Parkinson's disease results in a reduced b-wave and oscillatory potentials [13]. In autism spectrum disorder (ASD) and ADHD, differences in the balance between glutamate may be responsible for the elevated b-waves in ADHD compared to the reduced b-wave amplitudes reported in ASD [14]. In schizophrenia, the reduced a- and

b-wave wave amplitudes are thought to be due to increased GABAergic inhibition by the horizontal cells [15] and may help to distinguish schizophrenia from bipolar disorder [16]. Developing methods for the classification of ASD and potentially other conditions affecting the CNS through the analysis of the ERG could provide improved earlier diagnosis and management of these conditions to improve patient outcomes [17].

With respect to ASD, the search for a biomarker to detect this condition has been extensive, with currently no clinical diagnostic test able to reliably identify a child with ASD [18]. The ERG may be a potential new test that, with more extensive clinical trials, could provide a novel biomarker for ASD. Early studies have identified reduced DA- and LA-ERG responses in children with ASD [19], [20]. However, in adult populations, the results have been mixed [21], [22] with respect to the LA-ERG changes. There is some evidence in small study populations that the ERG changes may differ between ADHD and ASD groups [14], [23]. Still, these early findings require more extensive studies to replicate in younger clinical populations. Whilst developments in this field continue with the use of signal analysis of the ERG using variable frequency complex demodulation [24] showing potential to not only classify ASD but also to differentiate between ASD and ADHD [25], [26]. Other methods using aspects of Functional Data Analysis of the b-wave have also recently been reported [27] that may provide additional features that may contribute to the classification of retinal disorders [28].

B. APPLICATION OF MACHINE LEARNING (ML) AND ARTIFICIAL INTELLIGENCE TECHNIQUES AND THE ERG

The studies described in Table 1 have incorporated ML to enhance the diagnosis of ophthalmic or neurological conditions through the analysis of ERG recordings. Yapici et al. [29] explored obesity's correlation with ocular health, and achieved 94.1% and 92.9% classification accuracy for obesity with an artificial neural network based on discrete wavelet transform analysis of the LA- and DA-ERG waveforms in 47 subjects. Lopez et al. [30] investigated multiple sclerosis, using support vector machines (SVMs) to identify multifocal ERG (mfERG) feature differences based on continuous wavelet transform of the signal in 15 subjects. Zhdanov et al. [31] applied ML for the classification of adults compared to pediatric DA- and LA-ERGs with and without retinal disease based on wavelet transforms to improve the classification of the groups using classical time-domain and novel time-frequency features. Glinton et al. [32] using DA- and LA-ERGs time-domain features recorded from 597 cases with ABCA4 retinopathy and with regression models developed genotype-phenotype models to not only predict disease progression but also classify the phenotypes into three groups with up to 91.8% accuracy. Gajendran et al. [33] addressed early-stage glaucoma diagnosis in a mouse model based on analysis of advanced features from the DA- and LA-ERGs distributions using ML to identify early ganglion cell loss. Manjur et al. [34] explored ERG-based ASD detection, achieving 86% accuracy, and emphasized the potential for earlier diagnosis using decomposition of the LA-ERGs. Kulyabin et al. [35] determined optimal wavelet-DL model combinations for pediatric ERG signal analysis, providing insights into selecting appropriate mother wavelets. Posada-Quintero et al. [24] compared signal analytical methods, including time-domain and time-frequency domain features derived from the distributions for ASD classification, supporting the ERG waveform derived from a single flash strength in the right eye as a potential practical clinical biomarker for ASD. Manjur et al. [25] applied variable frequency complex demodulation and ML for ASD and ADHD classification, achieving 0.84 accuracy with gradient boosting. and more recently achieved a 70% overall accuracy for differentiating between ASD, ADHD and controsl [26]. Further studies may explore sensitivity and specificity with controls meeting both ASD and ADHD classifications [25]. Taken together, these studies using different features of the ERG signal based on full field and mfERG in human and mouse studies suggest that ML coupled with ERG features and clinical parameters will improve earlier diagnosis, prognosis, and management of conditions affecting the CNS.

Studies incorporating ML methods with the ERG waveform are increasing but still relatively uncommon, with neural networks encountered in only three publications relating to ASD and the ERG [24], [25], [35], signifying an interest in adopting ML techniques [36], [37]. These observations emphasize the substantial untapped potential offered by ML methods for the comprehensive analysis of the ERG waveform signal to improve classification between groups and potentially earlier identification of retinal disease processes [24], [25], [34], [38]. The wider adoption of ML methods may further improve the clinical utility of the ERG [39], [40] and with further technological advances such as smartphone-based devices could become more accessible [41].

Additionally, applying Deep Learning (DL) approaches could potentially improve the accuracy of ERG signal classification of retinal and CNS-based disorders, thereby enhancing not only the quality of ASD detection in the early stages and improving long-term outcomes for individuals with ASD [42] but also related disorders where the ERG is atypical [2], [39].

Moreover, in order to realize the full potential of the ERG in the classification or earlier detection of CNS disorders, the application of signal analysis using wavelets and variable frequency complex demodulation has been applied recently in studies involving individuals with ASD and ADHD [23], [25]. These preliminary studies in childhood have identified the potential for identifying features extracted from signal analysis to improve machine learning (ML) classification models. However, one limitation in such studies is the heterogeneity of the ASD population despite standardization of clinical assessments such as the Autism Diagnostic Observational schedule (ADOS) [43]. Heterogeneity in ASD may manifest in the severity and co-occurrence of additional conditions

TABLE 1. Comparative Table of Publications on the Application of Machine Learning and Artificial Intelligence Techniques for ERG

First author, year [reference]	Diagnosed Disease	Subjects / Signals	Recording	ML / AI Algorithm
López-Dorado, 2021 [30]	Multiple Sclerosis	21 / Not Specified	mfERG	Support Vector Machine (SVM)
Yapici, 2021 [29]	Obesity diseases	47 / Not Specified	ERG DA and LA	Artificial Neural Network (ANN)
Zhdanov, 2022 [31]	Retinal Dystrophy	Not Specified / 425	ERG DA and LA	Decision Tree (DT)
Glinton, 2022 [32]	ABCA4-related retinopathy	597 / Not Specified	ERG DA and LA	SVM
Gajendran, 2022 [33]	Early-Stage Glaucoma	60 / 540	ERG DA and LA	Feature Engineering, RFC*
Manjur, 2022 [34]	Autism Spectrum Disorder (ASD)	143 / Not Specified	ERG LA	RFC, GB**, DT, SVM
Kulyabin, 2023 [35]	Retinal Dystrophy	Not Specified / 542	ERG DA and LA	VGG, ResNet, DensNet, ResNext, ViT
Posada-Quintero,2023 [24]	ASD	217 / Not Specified	ERG LA	RFC, GB, SVM
Manjur, 2023 [25]	ASD	143 / Not Specified	ERG LA	GradBoost, XGBoost
Manjur, 2024 [26]	ASD, ADHD***	286 / Not Specified	ERG LA	RFC, XGBoost

* Random Forest Classifier
** Gradient Boosting
*** Attention Deficit Hyperactivity Disorder

such as ADHD with phenotypic overlap [44]. ML approaches may help to classify neurodevelopmental disorders based on a combination of phenotypic and biological markers [45], [46]. In addition, the ERG findings in ASD have not been replicated in older age groups, suggesting that the findings may not be fully generalizable to all age groups [22].

Our previous studies [35], [47] have shown the superiority of Transformer over classical architectures in the time-frequency domain with respect to ERG with the condition that Transformer training requires a large dataset, which is challenging to obtain due to field specificity in many cases. One solution to this problem would be to apply an alternative structure to a Transformer with less trainable parameters and more efficiency for a reasonably shorter signal. Thus, we propose using the Gated Multilayer Perceptron (gMLP) [48] for ERG signal classification. In this study, we apply for the first time using ERG waveform time-series analysis, the gMLP architecture, and compare the performance of gMLP with other architectures in the time-series domain.

II. DATA

Fig.1 shows an LA-ERG signal waveform of a control subject used in this study. By analyzing the parameters of the ERG waveform, such as the amplitude of the a- and b-waves (Va, Vb) and their respective time to peaks (Ta, Tb), clinicians can identify abnormalities that help diagnose a range of retinal disorders [1]. Fig.2 illustrates a further series of representative LA-ERG waveforms at four flash strengths: -0.367, 0.114, 0.799, and 1.204 $(log\ cd.s.m^{-2})$, for an ASD and control participant. The amplitude of the b-wave increases with flash strength reaching a peak before falling and forming a plateau phase that is described as the photopic hill [49]. Fig.2 shows this with the b-wave amplitude being smaller at the highest flash strength of 1.204 $(log cd.s.m^{-2})$ and maximal at a lower flash strength of 0.799 $(log\ cd.s.m^{-2})$ that forms the "peak" of the photopic hill in this instance.

Comparing the ASD and control waveforms, it is apparent that the amplitude of the b-wave is reduced in the ASD subject. Notable also is the absence of prominent oscillatory potentials in the ASD waveforms. The oscillatory potentials

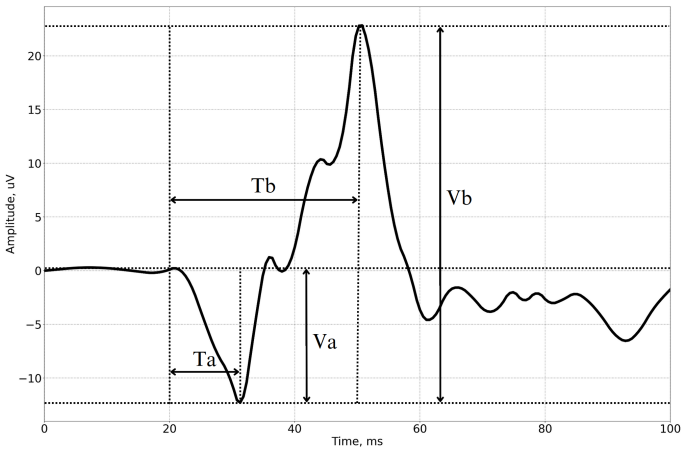
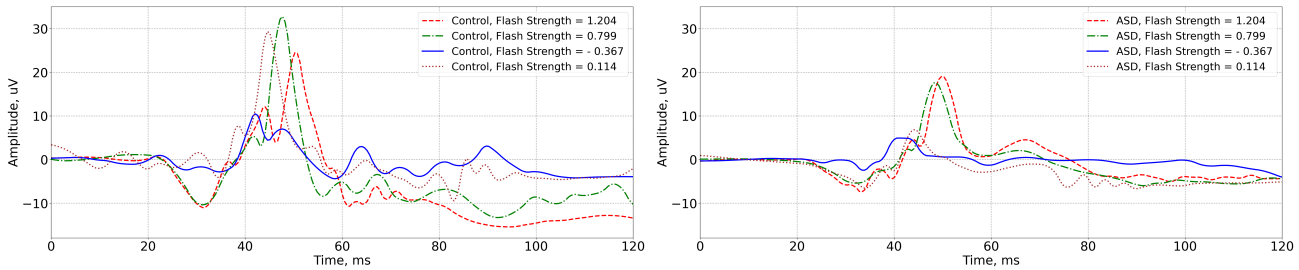


FIGURE 1. Light-adapted ERG waveform of a control individual. There are two prominent peaks in the waveform. The a-wave is the first negative deflection is mainly due to hyperpolarization of the photoreceptors, and the following positive b-wave is shaped by bipolar, amacrine, and glial cell currents. Small peaks are observed on the ascending limb of the b-wave, which are termed the oscillatory potentials that have their origins in the amacrine cells. Time domain features are indicated as Ta, Tb, Va, and Vb, corresponding to the time to peak and amplitudes of the a- and b-waves, respectively.

derive from amacrine cells [11] and are usually visible as small "ripples" or peaks on the ascending limb of the b-wave before the main peak and contribute to the high-frequency components of the ERG. These differences have may be due to a difference in the regulation of glutamate and/or dopamine that contribute to the amplitude of the b-wave and oscillatory potentials as described in Lee at al [14].

In this work, we re-analyzed the LA-ERG waveform recordings from previously reported studies [14], [20], [23]. This dataset contained signals from 20 control and 30 ASD individuals collected in two different locations: London (UK) and Adelaide (Australia). Full-field LA-ERG recordings were performed on each eye (always right first), following the guidelines of the ISCEV ERG standard [6]. A series of brief flashes of different strengths were applied to the eyes on a 40 $(cd.m^{-2})$ white background. Recordings were performed with the RETeval (LKC Technologies, Gaithersburg, MD,



(a) Control ERG signals. (b) ASD ERG signal.

FIGURE 2. **Examples of ERG waveforms recorded from a control (a) and ASD (b) individual to flash strengths -0.367, 0.144 0.799, and 1.204 ($log\ cd.s.m^{-2}$). The prominent b-wave positive peak is reduced in the ASD waveforms, with less noticeable oscillatory potentials visible on the ascending limb of the b-wave. Note the amplitude of the b-wave is maximal at the intermediate flash strength of 0.799 ($log\ cd.s.m^{-2}$)and reduces as flash strength increases with the 1.204 ($log\ cd.s.m^{-2}$) strength.**

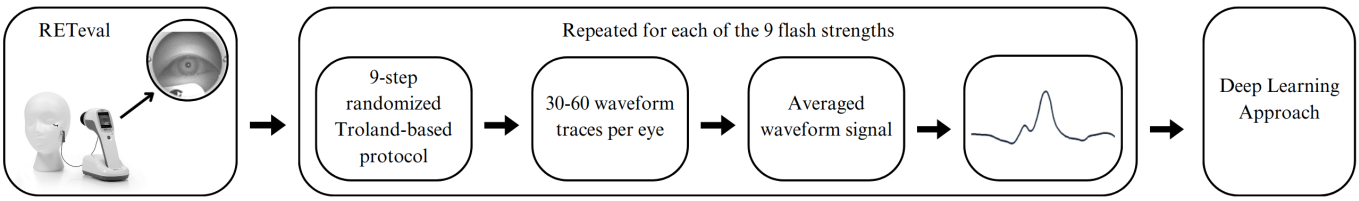


FIGURE 3. **Recordings were performed with the RETeval from each eye using a skin electrode with nine randomized flash strengths. 30 to 60 averages of the ERG were recorded to generate a signal average that was then used for classification of groups using Deep Learning methods.**

TABLE 2. **Dataset distribution**

	ASD	Control
	Individuals	
	30	20
	Signals	
Flash Strength		
1.204	58	59
1.114	60	51
0.949	56	51
0.799	56	57
0.602	58	56
0.398	60	55
0.114	53	50
-0.119	56	50
-0.367	52	53
Eye		
Left	255	252
Right	254	230
Total	509	482

USA) with a custom nine-step randomized Troland-based protocol with skin electrodes placed 2-3 mm below the lower eyelid. Flashes delivered at 2 Hz were averaged from 30-60 waveform traces per eye to generate the reported average waveform signal that was used in the analysis. Waveforms with artifacts such as blinks were automatically rejected from the average if they fell within the upper or lower quartile of the overall average. Two recordings were typically made in each eye and included in the dataset for analysis. Fig.3 shows the signal processing method. The dataset distribution of included ERG waveform signals recorded from the participants in each group and at each flash strength is shown in Table 2. Signals from one individual appear only in one subset (fold) for the cross-validation steps.

III. GATED MULTILAYER PERCEPTRON

The gMLP is a neural network architecture that aims to process sequential data using a novel design centered around Multi-Layer Perceptrons (MLPs) with gating mechanisms. Unlike Transformer-based models, which rely heavily on self-attention mechanisms, gMLP explores an alternative approach for sequence modeling: it simply consists of channel projections and spatial projections with static parameterization. It demonstrates high performance on time-series domain tasks and uses fewer trainable parameters than Transformer models in general. The main components of the structure are the gMLP main block and the Spatial Gating Unit (SGU), which are described below.

A. GMLP BLOCK

The MLP is a basic form of a neural network, consisting of a simple series of fully-connected layers or perceptrons. The overview of the gMLP model is shown in Figure 4. It consists of a stack of L blocks, each with identical size and structure. Each block L is defined as:

$$Z = \sigma(XU), \quad \tilde{Z} = s(Z), \quad Y = \tilde{Z}V \tag{1}$$

where $X \in R^{n\times d}$ is a token with sequence length n and dimension d, and σ is an Activation function. $U \in R^{d\times d_{ffn}}$ and $V \in R^{d_{ffn}\times d}$ define linear projections along the channel dimension, $Z \in R^{n\times d_{ffn}}$, and $s(\cdot)$ is a layer that captures spatial

interactions and is defined as a spatial depth-wise convolution. Unlike Transformers, gMLP does not require position embeddings because $s(\cdot)$ already contains this information [48].

B. SPATIAL GATING UNIT

Layer $s(\cdot)$ should contain a contraction operation over the spatial dimension to enable cross-token interactions that could be performed with linear projection:

$$f_{W,b}(Z) = WZ + b \tag{2}$$

where $W \in R^{n \times n}$ is independent of the input representations matrix for which the size is the same as the sequence length n, and b refers to token-specific biases. Layer $s(\cdot)$ is the output of the linear gating:

$$s(Z) = Z \odot f_{W,b}(Z) \tag{3}$$

where $s(\odot)$ is an element-wise multiplication. For training W is initialized as near-zero value and b as ones, consequently $f_{W,b}(Z) \approx 1$ and $s(Z) \approx Z$. For effectiveness, Z is split into two independent parts Z_1, Z_2 along the channel dimension for the gating function and for the multiplicative bypass:

$$s(Z) = Z_1 \odot f_{W,b}(Z_2) \tag{4}$$

For improvement of the stability of large models, the input is normalized to $f_{W,b}$. This unit was referred to the SGU.

The Algorithm 1 appears similar to the attention mechanism in Transformers. However, it is not identical. Here, the weights stay the same during the inference, independent of the input. Meanwhile, in the attention mechanism, the weights change depending on the input, which can lead to better performance during inference. On the other hand, this makes transformers more challenging to train.

Algorithm 1 Work of the gMLP block.

gMLP Block$(\mathbf{X}, \mathbf{d}, \mathbf{d_{ffn}})$
- $\mathbf{S}hortcut = \mathbf{X}$
- $\mathbf{X} \leftarrow Norm(\mathbf{X},\ axis = Channel)$
- $\mathbf{U} \leftarrow Proj(\mathbf{X}, \mathbf{d_{ffn}},\ axis = Channel)$
- $\mathbf{Z} \leftarrow GELU(\mathbf{XU})$
- $\tilde{\mathbf{Z}} \leftarrow Spatial\ Gating\ Unit(\mathbf{Z})$
- $\mathbf{V} \leftarrow Proj(\tilde{\mathbf{Z}}, \mathbf{d},\ axis = Channel)$
- **return** $\tilde{\mathbf{Z}}\mathbf{V} + \mathbf{S}hortcut$

Spatial Gating Unit$(\mathbf{Z})$
- $\mathbf{Z_1}, \mathbf{Z_2} \leftarrow Split(\mathbf{Z},\ axis = Channel)$
- $\mathbf{Z_2} \leftarrow Norm(\mathbf{Z_2},\ axis = Channel)$
- $\mathbf{n} \leftarrow Get\ Dimension(\mathbf{Z_2},\ axis = Spatial)$
- $\mathbf{Z_2} \leftarrow Proj(\mathbf{Z_2}, \mathbf{n},\ axis = Spatial,\ init\_bias = 1)$
- **return** $\mathbf{Z_1} \odot \mathbf{Z_2}$

IV. EVALUATION

In this work, we compared multiple time-series DL models, allowing for a systematic assessment of their respective performance characteristics and facilitating the selection of the most accurate model for the classification problem in the domain while advancing scientific understanding of their applicability and limitations.

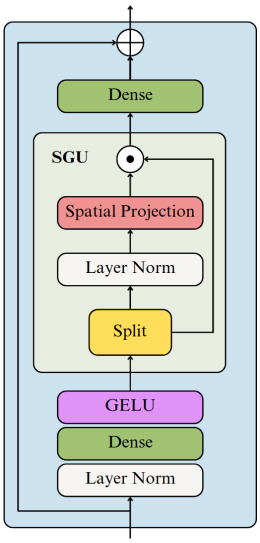


FIGURE 4. Overview of the gMLP architecture. The model consists of a stack of L blocks with identical structure and size. Each block consists of channel projections before and after the Spatial Gating Unit (SGU). Together with the activation function (GELU), they act as feedforward layers.

Bidirectional Long Short-Term Memory (BiLSTM) [50] is a type of recurrent neural network (RNN) that can capture temporal dependencies in data. They are well-suited for sequential data with long-range dependencies. Residual Network (ResNet) [51] is a deep convolutional neural network designed for image data, but it can also be adapted for time series tasks. It uses skip connections to mitigate the vanishing gradient problem [52] and allows for the training of very deep networks. InceptionTime [53] is a time series model inspired by Google's Inception architecture. It uses multiple parallel convolutional layers with different kernel sizes to capture various temporal patterns at different scales. OmniScale [54] is a model designed to handle a wide range of time series tasks, from short to very long time series. It uses a combination of dilated causal convolutions and self-attention mechanisms to capture temporal dependencies efficiently. A Time Series Transformer (TST) [55] is a transformer-based architecture adapted for time series data. It utilizes self-attention mechanisms to capture temporal dependencies and global patterns effectively. Time Series in Transformers (TSiT) [56] is another transformer-based model explicitly designed for time series tasks. It incorporates additional components like recurrence and autoregressive attention to capture sequential patterns. PatchTST [56] combines the concepts of patch-based processing with time series data. It divides the time series into smaller patches and applies transformer-based models to each patch.

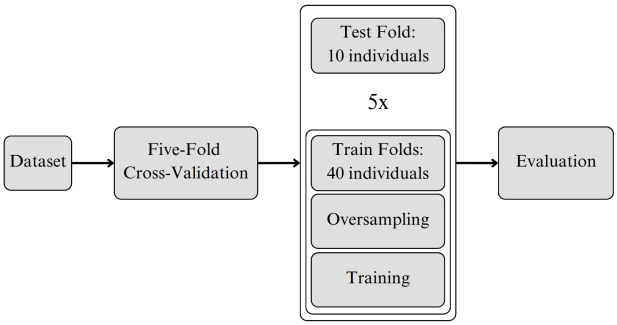


FIGURE 5. Evaluation pipeline. To perform the subject-wise cross-validation, all individuals from the dataset were randomly divided into five groups (folds). On each iteration four folds (40 ASD/Control individuals) were in the training subset, and one fold (10 ASD/Control individuals) was in the test subset. Each cross-validation step was repeated five times on the five folds. On each occasion within the training subset, oversampling was applied to avoid unbalancing. In this way, all models were trained, and classification metrics compared.

A. TRAINING

We performed a five-fold subject-wise cross-validation to evaluate each model: we randomly divided all 50 subjects on five folds, Fig.5. There were at least nine signals from both eyes per subject. This separation was necessary to avoid having signals from the same subject in different folds, as this would have falsely increased the accuracy of the prediction. Every time, four folds were used for training, and the last fold was used for testing. The models were trained using the entire dataset without dividing the ERG signals into flash strength classes, as it would have resulted in a reduced training subset, which is not suited for training computationally intensive models. To solve the unbalanced problem on the training subset, we applied oversampling by individual class and class weights. Oversampling was performed as upsampling of the data related to the minority class (control).

For all of the models, we used the CrossEntropyLossFlat loss function so that we could pass in a weight parameter. We used Adam as an optimizer with learning rates [0.0001, 0.001]. The validation metric was accuracy.

Since the objective was to reduce the number of parameters and efficiency, in the current work, we used a comparable TST gMLP "Tiny" version from the original study [48]: the model dimension (d_{model}) equaled 128, and the feed-forward dimension (d_{ffn}) was 768. The depth of the model was set to 12. To reduce the number of training parameters but still compete with TST, we reduced the model parameters to: d_{model} equal to 64, d_{ffn} equal to 512, and depth equal to 6 ("Nano"). We used GELU as an activation function. The models were trained until convergence with a maximum learning rate of 0.0001 with a batch size of 32.

B. METRICS

For a complete understanding of the model performance, several metrics were computed: Precision (P), Recall (R), and F1 Score:

$$Precision = \frac{TP}{TP + FP}, \tag{5}$$

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

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall}, \tag{7}$$

where

- $TP = True\ Positive$,
- $TN = True\ Negative$,
- $FP = False\ Positive$,
- $FN = False\ Negative$.

As the test subsets were not balanced, we considered the Balanced Accuracy (BA):

$$Balanced\ Accuracy = \frac{Sensitivity + Specificity}{2}, \tag{8}$$

where

$$Sensitivity = \frac{TP}{TP + FN}, \tag{9}$$

$$Specificity = \frac{TN}{TN + FP}. \tag{10}$$

C. RESULTS

The metrics presented in Table 3 demonstrate the high performance of the gMLP and transformer-based architecture models compared with other models for classification on this dataset. Fig.6a also shows the Receiver Operating Characteristic curves (ROC) with corresponding Areas Under the Curves (AUC) for the models tested. Notably, "Tiny" gMLP demonstrated better performance within our constrained dataset than Transformers across most metrics. Furthermore, the shorted version "Nano" also outperformed TST and secured the second position across most metrics despite having significantly fewer trainable parameters. For comparison, Fig.6b shows the ROC curves of gMLP and TST models. Specificity and Sensitivity, defined by the formulas 9 and 10, were used to construct the ROC curves with the following indicators: $FP_{TST} = 7$, $FN_{TST} = 18$; $FP_{gMLP_{Tiny}} = 7$, $FN_{gMLP_{Tiny}} = 12$; $FP_{gMLP_{Nano}} = 11$, $FN_{gMLP_{Nano}} = 10$.

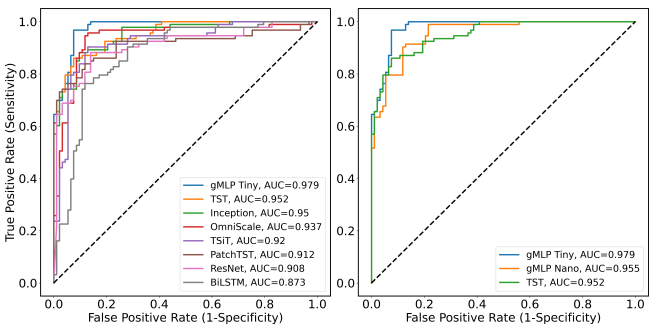
Table 3 shows the number of trainable parameters of each tested model. TSiT has 86 million trainable parameters, which is impractical for training on ERG signals. TST has 1.5 million parameters compared to 5.9 million for the "Tiny" gMLP. However, the shortened version "Nano" has only 930 thousand and is the leader in terms of the ratio with the other metrics.

The reason for these outcomes could be attributed to various factors. For instance, gMLP has better parameter efficiency compared to Transformers. Transformers typically require a large amount of data to train their numerous parameters, including attention mechanisms and positional encodings. On the other hand, gMLP relies mainly on MLP layers, which generally have fewer parameters, making them less dependent on large datasets for practical training. Additionally, transformers may be susceptible to overfitting when dealing

with smaller datasets due to excessive learning. In contrast, gMLP's architecture, which relies on MLP layers supported by gating mechanisms, helps mitigate overfitting due to its simple structure.

TABLE 3. Evaluation metrics of different DL models

Network	BA	P	R	F1	AUC	Param.
BiLSTM	0.822	0.865	0.763	0.811	0.873	82K
ResNet	0.833	0.867	0.752	0.818	0.908	478K
PatchTST	0.844	0.881	0.795	0.836	0.912	4.3M
Inception	0.860	0.838	0.891	0.864	0.950	388K
OmniScale	0.865	0.905	0.795	0.855	0.937	252K
TSiT	0.870	0.866	**0.903**	0.884	0.920	86M
TST	0.879	**0.915**	0.867	0.885	0.952	1.5M
gMLP Nano	**0.887**	0.882	**0.892**	**0.887**	**0.955**	930K
gMLP Tiny	**0.897**	**0.920**	0.870	**0.895**	**0.979**	5.9M



(a) (b)

FIGURE 6. Receiver operating characteristic curves (ROC) for binary classification of ERG signals with corresponding areas under curves (AUC) for all tested models (a) and for the best three models with highest AUC (b): gMLP "Tiny", gMLP "Nano", TST.

When the amplitude of the b-wave is plotted against flash strength, the function is termed the "photopic hill" with the peak dominated by OFF-retinal pathways and the later plateau phase the ON-retinal pathways [57]. The metrics were compared by performing an ablation analysis where triplets of strength were used with gMLP Nano to evaluate any differences in the range of flash strengths to the classification model. We applied the same training procedure independently for the three flash strength ranges. The test and training subsets were consequently reduced by a factor of three. However, in this way, we could compare the relative contribution of each triplet of strengths to the overall classification. Table 4 shows the overall performance of the gMLP Nano using the three triplet flash strengths ranges that corresponded to the early (-0.367 to 0.114), mid (0.398 to 0.799), and later (0.949 to 1.204) portions of the photopic hill.

V. DISCUSSION

In this work, we have built on previous studies that have used combinations of ML with ERG signal analysis [24]–[26], [34] or time domain parameters of the DA- and LA-ERG waveform [14], [20], [21], Here we investigated the complete time series of the ERG waveform signal that provides some advantages over previous studies by using DL models that can automatically learn relevant features from the raw time-series dataset. Time-series data of the ERG waveform can contain complex and hierarchical patterns that may not be evident through manual feature engineering. Analyzing the entire signal allows these models to uncover complex temporal relationships that may be missed when relying on waveform peak parameters such as amplitude and time or time-frequency analytical solutions.

TABLE 4. Evaluation metrics of gMLP Nano with three flash strength ranges.

Strength	BA	P	R	F1	AUC
0.949 & 1.114 & 1.204	0.856	0.800	0.933	0.861	0.965
0.398 & 0.602 & 0.799	0.851	0.811	0.901	0.865	0.901
-0.367 & -0.119 & 0.114	0.849	0.810	0.937	0.869	0.904

There is interest in expanding the clinical potential of the ERG as a biomarker for disorders affecting the CNS [3], [58]. The application of ML to identify phenotype-genotype correlations has been demonstrated in retinal disease [32], and this may also be beneficial in complex disorders where genotypic risk factors can be linked to a phenotype in other inherited retinal diseases [59]. While previous studies have utilized signal analysis of the ERG signal to extract salient features for classification [34], this subsequent analysis relied on features identified from the raw time series that may offer an alternative and additional method in the quest for the classification and identification of retinal and CNS disorders based on retinal functional biomarkers [3], [15], [16], [32], [60].

Previous studies using parameters from a Gaussian and logistic growth function to model the photopic hill as defined by Hamilton et al. [61] indicated a more likely loss of the ON-pathway associated with the higher flash strengths. In the ablation analysis, we found equivalent contributions of the three selected triplet flash strength ranges - although the higher flash strength series had a slightly superior overall balanced accuracy in keeping with previous suggestions that there is a more significant ON-pathway loss in ASD [19]–[21]. The use of regions of the photopic hill has been applied to discriminating schizophrenia from bipolar and control groups previously [16] and in this analysis, selecting a range of three strengths in the higher range provides similar performance to all strengths (three high strengths AUC = 0.965 compared with all strengths AUC = 0.955 and slightly lower BA = 0.856 compared with BA = 0.887) using gMLP Nano. This may help select the minimal number of flash strengths required in a test to classify ASD subjects accurately in future studies and support the potential application of ERG analysis in classifying ASD in this subject group.

For the analysis of the ERG signal to be validated, further studies will need to be performed in which sex and developmental age are matched between groups to minimize heterogeneity between study populations [62]. Additional DL

models using feature engineering techniques, such as distributional analysis of features, may also be advantageous in the future with larger and more complex clinical datasets to support more robust classification models [63], [64]. The ERG may form part of the classification of biotypes [65] or transdiagnostic endophenotypes [66] that could provide improved stratification of neurodevelopmental conditions [46] in conjunction with genotypic and phenotypic data [67].

The use of gMLP provides a powerful method to explore further and refine the diagnostic potential of the ERG signal in CNS disorders such as ASD. This may help with improved earlier interventions and better outcomes for individuals with a diagnosis of ASD [17], [42]. However, the current limitations with respect to ASD diagnosis, based on ERG recordings, is that, typically, children will be diagnosed before the age of 5 years with early indications of language delay, lack of declarative gestures, and eye contact commonly observed before a formal diagnosis [62]. In addition, ASD may present in combination with co-occurring developmental conditions such as ADHD, and as such, the specificity of the classification would need to be further evaluated in children with ASD plus an additional co-occurring neurodevelopmental condition(s) [68]. Reducing the cost and increasing the accessibility to recording the ERG in clinical populations may also be improved in the future with technological developments, including smartphone-based devices that can record and perform sophisticated analyses of the waveform [41]. Thus, further studies are required in younger cohorts with a wide spectrum of neurodevelopmental conditions to establish the LA-ERG as a specific biomarker for ASD. However, these early findings further support the use of ML in the potential classification of neurodevelopmental conditions [36], [37] and with larger datasets in clinical populations, there is the potential for the ERG to assist with triaging children that may require further clinical assessments or to monitor therapeutic interventions targeting the CNS in neurodevelopmental disorders.

VI. ETHICS

Clinical recordings were approved by local institutional ethics committees and were in accordance with the Declaration of Helsinki.

VII. CONCLUSIONS

gMLP is a novel architecture with the strengths of traditional MLP and challenges some aspects of Transformers. Our findings have demonstrated comparability to Transformers in the ERG time-series domain. gMLP has a reasonably simple structure while offering the ability to process long-range dependencies in sequential data: gMLP "Tiny" showed the highest balanced accuracy of 0.89 on the dataset and performed better than or equivalent to other time-series models. The next best was the "Nano" version, with far fewer parameters with the training process requiring a manageable number of signals. Therefore, its application to the ERG waveform is promising where clinical populations may be rare, such as in inherited retinal diseases and heterogeneic neurodevelopmental disorders. The analysis of the ERG waveform as a potential biomarker in conjunction with the gMLP could further improve the accuracy of ASD detection. The application of gMLP may also contribute to the field of ERG analysis in human and animal studies [2] and as illustrated in this study able to provide a robust method for detecting ASD within this clinical population.

REFERENCES

[1] A. G. Robson, J. Nilsson, S. Li, S. Jalali, A. B. Fulton, A. P. Tormene, G. E. Holder, and S. E. Brodie, "ISCEV Guide to visual electrodiagnostic procedures," *Documenta Ophthalmologica*, vol. 136, no. 1, p. 1–26, 2018.

[2] P. A. Constable, J. K. Lim, and D. A. Thompson, "Retinal Electrophysiology in Central Nervous System Disorders. A review of human and mouse studies," *Frontiers in Neuroscience*, vol. 17, 2023.

[3] M.-A. Dubois, C.-A. Pelletier, C. Mérette, V. Jomphe, R. Turgeon, R. E. Bélanger, S. Grondin, and M. Hébert, "Evaluation of electroretinography (erg) parameters as a biomarker for adhd," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, p. 110807, 2023.

[4] G. Gross, K. Tursini, E. Albuisson, K. Angioi-Duprez, J.-B. Conart, V. Louis Dorr, R. Schwan, and T. Schwitzer, "Bipolar disorders and retinal electrophysiological markers (bimar): Study protocol for a comparison of electroretinogram measurements between subjects with bipolar disorder and a healthy control group," *Frontiers in Psychiatry*, vol. 13, p. 960512, 2022.

[5] K. K. Tran, V. H. Wong, J. K. Lim, A. Shahandeh, A. Hoang, D. I. Finkelstein, B. V. Bui, and C. T. Nguyen, "Characterization of retinal function and structure in the mptp murine model of parkinson's disease," *Scientific reports*, vol. 12, no. 1, p. 7610, 2022.

[6] A. G. Robson, L. J. Frishman, J. Grigg, R. Hamilton, B. G. Jeffrey, M. Kondo, S. Li, and D. L. McCulloch, "Iscev standard for full-field clinical electroretinography (2022 update)," *Documenta Ophthalmologica*, vol. 144, no. 3, pp. 165–177, 2022.

[7] Y. Bhatt, D. M. Hunt, and L. S. Carvalho, "The origins of the full-field flash electroretinogram b-wave," *Frontiers in Molecular Neuroscience*, vol. 16, 2023.

[8] M. Thomas and T. Lamb, "Light adaptation and dark adaptation of human rod photoreceptors measured from the a-wave of the electroretinogram," *The Journal of physiology*, vol. 518, no. 2, pp. 479–496, 1999.

[9] J. H. Brandstätter and I. Hack, "Localization of glutamate receptors at a complex synapse: the mammalian photoreceptor synapse," *Cell and Tissue Research*, vol. 303, pp. 1–14, 2001.

[10] S. Barnes, J. C. Grove, C. F. McHugh, A. A. Hirano, and N. C. Brecha, "Horizontal cell feedback to cone photoreceptors in mammalian retina: novel insights from the gaba-ph hybrid model," *Frontiers in Cellular Neuroscience*, vol. 14, p. 595064, 2020.

[11] L. Wachtmeister, "Some aspects of the oscillatory response of the retina," *Progress in Brain Research*, vol. 131, pp. 465–474, 2001.

[12] T. Schwitzer, J. Lavoie, A. Giersch, R. Schwan, and V. Laprevote, "The emerging field of retinal electrophysiological measurements in psychiatric research: a review of the findings and the perspectives in major depressive disorder," *Journal of psychiatric research*, vol. 70, pp. 113–120, 2015.

[13] B. Nowacka, W. Lubiński, K. Honczarenko, A. Potemkowski, and K. Safranow, "Bioelectrical function and structural assessment of the retina in patients with early stages of parkinson's disease (pd)," *Documenta Ophthalmologica*, vol. 131, pp. 95–104, 2015.

[14] I. O. Lee, D. H. Skuse, P. A. Constable, F. Marmolejo-Ramos, L. R. Olsen, and D. A. Thompson, "The electroretinogram b-wave amplitude: a differential physiological measure for attention deficit hyperactivity disorder and autism spectrum disorder," *Journal of Neurodevelopmental Disorders*, vol. 14, no. 1, p. 30, 2022.

[15] D. L. Demmin, Q. Davis, M. Roché, and S. M. Silverstein, "Electroretinographic anomalies in schizophrenia." *Journal of abnormal psychology*, vol. 127, no. 4, p. 417, 2018.

[16] M. Hébert, C. Mérette, A.-M. Gagné, T. Paccalet, I. Moreau, J. Lavoie, and M. Maziade, "The electroretinogram may differentiate schizophrenia from bipolar disorder," *Biological psychiatry*, vol. 87, no. 3, pp. 263–270, 2020.

[17] A. J. Whitehouse, K. J. Varcin, G. A. Alvares, J. Barbaro, C. Bent, M. Boutrus, L. Chetcuti, M. N. Cooper, A. Clark, E. Davidson *et al.*, "Pre-emptive intervention versus treatment as usual for infants showing

early behavioural risk signs of autism spectrum disorder: a single-blind, randomised controlled trial," *The Lancet Child & Adolescent Health*, vol. 3, no. 9, pp. 605–615, 2019.

[18] M. Parellada, Á. Andreu-Bernabeu, M. Burdeus, A. San José Cáceres, E. Urbiola, L. L. Carpenter, N. V. Kraguljac, W. M. McDonald, C. B. Nemeroff, C. I. Rodriguez *et al.*, "In search of biomarkers to guide interventions in autism spectrum disorder: a systematic review," *American Journal of Psychiatry*, vol. 180, no. 1, pp. 23–40, 2023.

[19] E. R. Ritvo, D. Creel, G. Realmuto, A. S. Crandall, B. Freeman, J. B. Bateman, R. Barr, C. Pingree, M. Coleman, and R. Purple, "Electroretinograms in autism: a pilot study of b-wave amplitudes." *The American Journal of Psychiatry*, vol. 145, no. 2, pp. 229–232, 1988.

[20] P. A. Constable, E. R. Ritvo, A. R. Ritvo, I. O. Lee, M. L. McNair, D. Stahl, J. Sowden, S. Quinn, D. H. Skuse, D. A. Thompson *et al.*, "Light-adapted electroretinogram differences in autism spectrum disorder," *Journal of Autism and Developmental Disorders*, vol. 50, pp. 2874–2885, 2020.

[21] P. A. Constable, S. B. Gaigg, D. M. Bowler, H. Jägle, and D. A. Thompson, "Full-field electroretinogram in autism spectrum disorder," *Documenta Ophthalmologica*, vol. 132, pp. 83–99, 2016.

[22] E. B. Friedel, M. Schäfer, D. Endres, S. Maier, K. Runge, M. Bach, S. P. Heinrich, D. Ebert, K. Domschke, L. Tebartz van Elst *et al.*, "Electroretinography in adults with high-functioning autism spectrum disorder," *Autism Research*, vol. 15, no. 11, pp. 2026–2037, 2022.

[23] P. A. Constable, F. Marmolejo-Ramos, M. Gauthier, I. O. Lee, D. H. Skuse, and D. A. Thompson, "Discrete wavelet transform analysis of the electroretinogram in autism spectrum disorder and attention deficit hyperactivity disorder," *Frontiers in Neuroscience*, vol. 16, p. 890461, 2022.

[24] H. F. Posada-Quintero, S. M. Manjur, M. B. Hossain, F. Marmolejo-Ramos, I. O. Lee, D. H. Skuse, D. A. Thompson, and P. A. Constable, "Autism spectrum disorder detection using variable frequency complex demodulation of the electroretinogram," *Research in Autism Spectrum Disorders*, vol. 109, p. 102258, 2023.

[25] S. M. Manjur, B. M. Hossain, P. A. Constable, D. A. Thompson, F. Marmolejo-Ramos, I. O. Lee, and H. Posada-Quintero, "Spectral analysis of electroretinography to differentiate autism spectrum disorder and attention deficit hyperactivity disorder," in *IEEE-EMBS International Conference on Biomedical and Health Informatics*, 2023.

[26] S. M. Manjur, L. R. M. Diaz, I. O. Lee, D. H. Skuse, D. A. Thompson, F. Marmolejos-Ramos, P. A. Constable, and H. F. Posada-Quintero, "Detecting autism spectrum disorder and attention deficit hyperactivity disorder using multimodal time-frequency analysis with machine learning using the electroretinogram from two flash strengths," *Journal of Autism and Developmental Disorders*, pp. 1–14, 2024.

[27] M. Brabec, P. A. Constable, D. A. Thompson, and F. Marmolejo-Ramos, "Group comparisons of the individual electroretinogram time trajectories for the ascending limb of the b-wave using a raw and registered time series," *BMC Research Notes*, vol. 16, no. 1, p. 238, 2023.

[28] Y. A. Veturi, W. Woof, T. Lazebnik, I. Moghul, P. Woodward-Court, S. K. Wagner, T. A. C. de Guimarães, M. D. Varela, B. Liefers, P. J. Patel *et al.*, "Syntheye: Investigating the impact of synthetic data on artificial intelligence-assisted gene diagnosis of inherited retinal disease," *Ophthalmology Science*, vol. 3, no. 2, p. 100258, 2023.

[29] I. S. Yapici, O. Erkaymaz, and R. U. Arslan, "A hybrid intelligent classifier to estimate obesity levels based on erg signals," *Physics Letters A*, vol. 399, p. 127281, 2021.

[30] A. López-Dorado, J. Pérez, M. J. Rodrigo, J. M. Miguel-Jiménez, M. Ortiz, L. de Santiago, E. López-Guillén, R. Blanco, C. Cavalliere, E. M. S. Morla *et al.*, "Diagnosis of multiple sclerosis using multifocal erg data feature fusion," *Information Fusion*, vol. 76, pp. 157–167, 2021.

[31] A. Zhdanov, A. Dolganov, D. Zanca, V. Borisov, and M. Ronkin, "Advanced analysis of electroretinograms based on wavelet scalogram processing," *Applied Sciences*, vol. 12, no. 23, p. 12365, 2022.

[32] S. L. Glinton, A. Calcagni, W. Lilaonitkul, N. Pontikos, S. Vermeirsch, G. Zhang, G. Arno, S. K. Wagner, M. Michaelides, P. A. Keane *et al.*, "Phenotyping of abca4 retinopathy by machine learning analysis of full-field electroretinography," *Translational Vision Science & Technology*, vol. 11, no. 9, pp. 34–34, 2022.

[33] M. K. Gajendran, L. J. Rohowetz, P. Koulen, and A. Mehdizadeh, "Novel machine-learning based framework using electroretinography data for the detection of early-stage glaucoma," *Frontiers in Neuroscience*, vol. 16, p. 869137, 2022.

[34] S. M. Manjur, M.-B. Hossain, P. A. Constable, D. A. Thompson, F. Marmolejo-Ramos, I. O. Lee, D. H. Skuse, and H. F. Posada-Quintero, "Detecting autism spectrum disorder using spectral analysis of electroretinogram and machine learning: Preliminary results," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2022, pp. 3435–3438.

[35] M. Kulyabin, A. Zhdanov, A. Dolganov, and A. Maier, "Optimal combination of mother wavelet and ai model for precise classification of pediatric electroretinogram signals," *Sensors*, vol. 23, no. 13, p. 5813, 2023.

[36] C. Song, Z.-Q. Jiang, D. Liu, and L.-L. Wu, "Application and research progress of machine learning in the diagnosis and treatment of neurodevelopmental disorders in children," *Frontiers in Psychiatry*, vol. 13, p. 960672, 2022.

[37] C. Moreau, C. Deruelle, and G. Auzias, "Machine learning for neurodevelopmental disorders," *Machine Learning for Brain Disorders*, pp. 977–1007, 2023.

[38] R. Noguez Imm, J. Muñoz-Benitez, D. Medina, E. Barcenas, G. Molero-Castillo, P. Reyes-Ortega, J. A. Hughes-Cano, L. Medrano-Gracia, M. Miranda-Anaya, G. Rojas-Piloni *et al.*, "Preventable risk factors for type 2 diabetes can be detected using noninvasive spontaneous electroretinogram signals," *Plos one*, vol. 18, no. 1, p. e0278388, 2023.

[39] O. A. Mahroo, "Visual electrophysiology and "the potential of the potentials"," *Eye*, pp. 1–10, 2023.

[40] R. Hamilton, "Clinical electrophysiology of vision—commentary on current status and future prospects," *Eye*, vol. 35, no. 9, pp. 2341–2343, 2021.

[41] O. Huddy, A. Tomas, S. M. Manjur, and H. Posada-Quintero, "Prototype for smartphone-based electroretinogram," in *2023 IEEE 19th International Conference on Body Sensor Networks (BSN)*, 2023, pp. 1–4.

[42] A. Masi, C. Dissanayake, T. Alach, K. Cameron, K. Fordyce, G. Frost, R. Grove, H. Heussler, N. Silove, R. Sulek *et al.*, "Clinical outcomes and associated predictors of early intervention in autism spectrum disorder: A study protocol," *BMJ open*, vol. 11, no. 8, p. e047290, 2021.

[43] C. Lord, M. Rutter, P. DiLavore, S. Risi, K. Gotham, S. Bishop, R. Luyster, and W. Guthri, "Autism diagnostic observation schedule: Ados-2| pearson assessment," 2012.

[44] A. D. Krakowski, K. T. Cost, P. Szatmari, E. Anagnostou, J. Crosbie, R. Schachar, E. Duku, S. Georgiades, M. Ayub, E. Kelley *et al.*, "Characterizing the asd–adhd phenotype: Measurement structure and invariance in a clinical sample," *Journal of Child Psychology and Psychiatry*, vol. 63, no. 12, pp. 1534–1543, 2022.

[45] S. Jacob, J. J. Wolff, M. S. Steinbach, C. B. Doyle, V. Kumar, and J. T. Elison, "Neurodevelopmental heterogeneity and computational approaches for understanding autism," *Translational psychiatry*, vol. 9, no. 1, p. 63, 2019.

[46] C. Molloy and L. Gallagher, "Can stratification biomarkers address the heterogeneity of autism spectrum disorder?" *Irish Journal of Psychological Medicine*, vol. 39, no. 3, pp. 305–311, 2022.

[47] M. Kulyabin, A. Zhdanov, A. Dolganov, M. Ronkin, V. Borisov, and A. Maier, "Enhancing electroretinogram classification with multi-wavelet analysis and visual transformer," *Sensors*, vol. 23, no. 21, p. 8727, 2023.

[48] H. Liu, Z. Dai, D. So, and Q. V. Le, "Pay attention to mlps," *Advances in Neural Information Processing Systems*, vol. 34, pp. 9204–9215, 2021.

[49] M. Rufiange, J. Dassa, O. Dembinska, R. K. Koenekoop, J. M. Little, R. C. Polomeno, M. Dumont, S. Chemtob, and P. Lachapelle, "The photopic erg luminance-response function (photopic hill): method of analysis and clinical application," *Vision research*, vol. 43, no. 12, pp. 1405–1412, 2003.

[50] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[51] Z. Wang, W. Yan, and T. Oates, "Time series classification from scratch with deep neural networks: A strong baseline," in *2017 International joint conference on neural networks (IJCNN)*. IEEE, 2017, pp. 1578–1585.

[52] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *International conference on machine learning*. Pmlr, 2013, pp. 1310–1318.

[53] H. Ismail Fawaz, B. Lucas, G. Forestier, C. Pelletier, D. F. Schmidt, J. Weber, G. I. Webb, L. Idoumghar, P.-A. Muller, and F. Petitjean, "Inceptiontime: Finding alexnet for time series classification," *Data Mining and Knowledge Discovery*, vol. 34, no. 6, pp. 1936–1962, 2020.

[54] W. Tang, G. Long, L. Liu, T. Zhou, M. Blumenstein, and J. Jiang, "Omni-scale cnns: a simple and effective kernel size configuration for time series classification," *arXiv preprint arXiv:2002.10061*, 2020.

[55] G. Zerveas, S. Jayaraman, D. Patel, A. Bhamidipaty, and C. Eickhoff, "A transformer-based framework for multivariate time series representation learning," ser. KDD '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 2114–2124. [Online]. Available: https://doi.org/10.1145/3447548.3467401

[56] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.
[57] M.-L. Garon, A. Dorfman, J. Racine, R. Koenekoop, J. Little, and P. Lachapelle, "Estimating on and off contributions to the photopic hill: normative data and clinical applications," *Documenta Ophthalmologica*, vol. 129, pp. 9–16, 2014.
[58] I. Koychev, W. El-Deredy, T. Mukherjee, C. Haenschel, and J. Deakin, "Core dysfunction in schizophrenia: electrophysiology trait biomarkers," *Acta Psychiatrica Scandinavica*, vol. 126, no. 1, pp. 59–71, 2012.
[59] N. Schneider, Y. Sundaresan, P. Gopalakrishnan, A. Beryozkin, M. Hanany, E. Y. Levanon, E. Banin, S. Ben-Aroya, and D. Sharon, "Inherited retinal diseases: Linking genes, disease-causing variants, and relevant therapeutic modalities," *Progress in Retinal and Eye Research*, vol. 89, p. 101029, 2022.
[60] I. Moreau, M. Hébert, M. Maziade, A. Painchaud, and C. Mérette, "The electroretinogram as a potential biomarker of psychosis in children at familial risk," *Schizophrenia Bulletin Open*, vol. 3, no. 1, p. sgac016, 2022.
[61] R. Hamilton, M. Bees, C. A. Chaplin, and D. L. McCulloch, "The luminance–response function of the human photopic electroretinogram: A mathematical model," *Vision research*, vol. 47, no. 23, pp. 2968–2972, 2007.
[62] C. Lord, M. Elsabbagh, G. Baird, and J. Veenstra-Vanderweele, "Autism spectrum disorder," *The lancet*, vol. 392, no. 10146, pp. 508–520, 2018.
[63] L. Zhang, A. Rao, and M. Agrawala, "Adding conditional control to text-to-image diffusion models," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 3836–3847.
[64] A. Hussain, S. Ali, H.-C. Kim *et al.*, "Activity detection for the wellbeing of dogs using wearable sensors based on deep learning," *IEEE Access*, vol. 10, pp. 53 153–53 163, 2022.
[65] Y. Du, H. Hao, Y. Xing, J. Niu, and V. D. Calhoun, "A transdiagnostic biotype detection method for schizophrenia and autism spectrum disorder based on graph kernel," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2021, pp. 3241–3244.
[66] L. Waterhouse, "Heterogeneity thwarts autism explanatory power: A proposal for endophenotypes," *Frontiers in Psychiatry*, vol. 13, p. 947653, 2022.
[67] S. Li, Z. Guo, J. B. Ioffe, Y. Hu, Y. Zhen, and X. Zhou, "Text mining of gene–phenotype associations reveals new phenotypic profiles of autism-associated genes," *Scientific Reports*, vol. 11, no. 1, p. 15269, 2021.
[68] M. Micai, L. M. Fatta, L. Gila, A. Caruso, T. Salvitti, F. Fulceri, A. Ciaramella, R. D'Amico, C. Del Giovane, M. Bertelli *et al.*, "Prevalence of co-occurring conditions in children and adults with autism spectrum disorder: A systematic review and meta-analysis," *Neuroscience & Biobehavioral Reviews*, vol. 155, p. 105436, 2023.

MIKHAIL KULYABIN received an Engineer's degree in mechanical engineering from Bauman Moscow State Technical University (BMSTU), Russia, in 2018, and a M.Sc. degree in computational engineering from Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Germany, in 2022, where he is currently pursuing the Ph.D. in computer science. He joined FAU's Pattern Recognition Laboratory in 2023. He is a part of the "Artificial Intelligence for Diagnosing Retinal Diseases (AID)" project. His main research interest lies in applying AI to topics related to ophthalmology.

PAUL CONSTABLE studied optometry at the University of Melbourne and graduated in 1991. He completed his Ph.D. at City University, London 2007, investigating drug transporters in the outer blood-retinal barrier. He began exploring visual function in autism spectrum disorder and leading the recent studies on the electroretinogram in autism. Dr. Constable was the academic lead for optometry at Flinders University and the course coordinator for optometry at City University.

He is a member of the International Society for Clinical Electrophysiology of Vision and contributes to the clinical standard for the electrooculogram. Dr. Constable leads a team of collaborators investigating retinal biomarkers, including structural and functional measures for neurological and vascular conditions.

ALEKSEI ZHDANOV received M.Sc. degrees in Information and Communication Technology from the Ural Federal University, Yekaterinburg, and the University of Erlangen-Nuremberg, Erlangen, in 2018 as part of the Erasmus Mundus Joint Masters Scholarships program. He graduated from the Ph.D. program in Biomedical/Medical Engineering at Ural Federal University in 2022. From 2018, he working as a Research Engineer at the Engineering School of Information Technologies, Telecommunications, and Control Systems, Ural Federal University. In 2023, he was a Researcher at the Machine Learning and Data Analytics Lab, University of Erlangen-Nuremberg, as part of the Bi-nationally Supervised Doctoral Degrees/Cotutelle DAAD Research Grant. Since 2023, he has been a Project Engineer at Siemens Healthineers, Erlangen. He is the author of more than 53 articles and more than 10 inventions. His research interests include biomedical engineering and medical devices.

IRENE LEE is currently a Data Manager at the Great Ormond Street Institute of Child Health. Irene was previously the Scientific Officer (Imperial Cancer Research Fund) at Imperial College London. Irene completed her MSc (Neuropharmacology) in 2021 from Bristol University and has been a member of the Institute of Biomedical Science since 1992. Irene is currently a member of the project team for IMAGINE-2: Stratifying Genomic Causes of Intellectual Disability by Mental Health Outcomes in Childhood and Adolescence, funded by the Medical Research Council (UK).

Irene has expertise in recording biological signals, evaluating neurodevelopmental disorders, and linking genotypes with phenotypes. Irene is currently collaborating on retinal biomarkers in neurodevelopmental disorders.

DOROTHY THOMPSON is a Consultant Clinical Scientist working in the UK National Health Service of Great Ormond Street for Children, London, UK. She leads the Clinical Vision Science Services in the Ophthalmology Department, which includes providing specialist pediatric visual electrodiagnostic tests. Dorothy is also an Honorary Associate Professor at University College London, Great Ormond Street Institute for Child Health, Developmental Biology, and Cancer Department. After Ph.D. research into the retinal origins of the pattern electroretinogram and Wellcome post-doctoral Fellowships about mapping specialized cortical visual areas of the brain, Dorothy's research interest continues into the mechanisms of retinal and visual pathway signaling and their clinical application. Dorothy has authored more than 130 articles. She sits on the Board of the International Society for Visual Electrophysiology of Vision (ISCEV) as the Vice President for Europe and Africa. She is an associate editor of the journal Documenta Ophthalmologica.

ANDREAS MAIER (Senior Member, IEEE) was born in Erlangen, Germany, in November 1980. He graduated in computer science and the Ph.D. degree from Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, in 2005 and 2009, respectively. From 2005 to 2009, he was with the Pattern Recognition Laboratory, Computer Science Department, Friedrich-Alexander-Universität Erlangen-Nürnberg. In this period, he developed the first online speech intelligibility assessment tool—PEAKS—that has been used to analyze over 4000 patients and control subjects so far. From 2009 to 2010, he started working on flat-panel C-arm CT as a Post-Doctoral Fellow at the Radiological Sciences Laboratory, Department of Radiology, Stanford University, Stanford, CA, USA. From 2011 to 2012, he joined Siemens Healthcare, Erlangen, Germany, as an Innovation Project Manager and was responsible for reconstruction topics in the Angiography and X-ray business unit. In 2012, he returned to the Friedrich-Alexander-Universität Erlangen-Nürnberg as the Head of the Medical Reconstruction Group at the Pattern Recognition Laboratory, where he became a Professor and the Head of the Pattern Recognition Laboratory in 2015. His major research subject was medical signal processing in speech data. His research interests include medical imaging, image and audio processing, digital humanities, and interpretable machine learning, and the use of known operators.

Dr. Maier has been a member of the Steering Committee of the European Time Machine Consortium since 2016. In 2018, he was awarded an ERC Synergy Grant "4D nanoscope."