IEEE Journal of Translational Engineering in Health and Medicine

Received 14 February 2023; revised 16 July 2023 and 9 September 2023; accepted 21 September 2023.
Date of publication 29 September 2023; date of current version 24 November 2023.

Digital Object Identifier 10.1109/JTEHM.2023.3320713

A Robust Gaze Estimation Approach via Exploring Relevant Electrooculogram Features and Optimal Electrodes Placements

ZHENG ZENG[1], (Student Member, IEEE), LINKAI TAO[2], HANGYU ZHU[1], YUNFENG ZHU[1], LONG MENG[1], (Member, IEEE), JIAHAO FAN[1], CHEN CHEN[3], AND WEI CHEN[1,3], (Senior Member, IEEE)

[1]Center for Intelligent Medical Electronics, School of Information Science and Technology, Fudan University, Shanghai 200433, China
[2]Department of Industrial Design, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands
[3]Human Phenome Institute, Fudan University, Shanghai 201203, China

CORRESPONDING AUTHORS: C. CHEN (chenchen_fd@fudan.edu.cn) AND W. CHEN (w_chen@fudan.edu.cn)

This work was supported in part by the Shanghai Municipal Science and Technology International Research and Development Collaboration Project under Grant 20510710500, in part by the Shanghai Committee of Science and Technology under Grant 20S31903900, and in part by the National Natural Science Foundation of China under Grant 62001118.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Fudan University under Application No. FE23170I.

This article has supplementary downloadable material available at https://doi.org/10.1109/JTEHM.2023.3320713, provided by the authors.

ABSTRACT Gaze estimation, as a technique that reflects individual attention, can be used for disability assistance and assisting physicians in diagnosing diseases such as autism spectrum disorder (ASD), Parkinson's disease, and attention deficit hyperactivity disorder (ADHD). Various techniques have been proposed for gaze estimation and achieved high resolution. Among these approaches, electrooculography (EOG)-based gaze estimation, as an economical and effective method, offers a promising solution for practical applications. Objective: In this paper, we systematically investigated the possible EOG electrode locations which are spatially distributed around the orbital cavity. Afterward, quantities of informative features to characterize physiological information of eye movement from the temporal-spectral domain are extracted from the seven differential channels. Methods and procedures: To select the optimum channels and relevant features, and eliminate irrelevant information, a heuristical search algorithm (i.e., forward stepwise strategy) is applied. Subsequently, a comparative analysis of the impacts of electrode placement and feature contributions on gaze estimation is evaluated via 6 classic models with 18 subjects. Results: Experimental results showed that the promising performance was achieved both in the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) within a wide gaze that ranges from $-50°$ to $+50°$. The MAE and RMSE can be improved to 2.80° and 3.74° ultimately, while only using 10 features extracted from 2 channels. Compared with the prevailing EOG-based techniques, the performance improvement of MAE and RMSE range from 0.70° to 5.48° and 0.66° to 5.42°, respectively. Conclusion: We proposed a robust EOG-based gaze estimation approach by systematically investigating the optimal channel/feature combination. The experimental results indicated not only the superiority of the proposed approach but also its potential for clinical application. Clinical and translational impact statement: Accurate gaze estimation is a key step for assisting disabilities and accurate diagnosis of various diseases including ASD, Parkinson's disease, and ADHD. The proposed approach can accurately estimate the points of gaze via EOG signals, and thus has the potential for various related medical applications.

INDEX TERMS Channel selection, electrooculography (EOG), feature selection, gaze estimation, saccade.

IEEE Journal of Translational Engineering in Health and Medicine

I. INTRODUCTION

Gaze estimation, as a technique to estimate the points of gaze, can be used in a large variety of medical applications. Since it can reflect individual attention and abundant physiological information of the subjects, gaze estimation is of great potential for disability assistance [1], Parkinson's disease diagnosis [2], ASD diagnosis [3], [4], and ADHD diagnosis [5], [6], [7] So far, varieties of gaze estimation methods have been proposed, ranging from scleral search coils, video-oculography (VOG) to EOG. The VOG-based technique is the most popular approach to estimating gaze due to its high performance in terms of accuracy. The VOG-based technique can obtain the 2-D image of the eyes to compute the gaze point by employing the image processing algorithms [8]. However, it also faces several intrinsic limits. Firstly, redundant computing resources and storage are required. Secondly, a VOG-based system needs to be equipped with an external stable infrared lighting source, and it is sensitive to light and motion artifacts. Last, for severely nearsighted subjects, the reflection of the glasses can make the ocular pose unable to be tracked.

In contrast, the EOG-based technique can overcome the aforementioned flaws. Conventional EOG-based approaches can be roughly divided into polynomial methods, electric methods, and machine learning methods. All these methods mainly employed the standard 4-channel electrode placement to capture the eye movements and estimate the gaze points, accordingly [9], [10]. The polynomial-based method is mainly to find the regression relationship between EOG voltage and saccadic angle, including linear regression and nonlinear regression. Specifically, Kumar et al. [11] employed a linear regression to estimate gaze in a range of $\pm 45^\circ$ ($R^2 = 0.9921$) via the standard 4-channel placement. Barea et al. [12] proposed the derivative-based linear regression model, in which the error was around 2°. Favre et al. [13] exploited a nonlinear model for gaze estimation. The mean estimation error was 4.67°. Manabe et al. [14] employed different order polynomials for estimation, and the absolute angle error was less than 4°. The second method, electric models, estimated the gaze points from the perspectives of electrode position and eyeball physiological information. To illustrate, Barbara et al. [15] adopted the battery model (one of the electric models) to estimate gaze points via the standard 4-channel placement. The mean absolute error was about 2.42° across 6 subjects. Further, an extended Kalman filter has been applied in [16] to improve the accuracy. The third method, machine learning models, has shown great potential in eye movement detection. Bulling et al. [17] used a support vector machine (SVM) to detect repetitive patterns of eye movements. The EOG signals were acquired via the standard 4-channel electrode placement. Latifoglu et al. [18] employed random forests to detect the reading eye movements. Yan et al. [19] used the backpropagation network to distinguish the eye movements with an accuracy of 96%.

However, existing studies have not fully investigated the potential optimal electrodes around the eyes. Since multiple channels can provide relevant and irrelevant spatial information [20], [21], it is of great importance to further refine the optimal channels for performance improvement. In addition, excessive features may also bring the risk of overfitting, leading to model degradation. To provide a more compact and accurate EOG-based gaze estimation model, it requires a comprehensive investigation of discriminative features to exclude harmful features. To this end, several general solutions, including wrapper, filter, and embedded methods, have been widely applied for channel/feature selection [22]. The filter method mainly employs manual evaluations (i.e., dependency measure, distance measure, and consistency measure) to generate the candidate subset [23]. However, the manual evaluations following prior knowledge are often suboptimal, and difficult to capture the complementary information. By contrast, the wrapper method, which generates the candidate subset according to the model assessment, can directly search for an optimized candidate subset for the given model. Since the performance of the model is used as the evaluation criterion, impressive results can be obtained via the wrapper methods [22].

In this study, a robust gaze estimation approach was proposed by exploring relevant EOG features and optimal electrode placements. We investigated the performance of 7 pairs of differential channels around the orbital cavity instead of using the standard 4-channel electrode placement. Meanwhile, 13 widely used features in EOG fields were evaluated. We employed the wrapper method to further refine the channels and features. Specifically, a forward greedy strategy was applied to screen discriminative channels and features. The selected optimal channels/features were also evaluated on 6 classical models to indicate the superiority. The contributions of our work can be summarized as follows:

1) We systematically investigated the possible electrode locations around the orbital cavity to enhance the performance of EOG-based gaze estimation. Specifically, we employed the wrapper method combined with a forward stepwise strategy to eliminate the irrelevant channels, thus further boosting the performance.

2) Quantities of informative features to characterize physiological information of gaze estimation from the temporal-spectral domain are extracted. We further refine the feature via the wrapper method to propose a compact model which can take full advantage of the discriminative features.

3) We also compared the performance with prevailing EOG-based gaze estimation approaches. Our results demonstrated that favorable performance can be achieved by using the optimal channel/feature combination.

The rest of this paper is organized as follows. In Section II, we introduce the data collection and data preprocessing. In Section III, the extracted features and validation methods are described. The results are presented in Section IV. In Section V, we discuss the results and compare the proposed

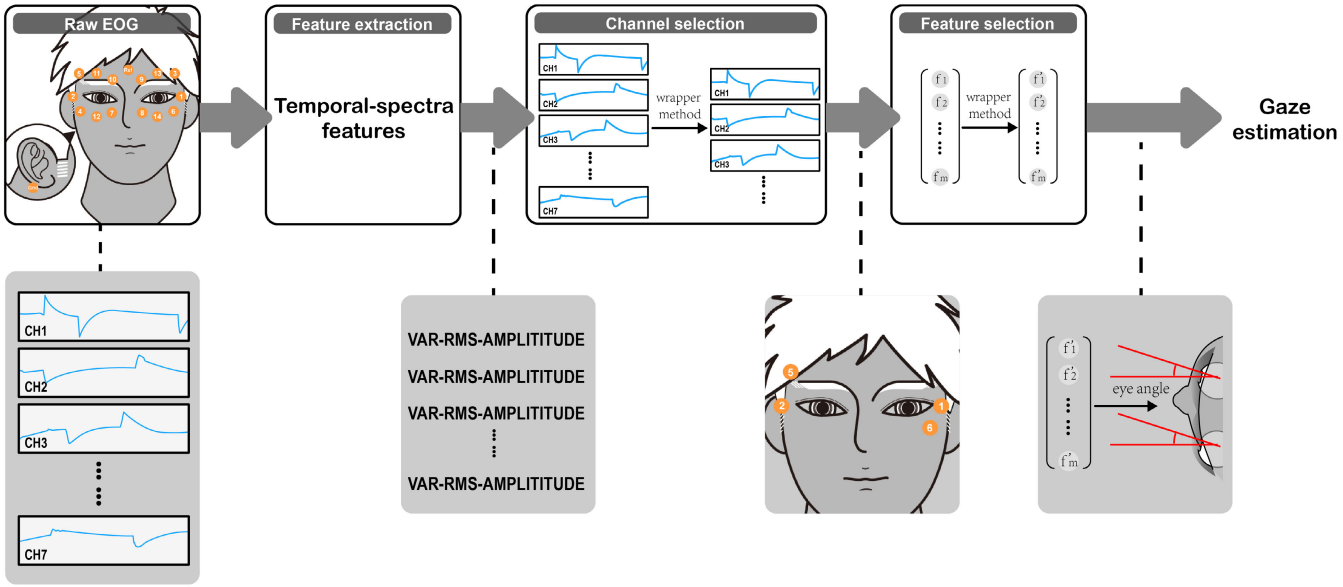

FIGURE 1. The overall framework of the proposed method includes feature extraction, channel selection, and feature selection. Feature extraction: temporal-spectra feature extraction. Channel selection: refining channel combinations via the wrapper technique. Feature selection: refining feature combinations via the wrapper technique.

method with other works. In the last section, we conclude the paper and present the future work of this study.

II. METHODOLOGY

A. OVERALL FRAMEWORK OF THE PROPOSED METHOD

The overall framework of the proposed method is shown in Fig. 1. Firstly, each channel extracted 13 hand-crafted features to represent one saccadic eye movement. Secondly, we applied the wrapper technique to refine the channel combinations. Finally, we also employed the wrapper technique to further reduce the computational burden and ineffective alternative features. The robustness of the selected channels and features was also evaluated on 6 classic models.

B. PARTICIPANTS AND DATA ACQUISITION

Experiments were conducted with 19 healthy subjects aged 25±4 years (9 males and 10 females). Before the experiment, all subjects were informed about the experiment's purpose and signed an informed consent.

Subjects sat approximately 70cm away from the gazing targets, with their head held immobile using a support frame. During the experiments, subjects were asked not to move their head. A Polysomnography (PSG) device with a sampling frequency of 256 Hz was applied for EOG data acquisition [24].

Fifteen Ag/AgCl electrodes are attached around the orbital cavity. The grounding electrode is placed at the mastoid process. Different channels yield seven pairs of differential EOGs in the horizontal (channel 1: 1-2, channel 4: 7-8, channel 5: 9-10), vertical (channel 6: 11-12, channel 7: 13-14) and cross (channel 2: 3-4, channel 3: 5-6) axes as shown in Fig. 2. Twenty-three targets are placed horizontally on a straight line from −50° to 50°, as shown in Fig. 3. The

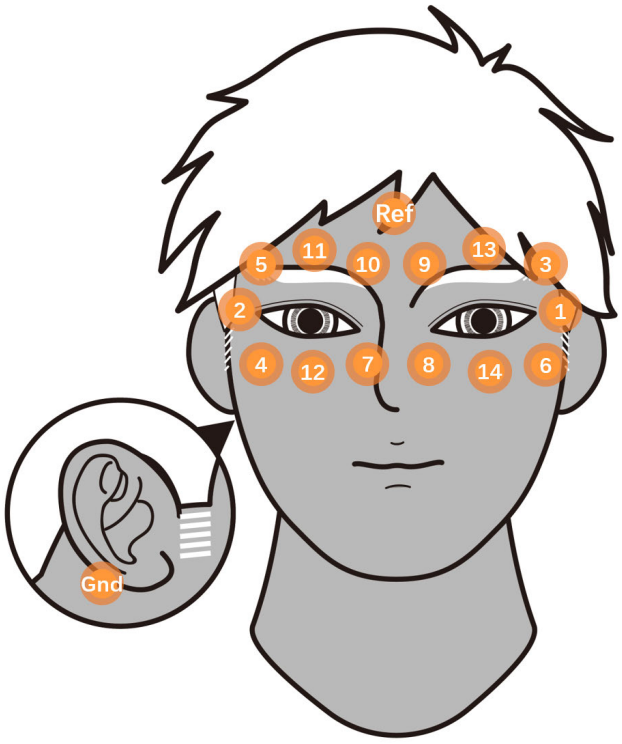

FIGURE 2. The 7 differential EOG channels. channel 1: 1-2, channel 2: 3-4, channel 3: 5-6, channel 4: 7-8, channel 5: 9-10, channel 6: 11-12, channel 7:13-14.

specific target angle sequence from left to right is: −50°, −48°, −46°, −44°, −42°, −40°, −38°, −35°, −30°, −20°, −10°, 0°, 10°, 20°, 30°, 35°, 38°, 40°, 42°, 44°, 46°, 48°, 50°. A graphical user interface (GUI) program was designed

IEEE Journal of Translational Engineering in Health and Medicine

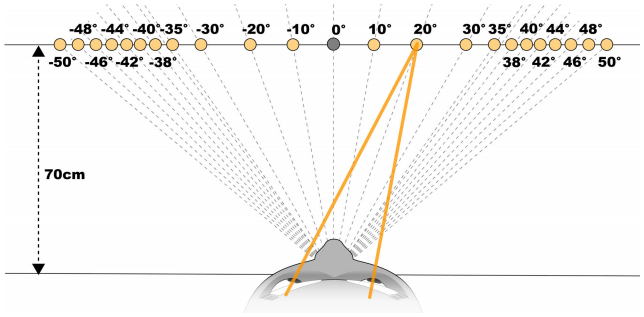

FIGURE 3. Arrangement of targets in the experiments. The electrical potential is captured by the electrodes around the orbital cavity. For a subject, the complete experiment includes 44 trials.

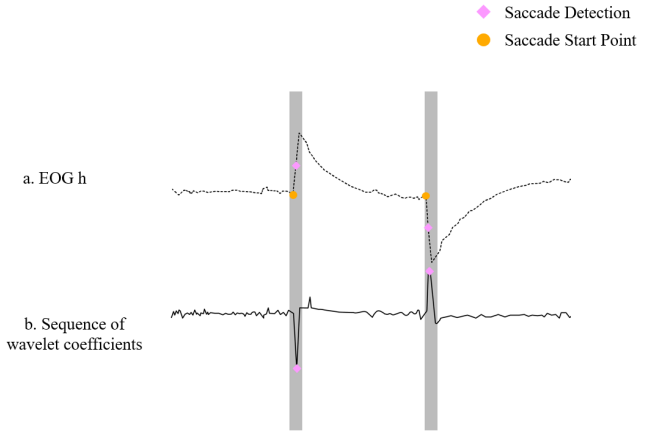

FIGURE 4. Saccade Detection algorithm. (a) Denoised and baseline drift removed EOG signals (b) The sequence of wavelet coefficients at scale 20 using a Haar mother wavelet.

to control the experiment process and to mark the saccade signal for subsequent signal processing.

C. EXPERIMENT PROCEDURE

The protocol started with the subject gazing at the center point (0° points). During each trial, auditory instruction was used to cue the subjects when to perform saccades and when to finish. Specifically, subjects perform a three-second fixation at the center point (0° points). Then subjects execute a saccade from the center point to the target point, followed by another three seconds of fixation. Last, a saccade back to the center point is performed. All the subjects were required to inform the experiment assistant if they missed any targets or performed a wrong saccade. After the data were captured, the EOG signals were manually examined, and the corrupted signal that clearly included large noise components such as blinking, head moving, or gazing at the wrong target was eliminated from the analysis. We recorded 19 subjects but one was withdrawn due to poor signal quality. For this subject, the EOG signal was completely distorted due to bad electrode placements. Finally, a total of 715 saccades were recorded.

D. DATA PREPROCESSING

EOG signals can be interfered with by diverse noise, such as power-line interference and high-frequency electromyographic noise. In this work, the raw EOG signals were lowpass filtered below 20Hz [25]. Also, a notch filter was applied to remove the 50Hz power line interference. Another noise factor is the baseline warder, which is caused by background signal interference or electrode polarization [26]. Here, a 0.1Hz high-pass filter was applied to lower the impact of the baseline wander. Further, we used the Continuous Wavelet Transform algorithm (CWT) for saccade detection [17]. The formula of continuous wavelet transform can be expressed as follows:

$$C_b^a = \int S(t)\frac{1}{\sqrt{a}}\gamma(\frac{t-b}{a})dt$$

where $S(t)$ is the EOG signal and C_b^a is the wavelet coefficient at scale a and position b. Based on the Haar wavelet, the coefficient can be used to detect the start point of saccades (as shown in Fig. 4). Then, saccades were segmented into the 1s window for further analysis. The typical duration of a saccade is between 10ms to 100ms [17], so the 1s window can effectively include the saccadic eye movement [21].

E. FEATURE EXTRACTION

Temporal-spectral domain features (as shown in Table 1) were extracted from each 1-s task from each EOG channel. These features were: Variance, Amplitude [19], Power spectra density (PSD) of δ (0.5-4Hz), θ (4-8Hz) and α (8-14Hz) frequency band, Nonlinear energy [27], Energy, Kurtosis, Sample entropy [28], Area under the curve (AUC) [29], Root mean square (RMS), Form factor (FF), Crest factor (CF). For a specific saccade, features were computed using the full 1-s task. Let $x_i^j(t)$ be the t^{th} sample (a 1s window contains N=256 samples) from the j^{th} saccade of channel i.

1) Variance: The variance of the EOG signal [30] is calculated as:

$$Var = \frac{1}{N}\sum_{t=1}^{N}(x_i^j(t) - \bar{x})^2$$

where Var is the Variance of the j^{th} saccade signal of channel i and $\bar{x}$ is the mean value of the 1s window.

2) Root mean square (RMS): RMS reflects the average power of the EOG signal which is given by the following formula:

$$RMS = \sqrt{\frac{\sum_{t=1}^{N} x_i^j(t)^2}{N}}$$

3) Amplitude: Previous study [31] has shown the correlation between amplitude and eye movement angle. Amplitude is given by the following formula:

$$Amplitude = max(|EOG_{crest} - EOG_{trough}|)$$

where the *Amplitude* is the difference between crest and trough.

4) Power spectra density (PSD) of δ (0.5-4Hz), θ (4-8Hz), and α (8-14Hz): The power spectral density was obtained by Welch's overlapped segment averaging estimator (Parameter: 50% segment overlap and 8 overlapped segments). If the length of the signal cannot be divided into an integer number of segments with 50% overlap, the signal is truncated accordingly.

5) Nonlinear energy: Nonlinear energy is designed to reflect the saccadic property. The detailed formula is shown as follows:

$$Nonlinear\_energy_i^j = \sum_{t=2}^{256}(x_i^j(t)^2 - x_i^j(t-1) * x_i^j(t+1))$$

6) Energy: Energy can be used to distinguish different saccade angles. The detailed formula is shown as follows:

$$Energy_i^j = \sum_{t=2}^{256} x_i^j(t)^2$$

7) Kurtosis: Kurtosis is a parameter reflecting the sharpness of the peak. The kurtosis value was computed using the following form:

$$Kurtosis_i^j = \frac{\frac{1}{N}\sum_{t=1}^{N}(x_i^j(t) - \bar{x})^4}{(\frac{1}{N}\sum_{t=1}^{N}(x_i^j(t) - \bar{x})^2)^2}$$

8) Sample entropy: Entropy reflects the degree of the chaos of the signal. For a full 1-s trial with 256 points, Sample entropy was calculated using the Chebyshev distance type as follows:

$$SmapEn(m, r) = -ln(\frac{K^{m+1}(r)}{K^m(r)})$$

where m represents the embedding dimension and r is the tolerance. $K^m(r)$ is the probability of two sequences matching m points under tolerance r. Similarly, $K^{m+1}(r)$ is the probability of two sequences matching m+1 points under tolerance r. Here, we set $m = 1$ and $r = 0.25$ [28].

9) Area under the curve (AUC): For a valid extracted EOG signal segment, we used the trapezoidal rule to find the approximate value of the AUC [29].

10) Form factor (FF), Crest factor (CF): Form factor (FF) and Crest factor (CF) were extracted to quantify the distortion and extreme of the waveform. These two features were calculated as follows:

$$FF = \frac{RMS}{ARV}, \quad CF = \frac{x_{peak}}{RMS}$$

where ARV is the rectified mean value and x_{peak} is the peak value of the EOG signal.

F. SELECTION STRATEGIES OF CHANNELS AND FEATURES

1) CHANNEL SELECTION AND FEATURE SELECTION

The computational efficiency depends on the number of alternative channels and features. Thus, due to the high computational burden and time-consuming solution, the global optimum solution may not be feasible for practical use.

TABLE 1. The 13 hand-crafted features.

No	Features	Abbreviation	Parameters
1	Variance	Var	-
2	Root mean square	RMS	-
3	Amplitude	-	-
4	Power spectra density of δ frequency band	-	$segmentoverlap = 50\%, overlapped = 8$
5	Power spectra density of θ frequency band	-	$segmentoverlap = 50\%, overlapped = 8$
6	Power spectra density of α frequency band	-	$segmentoverlap = 50\%, overlapped = 8$
7	Nonlinear energy	-	-
8	Energy	-	-
9	Kurtosis	$Kurt$	-
10	Sample entropy	$SampEn$	$m = 1, r = 0.25$
11	Area under the curve	AUC	-
12	Form factor	FF	-
13	Crest factor	CF	-

Regarding channel selection, we employed the wrapper method combined with the forward stepwise strategy to select the optimal channels [32]. First, we initially performed the model evaluation with each channel and fixed the channel satisfying the selecting criterion as the first selected channel. Second, we combined each of the rest channels with the first selected channel for model evaluation. The channel, together with the first selected channel, meeting the selecting criterion was fixed as the second selected channel. Finally, we repeated the above steps until all channels were selected. The optimal channel set was then determined based on the 7-round (7 pairs of differential channels) forward stepwise selection results.

Excessive ineffective features may induce computational burden or even dimension disaster. Thus, it may be not feasible to directly optimize all features from all the channels. Besides, excessive alternative also increases the difficulty in refining features, leading to subnormal feature combinations. Thus, we also conducted the forward stepwise strategy to further refine the features from the previously selected optimal channels. Specifically, to achieve N_f optimal features, we conducted 26 rounds (corresponding to 2 channels$\times$13 features) of feature selection. The N_f optimal features were determined if the performance no longer increased.

2) EVALUATION OF THE RECOMMENDED CHANNELS AND FEATURES

We refined the channels and features by employing the forward stepwise strategy combined with the random forest baseline model. To further evaluate the robustness, 6 regression models (i.e., support vector regression (SVR), random forest (RF), BP network (BP), ensemble model (ENS), generalized linear model (GLM), and decision trees (DT)) were also applied to the recommended channels and features. The corresponding parameter setting was shown in Table 2. Specifically, we compared the model performance before and after selection. MAE and RMSE were used as the metrics.

IEEE Journal of Translational Engineering in Health and Medicine

TABLE 2. Parameter setting for the regression models.

Regression models	Parameter setting for the corresponding models
SVR	SMO solver
RF	50 trees
BP	Double layes,10 neurons per layer
GLM	Normal distribution
ENS	LSBoost, 1000 training cycles
DT	Bayesian optimization

G. VALIDATION PROTOCOL

Leave-One-subject-Out Cross-Validation technique [33] was used to evaluate the performance. The evaluation indicators including MAE and RMSE were calculated by the following formula:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{\theta}_i - \theta_i|$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta_i)^2}$$

where N is the total saccades of a subject, θ_i is the true value of the i^{th} saccade and $\hat{\theta}_i$ is the estimated value of the i^{th} saccade. The experiments were performed on a hardware specification with an Intel core i5-10400F and GTX 1660 GPU.

H. STATISTICAL ANALYSIS

Statistical analyses on two factors (i.e., channels and features.) were explored in this study. A non-parametric test was applied when the obtained results with the same factor did not follow a Gaussian distribution. Otherwise, a one-way repeated-measures (RM) analysis of variance (ANOVA) was applied. For the non-parametric test, the Friedman test was performed to detect whether there was an overall significant difference. Suppose an overall significant difference was found, the Wilcoxon signed-rank test was further conducted to perform the pair-wise comparison. For the p values obtained with either parametric or non-parametric tests, the Holm-Bonferroni correction was adopted to avoid errors caused by multiple comparisons. The significance level was set to 0.05.

III. RESULTS

A. CHANNEL SELECTION PERFORMANCE

To investigate the available channel combinations, 7 rounds of channel selection processes were conducted to refine the channels. Fig. 5 shows the optimal channels for each round and corresponding performance. The optimal channel selection process is channel 1-3-2-6-7-5-4. One interesting observation is that the combination of channel 1 and channel 3 contributed to the highest performance of gaze estimation. Channel 1 was selected first and achieved an accuracy (MAE:3.06° and RMSE:4.02°). Then, in the second round, the performance was improved by adding channel 3. Subsequent channels failed to enhance performance after the second round. Based on the channel selection results, the optimal performance was achieved using channel 1 and channel 3 (MAE:2.95° and RMSE:3.89°)

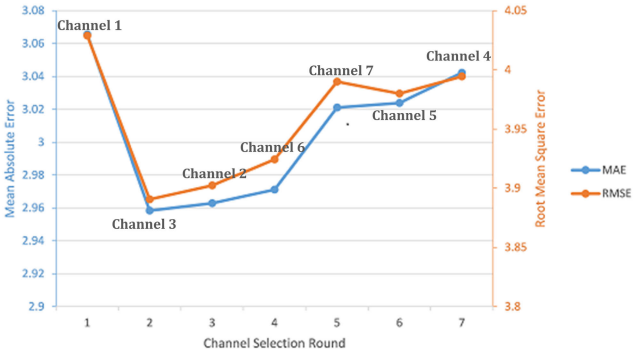

FIGURE 5. The optimal channels selected for each round and the performance achieved using these optimal channels. The optimal channel selection process is channel 1-3-2-6-7-5-4.

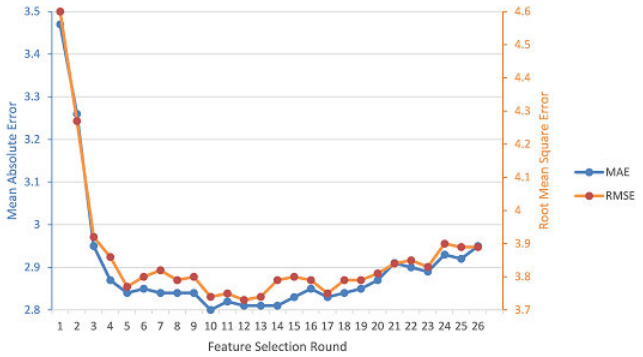

FIGURE 6. Feature selection results based on the random forest.

B. FEATURE SELECTION PERFORMANCE

Regarding the feature selection, here, two 13×1 feature vectors were extracted separately from channel 1 and channe 3 to concatenate a 26×1 feature vector. Overall, 26 rounds of feature selection processes were conducted as shown in Fig. 6.

In the first 10 rounds of feature selection, the MAE and RMSE both reduced to the lowest. For the subsequent selection rounds, the accuracy gradually decreased. The newly selected features failed to provide more discriminative information. Thus, the feature combination was determined.

By adopting the cross-subject validation technique, a group of feature combinations was recommended for EOG-based cross-subject gaze estimation (as shown in Table 3.) The selected feature combinations were: PSD of θ, Amplitude, Nonlinear energy, Var, Energy, PSD of α and CF (channel 1); CF, Amplitude, Var (channel 3). Ultimately, the accuracy (MAE:2.80° and RMSE: 3.74°) can be achieved. Among the selected features, CF, Amplitude, and Var were selected most, indicating their superiority in cross-subject gaze estimation.

TABLE 3. The optimal feature combination selected by the forward stepwise strategy.

Feature ranking	Feature	Channel
1	PSD of θ frequency band	1
2	Crest Factor	3
3	Amplitude	1
4	Nonlinear Energy	1
5	Amplitude	3
6	Var	1
7	Energy	1
8	PSD of α frequency band	1
9	Crest Factor	1
10	Var	3

C. STATISTICAL ANALYSIS FOR CHANNEL SELECTION AND FEATURE SELECTION

After channel selection, no significant differences were found on the MAE and RMSE ($p_{MAE} = 0.118$ and $p_{RMSE} = 0.130$). Thus, the utilization of fewer channels can achieve similar performance with the multichannel EOGs, which is conducive to both computational efficiency and convenient setup. Further, significant performance improvements were found after the feature selection ($p_{MAE} = 0.0006$ and $p_{RMSE} = 0.12$). Hence, screening out discriminating features can enhance not only the model compactness but also the model performance.

D. EVALUATION OF THE RECOMMENDED CHANNELS AND FEATURES

We evaluated the robustness of the selected channels and features. Here, 6 classic models were employed for the evaluation. Specifically, the feature vectors were concatenated by using: (1) all features from all channels (Before selection). (2) selected features from the selected channels (After selection). The evaluation results were shown in Fig. 7. The performance improvement was observed in all 6 regression models (i.e., support vector regression (SVR), random forest (RF), BP network, ensemble model (ENS), generalized linear model (GLM), and decision trees (DT), indicating the robustness of the recommended channel and feature combination.

IV. DISCUSSION

In this study, we evaluated the robustness of the EOG-based cross-subject gaze estimation approach. Specifically, electrode placements and feature combinations were investigated, which had several merits including model compactness, computational effectiveness, and novel electrode placement design. Finally, impressive accuracy (MAE: 2.80° and RMSE: 3.74°) was achieved, indicating the robustness in cross-subject gaze estimation.

A. PERFORMANCE OF CHANNEL SELECTION AND FEATURE SELECTION

Regarding the channel selection performance, selecting all the channels for gaze estimation may involve useless information, which is regarded as the noise for the estimation. Hence, it is of great importance to refine the optimal channel with the most contribution and exclude the irrelevant channels. One interesting observation is that the combination of channel 1 and channel 3 contributed to the highest performance of gaze estimation. Previous studies [21], [26] have reported that multi-channel information improves estimation accuracy. Thus, one possible explanation may fall into that the cross channels may provide supplemental spatial information to generate discriminative features since the cross channels can also capture the horizontal eye movement in the cross axis. As the number of channels further increased, the performance tended to degrade, which indicated that the spatial information from the subsequent channels may be redundant. Excessive channels may induce the computational burden or even dimension disaster. Also, the utilization of excessive channels may increase the risk of long time-consuming setup and acquisition costs. Hence, performance and practicability were required to trade off in the exploration of novel electrode placement.

Regarding the feature selection, an optimal feature combination was recommended for the EOG-based gaze estimation. By traversing the feature set, the MAE and RMSE both gradually reduced, which achieved a higher performance compared with using all the features. Besides, the elimination of ineffective features was also conducive to the model compactness. Thus, the utilization of the feature optimization technique was essential to further reduce redundancy.

For the features optimization techniques, different optimal feature combinations can be obtained due to the different selection strategies. Also, based on different models, the wrapper selection method can generate different optimal feature subsets. Here, we mainly investigated the feature combinations by employing the random forests, which had the highest performance in comparison. Thus, we progressively traverse all the features to search for a set of feature subsets.

Our results demonstrated that the 10 features from channel 1 and channel 3 can provide sufficient spatial information. Specifically, the amplitude-related features (amplitude and crest factor) both had high priority during the selection. Previous studies have reported that the saccadic angles were considered linear or nonlinear with the EOG amplitude. Thus, the amplitude-related features implicitly reflect the spatial information of eye movements. However, due to the individual differences, amplitude-related features can vary from each subject [9], leading to the degeneration in the cross-subject evaluation. Accordingly, amplitude-related features were not considered as the top-1 ranking during the selection.

In addition, the saccadic angles were also associated with the EOG energy (energy, nonlinear energy, and var), because the energy-related features implicitly reflected the intensity of eye movements. Hence, these features can provide discriminative information to further distinguish different saccadic angles. Among the energy-related features, nonlinear energy achieved the highest ranking during the selection, which is mainly because the nonlinear energy can be utilized to characterize the peak velocity. Traditionally, the EOG peak

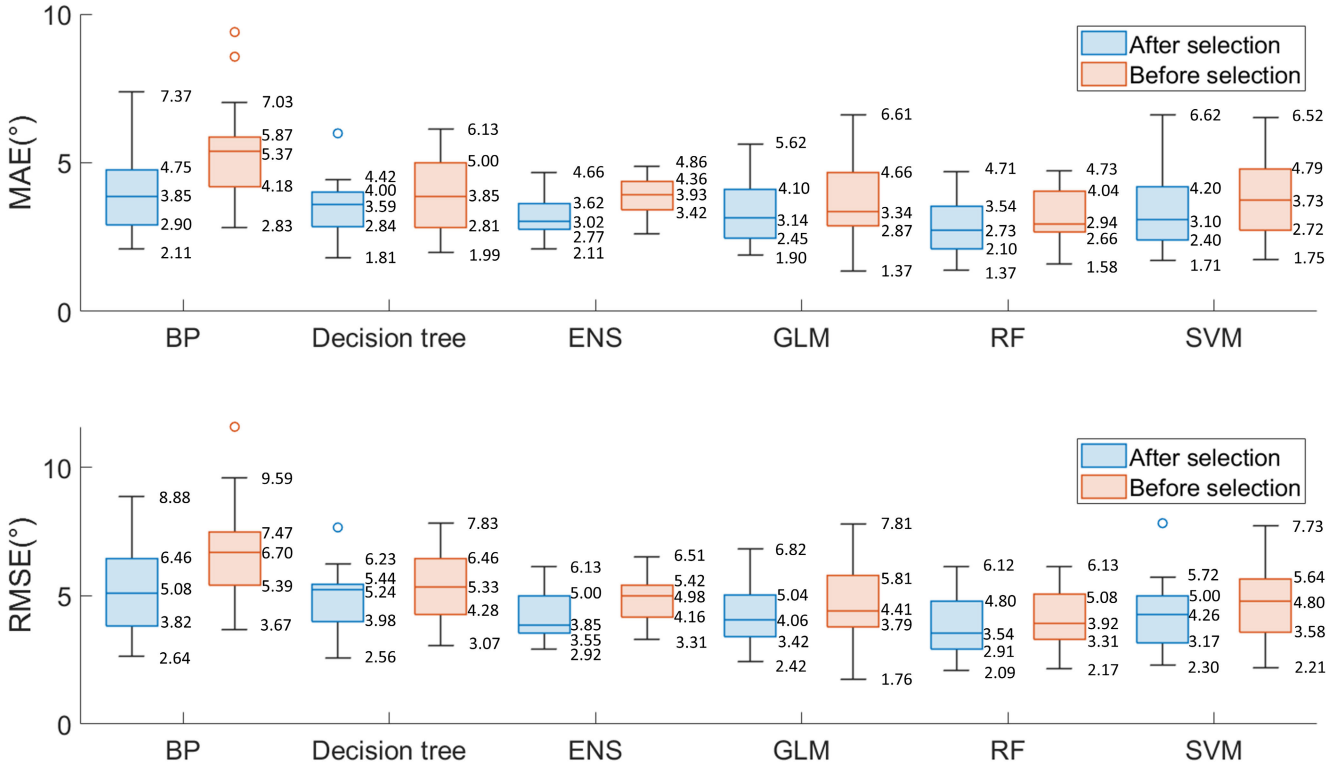

FIGURE 7. **Performance of different models based on the recommended channels and feature combination.**

TABLE 4. **The performance comparison with the prevailing methods.**

	Methods	Electrode placement	Feature combination	MAE	RMSE
ref.[14]	Third-order polynomial	Horizontal channel	Amplitude	3.50°	4.40°
ref.[14]	Fourth-order polynomial	Horizontal channel	Amplitude	3.50°	4.41°
ref.[15]	Electric method	Horizontal and vertical channels	Electrode positions +Amplitude	8.28°	9.16°
ref.[34]	Univariate linear	Horizontal	Amplitude	3.76	4.70
Our work	**Chanel and feature selection**	**Horizontal and cross channels**	**PSD of theta +CF+Amplitude +Nonlinear Energy+Var +Energy+PSD of alpha**	**2.80°**	**3.74°**

velocity was significantly different between large angles and small angles [35]. Thus, our results also demonstrated that different energy-related features that can distinguish the EOG waveforms were needed for further exploration. Regarding the frequency-related features (PSD of θ frequency band and PSD of α frequency band), our results also show consistency with the previous study [36]. Different directional eye movements contain different energy distributions on frequency bands, which can be integrated for gaze estimation. Specifically, the domain energy is 8-16Hz for left eye movements and 2-4Hz for right eye movements [37], implicitly representing the direction information of eye movement. Thus, the frequency-related features were also worth further investigation in cross-subject estimation. Besides, we tentatively attempted transfer learning for data alignment to further map the feature space by employing the recommended features. The results demonstrated that improved accuracy by 8% can be achieved.

Besides, we also investigated the performance of the selected channel and feature combination on multiple models. As described in the prior paragraph, searching for global optimization may require redundant time and computing resources. In addition, for the wrapper method, the global optimal parameters may highly depend on the specific model, lacking generalization. Here, we further assessed the selected features and channels on 6 widely-used models. The results demonstrated that the performance was improved compared with employing all features from all channels, indicating the necessity of the selection process. Besides, the selected channel and feature combination was available for multiple models, which is recommended for cross-subject gaze estimation.

B. COMPARISON WITH THE PREVAILING TECHNIQUES

In addition, we further compared several classic polynomial and electric models to evaluate the cross-subject robustness. Specifically, we selected the two polynomial methods (i.e., third-order and fourth-order polynomials) in [14] for comparison. For the electric methods, we selected the battery model and employed the two-stage optimization with the Levenberg-Marquardt algorithm as described in [15]. To ensure fairness, we re-implemented all these models in our dataset.

Our work achieved the highest performance compared with the polynomial and electric models (as shown in Table 4). To illustrate, for the polynomial models, polynomial coefficients were required as a prerequisite for the specific subject, which can be regarded as a calibration to construct the subject-dependent model. Due to the existence of individual variabilities, the lack of subject-dependent calibration may cause the degeneration of the model. Hence, it may be not feasible to directly apply cross-subject pre-training coefficients which are unable to characterize the individual's EOG signals effectively. Similarly, the electric method suffers from the same problem. During the optimization, electric methods implicitly required individual information to derive the spatial distribution of electrodes, which may significantly affect the accuracy of gaze estimation. Besides, the electric method also suffers from several natural flaws: 1) Electric model assumes the eyeball is a standard sphere, which is inconsistent with the fact that the eyeball is an ellipsoid. 2) Model parameters such as eye radius and inter-pupillary distance are difficult to measure directly. The utilization of empirical value may reduce performance. Thus, the evaluation of the electric model is not ideal.

C. FUTURE WORK

In this study, we enhanced the performance of EOG-based gaze estimation by comprehensively investigating the potential optimal channel/feature combination. Moreover, the robustness of the optimal channel/feature combination was further evaluated on 6 classic models. However, the works still can be enhanced by involving the following issues in our future work:

Although the number of recruited subjects is relatively large compared with related studies, the selected channels (features) may be limited to specific datasets. In the future, we also plan to construct a large-scale population estimation dataset to further evaluate the effectiveness of the selected channels (features). In this context, deep learning-based techniques, such as multi-scale CNN networks, transformer models, and deep transfer learning strategies, can also be applied to extract discriminative features. Besides, as a preliminary study to fully investigate the feasibility of the proposed method in electrode placement and feature selection, we mainly focus on horizontal gaze estimation. In the future, gaze estimation in all directions will be analyzed. Finally, in this paper, only healthy subjects are involved, while subjects with neurological diseases may have different characteristics in EOG signals. Consequently, it may be more challenging to estimate the gaze angle for these patients. More attention can be paid to establishing models with strong generalization abilities in the future.

V. CONCLUSION

In this study, a robust gaze estimation approach was proposed by exploring relevant EOG features and optimal electrode placements. The evaluation was conducted on 715 saccades from 18 healthy subjects in a wide range of saccadic angles from $-50°$ to $50°$. By employing the forward stepwise selection strategy, we investigated the effect of electrode placements around the orbital cavity. Experimental results demonstrated that the novel electrode placements with the highest performance were recommended for gaze estimation compared with the standard 4-channel electrode placement. Besides, we further reduced the feature dimension to search for the feature combination, which was also recommended and conducive for cross-subject gaze estimation. Ultimately, by using the 10 features from the novel electrode placements, a promising accuracy (MAE:2.80° and RMSE:3.74°) was obtained. The recommended combinations also contributed to higher performance on 6 widely-used regression models, indicating their robustness. Compared with the prevailing approaches, our proposed method achieved the highest performance, indicating its superiority. Ultimately, our work can not only serve as screening guidance for EOG-based gaze estimation but also contribute to EOG-based applications.

REFERENCES

[1] S. He and Y. Li, "A single-channel EOG-based speller," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 11, pp. 1978–1987, Nov. 2017.

[2] S. K. B. Sangeetha, "A survey on deep learning based eye gaze estimation methods," *J. Innov. Image Process.*, vol. 3, no. 3, pp. 190–207, Sep. 2021.

[3] J. Li et al., "Appearance-based gaze estimation for ASD diagnosis," *IEEE Trans. Cybern.*, vol. 52, no. 7, pp. 6504–6517, Jul. 2022.

[4] H. Cai et al., "Sensing-enhanced therapy system for assessing children with autism spectrum disorders: A feasibility study," *IEEE Sensors J.*, vol. 19, no. 4, pp. 1508–1518, Feb. 2019.

[5] C. Mauriello et al., "Dysfunctional temporal stages of eye-gaze perception in adults with ADHD: A high-density EEG study," *Biol. Psychol.*, vol. 171, May 2022, Art. no. 108351.

[6] A. M. Michalek, G. Jayawardena, and S. Jayarathna, "Predicting ADHD using eye gaze metrics indexing working memory capacity," in *Computational Models for Biomedical Reasoning and Problem Solving*. Pennsylvania, PA, USA: IGI Global, 2019, pp. 66–88.

[7] S. De Silva et al., "A rule-based system for ADHD identification using eye movement data," in *Proc. Moratuwa Eng. Res. Conf. (MERCon)*, Jul. 2019, pp. 538–543.

[8] N. Kumar, S. Kohlbecher, and E. Schneider, "A novel approach to video-based pupil tracking," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, Oct. 2009, pp. 1255–1262.

[9] S. Milanizadeh and J. Safaie, "EOG-based HCI system for quadcopter navigation," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 11, pp. 8992–8999, Nov. 2020.

[10] C. Topal, S. Gunal, O. Koçdeviren, A. Dogan, and Ö. N. Gerek, "A low-computational approach on gaze estimation with eye touch system," *IEEE Trans. Cybern.*, vol. 44, no. 2, pp. 228–239, Feb. 2014.

[11] D. Kumar and E. Poole, "Classification of EOG for human computer interface," in *Proc. 2nd Joint 24th Annu. Conf. Annu. Fall Meeting Biomed. Eng. Soc., Eng. Med. Biol.*, 2002, pp. 64–67.

IEEE Journal of Translational Engineering in Health and Medicine

[12] R. Barea, L. Boquete, S. Ortega, E. López, and J. M. Rodríguez-Ascariz, "EOG-based eye movements codification for human computer interaction," *Expert Syst. Appl.*, vol. 39, no. 3, pp. 2677–2683, Feb. 2012.

[13] A. Favre-Félix, C. Graversen, T. Dau, and T. Lunner, "Real-time estimation of eye gaze by in-ear electrodes," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 4086–4089.

[14] H. Manabe, M. Fukumoto, and T. Yagi, "Direct gaze estimation based on nonlinearity of EOG," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1553–1562, Jun. 2015.

[15] N. Barbara, T. A. Camilleri, and K. P. Camilleri, "EOG-based gaze angle estimation using a battery model of the eye," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 6918–6921.

[16] N. Barbara, T. A. Camilleri, and K. P. Camilleri, "EOG-based ocular and gaze angle estimation using an extended Kalman filter," in *Proc. ACM Symp. Eye Tracking Res. Appl.*, Jun. 2020, pp. 1–5.

[17] A. Bulling, J. A. Ward, H. Gellersen, and G. Tröster, "Eye movement analysis for activity recognition using electrooculography," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 4, pp. 741–753, Apr. 2011.

[18] F. Latifoglu, R. Ileri, E. Demirci, and Ç. G. Altintop, "Detection of reading movement from EOG signals," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2020, pp. 1–5.

[19] Y. Bei, Y. Sichun, L. Mengke, and L. Xiangkun, "Gaze estimation method based on EOG signals," in *Proc. 6th Int. Conf. Instrum. Meas., Comput., Commun. Control (IMCCC)*, Jul. 2016, pp. 443–446.

[20] A. López, F. J. Ferrero, M. Valledor, J. C. Campo, and O. Postolache, "A study on electrode placement in EOG systems for medical applications," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, May 2016, pp. 1–5.

[21] M. Yan, H. Tamura, and K. Tanno, "A study on gaze estimation system using cross-channels electrooculogram signals," in *Proc. Int. Multiconf. Eng. Comput. Sci.*, vol. 1, Jan. 2014, pp. 112–116.

[22] T. Alotaiby, F. E. A. El-Samie, S. A. Alshebeili, and I. Ahmad, "A review of channel selection algorithms for EEG signal processing," *EURASIP J. Adv. Signal Process.*, vol. 2015, no. 1, pp. 1–21, Dec. 2015.

[23] J. W. Li et al., "Single-channel selection for EEG-based emotion recognition using brain rhythm sequencing," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 6, pp. 2493–2503, Jun. 2022.

[24] M. Liu et al., "Overview of a sleep monitoring protocol for a large natural population," *Phenomics*, vol. 3, pp. 1–18, May 2023.

[25] S. N. Abbas and M. Abo-Zahhad, "Eye blinking EOG signals as biometrics," in *Biometric Security and Privacy*. Cham, Switzerland: Springer, 2017, pp. 121–140.

[26] N. Barbara, T. A. Camilleri, and K. P. Camilleri, "EOG-based eye movement detection and gaze estimation for an asynchronous virtual keyboard," *Biomed. Signal Process. Control*, vol. 47, pp. 159–167, Jan. 2019.

[27] C. Sun, C. Chen, W. Li, J. Fan, and W. Chen, "A hierarchical neural network for sleep stage classification based on comprehensive feature learning and multi-flow sequence learning," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 5, pp. 1351–1366, May 2020.

[28] N. Sho'ouri, "EOG biofeedback protocol based on selecting distinctive features to treat or reduce ADHD symptoms," *Biomed. Signal Process. Control*, vol. 71, Jan. 2022, Art. no. 102748.

[29] A. Kaur, "Wheelchair control for disabled patients using EMG/EOG based human machine interface: A review," *J. Med. Eng. Technol.*, vol. 45, no. 1, pp. 61–74, Jan. 2021.

[30] S. Aungsakul, A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Evaluating feature extraction methods of electrooculography (EOG) signal for human-computer interface," *Proc. Eng.*, vol. 32, pp. 246–252, Jan. 2012.

[31] F. Simini, A. Touya, A. Senatore, and J. Pereira, "Gaze tracker by electrooculography (EOG) on a head-band," in *Proc. 10th Int. Workshop Biomed. Eng.*, Oct. 2011, pp. 1–4.

[32] I. Rejer and K. Lorenz, "Genetic algorithm and forward method for feature selection in EEG feature space," *J. Theor. Appl. Comput. Sci.*, vol. 7, no. 2, pp. 72–82, 2013.

[33] L. Meng et al., "Exploration of human activity recognition using a single sensor for stroke survivors and able-bodied people," *Sensors*, vol. 21, no. 3, p. 799, Jan. 2021.

[34] J. Ha, K.-M. Choi, and C.-H. Im, "Feasibility of using electrooculography-based eye-trackers for neuromarketing applications," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–10, 2022.

[35] D. Boghen, B. Troost, R. Daroff, L. Dell'Osso, and J. Birkett, "Velocity characteristics of normal human saccades," *Investigative Ophthalmol. Vis. Sci.*, vol. 13, no. 8, pp. 619–623, 1974.

[36] W. M. Bukhari, W. Daud, and R. Sudirman, "A wavelet approach on energy distribution of eye movement potential towards direction," in *Proc. IEEE Symp. Ind. Electron. Appl. (ISIEA)*, Oct. 2010, pp. 181–185.

[37] W. M. B. W. Daud and R. Sudirman, "Time frequency analysis of electrooculograph (EOG) signal of eye movement potentials based on wavelet energy distribution," in *Proc. 5th Asia Model. Symp.*, May 2011, pp. 81–86.