Erciyes University Journal of Institute Of Science and Technology

Research Article
*10.65520/erciyesfen.1867956*

***Imprint:***
Volume: 42(3)
Year: 2026
Page: 161-186

Emre DEMIROZ[a]
Ayse Nur AY GUL[b*]

[a] Master's Student, Sakarya University of Applied Sciences, Institute of Graduate Studies, Department of Mechatronics Engineering, Sakarya, Türkiye., emredemiroz173@gmail.com
[b] Asst. Prof. Dr. Sakarya University of Applied Sciences, Biomedical Technologies Application and Research Center (BIYOTAM), Sakarya, Türkiye., ay@subu.edu.tr

[*] Corresponding Author

*Received: 1/21/2026*
*Accepted: 8/27/2026*

***Citation:***
Emre DEMIROZ, Ayse Nur AY GUL (2026). Subject-Independent EOG-Based Classification of Horizontal Eye Movements for Real-Time Robot Control. *Erciyes University Journal of Institute Of Science and Technology*, 42(3), 161-186. https://doi.org/10.65520/erciyesfen.*1867956*

Screened by iThenticate for Authors & Researchers

Except where otherwise noted, content in this article is licensed under a Creative Commons 4.0 International license. Icons by Font Awesome.

# Subject-Independent EOG-Based Classification of Horizontal Eye Movements for Real-Time Robot Control

## Abstract

**Background/Aim:** Neurodegenerative diseases and severe neurological trauma can substantially impair conventional communication and control channels, while ocular motor functions may remain preserved until advanced stages. This study proposes a subject-independent electrooculography (EOG)-based framework for classifying horizontal gaze directions and controlling an omnidirectional mobile robot without user-specific calibration.

**Methods:** EOG signals acquired from 12 healthy participants were filtered using a fourth-order Butterworth band-pass filter between 0.1 and 15 Hz and segmented into non-overlapping 5-second windows. Twelve time- and frequency-domain features were extracted from each window. Five classifiers—k-Nearest Neighbors (KNN), Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), and a multilayer perceptron-based artificial neural network (ANN)—were evaluated using Leave-One-Subject-Out cross-validation.

**Results:** SVM achieved the highest participant-wise mean accuracy of 79.61%, followed by KNN (74.47%), RF (72.19%), DT (68.69%), and ANN (66.11%). Pooled out-of-fold analysis further showed that SVM obtained a macro specificity of 0.911 and a macro one-vs-rest ROC-AUC of 0.907. The classified gaze commands were transmitted wirelessly to the mobile robot for real-time motion control.

**Conclusion:** The results demonstrate that the proposed framework provides a practical and computationally efficient solution for subject-independent assistive robotic control.

**Keywords:** Electrooculography (EOG), Human–Machine Interaction, Subject-Independent Classification, Gaze-Based Control, Machine Learning

# Gerçek Zamanlı Robot Kontrolü için Yatay Göz Hareketlerinin Elektrookülografi (EOG) Tabanlı, Katılımcı-Bağımsız Sınıflandırılması

## Öz

Nörodejeneratif hastalıklar ve ağır nörolojik travmalar, bireylerin geleneksel iletişim ve kontrol kanallarını önemli ölçüde kısıtlayabilirken oküler motor fonksiyonlar ileri evrelere kadar korunabilmektedir. Bu çalışmada, kullanıcıya özel kalibrasyon gerektirmeden yatay bakış yönlerini sınıflandıran ve omnidireksiyonel bir mobil robotu kontrol eden katılımcı-bağımsız bir Elektrookülografi (EOG) tabanlı yaklaşım önerilmiştir. On iki sağlıklı katılımcıdan alınan EOG sinyalleri, 0.1–15 Hz aralığında çalışan dördüncü dereceden Butterworth bant geçiren filtre ile işlenmiş ve örtüşmesiz 5 saniyelik pencerelere ayrılmıştır. Her pencereden zaman ve frekans alanlarına ait on iki öznitelik çıkarılmıştır. K-En Yakın Komşu (KNN), Rastgele Orman (RF), Destek Vektör Makinesi (SVM), Karar Ağacı (DT) ve çok katmanlı algılayıcı tabanlı

yapay sinir ağı (ANN) modelleri Leave-One-Subject-Out çapraz doğrulama yöntemiyle değerlendirilmiştir. SVM, %79.61 katılımcı bazlı ortalama doğrulukla en yüksek performansı göstermiş; KNN %74.47, RF %72.19, DT %68.69 ve ANN %66.11 doğruluk elde etmiştir. Birleştirilmiş fold-dışı tahminlerde SVM için makro özgüllük 0.911 ve makro bire-karşı-diğer ROC-AUC değeri 0.907 olarak hesaplanmıştır. Sınıflandırılan bakış komutları kablosuz olarak mobil robota iletilerek gerçek zamanlı hareket kontrolü gerçekleştirilmiştir. Bulgular, önerilen yaklaşımın katılımcı-bağımsız yardımcı robot kontrolü için uygulanabilir ve hesaplama açısından verimli bir çözüm sunduğunu göstermektedir.

**Anahtar kelimeler:** Elektrookülografi (EOG), İnsan–Makine Etkileşimi, Katılımcı-Bağımsız Sınıflandırma, Göz Hareketine Dayalı Kontrol, Makine Öğrenmesi

## 1. INTRODUCTION

Upper motor neuron diseases such as amyotrophic lateral sclerosis (ALS), primary lateral sclerosis (PLS), and hereditary spastic paraplegia (HSP), as well as neurological or trauma-induced conditions including traumatic spinal cord injury, multiple sclerosis (MS), cerebral palsy (CP), and stroke, severely impair the corticospinal and corticobulbar pathways responsible for voluntary motor control. Consequently, affected individuals experience partial or complete loss of essential motor functions such as hand–arm coordination, walking, and interaction with the environment. In many of these conditions, ocular muscles remain among the few voluntary motor functions preserved until advanced disease stages. Therefore, electrooculography (EOG)-based horizontal gaze classification systems offer significant potential for communication and environmental interaction for individuals with severe motor impairments [1,2].

EOG is a non-invasive technique that measures the natural biopotential difference between the retina and the cornea, enabling the acquisition of electrical signals generated by eye movements. These millivolt-level signals can be recorded using surface electrodes with low-cost and portable hardware systems [3]. The ability of EOG systems to operate independently of ambient lighting conditions, capture eye movements over wide visual angles, and allow relatively simple electrode placement makes this technique particularly attractive for human–computer interaction (HCI) applications [3,4]. These characteristics are especially relevant for real-time control scenarios where fast response and robustness are required. The scope and maturity of EOG-based HCI research have been systematically classified using a five-level model proposed in the literature (Figure 1) [3,5]. In addition, several comprehensive review studies have summarized the evolution, challenges, and application domains of EOG-based and multimodal biosignal-driven interaction systems, emphasizing their robustness and suitability for assistive technologies [5,6].

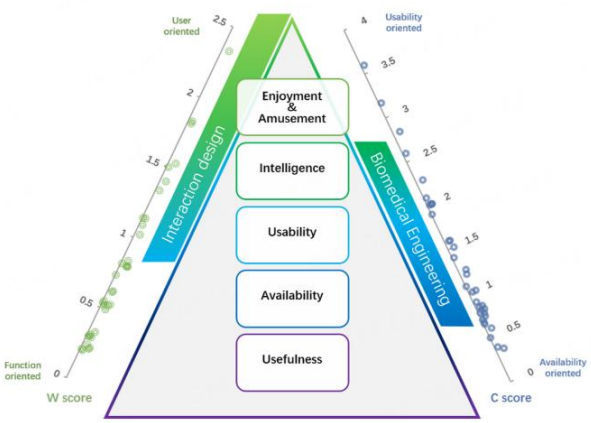



**Figure 1.** Conceptual five-level model for EOG-based HCI systems [5]

Among alternative control approaches, voice-command systems, facial-expression-based methods, and gesture-based interfaces may lose reliability, particularly in motor neuron diseases.

Electroencephalography (EEG)- and electromyography (EMG)-based brain–computer interfaces, on the other hand, often suffer from low signal-to-noise ratios, long calibration times, and hardware requirements that reduce user comfort [1,7]. In contrast, EOG stands out as a prominent biosignal in assistive technologies due to its high temporal resolution, fast response time, and suitability for long-term use [3]. Recent years have witnessed increasing interest in portable and wearable EOG systems. For instance, King et al. introduced the EarEOG system, a wearable EOG approach that acquires eye-movement-related potentials using periauricular electrodes placed around the ear and integrated into a headphone-like form factor, enabling natural interaction and high-accuracy eye-movement detection [8]. Similarly, Liu et al. proposed a real-time directional command generation framework based on EarEOG signals for mobility assistance [9].

EOG signals are typically acquired using electrode pairs positioned along the horizontal and vertical axes, and eye orientation is inferred from differential signal variations [3]. This approach has been successfully applied in various applications, including virtual keyboards, gaming interfaces, robotic manipulation, and directional control systems [3,10,11]. EOG-based interfaces have also been explored for safe electric wheelchair mobility using virtual environment–based training and control paradigms, emphasizing safety and usability for users with severe physical disabilities [12]. For example, Sharma et al. developed a robotic arm control system using EOG signals derived from eye blinks and gaze movements [4], while Huang et al. proposed a hybrid human–machine interface combining EOG and EEG signals for integrated wheelchair and robotic arm control [10].

Despite these advances, baseline drift and motion artifacts remain major technical challenges in EOG-based systems, often degrading classification accuracy [13]. To mitigate these effects, various filtering, normalization, and feature-extraction strategies have been proposed. Barbara et al. provided a comprehensive comparison of baseline drift removal techniques and analyzed their impact on EOG signal stability [13]. In recent studies, machine learning methods have been widely employed to discriminate eye-movement classes from EOG signals. Time- and frequency-domain features have been successfully classified using algorithms such as k-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Random Forest [14]. Hernández-Pérez et al. demonstrated that feature-based approaches using KNN and SVM can achieve high classification accuracy for EOG-based eye-movement recognition [14]. The real-time EOG-based virtual keyboard system developed by Barour and Ay Gul further highlights the growing interest and applicability of such approaches (Figure 2) [15].

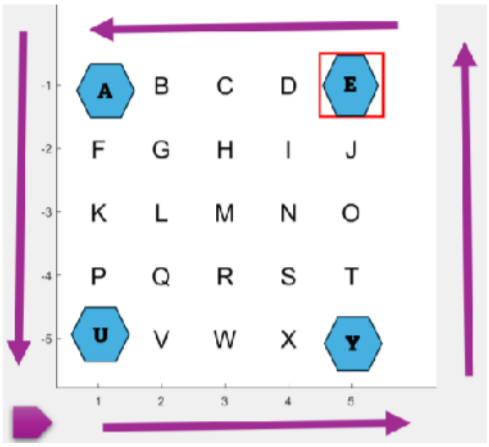



**Figure 2.** Conceptual layout of an EOG-based virtual keyboard interface [15]

The effective integration of human–machine interfaces with robotic systems depends not only on biosignal sensing accuracy but also on the reliable control of mobile robot platforms in real-world environments. Accordingly, evaluating autonomous motion, navigation, and control capabilities under realistic conditions is crucial for assessing the practical applicability of interaction-based systems. Atalı et al. investigated the ROS-based architecture of an autonomous mobile service robot developed for the HoReCa sector through both simulation and real-world experiments, demonstrating sufficient maneuverability, obstacle detection, and navigation performance for practical use [16].

While existing studies have established the feasibility of EOG-based interfaces, many approaches remain limited by their reliance on subject-dependent calibration, a primary focus on virtual communication tasks, or limited validation on physical robotic platforms. Furthermore, relatively few studies have systematically evaluated low-complexity feature representations using a rigorous subject-independent validation protocol. To address these gaps, the present study develops a subject-independent EOG-based horizontal gaze classification framework and demonstrates its applicability through the real-time control of an omnidirectional mobile robot. Five classifiers representing distance-based, ensemble-based, margin-based, tree-based, and neural-network-based learning approaches were evaluated under the same Leave-One-Subject-Out framework. In doing so, this work provides a practical, experimentally validated, and calibration-free approach for assistive robotic control.

## 2. MATERIALS AND METHODS

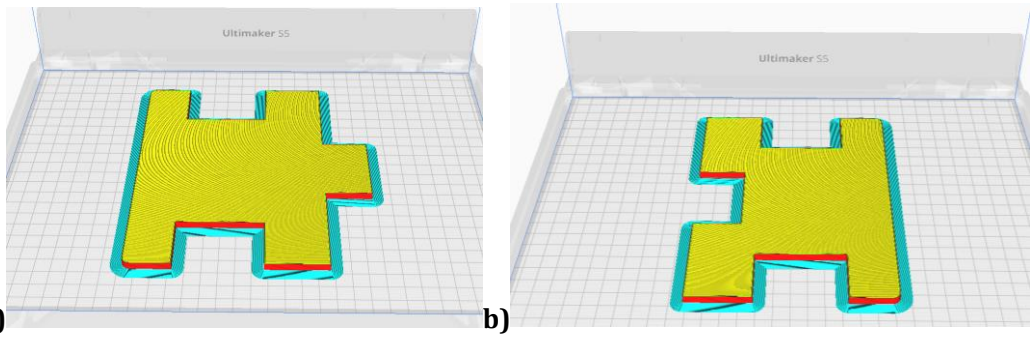
### 2.1. Participants

EOG data were acquired from 12 healthy male volunteers aged between 22 and 28 years (mean age: 24.0 ± 1.8 years). Recruitment was conducted on a voluntary basis, and written informed consent was obtained prior to data acquisition. Individuals with a known history of neurological disorders, visual impairments affecting eye movements, or any medical condition that could influence the measurements were excluded. All personal data were anonymized prior to analysis; no directly identifiable information was recorded, and each participant was assigned a unique coded identifier for data processing. The experimental protocol was approved by the Ethics Committee of Sakarya University of Applied Sciences on 15 January 2026 (Decision No: 64/42).

### 2.2. Design of the omnidirectional mobile robot

The developed mobile robot platform was designed with an omnidirectional mechanical structure to enable multi-directional motion and to provide a realistic testbed for evaluating EOG-based directional control strategies on a physical robotic system. Omnidirectional mobile robot architectures are widely adopted in assistive and HMI applications due to their ability to perform flexible maneuvers without requiring complex turning motions, thereby improving controllability and safety in constrained environments [17].

The robot chassis was modeled using a modular design approach, prioritizing lightweight construction, mechanical robustness, and ease of assembly. To this end, the chassis was designed for fabrication using a three-dimensional (3D) printing process, and interlocking male–female mechanical joints were employed to facilitate rapid assembly and disassembly. The designed chassis components prior to fabrication are illustrated in Figure 3.




**a)** **b)**

**Figure 3.** Components prior to fabrication; (a) male part, (b) female part

Following the fabrication process, the manufactured chassis components were obtained as shown in Figure 4.

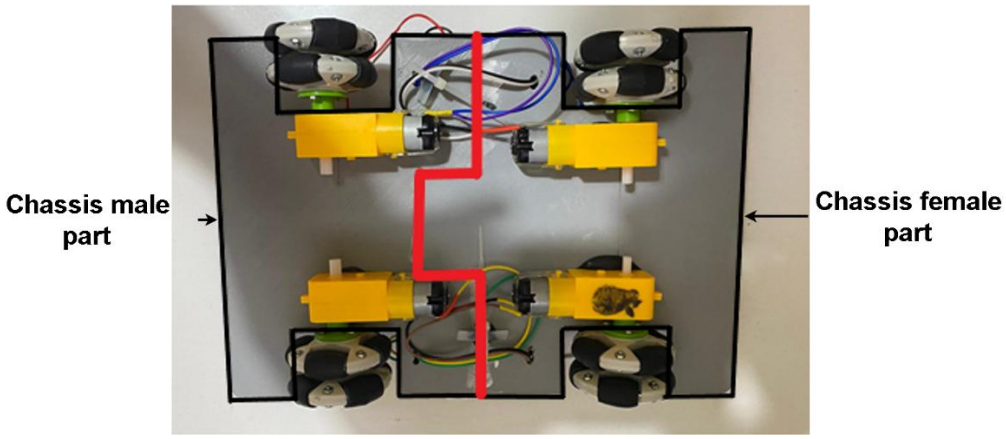



**Figure 4.** Manufactured chassis component

By integrating mecanum-based omnidirectional wheels into the platform, the robot is capable of performing forward–backward motion, lateral translation, diagonal movement, and in-place rotation simultaneously. Compared to conventional differential-drive mobile robots, this holonomic configuration provides significantly higher maneuverability and motion flexibility, which is particularly advantageous for evaluating gaze-based directional commands under realistic operating conditions [18].

After fabrication and electronic integration, the chassis components were assembled and inspected to ensure mechanical compatibility and structural integrity. The side view and physical appearance of the fully assembled omnidirectional mobile robot prototype are shown in Figure 5.

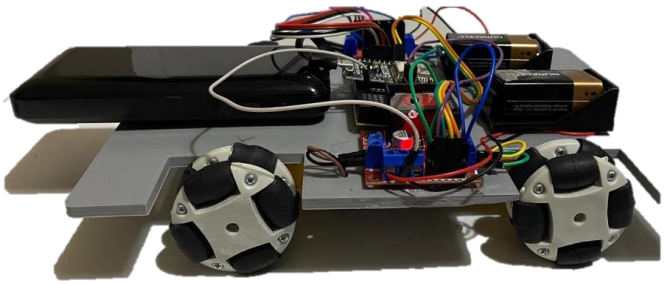


**Figure 5.** Side view of the assembled mobile robot

The overall electronic system architecture of the mobile robot, including sensing, processing, motor driving, and wireless communication units, is schematically presented in Figure 6. This architecture illustrates the interaction between the control unit, motor drivers, wireless communication modules, and peripheral components, providing an integrated view of the robot's operational workflow.


CC: BY-NC-ND

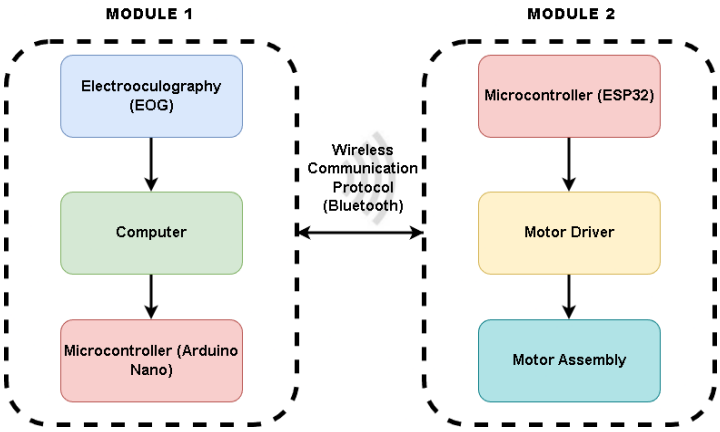



**Figure 6.** Block diagram illustrating the operation of the electronic system used in this study

The motion control subsystem of the mobile robot was implemented using an ESP32 microcontroller, which served as the central control unit. The ESP32 was selected due to its sufficient computational capability, real-time performance, and integrated Bluetooth communication module, making it well suited for mobile robotic applications. The speed and direction of the DC motors were controlled via motor driver circuits according to the received motion commands. Due to the use of omni-directional wheels, the mobile robot can execute multiple motion primitives without changing its orientation. The achievable motion directions including forward/backward translation, lateral motion, diagonal movement, and in-place rotation are schematically illustrated in Figure 7.

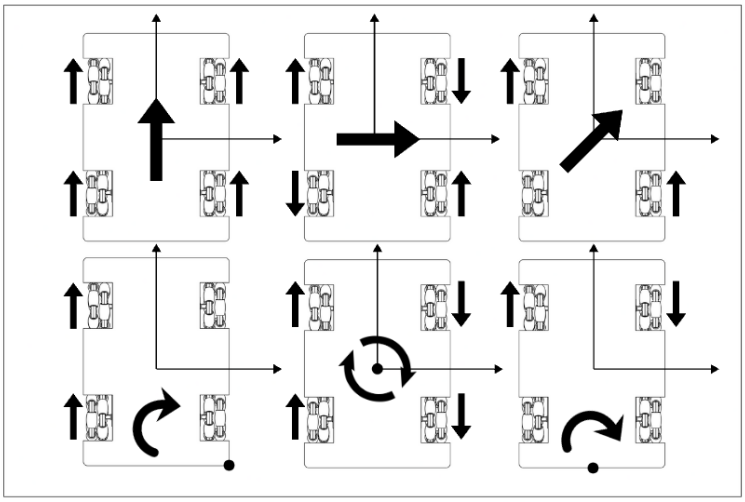


**Figure 7.** Appearance of the mobility capability of the omni-wheeled mobile robot

Within the scope of this study, horizontal gaze directions classified from EOG signals were directly mapped to robot motion commands. Specifically, a left gaze command resulted in lateral leftward motion, a right gaze command produced lateral rightward motion, and a center gaze command corresponded to forward motion or a neutral state depending on the operational mode. This direct gaze-to-motion mapping strategy leverages the holonomic mobility of the platform and enables intuitive directional control without requiring explicit rotational maneuvers, thereby simplifying user interaction and improving control intuitiveness.

In the proposed system architecture, EOG signals were acquired using an EOG sensor connected to an Arduino Nano microcontroller, which performed analog-to-digital conversion and preliminary data handling. The extracted directional commands were transmitted wirelessly to the ESP32-based mobile robot via an HC-05 Bluetooth module. The ESP32 received these commands and executed the corresponding motion actions on the omnidirectional platform. This wireless master–slave

communication structure enables flexible, untethered real-time operation and is consistent with previously reported Bluetooth-based EOG-controlled mobile robot systems [17].

### 2.3. EOG measurement setup

A multi-channel biosignal sensor board capable of EMG, EOG, and electrocardiography (ECG) measurements was used to acquire EOG signals, with only the EOG mode activated. The sensor board detects the corneo-retinal potential difference generated between the retina and the cornea during eye movements via Ag/AgCl surface electrodes. These microvolt-level signals are then amplified to a processable level through integrated analog pre-amplification and filtering stages.

During the data acquisition phase, recordings were made for 5 seconds each of center gaze, right gaze, left gaze, and eye blinks from each participant. Within the scope of this study, only EOG signals corresponding to horizontal gaze movements (center–right–left) were included in the analysis and classification process. Eye blink data were recorded solely for potential use in future studies aimed at expanding the command set. All experimental procedures were conducted in accordance with the ethical approval and with informed consent obtained from all participants.

The horizontal EOG electrodes were positioned at the outer canthi of the right and left eyes, while the reference electrode was placed on the forehead. This configuration represents a well-established standard for two-channel EOG-based human–computer interfaces, ensuring reliable detection of potential differences related to horizontal eye movements [19]. The analog EOG signals were digitized via the microcontroller's analog inputs at a sampling frequency of 100 Hz with 10-bit resolution. The acquired data were transmitted to a computer environment via serial communication and recorded in real time.

It is worth noting that the polarity of the recorded horizontal EOG signals in this study is inverted relative to the conventional representation found in some literature (where rightward gaze typically yields a positive deflection) [3,5]. This inversion results directly from the specific electrode configuration and reference orientation employed. However, all gaze labels, feature extraction procedures, and classification models were consistently defined and trained according to this polarity throughout the data processing pipeline. Therefore, this inversion does not affect the validity or classification performance of the proposed system. The EOG sensor board, electrode placement, and overall experimental setup are illustrated in Figure 8.

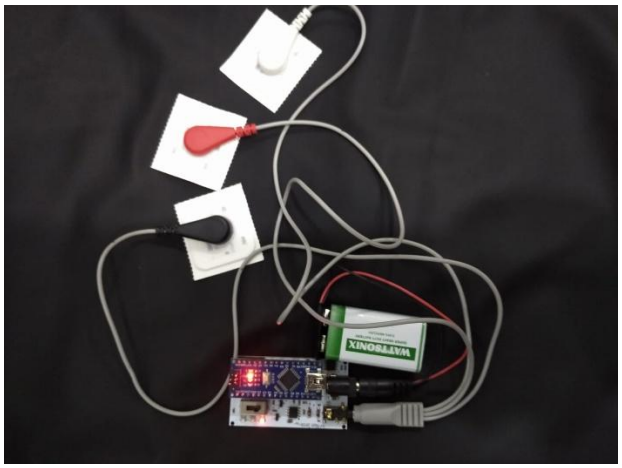



**Figure 8.** EOG sensor setup used in the study

### 2.4. Signal processing and time window generation

Raw EOG signals contain low-amplitude biopotential variations generated by eyeball movements; however, they are also significantly affected by undesired factors such as electrode–skin impedance variations, DC offset, baseline drift, and environmental noise components during acquisition [6]. Therefore, appropriate signal processing steps are required to obtain a reliable and


CC: BY-NC-ND

stable classification process. The horizontal EOG electrode placement and reference configuration adopted in this study, which form the basis for the acquired signal polarity and amplitude characteristics, are illustrated in Figure 9. The overall EOG signal acquisition and processing pipeline applied in this study is schematically illustrated in Figure 10.

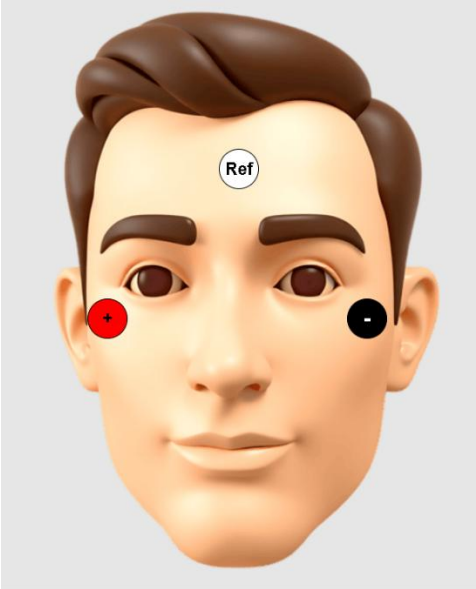



**Figure 9.** Horizontal EOG electrode placement and reference configuration used in the study

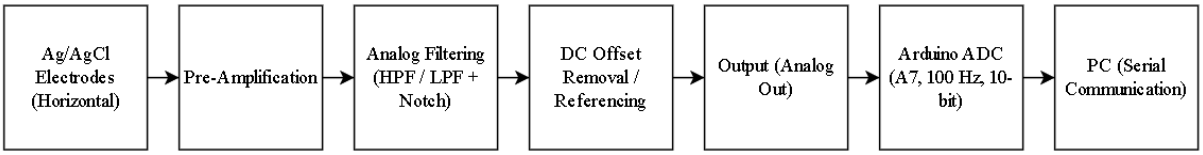



**Figure 10.** Block diagram of the EOG signal acquisition, analog front-end conditioning, and digital processing pipeline

In the software environment, the digitized EOG signals were further processed using a 4th-order Butterworth digital band-pass filter operating within the 0.1–15 Hz frequency range, which is widely adopted in EOG-based gaze analysis studies to suppress baseline drift and high-frequency noise while preserving horizontal eye movement–related signal components [13, 20]. To minimize phase distortion, filtering was performed using a forward–backward zero-phase filtering approach. This frequency band preserves the dominant spectral components associated with horizontal eye movements while effectively suppressing residual baseline drift and high-frequency noise. A time-domain comparison of the raw and digitally filtered EOG signals is presented in Figure 11, illustrating the effectiveness of the combined analog and digital signal processing stages.

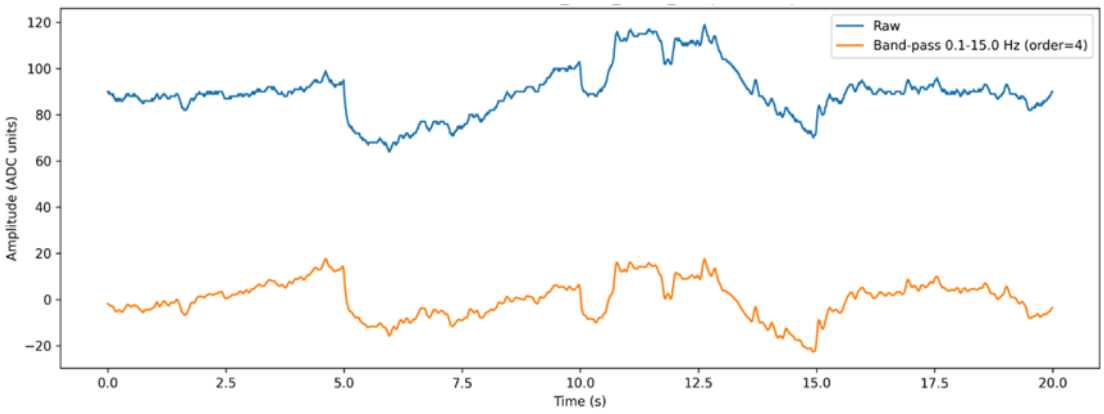



**Figure 11.** Time-domain comparison of the EOG signal before and after signal processing, illustrating the combined effect of analog front-end conditioning and digital band-pass filtering

In addition to horizontal eye movements, eye blink–related EOG activity was also recorded during the data acquisition process, as blink events naturally occur during voluntary gaze tasks. However, within the scope of this study, blink-related signals were not included in the signal processing, feature extraction, or classification stages. The analysis was intentionally restricted to horizontal gaze movements in order to focus on directional control commands and to avoid introducing additional variability associated with blink dynamics. Accordingly, time windows containing blink activity were recorded but excluded from subsequent processing and model training.

The filtered EOG signals were subsequently segmented into fixed-length 5-second time windows. This window length was selected to sufficiently capture the temporal dynamics of horizontal gaze movements while maintaining an acceptable latency for real-time applications. Each time window was labeled according to the horizontal gaze direction performed by the user (center, right, or left). Following the segmentation process, an additional quality control step was applied to ensure the reliability of the dataset.

To prevent hardware-induced noise from degrading the classification performance, a rigorous visual inspection was performed on the collected data. Segments heavily contaminated by severe baseline drift or involuntary artifacts were systematically rejected, alongside the intentional exclusion of blink-related segments. Following this strict quality control, exactly 71 valid epochs were retained per class. Consequently, the machine learning models were trained and evaluated on a balanced dataset comprising a total of 213 valid samples (71 instances × 3 classes).

The selection of a 5-second window length was a deliberate methodological choice driven by safety, signal resolution, and user intentionality. In gaze-based human-machine interaction, distinguishing deliberate directional commands from involuntary eye twitches, visual exploration, or pathological spasms (e.g., in patients with ALS or multiple sclerosis) is a fundamental challenge, widely known in the literature as the "Midas Touch" problem [21]. The 5-second epoch acts as a necessary dwell-time buffer to confirm command intentionality, which is critical for maintaining the user's sense of agency in human–machine interaction systems [22]. This approach helps prevent accidental false-positive classifications that could lead to unintended or unstable control actions in assistive systems. Furthermore, from a signal processing perspective, a 5-second duration provides optimal time-frequency resolution, ensuring that sufficient low-frequency EOG data points are captured to robustly compute spectral and statistical features.

### 2.5. Feature extraction

The non-stationary nature of EOG signals and significant inter-subject amplitude variability can negatively affect the performance of direct classification approaches based on raw signal samples. Therefore, feature-based analysis methods are commonly preferred in EOG-based human–computer interface studies due to their improved robustness and interpretability [6]. During model development and performance evaluation, feature extraction and classification were performed offline on the

recorded dataset. Following model development, the same signal-processing and classification pipeline was integrated into the real-time robot control interface.

From each 5-second time window, twelve (12) features representing the signal amplitude, statistical distribution, and band-based spectral content were extracted. Time-domain statistical features included the mean (μ), standard deviation (σ), minimum value (V_min), maximum value (V_max), peak-to-peak amplitude (V_pp), root mean square (RMS), skewness, and kurtosis. These descriptors characterize overall signal level variations, amplitude dispersion, and distribution properties of the EOG signal and are widely employed in eye-movement-related EOG analysis [6].

In addition, the zero-crossing rate (ZCR) was calculated to capture waveform variability. To incorporate frequency-domain information, the fast Fourier transform (FFT) of each signal segment was computed, and the total spectral energy (E), as well as the summed magnitude values within the 0–5 Hz and 5–15 Hz frequency bands, were extracted as separate features. These frequency bands were selected to distinguish dominant low-frequency eye movement components from higher-frequency signal content typically present in EOG recordings [6].

**As a result, the extracted features for each time window were concatenated to form a 1×12 feature vector:**

$$F_v = \mu\ , \sigma, V_{max}, V_{min}, V_{pp}, RMS, Skew, Kurt, ZCR, E, B_1, B_2 \quad (1)$$

To prevent features with different numerical scales from adversely affecting classification performance and to enhance subject-independent generalization capability, Z-score standardization was applied at each validation step. The normalization parameters were computed exclusively from the training data and subsequently applied to the validation and test datasets.

### 2.6. Classification algorithms and methodology

Supervised machine-learning algorithms were employed to classify the extracted EOG feature vectors into three horizontal gaze classes: center, right, and left. Five classifiers representing complementary learning paradigms were evaluated: k-Nearest Neighbors (KNN), Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), and a multilayer perceptron-based artificial neural network (ANN).

Prior to classification, Z-score standardization was applied to the KNN, SVM, and ANN models. The normalization parameters were calculated exclusively from the training subset within each fold and were subsequently applied to the corresponding validation and test subsets. This procedure prevented information leakage from the held-out participant and ensured that the reported results reflected subject-independent generalization.

KNN was employed as a distance-based reference classifier because of its simplicity and its ability to reflect the discriminative structure of the extracted feature space without imposing strong distributional assumptions. Previous studies have demonstrated that KNN can effectively classify EOG-based gaze and eye-movement patterns using statistical and frequency-domain features [6, 14].

RF was selected as an ensemble-based classifier because of its robustness to noise, ability to model nonlinear relationships, and reduced sensitivity to overfitting compared with individual decision trees. These characteristics make RF suitable for biosignal data affected by inter-participant variability and non-Gaussian feature distributions [6, 7, 14].

SVM with a radial basis function (RBF) kernel was included because of its strong generalization capability under limited-sample and nonlinear classification conditions. The RBF kernel enables nonlinear separation of the center, right, and left gaze classes in the standardized feature space, while the margin-maximization principle helps control model complexity [6, 14].

DT was added as an interpretable tree-based classifier that recursively partitions the feature space according to feature-specific decision rules. Its inclusion enabled direct evaluation of a single-tree model against the ensemble-based RF approach. Decision-tree models have previously been

applied to EOG-based eye-movement classification together with KNN and SVM, providing a relevant benchmark for comparing supervised classifiers in EOG analysis [14].

A small-scale multilayer perceptron ANN was also evaluated to examine whether a nonlinear neural-network model could capture interactions among the twelve extracted EOG features. ANN-based models have previously been used to identify directional eye movements from EOG signals, demonstrating their ability to learn nonlinear relationships between signal-derived features and eye-movement classes [23]. Because the present dataset was relatively small, a compact network architecture was preferred instead of a deep neural network in order to limit model complexity and reduce the risk of overfitting.

This combination of distance-based, ensemble-based, margin-based, tree-based, and neural-network-based learning approaches provided a broader and methodologically diverse comparison of subject-independent EOG classification performance.

Within each outer LOSO fold, all samples belonging to one participant were reserved exclusively as the test set. The samples from the remaining participants were randomly divided into training and validation subsets. The validation-set size was obtained by taking the floor of 20% of the available non-test samples, and the remaining samples were assigned to training. The split was generated using NumPy's default_rng function with a fold-specific seed calculated as 42 plus the fold number, where the fold number ranged from 1 to 12. Hyperparameter combinations were evaluated according to validation accuracy, and the first configuration was retained when multiple configurations achieved the same validation accuracy. The selected model was trained using only the corresponding training subset and was then evaluated once on the held-out participant; the training and validation subsets were not recombined before testing. The same procedure was applied to all five classifiers.

The hyperparameter search spaces and fixed settings used for all five classifiers are summarized in Table 1.


CC: BY-NC-ND

**Table 1.** Hyperparameter search spaces and fixed settings used for the evaluated classifiers

| Model | Hyperparameter | Values / Range |
|---|---|---|
| KNN | Number of neighbors, k | 1, 3, 5, …, 21 |
| KNN | Distance metric | Euclidean, Manhattan |
| KNN | Neighbor weighting | Uniform, distance |
| KNN | Feature preprocessing | Z-score standardization |
| RF | Number of trees | 200, 500 |
| RF | Maximum tree depth | None (unlimited), 10, 20 |
| RF | Minimum samples per leaf | 1, 2, 4 |
| RF | Minimum samples required to split | 2 |
| RF | Number of features considered at each split | Square root of the total number of features |
| RF | Class weighting | None |
| RF | Bootstrap sampling | Enabled |
| RF | Out-of-bag scoring | Disabled |
| RF | Random seed | 42 |
| SVM | Kernel | Radial basis function |
| SVM | Regularization parameter, C | 0.1, 1, 10, 100 |
| SVM | Kernel coefficient, $\gamma$ | scale, auto, 0.01, 0.1 |
| SVM | Class weighting | None |
| SVM | Feature preprocessing | Z-score standardization |
| DT | Split criterion | Gini impurity, entropy |
| DT | Maximum tree depth | None (unlimited), 3, 5, 10 |
| DT | Minimum samples required to split | 2, 5, 10 |
| DT | Minimum samples per leaf | 1, 2, 4 |
| DT | Class weighting | None, balanced |
| DT | Random seed | 42 |
| ANN | Hidden-layer architecture | (16), (32), (32, 16) neurons |
| ANN | Activation function | ReLU, tanh |
| ANN | (L2) regularization coefficient, $\alpha$ | 0.0001, 0.001, 0.01 |
| ANN | Optimization solver | L-BFGS |
| ANN | Maximum number of iterations | 2000 |
| ANN | Random seed | 42 |
| ANN | Feature preprocessing | Z-score standardization |

The hyperparameter ranges and fixed model settings were specified a priori and kept unchanged across all LOSO folds. Hyperparameter selection was performed independently within each fold using only the corresponding validation subset. Therefore, the selected configuration could differ depending on the held-out participant, while the search procedure and candidate values remained identical across folds. All fold-wise selected configurations were retained in the analysis records to ensure reproducibility.

#### 2.6.1. Validation strategy: leave-one-subject-out (LOSO)

Subject-independent performance was evaluated using Leave-One-Subject-Out (LOSO) cross-validation. In each of the twelve outer folds, all samples belonging to one participant were excluded from model development and used exclusively as the test set. The samples from the remaining

participants constituted the non-test dataset, from which the training and validation subsets were generated according to the procedure described in Section 2.6. Each participant therefore served as the unseen test participant exactly once. This subject-wise separation strategy is commonly used in biosignal classification studies to assess model generalization to previously unseen individuals and to prevent subject-specific characteristics from being shared between the training and test datasets [6, 24, 25].

The overall LOSO procedure adopted in this study is illustrated in Figure 12. At each iteration, one participant was reserved as the independent test participant, while samples from the remaining eleven participants were used for model development. The held-out participant was changed systematically until all twelve participants had served as the test participant once.

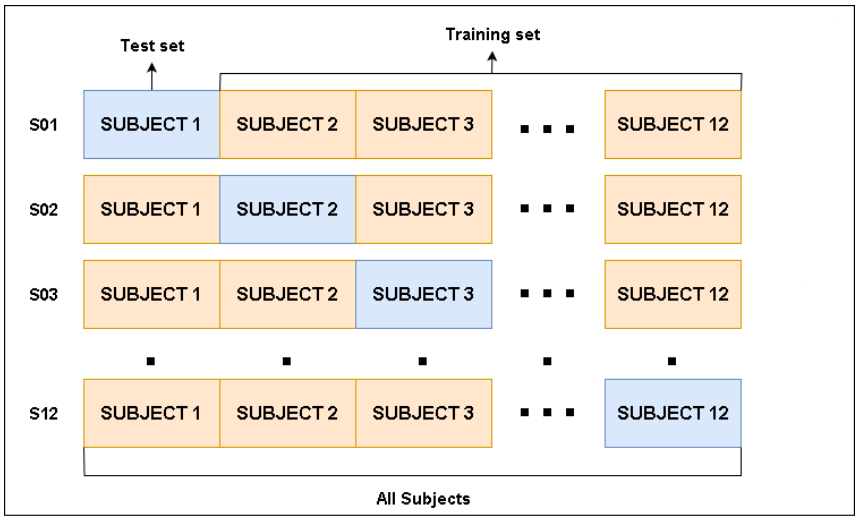



**Figure 12.** Schematic representation of the Leave-One-Subject-Out cross-validation strategy used in this study

Hyperparameter selection was conducted independently within each LOSO fold using validation accuracy. After the optimal configuration had been selected, the corresponding model was evaluated on the held-out participant without using any test samples during preprocessing, hyperparameter selection, or training. This nested separation between model development and final testing provided a realistic assessment of subject-independent generalization and reduced the risk of optimistic performance estimates caused by information leakage [25].

Participant-wise accuracy was calculated separately for each LOSO fold. The overall LOSO accuracy reported in Table 2 represents the unweighted mean of the twelve participant-wise fold accuracies, ensuring that each participant contributed equally regardless of the number of available samples. In contrast, the confusion matrices and class-wise performance metrics were calculated from the pooled out-of-fold predictions obtained across all twelve held-out participants.

### 2.6.2. Performance evaluation metrics

Classification performance was evaluated using accuracy, precision, recall, F1-score, specificity, and the area under the receiver operating characteristic curve (ROC-AUC). Accuracy represents the proportion of correctly classified samples among all evaluated samples. Precision indicates the proportion of samples predicted as a given class that were correctly classified, whereas recall represents the proportion of actual samples belonging to that class that were correctly identified. The F1-score was calculated as the harmonic mean of precision and recall.

**Specificity was calculated separately for each class using a one-vs-rest formulation:**

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$


CC: BY-NC-ND

where TN denotes the number of true-negative predictions and FP denotes the number of false-positive predictions for the corresponding class. For each gaze direction, the target class was treated as the positive class, while the remaining two classes were combined as the negative class.

Multiclass ROC-AUC was also calculated using the one-vs-rest approach. Class-probability outputs were used for KNN, RF, DT, and ANN, whereas continuous decision scores were used for SVM. Class-specific ROC curves were generated for the center, right, and left gaze classes. Macro-average specificity and ROC-AUC values were obtained by averaging the three class-specific results without weighting.

Participant-wise LOSO accuracy was calculated separately for each held-out participant. The mean, standard deviation, minimum, and maximum of the twelve participant-wise accuracies are reported in Table 2. In contrast, precision, recall, F1-score, specificity, ROC-AUC, and confusion matrices were calculated from the pooled out-of-fold predictions obtained across all 213 samples and are reported in Table 3.

## 3. RESULTS

The primary subject-independent evaluation was performed using a Leave-One-Subject-Out (LOSO) validation framework. A complementary subject-dependent evaluation is presented in Section 3.8. Classification performance was evaluated using feature vectors extracted from 5-second EOG signal windows and classified using KNN, RF, SVM, DT, and ANN models.

### 3.1. Overall classification accuracy

The overall classification accuracy results, including mean, standard deviation, and range values under LOSO validation, are presented in Table 2.

**Table 2.** Summary of overall classification accuracy results for each algorithm under LOSO validation

| Model | Mean Accuracy (%) | Std. (%) | Min. (%) | Max. (%) |
|---|---|---|---|---|
| SVM | 79.61 | 10.32 | 66.67 | 95.83 |
| KNN | 74.47 | 18.67 | 33.33 | 91.67 |
| RF | 72.19 | 16.31 | 33.33 | 88.89 |
| DT | 68.69 | 13.93 | 33.33 | 87.50 |
| ANN | 66.11 | 16.17 | 33.33 | 88.89 |

Among the evaluated models, SVM achieved the highest mean classification accuracy (79.61%) and the lowest standard deviation (10.32%), indicating the most stable performance across participants. KNN ranked second with a mean accuracy of 74.47%, although it exhibited the highest variability (18.67%). RF achieved an intermediate mean accuracy of 72.19%. DT and ANN produced lower subject-independent accuracies of 68.69% and 66.11%, respectively, suggesting that the limited dataset size and inter-participant variability restricted the generalization capability of the tree-based and neural-network-based models.

### 3.2. Class-based performance metrics

To provide a more comprehensive class-wise evaluation, precision, recall, F1-score, specificity, and one-vs-rest ROC-AUC were calculated separately for the center, right, and left gaze classes. These metrics were computed from the pooled out-of-fold predictions obtained across all twelve LOSO folds, corresponding to a total of 213 test predictions. The results are presented in Table 3.

**Table 3.** Class-wise performance metrics calculated from pooled out-of-fold LOSO predictions

| Model | Class | Precision | Recall | F1-Score | Specificity | ROC-AUC |
|---|---|---|---|---|---|---|
| KNN | Center | 0.818 | 0.761 | 0.788 | 0.915 | 0.884 |
| KNN | Right | 0.764 | 0.775 | 0.769 | 0.880 | 0.873 |
| KNN | Left | 0.787 | 0.831 | 0.808 | 0.887 | 0.900 |
| RF | Center | 0.737 | 0.789 | 0.762 | 0.859 | 0.890 |
| RF | Right | 0.743 | 0.732 | 0.738 | 0.873 | 0.888 |
| RF | Left | 0.821 | 0.775 | 0.797 | 0.915 | 0.905 |
| SVM | Center | 0.829 | 0.817 | 0.823 | 0.915 | 0.901 |
| SVM | Right | 0.756 | 0.831 | 0.792 | 0.866 | 0.896 |
| SVM | Left | 0.892 | 0.817 | 0.853 | 0.951 | 0.925 |
| DT | Center | 0.624 | 0.746 | 0.679 | 0.775 | 0.790 |
| DT | Right | 0.750 | 0.634 | 0.687 | 0.894 | 0.781 |
| DT | Left | 0.794 | 0.761 | 0.777 | 0.901 | 0.852 |
| ANN | Center | 0.681 | 0.690 | 0.685 | 0.838 | 0.857 |
| ANN | Right | 0.696 | 0.676 | 0.686 | 0.852 | 0.817 |
| ANN | Left | 0.736 | 0.746 | 0.741 | 0.866 | 0.854 |

Note: Each class contained 71 pooled out-of-fold test samples. ROC-AUC values were calculated using a one-vs-rest strategy.

The class-wise results indicate that SVM achieved the strongest overall discrimination across the three gaze directions. The highest class-level F1-score was obtained for the Left class using SVM (0.853), together with a specificity of 0.951 and a ROC-AUC of 0.925. KNN and RF also produced comparatively balanced class-wise results, whereas DT and ANN demonstrated lower F1-scores, particularly for the Center and Right classes. Across the evaluated models, the Right class was more frequently confused with the Center class, indicating greater overlap between the feature distributions of these two gaze conditions.

### 3.3. Participant-wise accuracy distribution

Figure 13 presents the participant-wise classification accuracies obtained under LOSO validation for the KNN, RF, SVM, DT, and ANN models. The results reveal noticeable inter-participant variability across all classifiers. Overall, SVM produced the most consistently high accuracy values, whereas DT and ANN showed comparatively lower performance for several held-out participants.

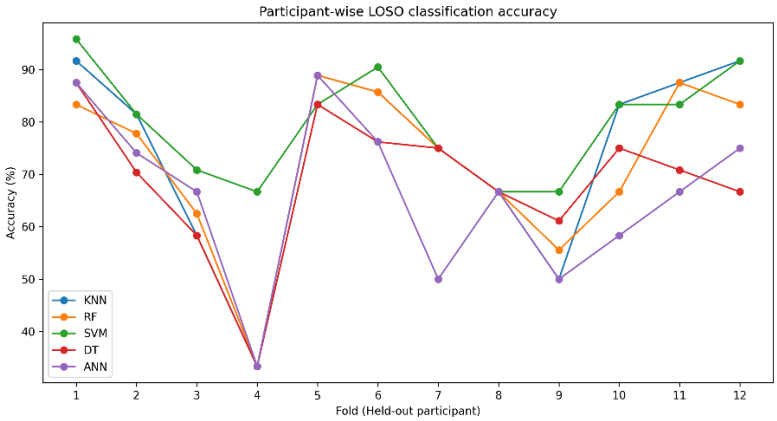



**Figure 13.** Participant-wise LOSO classification accuracies of the KNN, RF, SVM, DT, and ANN models across the twelve held-out participants


CC: BY-NC-ND

### 3.4. Confusion matrix analysis

The aggregated confusion matrices obtained from the pooled out-of-fold predictions of the KNN, RF, SVM, DT, and ANN models are presented in Figures 14–18, respectively. Each matrix summarizes the predictions generated across all twelve LOSO folds for the center, right, and left gaze classes.

SVM produced the strongest overall class separation, with comparatively high numbers of correct predictions along the main diagonal. KNN and RF also demonstrated balanced classification performance, whereas DT and ANN exhibited more misclassifications, particularly between the center and right classes. These findings are consistent with the class-wise performance metrics reported in Table 3.

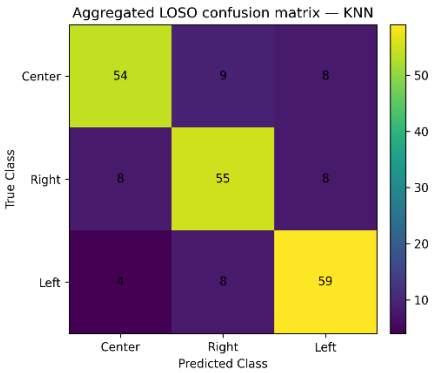



**Figure 14.** Aggregated LOSO confusion matrix for the KNN classifier

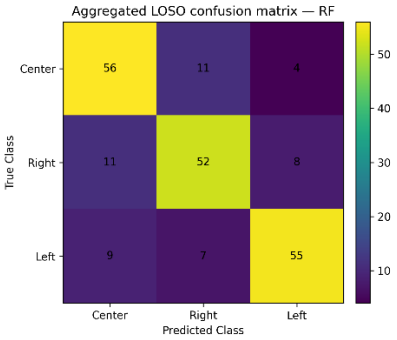



**Figure 15.** Aggregated LOSO confusion matrix for the RF classifier

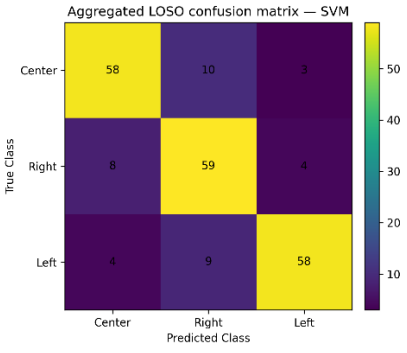



**Figure 16.** Aggregated LOSO confusion matrix for the SVM classifier.

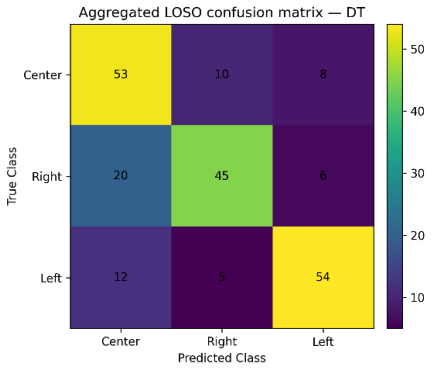



**Figure 17.** Aggregated LOSO confusion matrix for the DT classifier

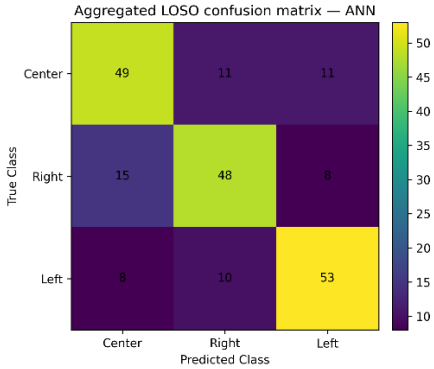



**Figure 18.** Aggregated LOSO confusion matrix for the ANN classifier

### 3.5. ROC-AUC analysis

The one-vs-rest ROC-AUC results reported in Table 3 further supported the overall classification findings. SVM achieved the strongest class-separation performance, with ROC-AUC values of 0.901, 0.896, and 0.925 for the center, right, and left classes, respectively. KNN and RF also demonstrated consistently high ROC-AUC values across all three classes. In contrast, DT and ANN produced comparatively lower ROC-AUC values, particularly for the right gaze class. These results indicate that SVM provided the most reliable discrimination among the three horizontal gaze directions under subject-independent conditions.

### 3.6. Window-based classification of a real-time test signal

Figure 19 illustrates the class assignments produced by the KNN, RF, and SVM models for a representative EOG signal. The raw EOG signal is shown in the upper panel, while the lower panels present the window-based predictions of each classifier over time. DT and ANN were included in the expanded LOSO comparison but were not part of the original real-time temporal visualization presented in this section.


CC: BY-NC-ND

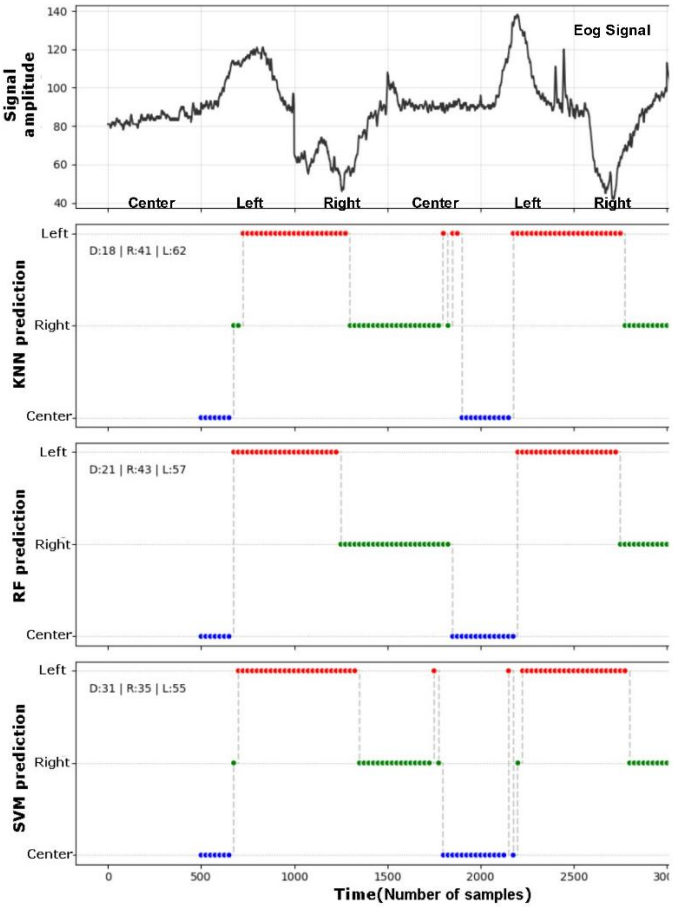



**Figure 19.** Class assignments of the KNN, RF, and SVM models for a representative real-time EOG test signal

The figure demonstrates that different classifiers exhibit varying levels of temporal stability and sensitivity, particularly around gaze transition regions, providing complementary insight beyond aggregate performance metrics. The signal shown in Figure 19 was used exclusively for qualitative visualization of real-time classification behavior and was not included in the training, validation, or statistical evaluation processes.

### 3.7. Real-time system implementation results

The proposed EOG-based classification framework was implemented as a real-time mobile robot control system. During operation, EOG signals acquired from the sensor were transferred to the computer via serial communication, preprocessed, and classified online using the selected machine-learning model. The resulting gaze-direction commands were transmitted wirelessly to the mobile robot controller for motion execution.

The graphical user interface developed for real-time visualization, model selection, online EOG classification, and command transmission is shown in Figure 20. In the representative interface screenshot, the KNN model was selected.

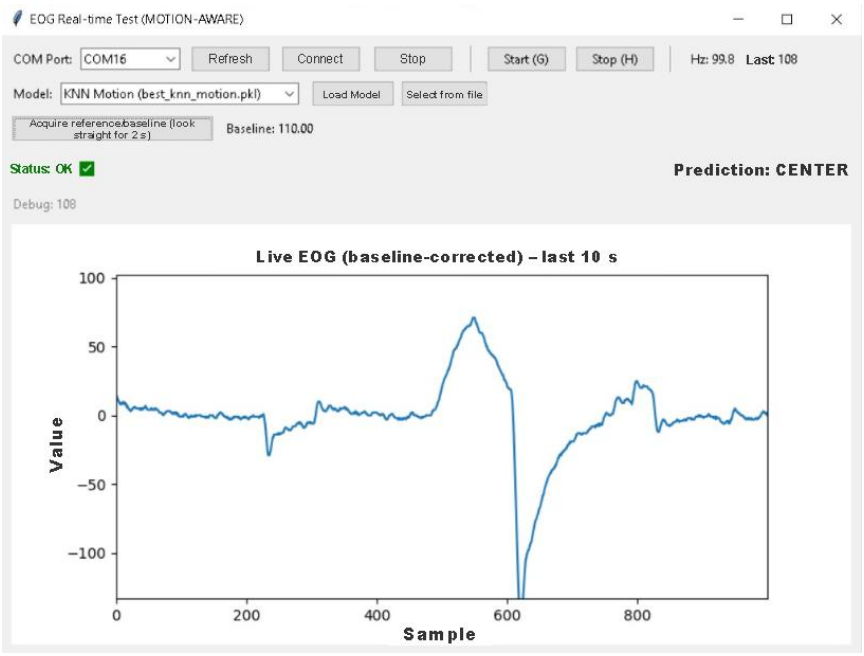



**Figure 20.** Graphical user interface for real-time EOG visualization, model selection, online classification, and command transmission; the KNN model is shown as a representative example

To ensure stable and safe control behavior, gaze classification was performed using non-overlapping 5-second windows. This interval functioned as a command-confirmation period that reduced the likelihood of transient artifacts, involuntary eye movements, and short-duration signal fluctuations being interpreted as intentional commands. Therefore, the 5-second interval represents the signal-acquisition and decision window rather than the computational latency of the system.

After completion of each decision window, feature extraction and classification required less than 2 ms on the host computer. The additional delay between command generation, wireless transmission, and robot actuation remained below 200 ms during repeated experimental trials. Thus, the total user-perceived response included the intentional 5-second command-confirmation interval, while the subsequent computational, communication, and actuation stages were completed within approximately 200 ms. This distinction avoids interpreting the window duration, inference time, and command-to-actuation delay as conflicting latency measurements.

### 3.8. Subject-dependent evaluation (train/validation/test split)

The expanded five-classifier comparison was conducted under the primary subject-independent LOSO framework. The complementary subject-dependent evaluation was retained for the original KNN, RF, and SVM models to enable direct comparison with the previously reported results. Therefore, DT and ANN were not included in this secondary analysis.

To provide a direct comparison with the subject-independent LOSO evaluation, a subject-dependent evaluation was also performed. In this approach, the dataset, consisting of 213 samples, was randomly divided into training (70%), validation (15%), and test (15%) sets. This split strategy is commonly adopted in machine learning studies to balance model training and unbiased evaluation by preserving separate validation and test sets [26, 27]. This corresponds to 149 training samples, 32 validation samples, and 32 test samples.

Unlike the LOSO framework, samples from the same subjects may appear in all subsets, allowing the models to learn subject-specific patterns. To ensure a fair and consistent comparison, the same 12 features used in the LOSO evaluation were consistently employed in this subject-dependent approach.

To optimize model performance, hyperparameter tuning was conducted using the validation set. The optimal hyperparameter configurations for each classifier are summarized in Table 4.


CC: BY-NC-ND

**Table 4.** Optimal hyperparameter configurations for the subject-dependent models

| Model | Selected Hyperparameters |
|---|---|
| KNN | n_neighbors = 9, metric = Euclidean, weights = distance |
| RF | n_estimators = 200, max_depth = None, min_samples_split = 5, min_samples_leaf = 1, max_features = sqrt |
| SVM | kernel = RBF, C = 0.1, gamma = scale |

The performance metrics obtained from the test set (32 samples) using these optimal hyperparameters are presented in Table 5. As expected, the subject-dependent evaluation demonstrated higher classification performance compared to the LOSO results. This improvement can be attributed to the models benefiting from the presence of subject-specific characteristics in both the training and testing phases.

Among the evaluated models, the KNN classifier achieved the highest performance with a test accuracy of 93.75% and a macro-averaged F1-score of 0.94, correctly classifying 30 out of 32 samples. Random Forest and SVM followed with accuracies of 87.50% (28 correct samples) and 84.38% (27 correct samples), respectively. All reported metrics were computed using macro-averaging across the three classes (center, right, and left).

The confusion matrix for the best-performing model (KNN) is illustrated in Figure 21, providing a detailed view of class-wise prediction performance under the subject-dependent evaluation.

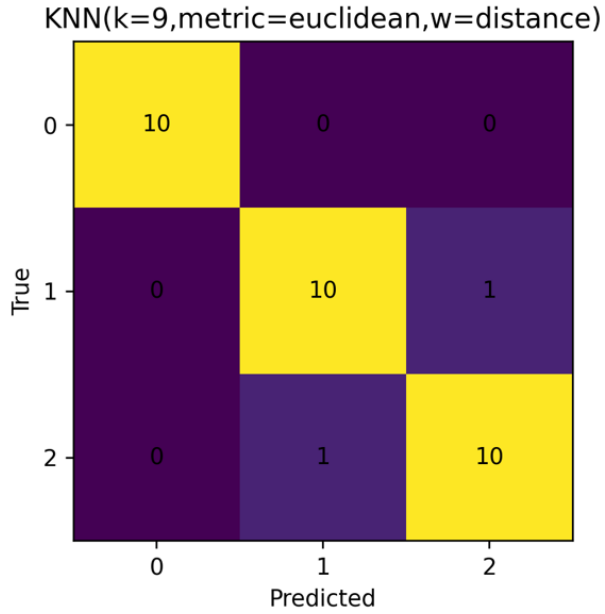



**Figure 21.** Confusion matrix of the best-performing model (KNN) in the subject-dependent evaluation

**Table 5.** Performance metrics of the subject-dependent evaluation on the test set

| Model | Accuracy (%) | Precision | Recall | F1-Score |
|---|---|---|---|---|
| KNN | 93.75 | 0.94 | 0.94 | 0.94 |
| RF | 87.50 | 0.88 | 0.88 | 0.88 |
| SVM | 84.38 | 0.85 | 0.84 | 0.84 |

## 4. DISCUSSION

The investigation focused on the subject-independent classification of horizontal gaze directions using EOG signals, demonstrating the feasibility of the proposed framework for real-time mobile robot control. Unlike subject-dependent approaches commonly reported in the literature, the adoption of the Leave-One-Subject-Out (LOSO) validation strategy facilitated a realistic assessment of the system's generalization capability across unseen individuals—a critical requirement for robust assistive HMI applications [6,14].

To contextualize the proposed approach within the existing literature, Table 6 provides a comparative summary of representative EOG-based gaze and control research. For investigations employing multiple classification or modeling strategies, the reported accuracy corresponds to the best-performing model as stated in the original publication. This comparison reveals that the majority of prior research relies on subject-dependent evaluation protocols and frequently reports higher accuracy values, whereas the current work emphasizes subject-independent performance validated through LOSO.

**Table 6.** Comparative summary of representative EOG-based gaze and control studies

| Study | Task | Features / Model | Validation | Accuracy (%) |
|---|---|---|---|---|
| Papetti et al. (2025) [24] | Stress detection (EOG-based) | Time-domain features + ML (RF, SVM, KNN) | Subject-independent (LOSO) | 85.8 (2-class) / 82.4 (3-class) |
| Huang et al. (2018) [28] | Ten-command wheelchair control (blink-based selection) | Single-channel vertical EOG + waveform detection (rule-based) | Subject-independent (training-free) | 96.7 (healthy) / 91.7 (SCI) |
| Martínez-Cerveró et al. (2020) [29] | HCI (4-direction gaze classification) | Time-domain statistical features + probabilistic classification | Subject-dependent (5-fold CV + online test) | 94.21 (model) / 90 (online) |
| Pérez-Reynoso et al. (2021) [30] | Omnidirectional robot control | Feature-based ML (MLP selected) | Subject-dependent (10-fold CV) | 79.5 (balanced accuracy) |
| Mulam et al. (2023) [26] | Bluetooth wheelchair control (5 commands) | Threshold-based EOG classification (rule-based) | Subject-dependent (small sample) | 100 |
| Zeng et al. (2025) [27] | Eye movement classification (multi-dataset) | CNN-based (MIDF-NET) | Subject-independent (LOSO) | 86.0–95.4 |
| Sharma et al. (2020) [4] | Blink + Gaze | Time-domain + Rule-based | Subject-dependent | 85–90 |
| Hernández-Pérez et al. (2023) [14] | Eye movement classification (5 classes) | Wavelet + KNN / SVM | Subject-dependent | 76.9 |
| Ogenga et al. (2023) [12] | Wheelchair control (3-class gaze) | Feature extraction + SVM | Subject-dependent | 96.54 |
| Zendehdel et al. (2025) [31] | UAV control | LSTM (end-to-end) | Subject-dependent | 97.6 |
| Sanjaya et al. (2017) [17] | Mobile robot control | Feature-based EOG + ML | Subject-dependent | 80–85 |
| Barour & Ay Gul (2024) [15] | Virtual keyboard | Time-domain EOG + ML | Subject-dependent | N/A |
| This study | Mobile robot control (Horizontal gaze) | Time- and frequency-band features + KNN/RF/SVM/DT/ANN (best: SVM) | LOSO (subject-independent) | 79.61 |


CC: BY-NC-ND

In the present study, five classifiers were evaluated under the same subject-independent LOSO protocol, and SVM achieved the highest mean accuracy.

Although the classification accuracy obtained in the current work is numerically lower than some subject-dependent benchmarks listed in Table 6, the results were derived under substantially more challenging subject-independent conditions. This distinction is fundamental, as subject-dependent evaluations often overestimate real-world performance due to shared subject-specific characteristics between training and testing datasets. As detailed in Table 2, while the SVM model maintained a robust mean accuracy of 79.61% with a relatively low standard deviation (10.32%), models like KNN exhibited a wide performance range (Min: 33.33% – Max: 91.67%). This significant variability is not an anomaly but rather reflects the inherent difficulty of the LOSO strategy; the minimum accuracy of 33.33% indicates instances where the model failed to generalize to a novel subject's unique EOG signal characteristics (e.g., skin impedance or potential amplitude), resulting in performance equivalent to random guessing. Consequently, the inter-subject variability highlights the realistic performance boundaries of biosignal interfaces when applied to unseen users, confirming that the results are consistent with the challenging nature of the validation method.

The expanded five-model comparison demonstrated that SVM provided the strongest subject-independent generalization, achieving the highest participant-wise mean accuracy (79.61%) together with the highest pooled macro specificity (0.911) and macro ROC-AUC (0.907). KNN and RF showed moderate and comparatively balanced performance, whereas DT and ANN achieved lower mean accuracies of 68.69% and 66.11%, respectively. The lower performance of DT may be associated with the sensitivity of individual tree structures to inter-participant variability, while the relatively small feature-based dataset may have limited the ability of ANN to learn more generalizable nonlinear patterns. These findings further support the suitability of margin-based SVM classification for limited-sample, subject-independent EOG analysis [6,14].

The class-wise analysis showed that the Left gaze class achieved the strongest discrimination, particularly with SVM, which produced an F1-score of 0.853, a specificity of 0.951, and a ROC-AUC of 0.925. In contrast, the Right class was more frequently confused with the Center class across several models. This result suggests greater overlap between the feature distributions of rightward and neutral gaze conditions and highlights the importance of class-wise evaluation in addition to overall accuracy.

Recent advancements have extended EOG-based gaze estimation beyond horizontal directions. Beyond feature-based classical models, contemporary research has shown that deep learning architectures, such as LSTM networks, can detect eye-movement patterns from EOG in real time for hands-free control applications, including unmanned aerial vehicle (UAV) command generation [31]. Furthermore, Wei et al. [32] demonstrated that horizontal and vertical gaze estimation can be achieved even under eyes-closed conditions by incorporating proprioceptive calibration, indicating that EOG signals contain richer spatial information than typically exploited. While the scope of the present investigation is limited to horizontal gaze for directional mobile robot control, these findings suggest that the proposed framework could be extended to multi-dimensional gaze interaction through appropriate feature representations and calibration strategies.

In terms of robotic applications, the developed system complements existing EOG-based control approaches by emphasizing a lightweight, low-cost, and sensor-efficient design. While Suhaimi et al. [33] combined EOG-based gaze estimation with camera-based object recognition for robotic arm control, demonstrating multimodal interaction, the current work prioritizes simplicity, portability, and real-time responsiveness—features particularly relevant for assistive mobile robotics and navigation tasks [3, 7, 14, 15].

Taken together, the findings confirm that subject-independent EOG-based gaze classification using low-cost hardware and conventional machine learning algorithms achieves reliable performance and can be effectively integrated into real-time robotic control architectures. The presented approach aligns with current trends in assistive robotics, emphasizing robustness, minimal calibration requirements, and ease of deployment in real-world environments [1, 2, 6].

In addition to the subject-independent evaluation, the results presented in Section 3.8 showed that classification performance increased under the subject-dependent split, with KNN outperforming SVM. This finding indicates that KNN benefited from subject-specific local similarity patterns, whereas SVM provided stronger generalization to unseen participants under LOSO. Therefore, model selection should be aligned with the intended deployment scenario, and LOSO remains the more realistic evaluation strategy for calibration-free use.

Additionally, unlike rule-based or single-channel approaches, the proposed method combines time-domain and frequency-band features with machine learning models, offering a balance between computational efficiency and classification performance. Although deep learning approaches such as those in [27, 31] achieve higher accuracy, they typically require larger datasets and higher computational resources. Therefore, the proposed framework provides a lightweight and generalizable alternative suitable for real-time assistive robotic systems.

## 5. CONCLUSIONS AND FUTURE WORK

This study demonstrated that horizontal gaze directions (left, center, and right) can be classified in a subject-independent manner using low-complexity time- and frequency-domain features extracted from EOG signals. Five machine-learning models—KNN, RF, SVM, DT, and ANN—were evaluated under the same LOSO framework. SVM achieved the strongest subject-independent performance, with a participant-wise mean accuracy of 79.61%, a pooled macro specificity of 0.911, and a macro one-vs-rest ROC-AUC of 0.907. KNN and RF provided moderate performance, whereas DT and ANN showed lower generalization across unseen participants.

The results indicate that margin-based SVM classification is particularly suitable for limited-sample EOG datasets affected by inter-participant variability. The successful wireless transmission of the classified gaze commands to the omnidirectional mobile robot further demonstrated the practical feasibility of the proposed framework for real-time assistive control. Although subject-dependent evaluation produced higher numerical accuracy, the LOSO results provide a more realistic assessment of performance for previously unseen users without requiring subject-specific retraining.

Future work will focus on extending the command set by incorporating vertical gaze and eye-blink signals, investigating overlapping and adaptive windowing strategies to improve the latency–stability trade-off, and conducting experiments with individuals with motor impairments. In addition, embedded implementation of the complete signal-processing and classification pipeline will be investigated to improve portability and real-world applicability.

**Peer-review:** External, Independent.

**Acknowledgements:**
The authors would like to thank the Sakarya University of Applied Sciences Biomedical Technologies Application and Research Center (BİYOTAM) and Mucit Akademisi for providing workshop facilities and a supportive working environment during the experimental and applied phases of this study.

**Declarations:**

***1. Statement of Originality:***
This work is original.

***2. Author Contributions:***
***Concept:*** ED; ***Conceptualization:*** ED; ***Literature Search:*** ED; ***Data Collection:*** ED; ***Data Processing:*** ED; ***Analysis:*** ED; ***Writing – original draft:*** ED,ANAG; ***Writing – review & editing:*** ED,ANAG.

***3. Ethics approval:***
The study was approved by the Ethics Committee of Sakarya University of Applied Sciences on 15 January 2026 (Decision No: 64/42). Written informed consent was obtained from all participants prior to data acquisition.

***4. Funding/Support:***


CC: BY-NC-ND

This work has not received any funding or support.

***5. Competing Interests:***

The authors declare no competing interests.

***6. GenAI Usage Statement:***

AI tools were used only for language editing and improvement. The authors take full responsibility for the content.

***7. Sustainable Development Goals:***

**REFERENCES**

[1] Esposito, D., Centracchio, J., Andreozzi, E., Gargiulo, G. D., Naik, G. R., & Bifulco, P. (2021). Biosignal-based human–machine interfaces for assistance and rehabilitation: A survey. Sensors, 21(20), 6863. https://doi.org/10.3390/s21206863

[2] Kaur, A. (2021). Wheelchair control for disabled patients using EMG/EOG-based human–machine interface: A review. Journal of Medical Engineering & Technology, 45(1), 61–74. https://doi.org/10.1080/03091902.2020.1853838

[3] Kim, D. Y., Han, C. H., & Im, C. H. (2018). Development of an electrooculogram-based human–computer interface using involuntary eye movement by spatially rotating sound for communication of locked-in patients. Scientific Reports, 8(1), 9505. https://doi.org/10.1038/s41598-018-27865-5

[4] Sharma, K., Jain, N., & Pal, P. K. (2020). Detection of eye closing/opening from EOG and its application in robotic arm control. Biocybernetics and Biomedical Engineering, 40(1), 173–186. https://doi.org/10.1016/j.bbe.2019.10.004

[5] Tao, L., Huang, H., Chen, C., Feijs, L., Hu, J., & Chen, W. (2023). Review of electrooculography-based human–computer interaction: Recent technologies, challenges and future trends. Connected Health and Telemedicine, 2, 2000010. https://doi.org/10.20517/chatmed.2023.05

[6] Belkhiria, C., Droit, C., Hurter, C., & Fanti, M. P. (2022). EOG-based human–computer interface: 2000–2020 review. Sensors, 22(13), 4914. https://doi.org/10.3390/s22134914

[7] Zhang, J., Wang, B., Zhang, C., Xiao, Y., & Wang, M. Y. (2019). An EEG/EMG/EOG-based multimodal human–machine interface to real-time control of a soft robot hand. Frontiers in Neurorobotics, 13, 7. https://doi.org/10.3389/fnbot.2019.00007

[8] King, T., Knierim, M., Lepold, P., Clarke, C., Gellersen, H., Beigl, M., & Röddiger, T. (2025). EarEOG via periauricular electrodes to facilitate eye tracking in a natural headphone form factor. Scientific Reports, 15(1), 32437. https://doi.org/10.1038/s41598-025-16839-z

[9] Liu, P., Puthusserypady, S., MacKenzie, I. S., Uyanik, C., & Hansen, J. P. (2025). EarEOG: Using headphones and around-the-ear EOG signals for real-time wheelchair control. Proceedings of

the ACM on Human-Computer Interaction, 9(3), Article ETRA08, 1–16. https://doi.org/10.1145/3725833

[10] Huang, Q., Zhang, Z., Yu, T., He, S., & Li, Y. (2019). An EEG-/EOG-based hybrid brain–computer interface: Application on controlling an integrated wheelchair robotic arm system. Frontiers in Neuroscience, 13, 1243. https://doi.org/10.3389/fnins.2019.01243

[11] Chang, W.-D., Choi, J.-H., & Shin, J. (2021). Recognition of eye-written characters using deep neural networks. Applied Sciences, 11(22), 11036. https://doi.org/10.3390/app112211036

[12] Ogenga, J. P. A., Njeri, P. W., & Muguro, J. (2023). Development of a virtual environment-based electrooculogram control system for safe electric wheelchair mobility for individuals with severe physical disabilities. Journal of Robotics and Control, 4(2), 165–178. https://doi.org/10.18196/jrc.v4i2.17165

[13] Barbara, N., Camilleri, T. A., & Camilleri, K. P. (2020). A comparison of EOG baseline drift mitigation techniques. Biomedical Signal Processing and Control, 57, 101738. https://doi.org/10.1016/j.bspc.2019.101738

[14] Hernández-Pérez, S. N., Pérez-Reynoso, F. D., González-Gutiérrez, C. A., Cosío-León, M. D. Á., & Ortega-Palacios, R. (2023). EOG signal classification with wavelet and supervised learning algorithms KNN, SVM and DT. Sensors, 23(9), 4553. https://doi.org/10.3390/s23094553

[15] Barour, L., & Ay Gul, A. N. (2024). Electrooculography signal acquisition and processing for real-time virtual keyboard. Journal of Smart Systems Research, 5(2), 66–75. https://doi.org/10.58769/joinssr.1499482

[16] Atalı, G., Değirmenci, Ş., & Demiröz, E. (2023). HoReCa sektörü için doğal navigasyon otonom mobil servis robotu: komiBot. Journal of Smart Systems Research, 4(2), 127–139. https://doi.org/10.58769/joinssr.1387566

[17] Sanjaya, W. S. M., Anggraeni, D., Multajam, R., Subkhi, M. N., & Muttaqien, I. (2017). Design and experiment of electrooculogram (EOG) system and its application to control mobile robot. IOP Conference Series: Materials Science and Engineering, 180(1), 012072. https://doi.org/10.1088/1757-899X/180/1/012072

[18] Alfiyan, M., & Puriyanto, R. D. (2023). Mecanum 4 omni wheel directional robot design system using PID method. Journal of Fuzzy Systems and Control, 1(1), 6–13. https://doi.org/10.59247/jfsc.v1i1.27

[19] Göktaş, M. S., & Aras, S. (2022). Two channel EOG circuit design and implementation for human computer interface. In 30th Signal Processing and Communications Applications Conference (SIU) (pp. 1–4). https://doi.org/10.1109/SIU55565.2022.9864907

[20] Mannan, M. M. N., Kamran, M. A., Kang, S., & Jeong, M. Y. (2018). Effect of EOG signal filtering on the removal of ocular artifacts and EEG-based brain–computer interface: A comprehensive study. Complexity, 2018, 4853741. https://doi.org/10.1155/2018/4853741

[21] Barea, R., Boquete, L., Mazo, M., & López, E. (2002). System for assisted mobility using eye movements based on electrooculography. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 10(4), 209–218. https://doi.org/10.1109/TNSRE.2002.806829

[22] Niwa, K., Nikaido, K., & Sugi, M. (2024). Effect of system response time on brain activity and sense of agency. IEEE Access, 12, 126819–126830. https://doi.org/10.1109/ACCESS.2024.3445158

[23] Erkaymaz, H., Ozer, M., & Orak, I. M. (2015). Detection of directional eye movements based on the electrooculogram signals through an artificial neural network. Chaos, Solitons & Fractals, 77, 225–229. https://doi.org/10.1016/j.chaos.2015.05.033


CC: BY-NC-ND

[24] Papetti, A., Ciccarelli, M., Manni, A., Caroppo, A., & Rescio, G. (2025). Investigating the use of electrooculography sensors to detect stress during working activities. Sensors, 25(10), 3015. https://doi.org/10.3390/s25103015

[25] Varoquaux, G., Raamana, P. R., Engemann, D. A., Hoyos-Idrobo, A., Schwartz, Y., & Thirion, B. (2017). Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. NeuroImage, 145, 166–179. https://doi.org/10.1016/j.neuroimage.2016.10.038

[26] Mulam, H., Mudigonda, M., Kumar, B. P. S., & Kuchulakanti, H. (2023). Electrooculogram-based wheelchair control in real-time. In Proceedings of the Second International Conference on Emerging Trends in Engineering (ICETE 2023) (pp. 55–67). Atlantis Press. https://doi.org/10.2991/978-94-6463-252-1_8

[27] Zeng, Z., Tao, L., Hu, J., Su, R., Meng, L., Chen, C., & Chen, W. (2025). Multi-scale inception-based deep fusion network for electrooculogram-based eye movements classification. Biomedical Signal Processing and Control, 103, 107377. https://doi.org/10.1016/j.bspc.2024.107377

[28] Huang, Q., He, S., Wang, Q., Gu, Z., Peng, N., Li, K., & Li, Y. (2018). An EOG-based human–machine interface for wheelchair control. IEEE Transactions on Biomedical Engineering, 65(9), 2023–2032. https://doi.org/10.1109/TBME.2017.2732479

[29] Martínez-Cerveró, J., Khalili Ardali, M., Jaramillo-Gonzalez, A., Wu, S., Tonin, A., Birbaumer, N., & Chaudhary, U. (2020). Open software/hardware platform for human–computer interface based on electrooculography (EOG) signal classification. Sensors, 20(9), 2443. https://doi.org/10.3390/s20092443

[30] Pérez-Reynoso, F. D., Rodríguez-Guerrero, L., Salgado-Ramírez, J. C., & Ortega-Palacios, R. (2021). Human–machine interface: Multiclass classification by machine learning on 1D EOG signals for the control of an omnidirectional robot. Sensors, 21(17), 5882. https://doi.org/10.3390/s21175882

[31] Zendehdel, N., Ghorbani Zadeh, K., Chen, H., Song, Y. S., & Leu, M. C. (2025). Hands-free UAV control: Real-time eye movement detection using EOG and LSTM networks. IEEE Access, 13, 101852–101868. https://doi.org/10.1109/ACCESS.2025.3578558

[32] Wei, X., Dollack, F., Kiyokawa, K., & Perusquía-Hernández, M. (2025). Electro-oculography and proprioceptive calibration enable horizontal and vertical gaze estimation, even with eyes closed. Sensors, 25(21), 6754. https://doi.org/10.3390/s25216754

[33] Suhaimi, M. S. A., Matsushita, K., Kitamura, T., Laksono, P. W., & Sasaki, M. (2023). Object grasp control of a 3D robot arm by combining EOG gaze estimation and camera-based object recognition. Biomimetics, 8(2), 208. https://doi.org/10.3390/biomimetics8020208