Journal of Translational Medicine

RESEARCH

Open Access

Diagnosis of Parkinson's disease by eliciting trait-specific eye movements in multi-visual tasks

Maosong Jiang[1†], Yanzhi Liu[2†], Yanlu Cao[1], Shufeng Xia[1], Fei Teng[6], Wenzhi Zhao[3], Yongzhong Lin[2,5*] and Wenlong Liu[1,4*]

Abstract

Background Parkinson's Disease (PD) is a neurodegenerative disorder, and eye movement abnormalities are a significant symptom of its diagnosis. In this paper, we developed a multi-task driven by eye movement in a virtual reality (VR) environment to elicit PD-specific eye movement abnormalities. The abnormal features were subsequently modeled by using the proposed deep learning algorithm to achieve an auxiliary diagnosis of PD.

Methods We recruited 114 PD patients and 125 healthy controls and collected their eye-tracking data in a VR environment. Participants completed a series of specific VR tasks, including gaze stability, pro-saccades, anti-saccades, and smooth pursuit. After the tasks, eye movement features were extracted from the behaviors of fixations, saccades, and smooth pursuit to establish a PD diagnostic model.

Results The performance of the models was evaluated through cross-validation, revealing a recall of 97.65%, an accuracy of 92.73%, and a receiver operator characteristic area under the curve (ROC-AUC) of 97.08% for the proposed model.

Conclusion We extracted PD-specific eye movement features from the behaviors of fixations, saccades, and smooth pursuit in a VR environment to create a model with high accuracy and recall for PD diagnosis. Our method provides physicians with a new auxiliary tool to improve the prognosis and quality of life of PD patients.

Keywords Parkinson's disease, Eye movement abnormalities, Virtual reality, Eye tracking, Deep learning

†Maosong Jiang and Yanzhi Liu contributed equally to this work.

*Correspondence:
Yongzhong Lin
lin19671024@163.com
Wenlong Liu
liuwl@dlut.edu.cn
[1] School of Information and Communication Engineering, Dalian University of Technology, No. 2 Linggong Road, 116024 Dalian, China
[2] Department of Neurology, The Second Affiliated Hospital of Dalian Medical University, No. 467 Zhongshan Road, 116024 Dalian, China
[3] Department of Orthopedic Surgery, The Second Affiliated Hospital of Dalian Medical University, No. 467 Zhongshan Road, 116024 Dalian, China
[4] Key Laboratory of Artificial Intelligence, DUT Artificial Intelligence Institute, 56A-22 Floor, Huoju Road, 116024 Dalian, China
[5] The Dalian Key Laboratory of Intelligent Brain Science, No. 467 Zhongshan Road, 116024 Dalian, China
[6] Department of Geriatrics, Central Hospital of Dalian University of Technology, No. 826 Southwest Road, 116024 Dalian, China

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Introduction

Parkinson's Disease (PD) is a complex, multi-systemic neurodegenerative disorder and the second most common neurodegenerative disease after Alzheimer's disease [1]. Due to the close relationship between the human visual system and structures such as the basal ganglia and brainstem, abnormal eye movements are frequently observed in PD patients [2]. Some literature show that PD patients exhibit eye movement abnormalities, which can manifest in the early stages of the disease, sometimes preceding motor symptoms such as tremors, rigidity, motor impairments, and postural instability [3–5]. Therefore, eye movements play an essential role in the early diagnosis and treatment of PD.

Existing diagnostic methods for PD include routine clinical examination [6], neuroimaging techniques [7], dopamine transporter (DAT) scans [8], and genetic testing [9], but these methods still have limitations. Routine clinical examinations rely on physicians assessing the patient's medical history and performing neurological examinations [6]. However, symptoms typically appear in the mid-to-late stages of the disease, making early detection challenging. Neuroimaging techniques, such as magnetic resonance imaging (MRI), brain ultrasound, and positron emission tomography (PET) scans, are valuable tools for diagnosing neurological diseases [7]. Nevertheless, they are expensive, complex, and difficult to server as the long-term monitors for tracking and recording the PD progresses. DAT scans are useful for assessing the functional status of dopaminergic neurons [8], but they are costly and have limited effectiveness in distinguishing non-motor symptoms, such as eye movement abnormalities. Furthermore, genetic testing can identify genetic predispositions related to PD, such as specific gene mutations, but it is limited in early-stage or non-familial cases due to the lack of clear genetic markers in many patients [9]. These limitations make early diagnosis and timely treatment of PD challenging, highlighting the urgent need to explore new methods to complement the existing PD diagnostic process.

In this work, we focus on the importance of eye movements in the diagnosis of PD.Previous research has indicated that abnormal eye movement is a significant clinical symptom in PD patients, and is a sensitive and specific marker for PD patients and high-risk populations [10–13]. These abnormalities can effectively distinguish between PD patients and healthy individuals. Research indicates that eye movement abnormalities in PD patients are associated with the behaviors of saccades, fixations, and smooth pursuit, are closely related to the severity and progression of PD, and may appear even before the onset of motor symptoms. However, current studies have focused mainly on a certain aspect of eye movement behaviors. One aspect of research has investigated saccadic tasks in PD patients, reporting that spontaneous saccadic impairments and decreased saccadic amplitudes occur in the early stage of the disease. Specifically, symptoms of early PD patients mainly manifest as impaired saccades in complex saccadic movements, such as anti-saccades [14]. As the disease progresses, the severity of anti-saccade involvement increases, resulting in decreased saccade accuracy and prolonged saccade latency [15]. Another aspect of research on fixation tasks has reported that PD patients have shorter fixation durations and are more susceptible to stimulus-induced attentional distraction [16, 17]. A further aspect suggests that PD patients have significantly lower gain values for smooth pursuit eye movement (SPEM, i.e., the ratio of eye speed to target speed) than healthy individuals, indicating significant differences in target tracking [2].

Previous research has often used a single task to elicit partial behaviors of specific eye movement. We designed multiple visual tasks to elicit specific eye movements related to the behaviors of fixations, saccades, and smooth pursuit, which contributes to improving the diagnosis of patients with PD. However, the completion of these tasks can significantly extend testing time, which can lead to increased fatigue and reduced attention among subjects, ultimately compromising the distinctiveness of eye movement features. Some studies have reported that elderly individuals exhibit a decrease in saccade duration and an increase in fixation duration during prolonged visual tasks [18]. As eye fatigue increases, fixation durations also tend to increase [19]. Moreover, with prolonged task duration, smooth pursuit performance decreases, and fixation position errors increase [20]. PD patients also exhibit decreased saccade duration and increased fixation duration during visual tasks [21], along with a decrease in smooth pursuit gain [2].

To address the issues above, we introduce virtual reality (VR) technology for the clinical diagnosis of PD. By providing eye-tracking-driven tasks with various attractions and interests in a VR environment, we can achieve a high level of ecological validity [22], maintain participants' attention, and increase their participation. Additionally, VR headsets can effectively mitigate external interference, such as changes in lighting conditions and movements of people or objects [23], thereby reducing the influence of external interference in eye-tracking tests and providing more reliable medical data for effective disease diagnosis.

We develop a series of tasks driven by eye movements in a VR environment to differentiate between healthy individuals and PD patients. The main contributions of this work are as follows:

- We designed four VR tasks to elicit specific eye movement features related to the behaviors of fixations, saccades, and smooth pursuits in PD patients.
- In the VR environment, participants performed four visual tasks while synchronously recording their eye movement data and extracting specific eye movement features.
- We propose a deep learning model to analyze specific eye movement behaviors for the automated diagnosis of PD. The results demonstrate that compared to traditional learning algorithms, the proposed model is more effective in capturing the dynamic features of PD patients' visual behavior and provides better accuracy in differentiating between healthy individuals and PD patients.

Methods

The framework of the entire system is shown in Fig. 1. The main process was as follows: first, eye-tracking data were collected from PD patients and healthy individuals engaged in VR visual tasks (Stage 1). Then, eye movement features related to the behaviors of fixations, saccades, and smooth pursuit were extracted from the eye-tracking data (Stage 2). A dataset was subsequently constructed on the basis of these eye movement features and labels. Finally, on the basis of the dataset, we integrated multiple features into one model by using learning algorithms to establish the correspondence between eye movement features and the disease, namely the PD diagnostic classification model (Stage 3). The classification model made it possible to distinguish between healthy individuals and PD patients.

Participants and apparatus

All participants, including healthy control (HC) and PD groups, were recruited from the Department of

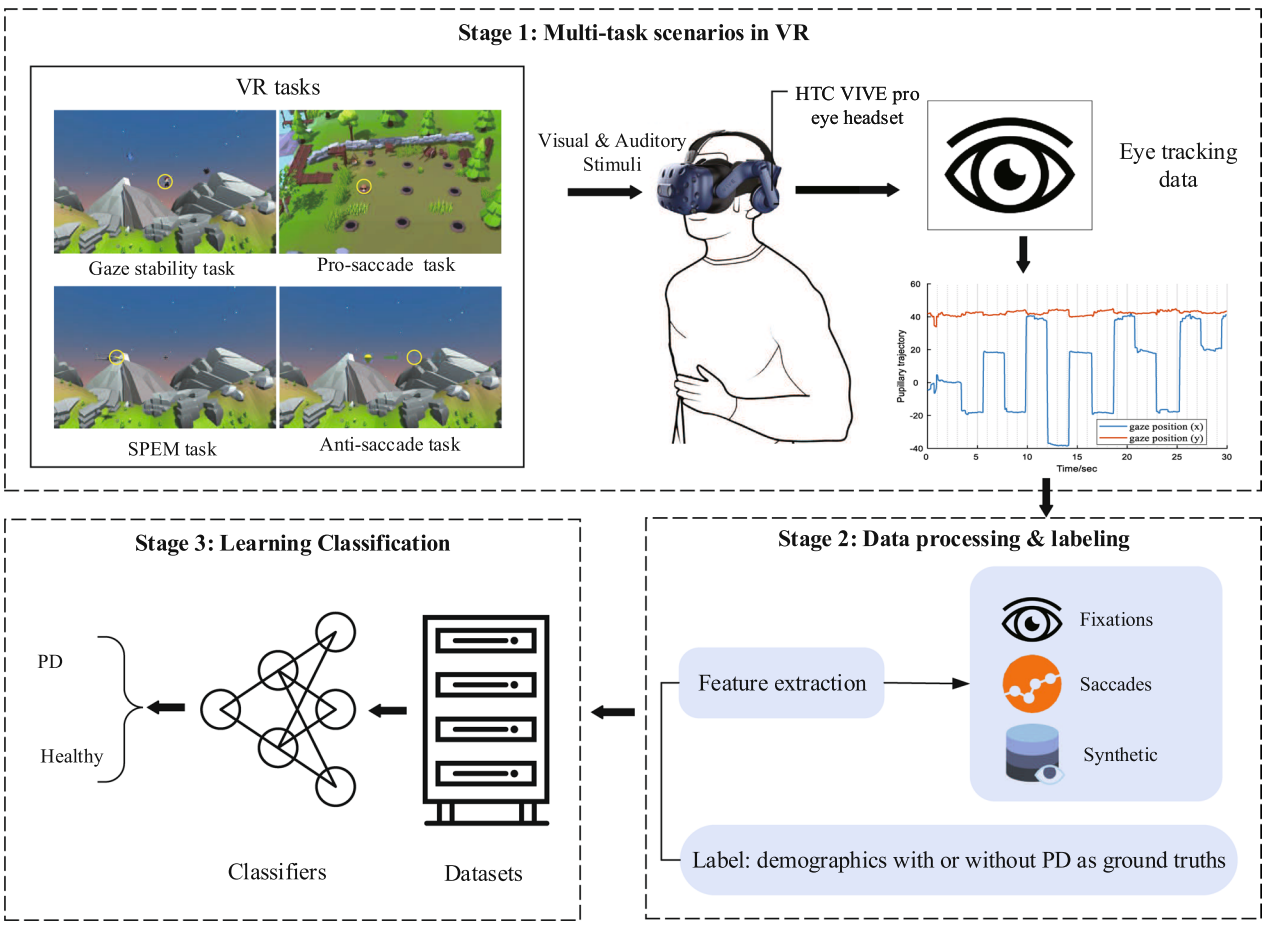

Fig. 1 Overall flowchart of the system for the automatic diagnosis of PD in a VR environment. The system was mainly divided into three steps, including multi-task scenarios, data processing and labeling, and learning classification. The yellow circle represents the location of the gaze

Neurology at the Second Affiliated Hospital of Dalian Medical University. The experimental procedures were reviewed and approved by the Biological and Medical Ethics Committee of the Second Affiliated Hospital of Dalian Medical University, ensuring compliance with the principles of the Helsinki Declaration. Participants in the control group had no history of neurodegenerative diseases or other neurological disorders and had normal or corrected-to-normal vision. Each participant in the study received an explanation of the informed consent form and provided their signature to confirm their consent. Each participant in the PD group underwent a motor examination via the motor section of the Unified PD Rating Scale (UPDRS III) [24]. Demographic and clinical information was recorded for the subjects in the PD and HC groups, as shown in Table 1.

In the experiment, we employed a generic Helmet Mounted Display (HMD) device (HTC Vive Pro Eye). The virtual environment was created and rendered using the Unity engine (version 2018.3.17f1). The device was connected to a computer (Intel Core i5-8400, 6 cores @ 2.8GHz, GeForce GTX 1060 6GB) and performed four visual tasks while monitoring the subjects' eye movements.

Multi-task scenario in VR: task scenario design

For the multi-task visual-searching scenario of VR, we designed four tasks to elicit PD-specific eye movement, as shown in Fig. 2A, B, C, and D. The four tasks include gaze stability, pro-saccade, anti-saccade, and smooth pursuit. The differences in eye movement behavior between PD patients and healthy individuals are described below.

Gaze stability task (Fig. 2A): At the beginning of the task, the participants were given instructions to focus their gaze on a central screen that displayed a cartoon character, as shown in Fig. 2A. During the task, they had to maintain their gaze on the target without being distracted by the sudden appearance of objects on the periphery. These distractors appeared randomly, and the participants did not know their location or timing. Compared with PD patients, HC individuals exhibit relatively good gaze stability, keeping the eyes relatively fixed when gazing at a static target. In contrast, PD patients show poor gaze stability, characterized by eye tremors or irregular movements, and are more susceptible to the influence of changes in the surrounding environment.

Table 1 Demographic and clinical characteristics of the PD and HC groups

Subjects	HC (n = 125)	PD (n = 114)	p-value
Sex (% Male)	59 (47.46%)	52 (45.54%)	0.845
Age	63.64 ± 6.9	66.95 ± 8.84	0.101
Education level (year)	9.30 ± 3.20	9.39 ± 3.13	0.742
PD duration	–	1.66 ± 1.98	–
UPDRS III	–	10.42 ± 13.49	–

Reported values are mean ± standard deviation

Pro-saccade task (Fig. 2B): This task was presented like a whack-a-mole game. In this task, participants were required to gaze at a target, depicted as a mole, as shown in Fig. 2B. When a target appeared, participants maintained their gaze on it for 300 milliseconds, after which the target disappeared. They waited for the next target to appear (which was different from the previous one). Compared with PD patients, HC individuals can perform rapid and accurate saccades, such as quickly shifting their gaze from one target to another, as seen in activities like a whack-a-mole game. In contrast, PD patients demonstrate slowed or irregular saccades, reducing saccade speed and accuracy.

Anti-saccade task (Fig. 2C): In this task, the participants were instructed to avoid focusing their gaze on the target, represented by yellow polyhedrons, as shown in Fig. 2C. Instead, they were required to look at the opposite side where the target appeared, essentially the mirror-image position of the target. The targets were randomly placed in four different locations. Each target remained visible for approximately 2 to 3 s before reappearing in a new location. The participants were expected to track the target in the opposite direction. Compared with PD patients, HC individuals typically exhibit accurate targeting of the mirrored position of the target. However, PD patients display impairments in backward saccades, including slowed speed, decreased accuracy, and erroneous anti-saccades.

Smooth pursuit task (Fig. 2D): At the beginning of the task, the participants focused on a black "cross" at the center of the screen. After two seconds, the black "cross" disappears. Participants were then asked to maintain their gaze on a designated target object, represented as a 3D helicopter, as shown in Fig. 2D. This target moved in a sinusoidal pattern at a constant velocity of either $20°/s$ or $30°/s$ along the horizontal axis, covering a range from -20 degrees to $+20$ degrees relative to the head's center. During the smooth pursuit trial, the participants were instructed to ensure that their eye position was aligned with the center of the screen. Compared with PD patients, HC individuals exhibit proficient smooth pursuit ability, accurately tracking and maintaining gaze on a moving target. In contrast, PD patients show reduced accuracy, decreased tracking speed, and instability in smooth pursuit, potentially resulting in difficulties when processing targets in dynamic environments.

Figure 2 shows an example of real eye-tracking data for one trial in the HC and PD groups. Differences in eye-tracking trajectories exist between healthy

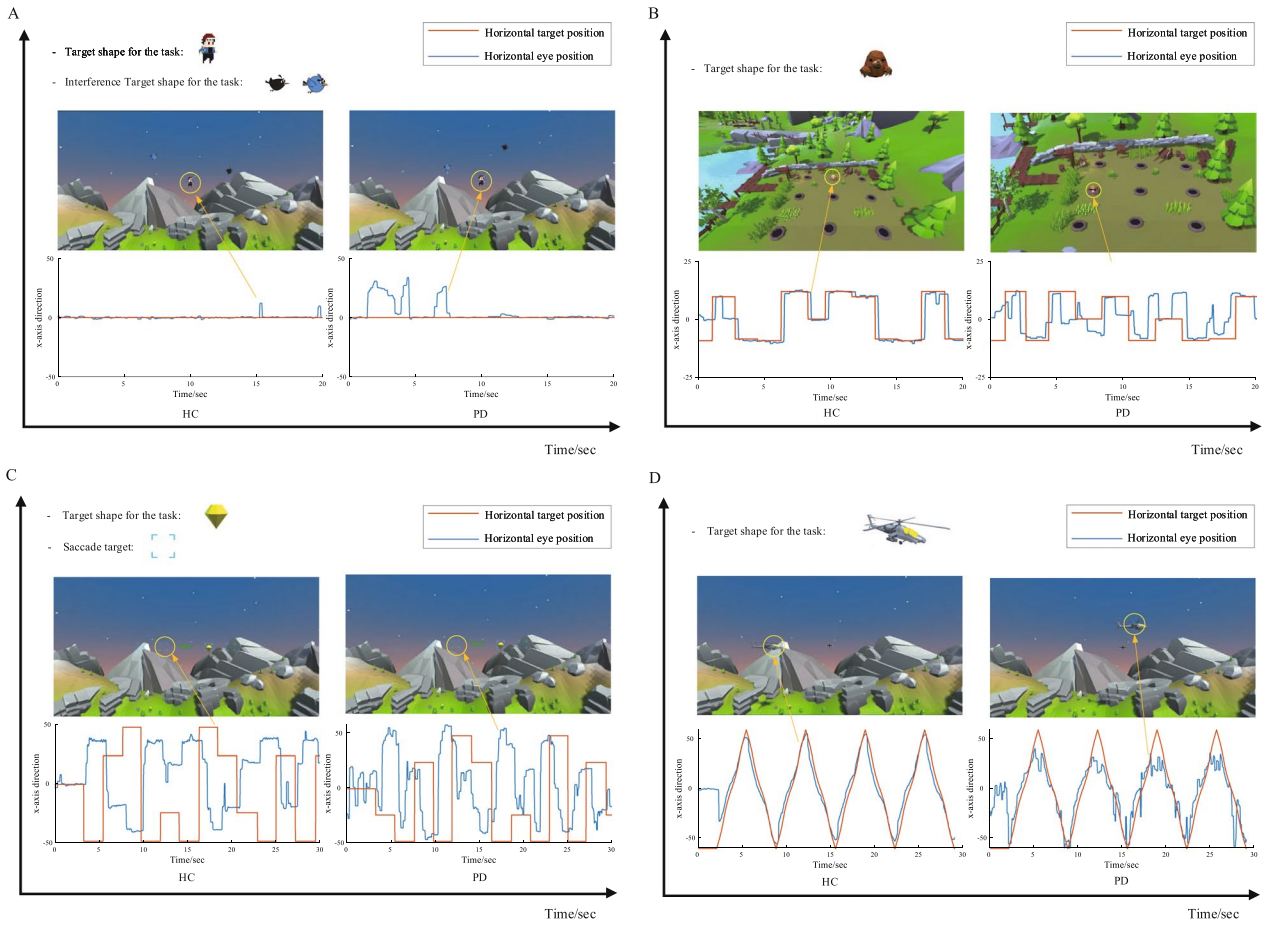

Fig. 2 Multi-task scenario in VR: task scenario design. The horizontal trajectories of the target and eye movements for the HC and PD groups, along with the experimental task scenarios. The **A** diagram represents the eye gaze with the constant position. During periods of continuous visual stimulation, the **B** diagram represents directional saccades, the **C** diagram represents oppositely directed saccades, and the D diagram represents smooth pursuit. The yellow circle represents the location of the gaze

individuals and PD patients when performing visual tasks. These results highlight specific differences in eye movement behavior, providing important insights for further understanding and diagnosing the eye-tracking features of PD.

Multi-task scenario in VR: experimental procedure

During the experiment, participants were tasked with completing the visual operations in VR environment. Participants were instructed to sit in a specific position and wear the VR headset with eye-tracking technology to record their eye movement trajectories. Before the experiment, operators calibrated each participant's headset, focal distance, and eye tracking. Participants were instructed to keep their heads as still as possible to minimize the impact of head movements on the experimental results.

Multi-task scenario in VR: data collection

Data collection was a significant step in the experiment, recording participants' eye movements as they completed tasks. Eye-tracking data (see Fig. 1, Multi-task scenarios in VR), including direction, velocity, and gaze position, were recorded in real-time during the execution of visual tasks. After the experiment ended, the eye movement data were stored on a computer and subjected to data organization and cleaning to remove any anomalies, thus ensuring data reliability and accuracy. These data serve as the foundation for subsequent data analysis and interpretation of the results.

Data was cleaned for quality. When conducting intergroup comparisons, it is crucial to maintain uniformity in eye-tracking quality across different clinical sites or testing operators. Therefore, data points with an average calibration error of 5 degrees or more were excluded, resulting in the exclusion of 12 PD patients and 7 healthy individuals. Independent t-tests were utilized to compare

the calibration accuracy measurements between groups. The results indicated no significant differences in calibration accuracy between PD patients and the HC group (t-test: average calibration accuracy, $p > 0.05$). This step was taken to ensure fairness in comparisons and the reliability of the data.

Extraction of eye movement features: processing of eye movement data

Raw eye-tracking data were obtained by recording eye movements via a built-in camera while completing multi-visual tasks in the VR environment. After that, we conducted data preprocessing operations in MATLAB (version 2021a, The MathWorks, Inc.), accurately identifying the behaviors of saccades, fixations, and smooth pursuit while ensuring data quality. We mainly detect fixations and saccades from raw eye-tracking data. A fixation is a period of focused attention on a single location, whereas a saccade is a quick jump between two fixations [25]. In this process, we also apply median filtering to the raw eye-tracking data. Median filtering helps reduce noise and smooth coordinate data, ensuring the accuracy of subsequent feature detection. The fixation-saccade detection algorithm is based on the Buscher method [26, 27]. Figure 3 shows the trajectories of participants' horizontal eye movements in both the HC and PD groups during the visual tasks, alongside the horizontal movement of targets. The behaviors of fixations, saccades, and smooth pursuit were identified within the eye-tracking trajectories.

Extraction of eye movement features: calculation of eye movement features

The key to establishing the PD diagnostic model is to adequately describe the dynamic eye neural network output as a set of numerical features (see Fig. 1, Data processing & labeling). We categorized eye movement behaviors into three types: saccades, fixations, and smooth pursuit, and described the features of these behaviors.

Fixation category [28]: Total Duration of Fixations (TDF), Total Number of Fixations (TNF), Mean Durations of Fixations (MDF). The recognition of fixations is based on the Buscher method [26, 27]. This method requires consecutive data points to be gathered in a limited area within the shortest time possible to be considered a fixation. A fixation is identified when the original data sample stays within a 1° diameter for at least 100 ms.

Saccade category: Mean Saccade Latency (MSLT) [29], saccade Error Rate (ER) [12], Scan path Length (SL), Scan path Duration (SD) [30], and Rate of Saccade/Fixation (Rsf) [31].

- The MSLT represents the average time interval from the onset of the target to the eyes' gaze on the target.

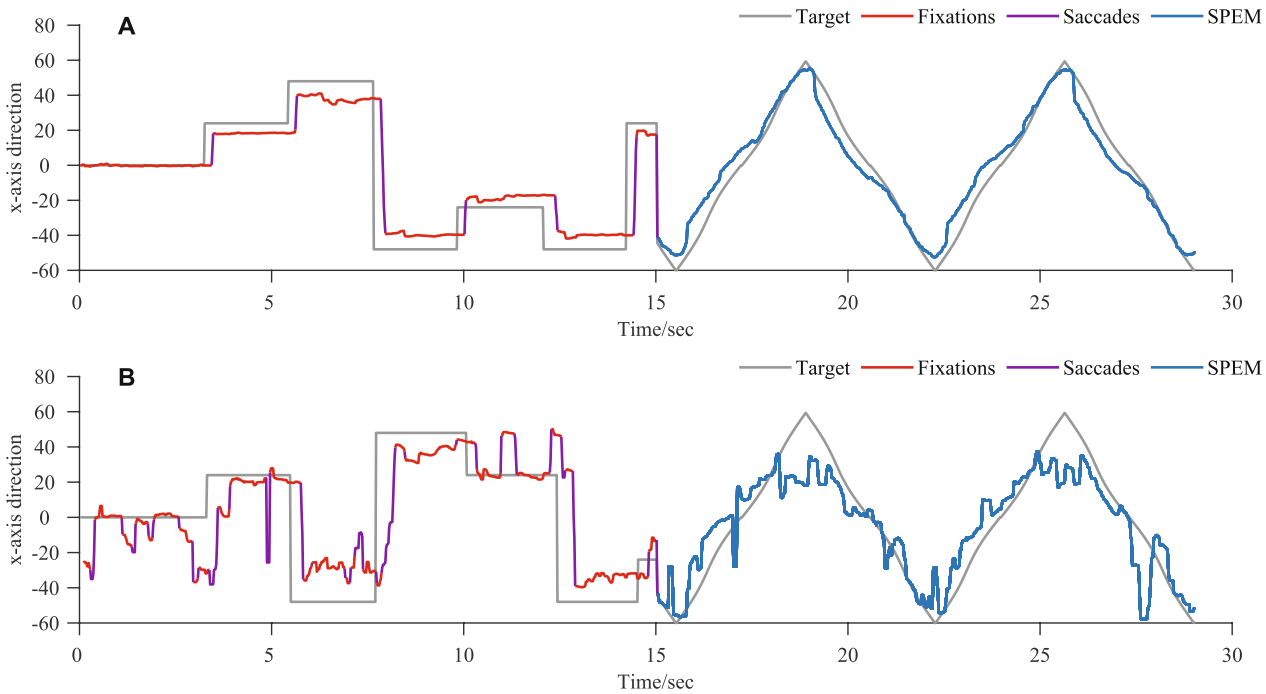

Fig. 3 Extraction of eye movement features: processing of eye movement data. The horizontal trajectories of the target and eye tracking, where (**A**) and (**B**) correspond to the HC and PD groups, respectively. The red lines represent fixations, the purple lines represent saccades, and the blue lines represent smooth pursuit movement. The gray lines represent the trajectory of the target movement

Let us denote the total number of target appearances in a visual task as N_{total}. For the i-th target event, let T_i denote the time of target appearance and T_i' denote the time when the eyes first gaze at it. The interval of time between them is $T_i' - T_i\left(T_i < T_i'\right)$, then the MSLT can be described via the following mathematical formula:

$$T_{MSLT} = \frac{1}{N_{total}} \sum_{i=1}^{N} \left(T_i' - T_i\right) \tag{1}$$

- The ER indicates the frequency at which the eye fails to gaze at the target when it appears successfully. N_{error} is the number of instances where the eyes fail to gaze at the target. The ER is represented as

$$Q_{ER} = N_{error}/N_{total} \tag{2}$$

- The SL is the trajectory obtained by connecting fixation points in any saccade sequence with straight lines, describing the static coverage pattern of the eyes when observing a stimulus [32]. This metric reflects the trajectory of eye movement on the stimulus surface, providing information on how visual attention is distributed on the stimulus. N is denoted as the number of fixation points generated during the task completion. The coordinates of the i-th fixation point are denoted as (x_i, y_i) $(i = 1, 2,..., N)$. The distance between the i-th fixation point and the $(i + 1)$-th fixation point can be expressed as

$$s_i = \sqrt{(x_{i+1} - x_i)^2 + \left(y_{i+1} - y_i\right)^2} \tag{3}$$

The SL can be calculated by summing the distances between all adjacent fixation points, expressed as follows:

$$S_{SL} = \sum_{i=1}^{N-1} s_i \tag{4}$$

- The SD represents the time it takes for the eyes to complete the total scan path. The time difference between the i-th fixation point at time T_i and the $(i + 1)$-th fixation point at time T_{i+1} is denoted as $(T_{i+1} - T_i)$. Thus, the SD can be expressed as:

$$T_{SD} = \sum_{i=1}^{N-1} \left(T_{i+1} - T_i\right) \tag{5}$$

- The Rsf represents the proportion of time allocated between saccades and fixations, indicating the distribution of eye movement patterns. Let the SD be denoted as T_{SD} and the TDF as T_{TDF}, and then the Rsf can be expressed as

$$Q_{Rsf} = T_{SD}/T_{TDF} \tag{6}$$

Smooth pursuit category: Smooth Pursuit Gain (SG) [33]. The SG value is the ratio of eye speed to target speed, used to measure the efficiency of eye movements for tracking moving targets. In the smooth pursuit task, when the target is moving at velocity $V_{target}(t)$, and the eye tracking velocity is $V_{eye}(t)$, the SG can be expressed as:

$$G_{SG} = E\left\{V_{eye}(t)/V_{target}(t)\right\} \tag{7}$$

where, $E\{\cdot\}$ represents the mean value.

Model establishment: establishment of a classification model

In this paper, we propose a hybrid model to build a classifier with robust generalization performance, which combines Inception v3 [34], ECA (Efficient Channel Attention) mechanism [35], and a convolutional neural network (CNN), called Inception-ECA-CNN. The main structure of the hybrid model includes an Inception v3 module, an ECA module, a CNN module, and a fully connected module, as shown in Fig. 4.

Firstly, we utilized the Inception v3 module to extract multi-scale features using one-dimensional convolution with different kernel sizes. The parallel convolution operations increase the network's width and enhance the richness of feature representation. The convolution kernel sizes are set to 1×1 or 1×3.

Secondly, the output of the Inception v3 module is fed into an ECA mechanism to enhance attention weights on feature dimensions. The ECA module adjusts feature weights through Global Average Pooling (CAP) and one-dimensional convolution operations, strengthening the representation of important features and suppressing irrelevant ones, thereby improving the model's focus on critical features.

Thirdly, the features adjusted by the ECA module are further processed using one-dimensional convolution and pooling layers to extract more advanced feature representations. The convolution kernel sizes are set to 1×3. Finally, the extracted features are flattened and fed into a fully connected layer for binary classification. The fully connected layer learns the mapping between the final features and the categories, and outputs binary classification probabilities using a sigmoid activation function to complete the entire classification task.

In summary, we established an efficient one dimensional binary classification model by combining the multi-scale feature extraction capability of the Inception v3 module, the feature weighting capability of the ECA module, and the feature extraction and classification capabilities of CNN. The design of the proposed model leverages the advantages of each module, facilitating

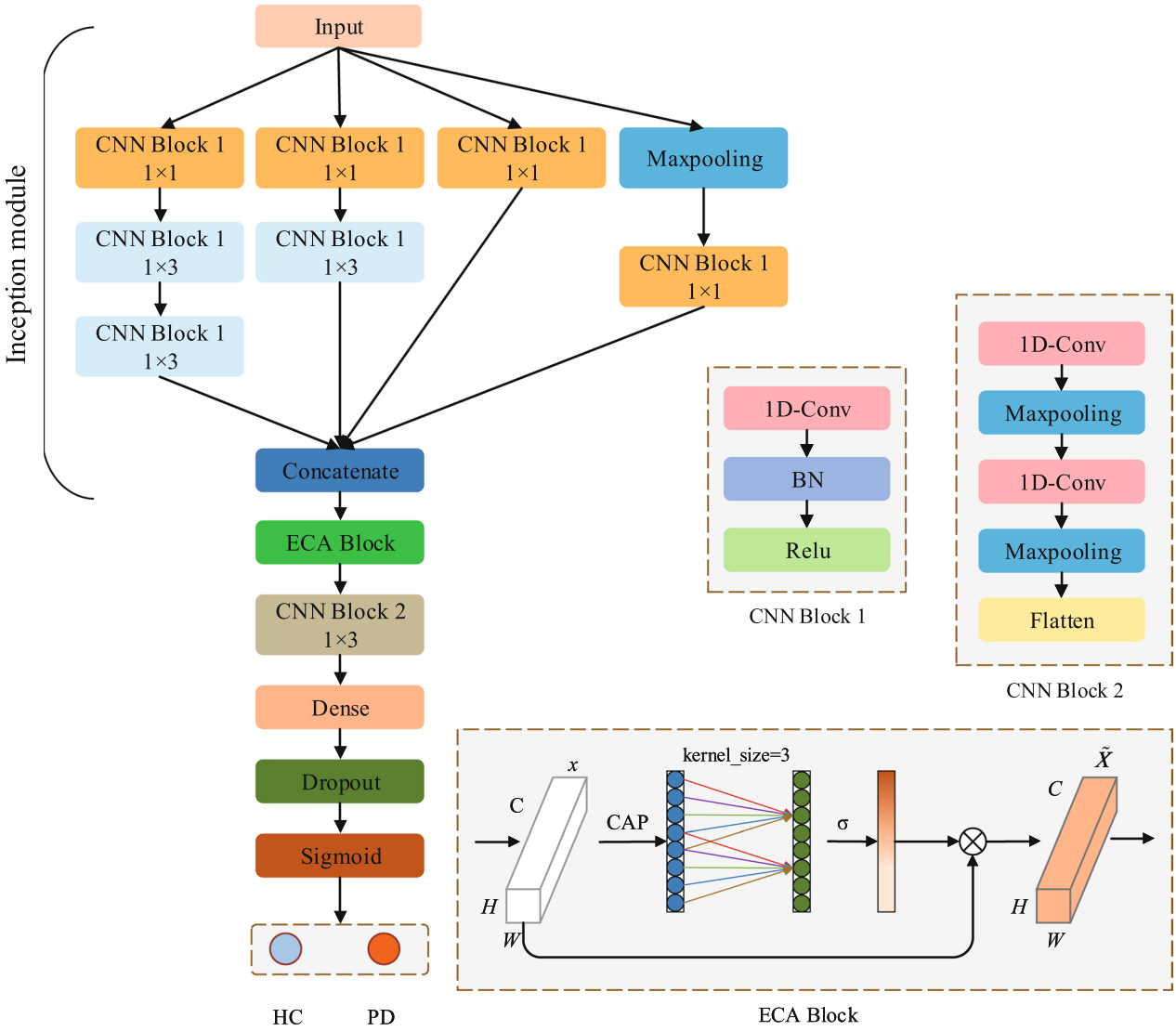

Fig. 4 Model establishment: establishment of a classification model. The structure of the Inception-ECA-CNN model

in-depth exploration of important features and thereby improving classification accuracy.

Model establishment: model training

The classification models were implemented in Python using the sklearn [36], Tensorflow, and Keras [37] libraries. We focused here on building a classifier that can detect PD patients effectively. A 5-fold nested cross-validation method [38] was employed on the training set to evaluate the performance of the classifier with the repeated creation of validation samples. The training set was partitioned into five disjoint subsets of approximately equal size. The model was then trained on four subsets, while the remaining subset was used as the validation set for performance evaluation. This procedure was repeated five times, with each subset serving as the validation set once. The average of the five performance measurements on the validation sets provided the cross-validated performance metric. Additionally, another independent cohort (15 PD patients and 11 healthy controls) was used as a validation set in each fold to evaluate the performance of the classifier [39–41].

Model establishment: performance evaluation of the classifiers

We compared the proposed model with various machine learning and deep learning algorithms during the evaluation process. Specifically, we compared the following machine learning algorithms: Logistic Regression (LR), Linear Discriminant Analysis (LDA), K-Nearest Neighbors

(KNN), Decision Trees (DT), Neural Networks (NN), Naive Bayes (NB), Gradient Boosting (GB), Random Forests (RF), and Support Vector Machines (SVM). Additionally, we also considered deep learning algorithms, including CNN, Inception and Inception-ECA. By comparing the performance of these algorithms, we can assess the advantages and effectiveness of the proposed hybrid model in PD diagnosis.

Evaluation metrics

In this paper, we introduced several evaluation metrics [42], including accuracy, precision, recall, and f1-score, to evaluate the performance of the classifier. These metrics are calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TN}{TP + FN} \tag{10}$$

$$F1{-}score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{11}$$

where TP, FP, FN, and TN are the calculated true positives, false positives, false negatives, and true negatives, respectively. We take PDs as the positive samples and HCs as the negative samples. In addition, we derived the receiver operating characteristic (ROC) curves and area under the curve (AUC) values to show specific details for the experimental results.

Experimental results

Statistical analysis of eye movement features

Independent sample t-tests were performed on the eye movement features of all participants, revealing significant differences between the PD group and the HC group under different visual tasks, including gaze stability, pro-saccades, anti-saccades, and smooth pursuit, as shown in Fig. 5.

In the gaze stability task (see Fig. 5A), MDF, TDF, and SD in the PD group significantly decreased, while TNF, SL, and Rsf significantly increased. In the pro-saccades task (see Fig. 5B), the PD group exhibited significant differences in TNF, TDF, SL, SD, Rsf, ER, and MLTP. Specifically, TNF, TDF, SL, and SD decreased in the PD

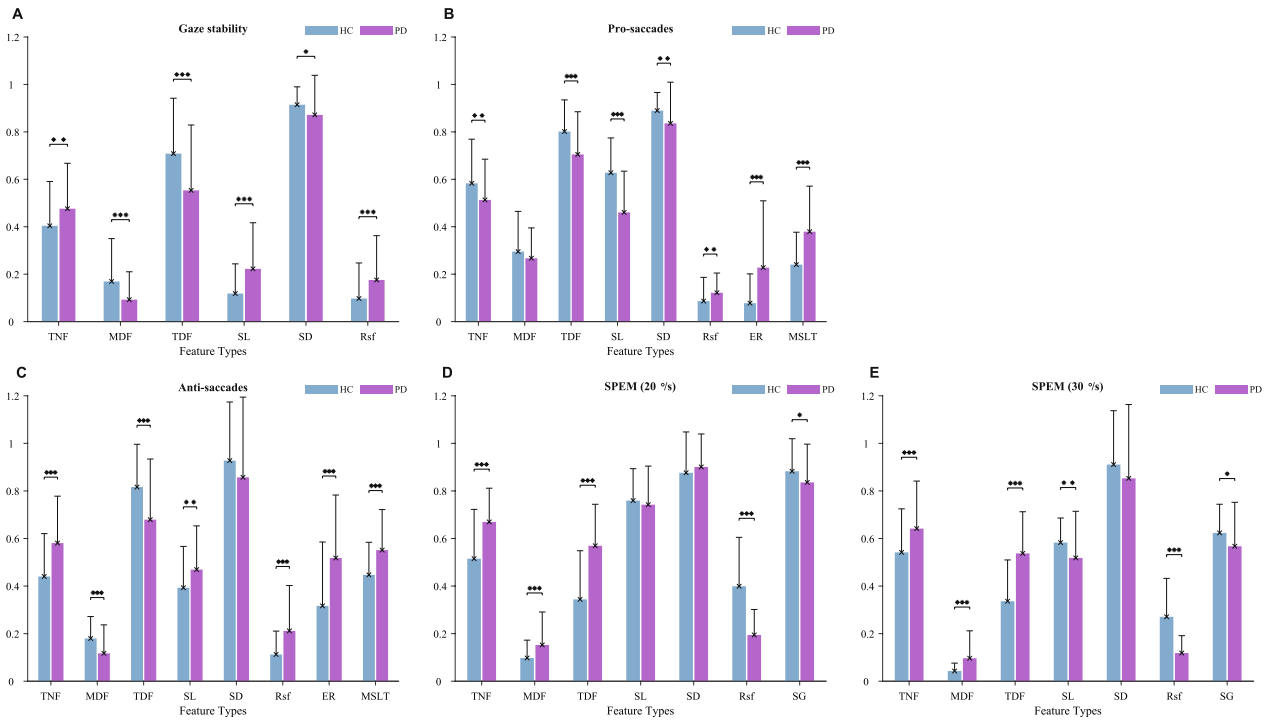

Fig. 5 Statistical analysis of eye movement features. The significant relationship of eye movement features in various tasks. We designed four typical eye movement tasks: **A**—gaze stability task, **B**—pro-saccades task, **C**—anti-saccades task, and SPEM task (with target speeds of 20°/s and 30°/s, denoted as **D** and **E**, respectively). Compared with the HC group, differences between the groups were marked with an asterisk (∗), with significance levels of $^{*}p < 0.05$,$^{**}p < 0.01$,$^{***}p < 0.001$, and a confidence level of 95%

group, while Rsf, ER, and MLTP increased. Additionally, the results of the anti-saccade task (see Fig. 5C) revealed that MDF and TDF decreased, SL, Rsf, ER, and MLTP increased in the PD group, while TNF significantly increased and was different from the pro-saccades. In the smooth pursuit task (see Fig. 5D), when the target speed was 20o/s, the PD group exhibited differential performance in terms of TNF, MDF, TDF, Rsf, and SG. Specifically, TNF, MDF, and TDF increased in the PD group, while Rsf and SG decreased. When the target speed was 30o/s (see Fig. 5E), SL, Rsf, and SG decreased significantly in the PD group, while TNF, MDF, and TDF increased significantly. In summary, the results of these tasks suggested that the HC and PD groups may have different characteristics and mechanisms in gaze stability, pro-saccades, anti-saccades, and smooth pursuit tasks.

Detection and prediction

As shown in Fig. 5, there were differences in eye movement features between the HC group and the PD group. We separately employed machine learning and deep learning algorithms to establish a PD diagnostic model by integrating multiple differential eye movement features. We compared nine machine learning algorithms (KNN, SVM, RF, LR, LD, DT, NN, NB, and GB), and three deep learning algorithms (CNN, Inception, and Inception-ECA). Finally, we used sample data to evaluate the performance of the classification model, thereby verifying its effectiveness and accuracy.

Table 2 shows the accuracy of each model in diagnosing PD. The results showed that the model established by the proposed model has higher accuracy, with an accuracy rate of 92.73%, which is better than those of the other methods. To further validate the performance of the model, the accuracy results on the independent set were presented in Table.3. In addition, ROC curves and AUC values were employed to evaluate the performance of various models. As shown in Fig. 6, the proposed model quickly reaches optimal performance during early training, consistently outperforming other methods across the entire range. The AUC value of 97.08% further confirms the excellent performance of the proposed model in PD diagnosis. These results imply that the proposed model demonstrates superior accuracy in distinguishing PD patients from healthy individuals, making it a highly reliable method.

Discussion

In this work, we identified PD-specific eye movement features by analyzing the significant differences between the eye movement behaviors of PD patients and healthy individuals. In the VR scenario, we incorporated four

Table 2 Experimental results of different classifiers for PD patients and healthy controls

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	AUC (%)
KNN	81.28 ± 3.01	83.42 ± 2.67	74.33 ± 7.64	78.37 ± 4.29	87.90 ± 3.23
SVM	78.96 ± 4.12	77.82 ± 6.18	77.29 ± 7.23	77.20 ± 4.04	88.43 ± 2.31
RF	78.99 ± 3.96	78.22 ± 5.17	76.38 ± 10.12	76.76 ± 5.17	86.28 ± 2.55
LR	77.62 ± 4.42	75.99 ± 2.30	75.43 ± 10.66	75.34 ± 6.05	88.78 ± 2.33
LD	76.70 ± 5.31	77.05 ± 4.68	71.52 ± 13.71	73.26 ± 8.01	87.03 ± 4.21
DT	74.85 ± 6.27	73.76 ± 7.08	71.33 ± 7.10	72.39 ± 6.44	76.58 ± 8.20
NN	81.30 ± 4.09	80.87 ± 4.64	78.33 ± 9.37	79.24 ± 5.09	85.30 ± 3.17
NB	78.51 ± 6.16	78.68 ± 8.12	74.33 ± 6.96	76.23 ± 6.32	86.10 ± 2.96
GB	79.42 ± 6.41	78.70 ± 6.44	76.33 ± 9.85	77.28 ± 6.98	86.94 ± 3.90
CNN	89.09 ± 3.91	87.17 ± 2.40	84.71 ± 14.22	85.06 ± 7.17	95.29 ± 4.32
Inception	90.91 ± 1.44	81.68 ± 2.75	**98.82 ± 2.35**	89.38 ± 1.51	96.34 ± 0.37
Inception-ECA	90.91 ± 1.44	82.22 ± 1.99	97.65 ± 2.88	89.24 ± 1.72	96.08 ± 0.79
Our	**92.73 ± 1.70**	**85.70 ± 3.15**	97.65 ± 2.88	**91.22 ± 1.97**	**97.08 ± 0.88**

The best performance is indicated in bold

Table 3 Experimental results of using different classifiers to classify PD patients and healthy controls in the validation set

	Accuracy (%) on different methods, mean and standard deviation (Std)												
Methods	KNN	SVM	RF	LR	LD	DT	NN	NB	GB	CNN	Inception	Inception-ECA	Our
Mean	79.23	83.85	84.62	83.08	75.38	77.69	86.92	83.08	88.00	89.09	90.00	90.00	92.31
Std	1.88	1.54	4.87	3.92	7.13	10.43	4.62	1.88	5.22	1.88	5.22	3.92	1.88

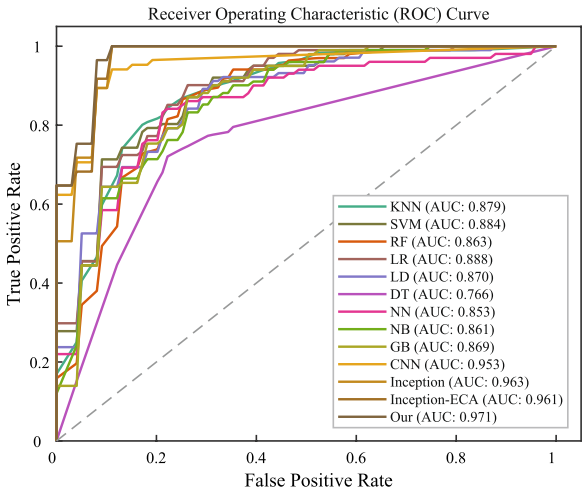

Fig. 6 Detection and prediction. ROC curves and AUC values of machine learning and deep learning classifiers

typical eye movement tasks: gaze stability, pro-saccades, anti-saccades, and smooth pursuit. Additionally, eye movement features were extracted from the behaviors of fixations, saccades, and smooth pursuit. These tasks provided valuable insights into the neural mechanisms underlying eye movement control. By focusing on the significant differences in eye movement features between the two groups, the model was able to capture the underlying patterns of eye movement, improving its accuracy in distinguishing PD patients from healthy individuals.

Relationships between abnormal eye movement and clinical symptoms in PD patients

In the gaze stability task (see Fig. 5A), PD patients exhibited significantly lower MDF and TDF, indicating a reduced ability to maintain stable gaze. Concurrently, increased TNF and Rsf suggest the activation of compensatory mechanisms to mitigate oculomotor instability. These changes are likely linked to basal ganglia dysfunction, which hampers precise eye movement control and disrupts visual stability [43]. These findings underscore the critical importance of gaze stability in preserving visual coherence for individuals with PD.

In the pro-saccade and anti-saccade tasks (see Fig. 5B and C), the PD group demonstrated decreased MDF and TDF alongside increased Rsf, ER, and MLTP. These results reflect challenges in initiating and executing accurate saccades, with anti-saccades showing more pronounced deficits. The greater difficulty with anti-saccades suggests a specific vulnerability in tasks requiring inhibitory control and voluntary redirection of gaze, processes reliant on the cortical-basal ganglia-superior colliculus pathway and prefrontal dopamine regulation [44]. These results underscore the progressive cognitive challenges faced by PD patients in complex visual tasks.

In the smooth pursuit task (see Fig. 5D and E), the results indicated that the PD group had significantly lower SG values ($p < 0.05$), suggesting difficulty in estimating the speed and direction of moving targets. This impairment is likely due to dysfunction in the cortical visual areas and disrupted pathways between the pontine nucleus and cerebellum [45]. The decrease in SPEM gain may result from impaired coordination among multiple brain regions. Specifically, the caudate nucleus receives the fiber region of the frontal eye field and the feedback loop between the basal ganglia and thalamus, and pathophysiological changes of these structures may lead to decreased SPEM gain in PD patients [46]. Additionally, the similarity between the functional structure of the SPEM task and the saccadic eye movement system further supports the notion that these structural impairments contribute to the decreased SPEM gain in PD patients.

In summary, these findings highlight the diverse impact of PD on oculomotor tasks. Anti-saccades showed more pronounced deficits than pro-saccades, suggesting that the additional cognitive demands of suppressing reflexive responses contribute to the impairment. Similarly, the reduced SG during smooth pursuit indicates difficulties in coordination across interconnected neural pathways. These task-specific abnormalities offer valuable insights into the neuropathological mechanisms of PD and emphasize the potential of specific eye movement features as biomarkers for early diagnosis, disease monitoring, and therapeutic intervention.

Performance analysis of learning algorithms

In this paper, we aimed to distinguish between PD patients and healthy individuals by assessing their visual behavioral characteristics in VR scenarios. We compared the proposed algorithm with machine learning and deep learning algorithms to assess their respective classification accuracies. When selecting a suitable diagnostic model, we prioritized having a high recall rate before focusing on classification accuracy. In this way, PD patients can be distinguished from healthy individuals as accurately as possible, minimizing the risk of misdiagnosis. As shown in Table 2, we found that the proposed model has higher classification accuracy than the other algorithms, with an accuracy rate of 92.73%, and a high-level ROC-AUC score of 97.08% (see Fig. 6). Additionally, the classification performance shown in Table 3 further validates that the proposed model consistently outperforms other methods on the independent set. These results indicate that the proposed model can more effectively diagnose and identify PD patients, even though

some healthy people may be misdiagnosed as PD patients without missing actual patients and delaying their treatment. Therefore, the proposed model is a promising option to achieve accurate prediction of PD by maintaining a high accuracy rate.

Although our research demonstrated the potential of using eye movement abnormalities in a VR environment to assist in the accurate diagnosis of PD, it still has some limitations. The study was based on cross-sectional data. However, this approach may not capture the evolution of eye movement features as the disease progresses. Therefore, future research should focus on collecting longitudinal data, and tracking and recording changes in eye movement patterns, aiming to get a deeper understanding of how the patterns evolve as the disease progresses. Furthermore, our research approach may also provide potential value in the assessment of other neurodegenerative diseases (i.e. Alzheimer's disease), and brain health. This indicates that subsequent studies could explore the application of eye movement abnormalities in these diseases, thereby providing more comprehensive tools for early screening, monitoring disease progression, and evaluating the effectiveness of therapeutic interventions.

Conclusion

We utilized eye movement features of visual behavior in a VR environment for predicting PD. We designed four basic visual tasks to obtain eye movement features and model them using the proposed model to establish a diagnostic model for PD. Compared with binary classification provided by machine learning and deep learning algorithms, the proposed model can accurately predict PD with an accuracy rate of 92.73%, and a high-level ROC-AUC score of 97.08%, eliminating subjectivity. This achievement is of great significance to neurosurgeons, as it not only simplifies the diagnostic process but also provides the possibility of further personalized treatment. By utilizing the established PD model, physicians can accurately determine whether a patient has PD based on eye movement data, enabling them to devise more effective treatment plans and better address the needs of patients.

Abbreviations

AUC	Area under the curve
CAP	Global average pooling
CNN	Convolutional neural networks
DT	Decision trees
ECA	Efficient channel attention
ER	Error rate
GB	Gradient Boosting
HC	Healthy control
KNN	K-nearest neighbors
LDA	Linear discriminant analysis
LR	Logistic regression
MDF	Mean durations of fixations
MSLT	Mean saccade latency
NB	Naive Bayes
NN	Neural networks
PD	Parkinson's disease
RF	Random forests
ROC	Receiver operating characteristic
Rsf	Rate of saccade/fixation
SD	Scan path duration
SG	Smooth pursuit gain
SL	Scan path length
SPEM	Smooth pursuit eye movement
SVM	Support vector machines
TDF	Total duration of fixations
TNF	Total number of fixations
UPDRS	Unified PD rating scale
VGS	Visual guided saccades
VR	Virtual reality

Acknowledgements

This research was supported by the "1+X" Program for Clinical Competency Enhancement-Interdisciplinary Innovation Project, The Second Hospital of Dalian Medical University under Grant 2022JCXKZD03, the Fundamental Research Funds for the Central Universities under Grant DUT24YG132 and DUT23YG223.

Author contributions

Maosong Jiang, Yanzhi Liu, Yongzhong Lin, Wenlong Liu conceived and supervised the research. Maosong Jiang, Yanzhi Liu: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data curation and Visualization. Yanlu Cao, Shufeng Xia: Formal analysis and Software. Yanzhi Liu, Wenzhi Zhao, Yongzhong Lin, Fei Teng: Resources and Data curation. Yongzhong Lin, Wenlong Liu: Conceptualization, Methodology, Formal analysis, and Funding acquisition. Maosong Jiang wrote the initial draft, with the other authors providing editorial comments.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 16 June 2024 Accepted: 25 December 2024
Published online: 14 January 2025

References

1. Chen B, Xu M, Yu H, He J, Li Y, Song D, Fan GG. Detection of mild cognitive impairment in Parkinson's disease using gradient boosting decision tree models based on multilevel DTI indices. J Transl Med. 2023;21(1):310.
2. Fukushima K, Fukushima J, Barnes GR. Clinical application of eye movement tasks as an aid to understanding Parkinson's disease pathophysiology. Exp Brain Res. 2017;235(5):1309–21.
3. Hikosaka O, Takikawa Y, Kawagoe R. Role of the basal ganglia in the control of purposive saccadic eye movements. Physiol Rev. 2000;80(3):953–78.
4. Zhou M-X, Wang Q, Lin Y, Xu Q, Wu L, Chen Y-J, Jiang Y-H, He Q, Zhao L, Dong Y-R, et al. Oculomotor impairments in de novo Parkinson's disease. Front Aging Neurosci. 2022;14: 985679.
5. Waldthaler J, Stock L, Student J, Sommerkorn J, Dowiasch S, Timmermann L. Antisaccades in Parkinson's disease: a meta-analysis. Neuropsychol Rev. 2021;31:628–42.
6. Jankovic J. Parkinson's disease: clinical features and diagnosis. J Neurol Neurosurg Psychiatry. 2008;79(4):368–76.
7. Postuma RB, Berg D. Advances in markers of prodromal Parkinson disease. Nat Rev Neurol. 2016;12(11):622–34.

8. Brooks DJ. Imaging dopamine transporters in Parkinson's disease. Biomark Med. 2010;4(5):651–60.
9. El Otmani H, Daghi M, Tahiri Jouti N, Lesage S. An overview of the worldwide distribution of lrrk2 mutations in Parkinson's disease. Neurodegener Dis Manag. 2023;13(6):335–50.
10. Weil RS, Schrag AE, Warren JD, Crutch SJ, Lees AJ, Morris HR. Visual dysfunction in Parkinson's disease. Brain. 2016;139(11):2827–43.
11. Pinkhardt EH, Kassubek J. Ocular motor abnormalities in Parkinsonian syndromes. Parkinsonism Relat Disord. 2011;17(4):223–30.
12. Antoniades C, Kennard C. Ocular motor abnormalities in neurodegenerative disorders. Eye. 2015;29(2):200–7.
13. Blekher T, Weaver M, Rupp J, Nichols WC, Hui SL, Gray J, Yee RD, Wojcieszek J, Foroud T. Multiple step pattern as a biomarker in Parkinson disease. Parkinson Relat Disord. 2009;15(7):506–10.
14. van Stockum S, MacAskill M, Anderson T, Dalrymple-Alford J. Don't look now or look away: two sources of saccadic disinhibition in Parkinson's disease? Neuropsychologia. 2008;46(13):3108–15.
15. MacAskill MR, Graham CF, Pitcher TL, Myall DJ, Livingston L, van Stockum S, Dalrymple-Alford JC, Anderson TJ. The influence of motor and cognitive impairment upon visually-guided saccades in Parkinson's disease. Neuropsychologia. 2012;50(14):3338–47.
16. Tsitsi P, Benfatto M, Seimyr G, Larsson O, Svenningsson P, Markaki I. Fixation duration and pupil size as diagnostic tools in Parkinson's disease. J Parkinsons Dis. 2021;11:865–75.
17. Gorges M, Müller H-P, Lulé D, Consortium L, Pinkhardt EH, Ludolph AC, Kassubek J. The association between alterations of eye movement control and cerebral intrinsic functional connectivity in Parkinson's disease. Brain Imaging Behav. 2016;10:79–91.
18. Zargari Marandi R, Madeleine P, Omland Ø, Vuillerme N, Samani A. Eye movement characteristics reflected fatigue development in both young and elderly individuals. Sci Rep. 2018;8(1):13148.
19. Lanthier S, Risko E, Smilek D, Kingstone A. Measuring the separate effects of practice and fatigue on eye movements during visual search. In: Proceedings of the Annual Meeting of the Cognitive Science Society. 2013;35(35):pp 2820–2825.
20. Bafna-Rührer T, Bækgaard P, Hansen JP. Smooth-pursuit performance during eye-typing from memory indicates mental fatigue. J Eye Mov Res. 2022;15(4):1–16.
21. Matsumoto H, Terao Y, Furubayashi T, Yugeta A, Fukuda H, Emoto M, Hanajima R, Ugawa Y. Small saccades restrict visual scanning area in Parkinson's disease. Mov Disord. 2011;26(9):1619–26.
22. Aubin G, Béliveau M-F, Klinger E. An exploration of the ecological validity of the virtual action planning-supermarket (vap-s) with people with schizophrenia. Neuropsychol Rehabil. 2018;28(5):689–708.
23. Ioannou A, Papastavrou E, Avraamides MN, Charalambous A. Virtual reality and symptoms management of anxiety, depression, fatigue, and pain: a systematic review. SAGE Open Nursing. 2020;6:2377960820936163.
24. Goetz CG, Poewe W, Rascol O, et al. Movement disorder society task force on rating scales for Parkinson's disease: the unified Parkinson's disease rating scale (updrs): status and recommendations. Mov Disord. 2003;18(7):738–50.
25. Bulling A, Ward JA, Gellersen H, Tröster G. Eye movement analysis for activity recognition. In: Proceedings of the 11th International Conference on Ubiquitous Computing, 2009;pp. 41–50.
26. Buscher G, Dengel A, van Elst L. Eye movements as implicit relevance feedback. In: CHI'08 Extended Abstracts on Human Factors in Computing Systems, 2008;pp. 2991–2996.
27. Rigaud C, Le T-N, Burie J-C, Ogier J-M, Ishimaru S, Iwata M, Kise K. Semi-automatic text and graphics extraction of manga using eye tracking information. In: 2016 12th IAPR Workshop on Document Analysis Systems (DAS), 2016;pp. 120–125. IEEE.
28. Marandi RZ, Gazerani P. Aging and eye tracking: in the quest for objective biomarkers. Future Neurol. 2019;14(4):33.
29. Terao Y, Fukuda H, Tokushuge S, Nomura Y, Hanajima R, Ugawa Y. Saccade abnormalities associated with focal cerebral lesions-how cortical and basal ganglia commands shape saccades in humans. Clin Neurophysiol. 2016;127(8):2953–67.
30. Tao L, Wang Q, Liu D, Wang J, Zhu Z, Feng L. Eye tracking metrics to screen and assess cognitive impairment in patients with neurological disorders. Neurol Sci. 2020;41:1697–704.
31. Berges AJ, Vedula SS, Chara A, Hager GD, Ishii M, Malpani A. Eye tracking and motion data predict endoscopic sinus surgery skill. Laryngoscope. 2023;133(3):500–5.
32. Nakamura T, Kanayama R, Sano R, Ohki M, Kimura Y, Aoyagi M, Koike Y. Quantitative analysis of ocular movements in Parkinson's disease. Acta Otolaryngol. 1991;111(sup481):559–62.
33. Andrienko G, Andrienko N, Burch M, Weiskopf D. Visual analytics methodology for eye movement studies. IEEE Trans Visual Comput Graphics. 2012;18(12):2889–98.
34. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016;pp. 2818–2826.
35. Wang Q, Wu B, Zhu P, Li P, Zuo W, Hu Q. Eca-net: efficient channel attention for deep convolutional neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020;pp. 11534–11542.
36. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. Scikit-learn: machine learning in python. J Mach Learn Res. 2011;12:2825–30.
37. Vasilev I, Slater D, Spacagna G, Roelants P, Zocca V. Python deep learning: exploring deep learning techniques and neural network architectures with Pytorch, Keras, and TensorFlow. Packt Publishing Ltd. 2019.
38. Cawley GC, Talbot NL. On over-fitting in model selection and subsequent selection bias in performance evaluation. J Mach Learn Res. 2010;11:2079–107.
39. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. Morgan Kaufman Publishing;1995.
40. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. Statist Surv. 2010;4:40–79.
41. Dietterich TG. Approximate statistical tests for comparing supervised classification learning algorithms. Neural Comput. 1998;10(7):1895–923.
42. Meng Z, Chen Y, Li H, Zhang Y, Yao X, Meng Y, Shi W, Liang Y, Hu Y, Liu D, et al. Machine learning and optical coherence tomography-derived radiomics analysis to predict persistent diabetic macular edema in patients undergoing anti-vegf intravitreal therapy. J Transl Med. 2024;22(1):358.
43. Wurtz RH, Hikosaka O. Role of the basal ganglia in the initiation of saccadic eye movements. Prog Brain Res. 1986;64:175–90.
44. Malhotra PA, Bronstein AM. Antisaccades and executive dysfunction in PD: Two sides of the same coin? Wiley Online Library;2015.
45. Bakst L, Fleuriet J, Mustari MJ. Fefsem neuronal response during combined volitional and reflexive pursuit. J Vis. 2017;17(5):13–13.
46. Ono S. The neuronal basis of on-line visual control in smooth pursuit eye movements. Vision Res. 2015;110:257–64.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.