

Article

Human-Centric Spatial Cognition Detecting System Based on Drivers' Electroencephalogram Signals for Autonomous Driving

Yu Cao ^{1,2} , Bo Zhang ^{1,2}, Xiaohui Hou ^{1,2,*}, Minggang Gan ^{1,2} and Wei Wu ^{1,2}

¹ School of Automation, Beijing Institute of Technology, Beijing 100081, China; 3220231086@bit.edu.cn (Y.C.); zbpt516@163.com (B.Z.); agan@bit.edu.cn (M.G.); wuwei@bit.edu.cn (W.W.)

² National Key Lab of Autonomous Intelligent Unmanned Systems, Beijing Institute of Technology, Beijing 100081, China

* Correspondence: houxh1220@163.com

Abstract: Existing autonomous driving systems face challenges in accurately capturing drivers' cognitive states, often resulting in decisions misaligned with drivers' intentions. To address this limitation, this study introduces a pioneering human-centric spatial cognition detecting system based on drivers' electroencephalogram (EEG) signals. Unlike conventional EEG-based systems that focus on intention recognition or hazard perception, the proposed system can further extract drivers' spatial cognition across two dimensions: relative distance and relative orientation. It consists of two components: EEG signal pre-processing and spatial cognition decoding, enabling the autonomous driving system to make more contextually aligned decisions regarding the targets drivers focus on. To enhance the detection accuracy of drivers' spatial cognition, we designed a novel EEG signal decoding method called a Dual-Time-Feature Network (DTFNet). This approach integrates coarse-grained and fine-grained temporal features of EEG signals across different scales and incorporates a Squeeze-and-Excitation module to evaluate the importance of electrodes. The DTFNet outperforms existing methods, achieving 65.67% and 50.65% accuracy in three-class tasks and 84.46% and 70.50% in binary tasks. Furthermore, we investigated the temporal dynamics of drivers' spatial cognition and observed that drivers' perception of relative distance occurs slightly later than their perception of relative orientation, providing valuable insights into the temporal aspects of cognitive processing.

Keywords: electroencephalogram (EEG); automatic driving; spatial cognition; human-machine cooperation



Academic Editors: Felipe Jiménez and Steven L. Waslander

Received: 3 December 2024

Revised: 29 December 2024

Accepted: 9 January 2025

Published: 10 January 2025

Citation: Cao, Y.; Zhang, B.; Hou, X.; Gan, M.; Wu, W. Human-Centric Spatial Cognition Detecting System Based on Drivers' Electroencephalogram Signals for Autonomous Driving. *Sensors* **2025**, *25*, 397. <https://doi.org/10.3390/s25020397>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the rapid development of autonomous driving technology, its application has become feasible in relatively simple environments. However, there are still many problems with autonomous driving technology, and it is still a long way from being fully applied [1–5]. For instance, current autonomous driving systems are highly susceptible to adverse weather conditions [1,2]. Moreover, when confronted with unexpected obstacles or emergencies on the road, these systems often struggle to make accurate judgments and respond promptly [3]. Failures in software or hardware components can further lead to severe safety risks [4]. These technical limitations necessitate human intervention in current autonomous driving systems [5]. In addition to safety challenges, existing autonomous driving technologies frequently fall short in ensuring driver comfort, particularly in achieving “human-like” driving behaviors. This disconnect can result in discrepancies between

system decisions and driver expectations, leading to discomfort, anxiety, or even traffic accidents [6].

To address these challenges, a promising solution involves developing human-centric assistant driving systems based on electroencephalogram (EEG) signals [7]. By introducing the cognition of drivers into the driving system, it is possible to enhance both the safety and comfort of autonomous driving [8,9]. EEG, as a commonly utilized physiological signal, captures valuable cognitive information about drivers. It reflects their perception and understanding of the environment, offering insights into processes such as perception, judgment, and decision-making. As a reliable source of reference for autonomous driving systems, EEG data can effectively compensate for the limitations of onboard sensors [10].

Several studies about the EEG-based human-centric assistant driving system have been published in recent years [11–26], and most of them focus on drivers' intention detection and hazard perception. Haufe et al. achieved a prediction 130 ms before the actual braking event using a linear discriminant analysis (LDA) classifier [11]. Zhang et al. proposed a model that includes hypergraph convolution for feature extraction, which outperforms the baseline in detecting various types of driving intentions, with an accuracy rate of 74.40% [13]. Teng et al. proposed an LDA model based on spatial-frequency features of EEG for detecting emergency braking intentions, achieving an accuracy of up to 94% [16]. As for driver's hazard perception, a hazard perception classifier based on EEG signals was developed for the scenario where pedestrians cross the road [19]. The low-frequency activities in EEG signals were found to increase when a vehicle cut in, which could be used to predict the cutting-in behavior of other vehicles [21]. Zhang et al. proposed an improved neural network intersection collision prediction model based on EEG indicators and driving behaviors [24].

However, focusing only on these two aspects is far from enough to help the driving system make effective decisions and improve ride comfort. With the development of relevant research, researchers found that hazards occurring in different locations can also trigger different neural responses. Li et al. found that hazards in the peripheral visual area can induce larger amplitude EEG signal components than those in the central visual area [27]. Wang et al. found that the brain activities triggered by lateral risks are stronger than those triggered by longitudinal risks using fNIRS [28]. These studies indicate that the spatial cognition of drivers can be reflected in neuroimaging techniques.

This study established a human-centric spatial cognition detecting system based on drivers' electroencephalogram signals for autonomous driving. By decoding the drivers' spatial cognition, the driving system can set a more comfortable following distance and keep the vehicle within the safe range perceived by the drivers based on the position of the target that the driver is paying attention to. And it is no longer necessary to decode the drivers' intention and hazard perception. However, there is no current research on spatial cognition detecting in driving scenarios based on EEG signals. Most of the EEG-based spatial cognition studies are based on simple scenes [29–32]. Kastrati et al. proposed EEGEyeNet, which is an EEG signal dataset for decoding gaze positions; by establishing the mapping relationship between EEG signals and annotated positions, it is possible to initially detect human spatial cognition [29]. Himmelberg et al. discovered that EEG signals can be utilized to decode the 3D movement direction of the observed target [31]. These studies have confirmed the feasibility of detecting spatial cognition from EEG signals.

To the best of our knowledge, this study is the first to explore and investigate the extraction of drivers' spatial cognition based on EEG signals, addressing a critical gap in this research domain. We demonstrate the feasibility of detecting drivers' spatial cognition, establishing a foundation for future advancements in this field. The key contributions of this study are as follows:

- **Pioneering a human-centric spatial cognition detecting system based on EEG signals:** This study introduces an EEG-based driver spatial cognition detecting system for the first time, which can equip the autonomous driving technology with high-level human spatial cognition to enhance its human likeness and comfort. The system can automatically preprocess EEG signals and decode drivers' spatial cognition along two dimensions, namely relative orientation and relative distance, with two components: EEG signal preprocessing and spatial cognition decoding.
- **Proposing an innovative EEG decoding method called a Dual-Time-Feature Network:** We designed a novel Dual-Time-Feature Network (DTFNet), which employs a convolutional module and a gated recurrent unit (GRU) module to extract local and global temporal features from EEG signals, respectively. Furthermore, DTFNet incorporates a Squeeze-and-Excitation (SE) module to assess the importance of different electrodes, enhancing its ability to capture relevant spatial cognition features.
- **Comprehensive evaluation:** We conducted comparative experiments across different traffic environments with varying directions and distances, exploring the temporal dynamics of drivers' spatial cognition. To rigorously validate the performance of the proposed method, we applied a five-fold cross-validation technique to evaluate its classification effectiveness against multiple baseline approaches.

The rest of this paper is structured as follows: Section 2 introduces the experimental setup and data process. Section 3 designs the methodology. Section 4 analyzes the experimental results. Section 5 draws the conclusion and discusses future research.

2. Experimental Procedure and Data Processing

The complete experimental process is shown in Figure 1. Firstly, we recorded both EEG signals and operational data from participants while they engaged in the spatial cognition task within a simulated traffic environment. Next, the EEG signals were preprocessed and divided into epochs, with labels generated based on the operation data and the stimuli. Finally, we classified the EEG signals using both traditional machine learning models and deep learning models. For the traditional machine learning models, additional feature extraction was performed on the EEG signals prior to classification, whereas deep learning models classified the signals directly. Each procedure of the experimental process is detailed in the following content.

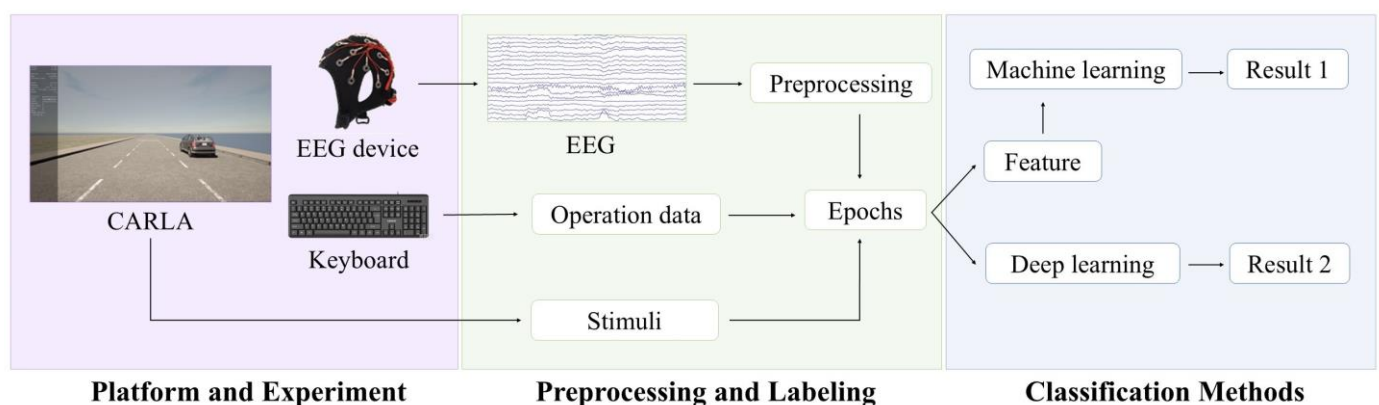


Figure 1. Flow diagram of experiment.

2.1. Experimental Platform and Experiment Design

2.1.1. Experimental Platform

As shown in Figure 2a, the experimental platform consists of two computers, an EEG acquisition device, and a keyboard. Computer 1 operates the driving simulation

environment, where participants interact using a keyboard to record operational data. The participants' EEG signals, which provide insights into their cognitive states, are collected via the EEG acquisition device. The EEG data are then transmitted to Computer 2 via Bluetooth for storage and further analysis. To ensure data synchronization between the two computers and the EEG device, we used the NTP (Network Time Protocol). For this study, we utilized the CARLA software (Version: Carla 0.9.15) [33] as the driving simulation environment and designed a circular three-lane map as the experimental setting, as shown in Figure 2b.

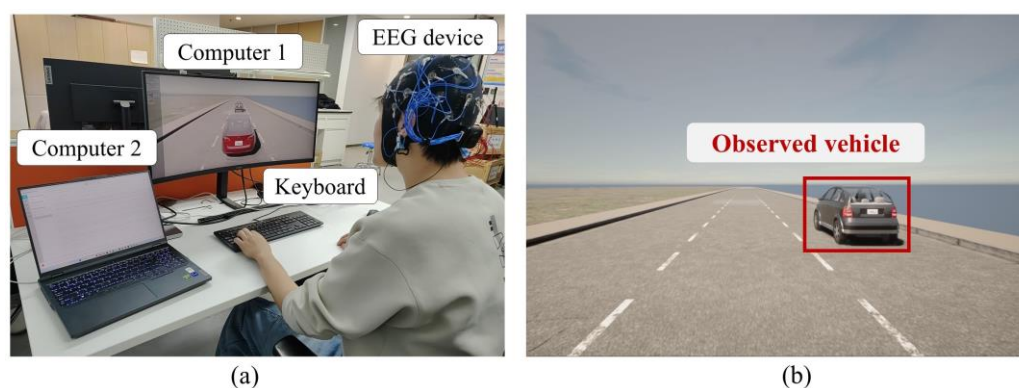


Figure 2. Driver-in-loop experimental platform. (a) Components of experimental platform and (b) driving simulation environment in CARLA.

Our EEG acquisition device is the FLEX 2 Gel-32 Channel Wireless EEG Head Cap System by EMOTIV, which operates at a sampling frequency of 128 Hz and includes 32 acquisition channels. The accompanying software system incorporates a timestamp correction algorithm that, under optimal wireless connection conditions, reduces corrected timestamp errors to within ± 0.5 ms, with a systemic error of approximately ± 5 ms.

2.1.2. Experimental Scene Design

We developed a spatial cognition experiment to collect and analyze participants' EEG signals during the driving process. In the experiment, participants drove an autonomous vehicle from a first-person perspective while maintaining a speed of 30 km/h on the circular three-lane map for 10 min. The first-person driving perspective is depicted in Figure 2b. During the experiment, a randomly generated observed vehicle appeared on one of the three lanes ahead of the autonomous vehicle. Participants were instructed to record their spatial cognition of the observed vehicle using the keyboard.

The experiment focused on recording data only when the autonomous vehicle was traveling on straight sections of the map. Participants were allowed to rest while navigating the curved sections. Each observed vehicle was displayed for 2 s before disappearing, and the next vehicle was generated after a 0.1 s interval.

As shown in Figure 3, a Cartesian coordinate system was established with the map's center as the origin. The coordinates corresponding to the centerlines of the straight sections of the three lanes on the map are designated as y_0 , y_1 , and y_2 , respectively. The position of the autonomous vehicle occupied by the participant is designated as (x_0, y_1) , indicating that the vehicle is always driving in the middle lane. Let the relative distance between the generated observed vehicle and the autonomous vehicle be denoted as Δx , and let the lane be any of the three lanes. The position of the observed vehicle can be represented as $(x_0 + \Delta x, y_k, k \in \{1, 2, 3\})$.

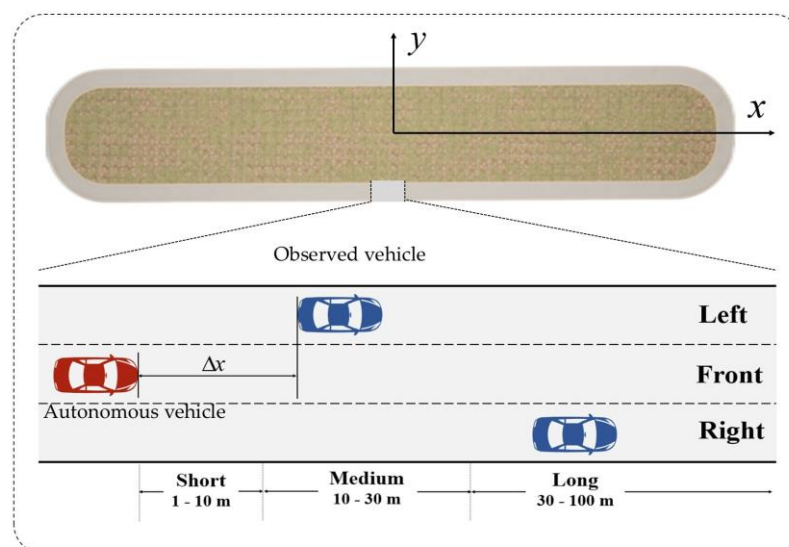


Figure 3. Experiment scene design in simulation environment.

To balance the occurrence probabilities among the three relative distance groups—short, medium, and long—we defined these categories based on factors such as braking distance, safe following distance, and observation range during driving [34–37]. The relative distances were categorized as follows: short range (1–10 m), medium range (10–30 m), and long range (30–100 m). For each generated observed vehicle, one of these groups was randomly selected, and the specific relative distance (Δx) was drawn from a uniform distribution within the selected range.

In this paper, we decoded two spatial cognition dimensions, namely relative distance and relative orientation simultaneously. Considering that EEG signals contain significant noise and have a relatively low signal-to-noise ratio [38], we simplified the decoding tasks to reduce complexity. The relative distance decoding task was framed as a three-class classification problem—short, medium, and long distance. Similarly, the relative orientation decoding task was reduced to a three-class classification problem involving the left side, front, and right side.

2.1.3. Participants

A total of 20 individuals participated in this study. The group consisted of 12 males and 8 females, aged 23 to 30 years. All participants had normal vision and hearing or corrected-to-normal vision and hearing, with no history of psychiatric or neurological disorders. Prior to the experiment, participants abstained from consuming any medications, tobacco, alcohol, or caffeine. Participation was entirely voluntary and conducted during regular working hours. Participants were informed that they had the right to withdraw from the study at any time without facing any penalties.

2.1.4. Data Collection

Given that different participants possess varying levels of spatial cognition, we required participants to indicate their perception of the current relative distance after each observed vehicle was generated by pressing a key on the keyboard. Keys 1, 2, and 3 corresponded to perceptions of short, medium, and long distances, respectively. Additionally, we recorded the generation time, disappearance time, and the position coordinates of both the autonomous vehicle and the observed vehicle for each instance. On average, the number of valid data collected from each participant is 170.5 epochs.

2.2. Data Preprocessing and Annotation

The raw EEG data collected during the experiment contained substantial noise, necessitating preprocessing to enhance data quality [39,40]. Additionally, the EEG signals were segmented into epochs and annotated to facilitate subsequent analyses and related experiments.

2.2.1. EEG Data Preprocessing

The preprocessing of EEG data in this study involved data selection, electrode localization, bandpass filtering to remove noise, re-referencing, and independent component analysis (ICA) for artifact removal.

- (1) **Data selection and electrode localization:** The data exported from our EEG device includes 116 channels, covering electrode connection quality, frequency data, and other metrics. For our analysis, we removed extraneous information and selected only the voltage data from 32 channels, with electrode positions determined according to the International 10–20 System.
- (2) **Band-pass filtering for noise removal:** A finite impulse response (FIR) filter was applied to perform bandpass filtering in the range of 0.5 to 40 Hz, reducing noise interference. For a given input time series, the output of an N-order finite impulse response filter was calculated in the following manner:

$$y(n) = \sum_{k=0}^{N-1} h(k)x(n-k) \quad (1)$$

where $h(k)$ represents the filter coefficients.

- (3) **Re-referencing:** All electrodes in this study were symmetrically distributed. To minimize the impact of the reference electrode on the experimental data, the EEG data were re-referenced using the average signal from all electrodes.
- (4) **ICA for artifact removal:** ICA, based on the statistical properties of signals, effectively separates overlapping EEG signals, removing noise and artifacts from the original EEG data to enhance data quality.

2.2.2. Data Annotation

After preprocessing the EEG signals, the data were segmented and annotated for subsequent spatial cognition detecting experiments. The EEG signals were segmented using the disappearance time of each observed vehicle as the starting time step. A 1.5 s segment, consisting of 192 time steps including the starting step, was treated as a single dataset. Next, relative distance labels (short, medium, long) were assigned to the EEG data based on the participants' marked distance perceptions. Additionally, relative orientation labels (left, front, right) were assigned using the recorded positional data of the observed vehicles.

3. Methodology

As described above, our spatial cognition recognition experiment involved two classification tasks: a three-class classification of relative distance (short, medium, long) and a three-class classification of relative orientation (left, front, right). To address these tasks, we designed a novel neural network, DTFNet, to decode the EEG signals effectively. For comparison, we selected three traditional machine learning methods and three deep learning-based methods as baselines for the experiments.

3.1. EEG Feature Extraction

Since traditional machine learning algorithms face challenges in directly processing high-dimensional data such as raw EEG signals [41], we performed feature extraction on the EEG signals to reduce the dimensionality of the data. As shown in Figure 4, the EEG electrodes were evenly distributed across the brain, with each electrode corresponding to a specific brain region. Based on their positions, the electrodes were categorized into four regions: frontal, parietal, occipital, and temporal lobes [42]. The specific electrodes corresponding to each region are listed in Table 1.

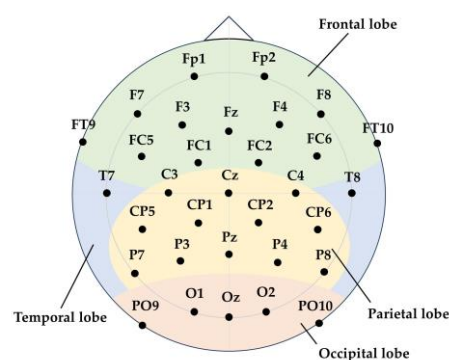


Figure 4. Electrode distribution of the EEG acquisition device.

Table 1. Electrode partition.

Region	Electrode
Frontal lobe	'Fp1', 'Fp2', 'Fz', 'F3', 'F4', 'F7', 'F8', 'FC1', 'FC2', 'FC5', 'FC6', 'FT9', 'FT10'
Parietal lobe	'Cz', 'C3', 'C4', 'CP1', 'CP2', 'CP5', 'CP6', 'Pz', 'P3', 'P4', 'P7', 'P8'
Occipital lobe	'Oz', 'O1', 'O2', 'PO9', 'PO10'
Temporal lobe	'T7', 'T8'

EEG signals can be divided into frequency bands, including Theta waves (3–8 Hz), Alpha waves (8–12 Hz), Beta waves (12–27 Hz), and Gamma waves (27 Hz and above) [43]. Among these, Alpha waves have been shown to correlate with spatial cognition and spatial attention abilities [29]. Therefore, we calculated the power spectral density (PSD) of the Alpha frequency band for the four brain regions to extract features relevant to the driver's spatial cognition recognition task. This resulted in the extraction of four features for each set of EEG signals [44].

We took 96 or 128 time steps as a sampling window and calculated the power spectral density of each channel for the data in each sampling window. For each partition, the average power spectral density of all channels within the partition was used as a feature. Regarding a discrete-time signal that has a length of N , the approach for computing its PSD is as follows:

$$\hat{S}_x(f) = \frac{1}{N} |X_N(f)|^2 \quad (2)$$

Here, $X_N(f)$ refers to the discrete Fourier transform (DFT) of $x[n]$.

In this research, in order to mitigate the influence of spectral leakage, each segment of the signal was dealt with by applying a Hann window prior to calculating the PSD. The Hann window function is defined in this way:

$$w(n) = 0.5(1 - \cos(\frac{2\pi n}{N-1})), \quad n = 0, 1, \dots, N-1 \quad (3)$$

The overall power within a particular frequency band can be acquired by adding up the PSD values across the intended frequency range, which is expressed as

$$P_{band} = \sum_{f_1}^{f_2} \hat{S}_x(f) \Delta f \quad (4)$$

In the above formula, f_1 and f_2 stand for the lower and upper frequency limits, respectively, and Δf represents the frequency resolution.

3.2. Traditional Machine Learning Algorithms

Using the extracted features as input, we employed three traditional machine learning algorithms to perform the driver's spatial cognition detecting task: K-Nearest Neighbors (KNNs) [45], Support Vector Machine (SVM) [46], and Random Forest (RF) [47]. The parameters of these algorithms are detailed in Table 2.

Table 2. Parameters of the three methods.

Method	Parameters
K-Nearest Neighbors (KNNs)	K = 30
Support Vector Machine (SVM)	Kernel = 'poly', C = 10, random_state = 42
Random Forest (RF)	Trees = 1000, Max depth = None, random_state = 42

3.3. Deep Learning Algorithms

With the advancement of deep learning algorithms, many studies on EEG signal decoding have shifted away from manual feature extraction. Instead, they focus on designing neural networks to automatically extract features and perform decoding. In this study, three neural networks specifically designed for EEG signal processing were employed for feature extraction and classification: MLP [48], EEGNet [49], and ConvNet [50]. Additionally, we propose a novel neural network based on dual-time feature fusion, which achieves the best decoding performance. The following sections provide a detailed introduction to these neural networks.

3.3.1. MLP-Based Model

This study constructs two spatial cognition decoding models based on the Multi-Layer Perceptron (MLP). The primary distinction between the two models lies in whether the EEG signals in the input dataset have undergone frequency domain feature extraction. Both models share a similar basic architecture, consisting of a fully connected network with four hidden layers. To prevent overfitting and enhance generalization, the dropout technique is applied for regularization. Assume that the input of the model is $x \in \mathbf{R}^n$ and the output of the model is $y \in \mathbf{R}^m$. Then, the MLP model can be represented as

$$y = f(Wx + b) \quad (5)$$

where $W \in \mathbf{R}^{n \times m}$ is the weight matrix, b is the bias, and f is the nonlinear activation function.

The input size of the raw EEG signals is either 32×96 or 32×128 , representing the number of channels and time steps, respectively. For manually extracted features, the input size is 1×4 , corresponding to the four features derived from the Alpha band power spectral density of different brain regions. These two MLP models were used to compare the impact of manual feature extraction versus direct feature extraction using deep learning models on the final classification performance of EEG signals. This comparison

highlights the effectiveness of automated feature extraction in decoding spatial cognition from EEG data.

3.3.2. EEGNet

The spatial cognition decoding from the EEG task introduced in this paper is a novel task. As no neural network models specifically designed for this task currently exist, we have chosen general EEG signal decoding models for evaluation.

The first model selected is EEGNet, a compact convolutional neural network tailored for EEG signal decoding tasks. EEGNet features a unique convolutional structure, incorporating depthwise convolution, separable convolution, and pointwise convolution, enabling it to efficiently extract EEG signal features. This model combines high efficiency—characterized by a small number of parameters, low computational complexity, and suitability for real-time applications—with strong generalization capabilities, performing effectively across diverse brain–computer interface paradigms. EEGNet is considered a baseline in many EEG signal decoding studies. The specific structure and parameters of EEGNet utilized in this study are presented in Table 3.

Table 3. The model structure and parameters of EEGNet.

Layer	Parameters	Activation
Conv2D	input = 1, output = 16, kernel_size = (1, 64), padding = ‘same’	-
BatchNorm	8	-
Conv2D	input = 16, output = 2×16 , kernel_size = (32, 1), groups = 8, max_norm = 1	-
BatchNorm	16	ELU
AveragePool2D	kernel_size = (1, 4)	-
Dropout	0.25	-
Conv2D	input = 2×16 , output = 32, kernel_size = (1, 16), groups = 16, padding = ‘same’	-
Conv2D	input = 32, output = 32, kernel_size = 1	-
BatchNorm	32	ELU
AveragePool2D	kernel_size = (1, 8)	-
Dropout	0.25	-
Flatten	-	-
FC	input = 64, output = 3, max_norm = 0.25	Softmax

3.3.3. ConvNet

ConvNet, proposed by R. Schirrmeister et al. [50], is a neural network designed for EEG signal classification tasks. The authors introduced two architectures, ShallowConvNet and DeepConvNet, which achieved state-of-the-art performance in distinguishing pathological EEG signals from normal ones. These architectures have since been established as general-purpose neural networks for EEG signal decoding.

In this study, both ShallowConvNet and DeepConvNet were employed to perform the spatial cognition decoding task. The detailed parameters of these two networks are presented in Tables 4 and 5, respectively.

Table 4. The model structure and parameters of ShallowConvNet.

Layer	Parameters	Activation
Conv2D	input = 1, output = 25, kernel_size = (1, 25)	-
Conv2D	input = 25, output = 25, kernel_size = (32, 1)	-
BatchNorm	25	ELU

Table 4. *Cont.*

Layer	Parameters	Activation
AveragePool2D	kernel_size= (1, 15), stride = 5	-
Dropout	0.4	-
Flatten	-	-
FC	input = 450, output = 3	Softmax

Table 5. The model structure and parameters of DeepConvNet.

Layer	Parameters	Activation
Conv2D	input = 1, output = 25, kernel_size = (1, 5), stride = (1, 2)	-
Conv2D	input = 25, output = 25, kernel_size = (32, 1),	-
BatchNorm	25	ELU
MaxPool2D	kernel_size = (1, 3), stride = 1	-
Dropout	0.4	-
Conv2D	input = 25, output = 50, kernel_size = (1, 5), stride = (1, 2)	-
BatchNorm	50	ELU
MaxPool2D	kernel_size = (1, 3), stride = 1	-
Dropout	0.4	-
Conv2D	input = 50, output = 100, kernel_size = (1, 5), stride = (1, 2)	-
BatchNorm	100	ELU
MaxPool2D	kernel_size = (1, 3), stride = 1	-
Dropout	0.4	-
Flatten	-	-
FC	input = 200, output = 3	Softmax

3.3.4. DTFNet

In this study, we developed a novel neural network, Dual-Time-Feature Net (DTFNet), to address the spatial cognition decoding task. The model is divided into two main modules: a temporal processing module and a spatial processing module. The structure and components of these two modules are introduced below.

- (1) **Temporal Processing Module:** The temporal processing module comprises a time-dimension convolutional module and a GRU module. Considering that convolutional modules often struggle to capture long-term dependencies in EEG signals, the GRU module [51] is incorporated to extract fine-grained temporal features for each EEG channel. These features are then fused with the coarse-grained temporal features extracted by the convolutional module, enabling the model to effectively handle both short-term and long-term temporal dependencies.
- (2) **Spatial Processing Module:** The spatial processing module is designed to address the challenge of determining which of the 32 EEG channels contain the most relevant information. To achieve this, we utilize the Squeeze-and-Excitation (SE) module [52], which computes channel attention to distinguish the contributions of different channels. Afterward, the features from all channels are fused using a spatial dimension convolutional module to generate the final spatiotemporal features. These features are passed to a fully connected layer to produce the final prediction results.

By integrating temporal features at two scales and leveraging spatial attention mechanisms, DTFNet demonstrates superior performance compared to general EEG decoding models in spatial cognition decoding tasks, making it a robust and efficient solution for decoding drivers' spatial cognition from EEG signals. The architecture of DTFNet is illustrated in Figure 5, and the parameters of the model are detailed in Table 6, offering a clear understanding of the network's design and implementation.

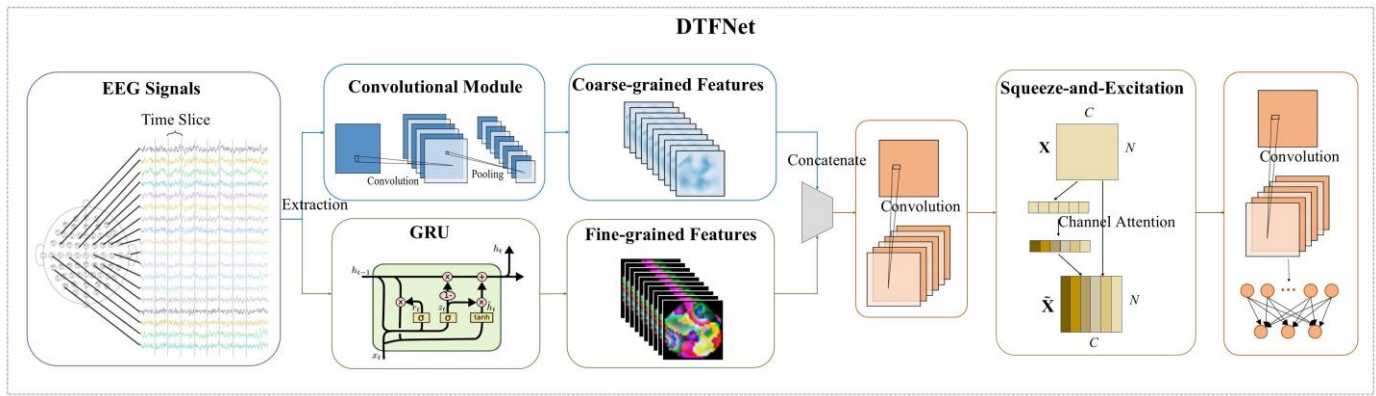


Figure 5. The structure of the DTFNet.

Table 6. The parameters of the DTFNet.

Layer	Parameters	Activation
GRU	input = 1, hidden_size = 64, num_layers = 1	ELU
Conv2D_1	Input = 1, output = 16, kernel_size = (1, 64), padding = 'same'	-
BatchNorm	16	ELU
AveragePool2D	(1, 2)	-
Dropout	0.5	-
Conv2D_2	input = 1, output = 16, kernel_size = (1, 128)	-
BatchNorm	16	ELU
Dropout	0.5	-
AveragePool2D	(None, 1)	-
Flatten	-	-
Dense	input = 32, output = 16	ReLU
Dense	input = 16, output = 32	Sigmoid
Dropout	0.5	-
Conv2D_3	input = 1, output = 16, kernel_size = (32, 1)	-
BatchNorm	16	ELU
Flatten	-	-
FC	input = 256, output = 3	Softmax

For a given input $x \in \mathbf{R}^{C \times T}$, it is first processed through two temporal feature extraction modules to obtain coarse-grained temporal features and fine-grained temporal features.

$$\text{Feature}_{\text{coarse}} = \text{ELU}(\text{Avg}(\text{Conv2d}_1(x))) \quad (6)$$

$$\text{Feature}_{\text{fine}} = \text{ELU}(\text{GRU}(x)) \quad (7)$$

Next, the coarse-grained and fine-grained temporal features are concatenated and passed through a convolutional layer to perform feature fusion. This process integrates the complementary information from both feature types, resulting in a unified representation of the temporal features.

$$\text{Feature}_{\text{temporal}} = \text{ELU}(\text{Conv2d}_2(\text{Concat}(\text{Feature}_{\text{coarse}}, \text{Feature}_{\text{fine}}))) \quad (8)$$

To effectively fuse the features of the EEG signals along the spatial dimension, we employ the SE module. This module starts by averaging the features of each channel through an average pooling layer, generating a global descriptor for each channel. These descriptors are then passed through two fully connected layers with an activation function in between, producing channel-specific weights. Finally, these weights are applied to the

original features via element-wise multiplication, amplifying the contributions of relevant channels and reducing the influence of less informative ones.

$$\text{Weight}_{channel} = \text{Sigmoid}(\text{FC}(\text{ReLU}(\text{FC}(\text{Avg}(\text{Feature}_{temporal})))))) \quad (9)$$

$$\text{Feature}_{channel} = \text{Weight}_{channel} \odot \text{Feature}_{temporal} \quad (10)$$

After the SE module, a convolutional layer is applied to fuse the features across the channel scale, resulting in the final spatiotemporal feature representation. Finally, these features are passed through a fully connected layer to generate the prediction results.

$$\text{Feature}_{spatiotemporal} = \text{ELU}(\text{Conv2d}_3(\text{Feature}_{channel})) \quad (11)$$

$$\text{Output} = \text{FC}(\text{Feature}_{spatiotemporal}) \quad (12)$$

4. Experimental Results

This section presents the classification results for the spatial cognition decoding task. For traditional machine learning models and one of the MLP models, the inputs are EEG signals that have undergone feature extraction. In contrast, the inputs for the other deep learning models consist of EEG signals that have been preprocessed but not feature-extracted. We employed the five-fold cross-validation method to rigorously evaluate the classification performance of all models in this study. Five-fold cross-validation divides the dataset into five subsets. For each iteration, one subset is used as the test set, while the remaining four subsets are used to train the model. This approach provides a robust evaluation of model performance by ensuring that all data are used for both training and testing across the five iterations.

The models were implemented using the PyTorch (Version 1.8.1) framework and the Scikit-learn library. All training was conducted on an NVIDIA GeForce GTX TITAN GPU. A learning rate of 0.0001 was applied to all models, and each model was trained for 500 epochs. To evaluate the classification performance, we adopted four standard metrics: accuracy, precision, recall, and F1 score. These metrics provide a comprehensive assessment of the models' effectiveness in the spatial cognition decoding task.

4.1. Relative Orientation Classification Task

We first conducted the relative orientation classification task, and the experimental results are presented in Table 7. This task evaluates the spatial cognition of drivers in the orientation dimension. Successfully decoding drivers' relative orientation cognition of a target from EEG signals can provide valuable insights for autonomous driving systems. By accurately understanding the orientation of the target that the driver is focusing on, the driving system can make informed and reasonable decisions that align with the driver's intentions. This helps to prevent situations where the autonomous driving system's actions contradict the driver's expectations, thereby enhancing system reliability and driver satisfaction.

Table 7. The results of the relative orientation classification task.

Models	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
KNN	37.30	37.62	37.30	37.01
SVM	38.22	40.31	38.22	31.46
RF	34.63	35.23	34.62	34.59
MLP_feature	32.78	32.70	32.79	32.18

Table 7. Cont.

Models	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
MLP_raw	49.32	48.41	49.30	46.92
EEGNet	62.63	62.43	62.25	62.03
ShallowConvNet	58.85	59.18	58.90	58.78
DeepConvNet	54.19	55.75	54.53	54.00
DTFNet (ours)	65.67	65.90	65.72	65.61

The relative orientation decoding task was simplified into a three-class classification task involving the categories left, front, and right, resulting in a chance-level probability of 33.33% for the classification task. For this task, a time window size of 96 was selected for the EEG signals, meaning that the input size is $x \in \mathbf{R}^{32 \times 96}$. The rationale for selecting this specific time window size will be discussed in detail in the Section 4.3.

From the results presented in Table 7, it can be observed that, with the exception of the feature-based MLP classifier, the classification performance of all other models surpasses the chance probability. We attribute the failure of the feature-based MLP classifier to the low dimensionality of the extracted features, which likely hindered effective training of the neural network. When comparing traditional machine learning algorithms with deep learning algorithms, it is evident that deep learning approaches achieve superior classification performance. This difference can be explained by two factors:

- (1) **Limited Relative Spatial Cognition Information in Extracted Features:** The features obtained through manual extraction contain relatively sparse information about relative orientation, which negatively impacts classification performance.
- (2) **Superior Feature Extraction and Learning Capabilities of Deep Learning Models:** Deep learning algorithms demonstrate more robust capabilities in automatically extracting and learning relevant features directly from EEG signals, leading to improved classification outcomes.

Our proposed algorithm achieves the best classification performance, with an accuracy of 65.67%, precision of 65.90%, recall of 65.72%, and an F1 score of 65.61%. These results highlight the effectiveness of our approach for the relative orientation decoding task.

4.2. Relative Distance Classification Task

The relative distance decoding task, similar to the relative orientation decoding task, was also simplified into a three-class classification problem, involving the categories close, medium, and far. This task, however, is more challenging than the relative orientation decoding task because the perception of relative distance is inherently more subjective. Upon analyzing the collected data, we identified a certain degree of overlap between the actual distances corresponding to the perceived categories of close, medium, and far. To address this issue, we excluded data with cognitive ambiguity before proceeding with the decoding task. Successfully decoding individuals' perception of relative distance from EEG signals can assist autonomous driving systems in better adjusting safe distances and improving driver comfort. For this task, a time window size of 128 was selected for the EEG signals; that is, the input is $x \in \mathbf{R}^{32 \times 128}$.

From the experimental results presented in Table 8, it can be observed that although decoding relative distance from EEG signals is inherently challenging, the classification accuracy of most models exceeds the chance probability, demonstrating the feasibility of decoding spatial cognition from EEG signals. Among the models, deep learning approaches consistently outperform traditional machine learning models, with the exception of the MLP model, which failed to achieve competitive classification performance. This highlights the superior capability of deep learning algorithms in feature extraction and learning from

raw EEG signals. Our proposed model achieved the best classification performance, with an accuracy of 50.65%, precision of 50.68%, recall of 50.64%, and an F1 score of 49.91%. These results validate the effectiveness of our model for addressing the relative distance decoding task.

Table 8. The results of the relative distance classification task.

Models	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
KNN	40.10	41.55	40.10	41.07
SVM	36.29	41.52	36.29	29.55
RF	39.99	40.87	39.99	39.91
MLP_feature	33.89	34.27	33.89	33.40
MLP_raw	35.09	35.02	35.01	34.90
EEGNet	47.02	46.68	46.98	45.99
ShallowConvNet	45.33	45.31	45.26	45.04
DeepConvNet	43.59	48.80	43.81	40.99
DTFNet (ours)	50.65	50.68	50.64	49.91

4.3. Further Discussion

To gain deeper insights into the experimental results, we decomposed the three-class classification problem into three binary classification problems. This approach allowed us to analyze the classification performance for each pair of categories individually, providing a more granular understanding of the challenges and strengths associated with decoding spatial cognition from EEG signals.

Furthermore, we explored the temporal dynamics of drivers' spatial detection, examining how drivers' spatial cognition evolves over time. This temporal analysis offers valuable insights into the relationship between EEG signals and drivers' perception of relative distances, which may further inform the development of adaptive and responsive autonomous driving systems.

We also provided an in-depth analysis of the challenges this system may face in real-world applications. Additionally, we proposed potential solutions to these challenges and identified promising avenues for future research.

4.3.1. Binary Classification Results

To further analyze the two spatial cognition decoding tasks, we simplified the three-class classification problems into multiple binary classification tasks. For instance, in the relative orientation classification problem, the task was divided into the following binary classification problems: Left vs. Front, Left vs. Right, and Front vs. Right. This division allows for a more detailed assessment of the performance of various methods on specific pairwise comparisons, offering deeper insights into their strengths and weaknesses. The classification accuracies of different methods for these binary classification problems are presented below.

(1) Binary classification results of relative orientation

The binary classification results for the relative orientation decoding task are displayed on Figure 6a. The results indicate that all models achieved classification accuracies exceeding the chance probability of 50%. From the figure, it is evident that the left-right binary classification task yields the best performance. This aligns with our hypothesis, as the cognitive distinction between left and right orientations is significantly greater than that for the other two binary classification tasks. For the left-front and front-right binary classification tasks, the classification performances are nearly identical, reflecting the symmetrical experimental scenarios used in these cases.

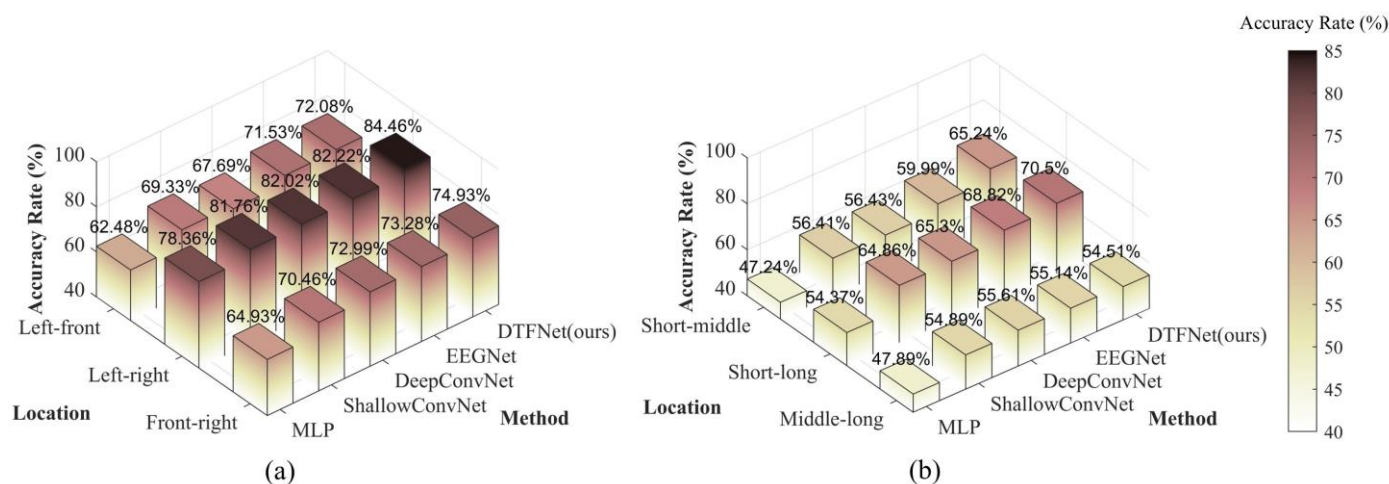


Figure 6. Binary classification results. (a) Relative orientation and (b) relative distance.

In summary, the results demonstrate that drivers' spatial cognition of relative orientation can be effectively decoded from EEG signals. Moreover, the findings indicate that drivers' cognitive processing of left and right orientations does not exhibit a strong inherent bias. Finally, while the performance differences among several deep learning models are minimal across the binary classification tasks, our proposed model consistently achieves the best classification performance, highlighting its effectiveness in detecting drivers' spatial cognition.

(2) Binary classification results of relative distance

The binary classification results for the relative distance decoding task are displayed in Figure 6b. As shown, the decoding performance for relative distance is inferior to that for relative orientation, consistent with the findings of the three-class classification study discussed earlier. Similar to the binary classification results for relative orientation, the short-long group achieves the best classification performance among the relative distance tasks. However, unlike relative orientation, relative distance is not a symmetrical scenario, and the decoding performance for the short-medium and medium-long groups differs significantly. The figure shows that the decoding performance for the short-medium group is better than that for the medium-long group. This result aligns with the cognitive patterns of human drivers. During driving, drivers tend to focus more on targets that are closer to them, as closer targets pose a higher potential danger. Conversely, targets at medium and long distances are less likely to pose an immediate threat, resulting in less pronounced differences in EEG signals for these groups.

4.3.2. Temporal Dynamics of Drivers' Spatial Cognition

We analyzed the temporal dynamics of drivers' spatial cognition using EEG signals. The sampling frequency of our EEG acquisition instrument is 128 Hz, meaning that 128 time steps correspond to 1 s in real time. EEG decoding began with 64 time steps and was repeated at intervals of 32 additional time steps, up to 192 time steps. The decoding results for relative orientation and relative distance are shown in Figure 7a,b, respectively. From the figures, it is evident that when the time step length is 64, it is difficult to extract effective spatial cognition information from EEG signals. This is likely because spatial cognition requires a certain amount of processing time. After visual stimuli are received by the retina, they must be transmitted to the visual processing cortex, where spatial cognition is formed [53]. And by comparing the results for the two tasks, the following observations can be made:

- (1) **Relative Orientation Decoding Task:** The best decoding performance is achieved when the EEG signal time step length is 96. Increasing the time window size beyond 96 does not significantly enhance classification performance, suggesting that orientation cognition stabilizes within this time frame.
- (2) **Relative Distance Decoding Task:** The optimal decoding performance is achieved at a time step length of 128, indicating that the perception of relative distance occurs slightly later than that of orientation.

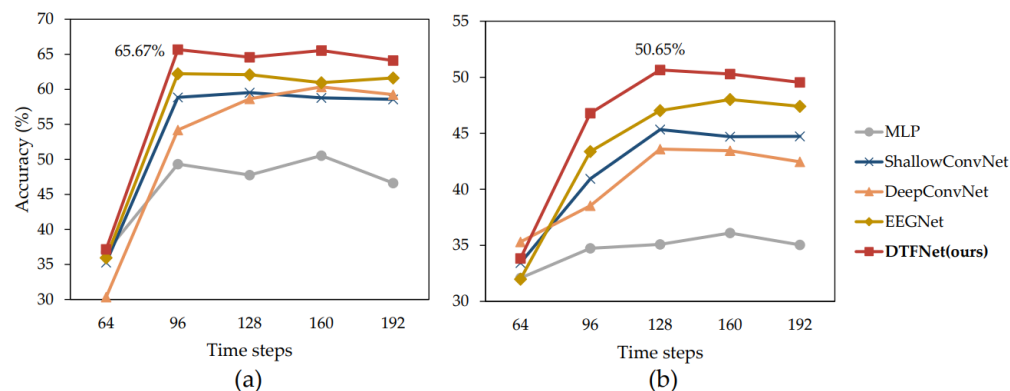


Figure 7. Detecting results with different time window sizes. (a) Relative Orientation and (b) relative Distance.

Additionally, across all time step lengths, our proposed model consistently outperforms the other models in classification accuracy. This demonstrates the robustness and effectiveness of our model for decoding drivers' spatial cognition in both relative orientation and relative distance tasks.

4.3.3. Application Analysis and Limitations

Although our designed human-centric spatial cognition detecting system can detect the spatial perception of drivers, enabling the driving system to establish a more comfortable following distance and maintain the vehicle within the safe range perceived by the drivers based on the position of the target they are focusing on, there is still a gap before it can be applied in real-world scenarios, and several challenges remain to be addressed.

The first challenge lies in the significant gap between the simulated scenarios used in our study and real-world driving conditions. In real driving scenarios, drivers encounter multiple targets moving continuously, and the resulting neural signals may differ from those observed in our simulated environment. However, even in real-world scenarios, drivers still engage in relative orientation and distance cognition of targets. Therefore, we believe that the EEG signals should not differ substantially. In the future, we plan to further refine our simulated scenarios to make them more closely resemble real-world conditions and conduct experiments in real driving environments to collect and analyze data for further research.

Another challenge is the real-time decoding of EEG signals. Unlike the current experiments, where EEG signals are segmented and then decoded, practical applications require real-time analysis. Real-time EEG decoding is an active area of research in the field of brain–computer interfaces, but it is not the primary focus of this study. In the future, we could explore integrating a real-time EEG decoding system into our framework. This would enable continuous analysis of EEG signals, allowing us to provide ongoing outputs of the driver's spatial perception, thereby contributing to the decision-making process of the driving system.

The driver's state can also affect the detecting performance. In real-world scenarios, drivers may experience fatigue during driving. In cases of mild fatigue, the detection of

spatial cognition—being a fundamental cognitive ability—may not be significantly affected. However, in cases of moderate to severe fatigue, further processing of the EEG signals will be necessary before detection. We consider using preprocessing algorithms, such as wavelet transform, or integrating existing driver fatigue detection algorithms to design a denoising module based on deep learning. Additionally, during actual driving, the driver's head movements and driving actions can introduce more artifacts, degrading the quality of the EEG signals. We plan to apply ICA or deep learning algorithms to remove these artifacts and improve the decoding algorithm to enhance performance on low-quality EEG signals.

In the future, we will conduct further research on the three challenges mentioned above, improving the existing simulation environment to make it more closely resemble real-world scenarios. At the same time, we will strive to enhance the robustness and accuracy of the driver's spatial cognition detecting system.

5. Conclusions

In response to the limitations of existing autonomous driving systems, particularly their limited ability to mimic human-like decision-making, this paper proposes a human-centric spatial cognition detection system based on drivers' signals for autonomous driving. The system decodes two critical dimensions of drivers' spatial cognition—relative distance and relative orientation—from EEG signals. Our findings demonstrate that spatial cognition can be effectively extracted from EEG data. To achieve this, we propose a novel EEG signal decoding neural network, DTFNet, which integrates temporal features at multiple scales. The proposed network achieves superior performance compared to other general-purpose EEG decoding models in spatial cognition tasks. Binary classification experiments reveal that drivers focus more on closer targets, while exhibiting no significant directional preference. Furthermore, our exploration of the temporal dynamics of drivers' spatial cognition indicates that drivers perceive relative orientation slightly earlier than relative distance.

In the future, our work will be carried out in two main directions. On the one hand, we will enhance the complexity of experimental scenarios to better simulate real-world driving conditions and collect a larger and more diverse dataset of drivers' spatial cognition to improve the system's robustness and generalizability. On the other hand, we will develop EEG signal decoding algorithms specifically tailored to spatial cognition tasks, focusing on improving the accuracy and effectiveness of decoding drivers' spatial cognition and addressing the unique challenges posed by this application.

Author Contributions: Conceptualization, Y.C. and M.G.; methodology, Y.C.; software, Y.C.; validation, Y.C. and W.W.; formal analysis, Y.C.; investigation, Y.C. and B.Z.; resources, Y.C.; data curation, Y.C.; writing—original draft preparation, Y.C. and B.Z.; writing—review and editing, X.H.; visualization, B.Z. and W.W.; supervision, X.H.; project administration, X.H.; funding acquisition, X.H. and M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 62303058.

Institutional Review Board Statement: This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by the Beijing Institute of Technology Research Ethics Committee and performed in line with the 2013 Declaration of Helsinki.

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

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

- Li, Y.; Moreau, J.; Ibanez-Guzman, J. Emergent Visual Sensors for Autonomous Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 4716–4737. [\[CrossRef\]](#)
- Wu, W.; Deng, X.; Jiang, P.; Wan, S.; Guo, Y. CrossFuser: Multi-Modal Feature Fusion for End-to-End Autonomous Driving Under Unseen Weather Conditions. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 14378–14392. [\[CrossRef\]](#)
- Lin, P.; Tsukada, M. Model Predictive Path-Planning Controller With Potential Function for Emergency Collision Avoidance on Highway Driving. *IEEE Robot. Autom. Lett.* **2022**, *7*, 4662–4669. [\[CrossRef\]](#)
- Huang, Z.; Sun, S.; Zhao, J.; Mao, L. Multi-modal policy fusion for end-to-end autonomous driving. *Inf. Fusion* **2023**, *98*, 165–195. [\[CrossRef\]](#)
- Fu, Y.; Li, C.; Yu, F.R.; Luan, T.H.; Zhang, Y. A Survey of Driving Safety With Sensing, Vehicular Communications, and Artificial Intelligence-Based Collision Avoidance. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 6142–6163. [\[CrossRef\]](#)
- Zhang, T.; Zhan, J.; Shi, J.; Xin, J.; Zheng, N. Human-Like Decision-Making of Autonomous Vehicles in Dynamic Traffic Scenarios. *IEEE/CAA J. Autom. Sin.* **2023**, *10*, 1905–1917. [\[CrossRef\]](#)
- Ju, J.; Bi, L.; Feleke, A.G. Noninvasive neural signal-based detection of soft and emergency braking intentions of drivers. *Biomed. Signal Process. Control* **2022**, *72*, 103330. [\[CrossRef\]](#)
- Malik, M.; Nandal, R.; Dalal, S.; Jalglan, V.; Le, D.-N. Deriving Driver Behavioral Pattern Analysis and Performance Using Neural Network Approaches. *Intell. Autom. Soft Comput.* **2022**, *32*, 87–99. [\[CrossRef\]](#)
- Sumit; Chhillar, R.S.; Dalal, S.; Dalal, S.; Lilhore, U.K.; Samiya, S. A dynamic and optimized routing approach for VANET communication in smart cities to secure intelligent transportation system via a chaotic multi-verse optimization algorithm. *Clust. Comput.* **2024**, *27*, 7023–7048. [\[CrossRef\]](#)
- Ju, J.; Li, H. A Survey of EEG-Based Driver State and Behavior Detection for Intelligent Vehicles. *IEEE Trans. Biom. Behav. Identity Sci.* **2024**, *6*, 420–434. [\[CrossRef\]](#)
- Haufe, S.; Treder, M.S.; Gugler, M.F.; Sagebaum, M.; Curio, G.; Blankertz, B. EEG potentials predict upcoming emergency brakings during simulated driving. *J. Neural. Eng.* **2011**, *8*, 56001. [\[CrossRef\]](#) [\[PubMed\]](#)
- Liang, X.; Yu, Y.; Liu, Y.; Liu, K.; Liu, Y.; Zhou, Z. EEG-Based Detection of Braking Intention During Simulated Driving. In Proceedings of the 2022 China Automation Congress (CAC), Xiamen, China, 25–27 November 2022; pp. 4356–4361.
- Zhang, W.; Qin, Y.; Tao, X. MMPHGCN: A Hypergraph Convolutional Network for Detection of Driver Intention on Multimodal Physiological Signals. *IEEE Signal Process. Lett.* **2024**, *31*, 1730–1734. [\[CrossRef\]](#)
- Lee, S.-M.; Kim, J.-W.; Lee, S.-W. Detecting Driver's Braking Intention Using Recurrent Convolutional Neural Networks Based EEG Analysis. In Proceedings of the 2017 4th IAPR Asian Conference on Pattern Recognition (ACPR), Nanjing, China, 26–29 November 2017; pp. 840–845.
- Liang, X.; Yu, Y.; Liu, Y.; Liu, K.; Liu, Y.; Zhou, Z. EEG-based emergency braking intention detection during simulated driving. *Biomed. Eng. Online* **2023**, *22*, 65. [\[CrossRef\]](#) [\[PubMed\]](#)
- Teng, T.; Bi, L.; Liu, Y. EEG-Based Detection of Driver Emergency Braking Intention for Brain-Controlled Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 1766–1773. [\[CrossRef\]](#)
- Bi, L.; Wang, H.; Teng, T.; Guan, C. A Novel Method of Emergency Situation Detection for a Brain-Controlled Vehicle by Combining EEG Signals With Surrounding Information. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2018**, *26*, 1926–1934. [\[CrossRef\]](#)
- Hernandez, L.G.; Mozos, O.M.; Ferrandez, J.M.; Antelis, J.M. EEG-Based Detection of Braking Intention Under Different Car Driving Conditions. *Front. Neurosci.* **2018**, *12*, 29. [\[CrossRef\]](#) [\[PubMed\]](#)
- Fu, J.; Zhang, X.; Yu, W.; Li, J.; Atia, M.M.; Wang, H.; Li, C.; Hao, Z. Decoding Passenger's EEG Signals From Encountering Emergency Road Events. In Proceedings of the 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), Macau, China, 8–12 October 2022; pp. 2214–2219.
- Tan, J.; Bi, R.; Guo, W.; Li, L.; Wang, Y. On Driver's Workload in Dangerous Scenes Based on EEG Data. In Proceedings of the 2020 IEEE 9th Data Driven Control and Learning Systems Conference (DDCLS), Liuzhou, China, 20–22 November 2020; pp. 1226–1233.
- Zhang, X.; Xiao, X.; Yang, Y.; Hao, Z.; Li, J.; Wang, H. EEG Signal Analysis for Early Detection of Critical Road Events and Emergency Response in Autonomous Driving. In Proceedings of the 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC), Bilbao, Spain, 24–28 September 2023; pp. 1706–1712.
- Duma, G.M.; Mento, G.; Manari, T.; Martinelli, M.; Tressoldi, P. Driving with Intuition: A Preregistered Study about the EEG Anticipation of Simulated Random Car Accidents. *PLoS ONE* **2017**, *12*, e0170370. [\[CrossRef\]](#) [\[PubMed\]](#)
- Liang, J.; Wang, Z.; Han, J.; Zhang, L. EEG-based driving intuition and collision anticipation using joint temporal-frequency multi-layer dynamic brain network. *Front. Neurosci.* **2024**, *18*, 1421010. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhang, X.; Yan, X. Predicting collision cases at unsignalized intersections using EEG metrics and driving simulator platform. *Accid. Anal. Prev.* **2023**, *180*, 106910. [\[CrossRef\]](#)

25. Zheng, C.; Zi, M.; Jiang, W.; Chu, M.; Zhang, Y.; Yuan, J.; Zhou, G.; Gong, J. Annotating Covert Hazardous Driving Scenarios Online: Utilizing Drivers' Electroencephalography (EEG) Signals. In Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA), London, UK, 29 May–2 June 2023; pp. 1456–1462.
26. Guo, Z.; Pan, Y.; Zhao, G.; Zhang, J.; Dong, N. Recognizing Hazard Perception in a Visual Blind Area Based on EEG Features. *IEEE Access* **2020**, *8*, 48917–48928. [CrossRef]
27. Li, H.; Chang, R.; Sui, X. The effect of the degree and location of danger in traffic hazard perception: An ERP study. *NeuroReport* **2022**, *33*, 215–220. [CrossRef] [PubMed]
28. Wang, H.; Zhang, X.; Li, J.; Li, B.; Gao, X.; Hao, Z.; Fu, J.; Zhou, Z.; Atia, M. Driving risk cognition of passengers in highly automated driving based on the prefrontal cortex activity via fNIRS. *Sci. Rep.* **2023**, *13*, 15839. [CrossRef]
29. Kastrati, A.; Płomecka, M.B.; Pascual, D.; Wolf, L.; Gillioz, V.; Wattenhofer, R.; Langer, N. EEGEyeNet: A simultaneous electroencephalography and eye-tracking dataset and benchmark for eye movement prediction. *arXiv* **2021**, arXiv:2111.05100.
30. Fuhl, W.; Zabel, S.; Harbig, T.; Moldt, J.-A.; Wietek, T.F.; Herrmann-Werner, A.; Nieselt, K. One step closer to EEG based eye tracking. In Proceedings of the 2023 Symposium on Eye Tracking Research and Applications, Tübingen, Germany, 30 May–2 June 2023; pp. 1–7.
31. Himmelberg, M.M.; Segala, F.G.; Maloney, R.T.; Harris, J.M.; Wade, A.R. Decoding Neural Responses to Motion-in-Depth Using EEG. *Front. Neurosci.* **2020**, *14*, 581706. [CrossRef] [PubMed]
32. Fazlyyyakhmatov, M.; Zvezdochkina, N.; Antipov, V. The EEG Activity during Binocular Depth Perception of 2D Images. *Comput. Intell. Neurosci.* **2018**, *2018*, 5623165. [CrossRef]
33. Dosovitskiy, A.; Ros, G.; Codevilla, F.; Lopez, A.; Koltun, V. CARLA: An Open Urban Driving Simulator. In Proceedings of the 1st Annual Conference on Robot Learning, Proceedings of Machine Learning Research, Mountain View, CA, USA, 13–15 November 2017; pp. 1–16.
34. Greibe, P.J.; Scion-DTU. Braking Distance, Friction and Behaviour. *Trafitec*. 2007. Available online: <https://api.semanticscholar.org/CorpusID:145815000> (accessed on 29 December 2024).
35. Brackstone, M.; McDonald, M. Car-following: A historical review. *Transp. Res. Part F Traffic Psychol. Behav.* **1999**, *2*, 181–196. [CrossRef]
36. Cheng, G.; Sun, X.; Han, J. Modeling driver distance recognition and speed perception at night for freeway speed limit selection in China. *Int. J. Saf. Secur. Eng.* **2014**, *4*, 143–153. [CrossRef]
37. Lertworawanich, P. Safe-following distances based on the car-following model. In Proceedings of the PIARC International seminar on Intelligent Transport System (ITS) in Road Network Operations, Kuala Lumpur, Malaysia, 14–16 August 2006.
38. Alharbi, N. A novel approach for noise removal and distinction of EEG recordings. *Biomed. Signal Process. Control.* **2018**, *39*, 23–33. [CrossRef]
39. Sun, C.; Mou, C. Survey on the research direction of EEG-based signal processing. *Front. Neurosci.* **2023**, *17*, 1203059. [CrossRef]
40. Chaddad, A.; Wu, Y.; Kateb, R.; Bouridane, A. Electroencephalography Signal Processing: A Comprehensive Review and Analysis of Methods and Techniques. *Sensors* **2023**, *23*, 6434. [CrossRef] [PubMed]
41. Jafari, M.; Shoeibi, A.; Khodatars, M.; Bagherzadeh, S.; Shalbaf, A.; Garcia, D.L.; Gorriz, J.M.; Acharya, U.R. Emotion recognition in EEG signals using deep learning methods: A review. *Comput. Biol. Med.* **2023**, *165*, 107450. [CrossRef] [PubMed]
42. Leong, D.; Do, T.T.; Lin, C.T. Ventral and Dorsal Stream EEG Channels: Key Features for EEG-Based Object Recognition and Identification. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2023**, *31*, 4862–4870. [CrossRef]
43. Rehman, M.; Anwer, H.; Garay, H.; Alemany-Iturriaga, J.; Diez, I.T.; Siddiqui, H.U.R.; Ullah, S. Decoding Brain Signals from Rapid-Event EEG for Visual Analysis Using Deep Learning. *Sensors* **2024**, *24*, 6965. [CrossRef] [PubMed]
44. Singh, A.K.; Krishnan, S. Trends in EEG signal feature extraction applications. *Front. Artif. Intell.* **2022**, *5*, 1072801. [CrossRef] [PubMed]
45. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [CrossRef]
46. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; Scholkopf, B. Support vector machines. *IEEE Intell. Syst. Their Appl.* **1998**, *13*, 18–28. [CrossRef]
47. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
48. Taud, H.; Mas, J.F. Multilayer Perceptron (MLP). In *Geomatic Approaches for Modeling Land Change Scenarios*; Camacho Olmedo, M.T., Paegelow, M., Mas, J.-F., Escobar, F., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 451–455.
49. Lawhern, V.J.; Solon, A.J.; Waytowich, N.R.; Gordon, S.M.; Hung, C.P.; Lance, B.J. EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces. *J. Neural. Eng.* **2018**, *15*, 056013. [CrossRef] [PubMed]
50. Schirrmeister, R.T.; Springenberg, J.T.; Fiederer, L.D.J.; Glasstetter, M.; Eggensperger, K.; Tangermann, M.; Hutter, F.; Burgard, W.; Ball, T. Deep learning with convolutional neural networks for EEG decoding and visualization. *Hum. Brain Mapp.* **2017**, *38*, 5391–5420. [CrossRef] [PubMed]
51. Dey, R.; Salem, F.M. Gate-variants of gated recurrent unit (GRU) neural networks. In Proceedings of the 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS), Boston, MA, USA, 6–9 August 2017; pp. 1597–1600.

52. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
53. Bo, K.; Cui, L.; Yin, S.; Hu, Z.; Hong, X.; Kim, S.; Keil, A.; Ding, M. Decoding the temporal dynamics of affective scene processing. *Neuroimage* **2022**, *261*, 119532. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.