POLYTECHNIC UNIVERSITY OF TURIN

Master's Degree in Data Science and Engineering

Master's Degree Thesis

Reconstructing Human Gaze Behavior in Stroop Test from EEG Data Using Inverse Reinforcement Learning

Supervisors

Prof. Paolo GARZA

Prof. Soroush KORIVAND

Candidate

Ali ABBASI

December 2024

Abstract

This research investigates the reconstruction of human gaze behavior during Stroop tests by integrating EEG data with Inverse Reinforcement Learning (IRL). Centered on cognitive neuroscience, the study seeks methods that not only predict gaze patterns but also elucidate the underlying motivations guiding attention shifts in cognitively demanding tasks. Utilizing the Stroop test—a standard measure of executive function and selective attention—eye-tracking and EEG data were collected from ten participants across congruent and incongruent task conditions. The study employed an IRL framework enhanced with Generative Adversarial Imitation Learning (GAIL) and Proximal Policy Optimization (PPO) to model gaze behavior through a reward-based approach. While the integration of EEG signals aimed to capture cognitive load and inform fixation choices within a dynamic visual model(Dynamic Contextual Beliefs), evaluation using metrics such as Target Fixation Probability and MultiMatch Sequence Score revealed that models utilizing fixation data alone (IRL-Image) performed comparably to those incorporating EEG data (IRL-EEG). These findings suggest that fixation patterns possess substantial predictive power in directing gaze toward task-relevant areas, even under high cognitive load conditions, without the need for additional EEG data. This research contributes to the development of efficient gaze prediction models in cognitive neuroscience and highlights the potential for simplified approaches in clinical diagnostics for conditions affecting executive function.

Summary

In the realm of cognitive neuroscience, eye-tracking and electroencephalography (EEG) technologies have become invaluable tools for understanding how individuals allocate attention in response to complex cognitive demands. The Stroop test, extensively used in cognitive psychology, challenges participants' executive functions by presenting tasks that require the inhibition of automatic responses—essentially testing selective attention, processing speed, and cognitive flexibility. Eye-tracking captures observable attention shifts, while EEG provides neural correlates, together offering a multidimensional view of cognitive engagement.

This study combines these technologies with an Inverse Reinforcement Learning (IRL) framework to explore the underlying motivations driving gaze behavior in the Stroop test. Traditional approaches often focus on predicting gaze patterns through machine learning but lack interpretability concerning the reward structures motivating these behaviors. In contrast, IRL enables the inference of reward-based motivations, which is essential for understanding the dynamic adjustments in gaze patterns as cognitive demands shift.

Participants and Experimental Setup: The study involved 10 participants who completed 20 Stroop trials, designed with both congruent and incongruent task conditions to induce varying cognitive loads. EEG data was captured at 250 Hz using a 64-channel system, while gaze data was recorded at 60 Hz. Each Stroop stimulus was displayed for five seconds, followed by a five-second interval, balancing cognitive engagement and resting periods.

Data Preprocessing: EEG and eye-tracking data underwent extensive preprocessing. EEG signals were filtered to remove artifacts, and fixation points were mapped to a grid of 640 patches. EEG feature extraction focused on power spectral density and artifact suppression, while eye-tracking data was processed to identify exact fixation locations and durations.

IRL Framework and Model Architecture: The IRL framework is designed to model gaze behavior based on inferred reward functions within a structured environment. The model employs Generative Adversarial Imitation Learning (GAIL), with a generator that synthesizes gaze paths and a discriminator that evaluates their alignment with actual human gaze patterns. Proximal Policy Optimization

(PPO) refines the generator's policy by updating fixation probabilities based on EEG-derived cognitive cues, enabling the model to adjust gaze predictions dynamically. A Dynamic Contextual Belief (DCB) structure organizes the visual field into low- and high-resolution regions, representing central (foveal) and peripheral vision areas, respectively. EEG data further refines the gaze prediction model, making it responsive to real-time cognitive shifts.

The model's core functionality centers on accurately predicting human-like scanpaths in response to EEG cues. States in the IRL environment are defined by visual context and EEG data, dynamically updated to mirror human fixation strategies. Actions represent fixation points within the grid, guided by EEG-informed cues and reinforced by task-aligned rewards.

EEG Feature Integration: EEG data informs the policy network, allowing the model to prioritize task-relevant areas during periods of high cognitive load. By adjusting fixation probabilities according to EEG-derived attention markers, the model emulates real-time cognitive adjustments, crucial for tasks involving cognitive control like the Stroop.

Reward and Policy Learning: The discriminator provides reward feedback by assessing the "realness" of generated scanpaths, reinforcing actions that mirror human attention shifts. PPO facilitates stable learning by constraining updates to fixation probabilities, allowing the model to replicate naturalistic gaze patterns across iterations.

Model Evaluation: The model's performance was evaluated using metrics such as Target Fixation Probability Area Under the Curve (TFP-AUC), Probability Mismatch, Sequence Score, MultiMatch Analysis, and Scanpath Ratio (SP Ratio). Comparisons were made between EEG-enhanced models (IRL-EEG) and non-EEG baselines (IRL-Image), highlighting the significance of EEG cues in predicting gaze behavior under varying cognitive loads. However, the results indicated that integrating EEG data did not lead to significant improvements in gaze prediction accuracy compared to models utilizing fixation data alone. Both the IRL-EEG and IRL-Image models exhibited comparable performances across most metrics, suggesting that fixation patterns alone possess substantial predictive power in directing gaze toward task-relevant areas, even under high cognitive load conditions like incongruent Stroop trials.

Comparison with Baseline Models: The EEG-enhanced IRL model showed similar or slightly better alignment with human gaze patterns in specific metrics, such as Probability Mismatch and Sequence Score. However, overall, the inclusion of EEG data did not provide a substantial advantage over the fixation-only model, indicating that fixation data alone is sufficient for accurate gaze behavior reconstruction in the context of the Stroop test.

Implications and Future Directions: The findings suggest that while EEG data offers additional cognitive state information, its integration within the IRL framework does not significantly enhance gaze prediction in structured tasks like the Stroop test. This has important implications for the design of gaze prediction models, indicating that simpler models using only eye-tracking data can achieve comparable accuracy without the added complexity of EEG integration. From a clinical perspective, eye-tracking metrics alone could serve as an efficient and cost-effective tool for assessing cognitive functions and detecting neurodegenerative or attention-related disorders. Future research should explore alternative methods for integrating neural data, expand participant diversity, and apply the framework to a broader range of cognitive tasks to fully ascertain the potential benefits of multimodal data integration in gaze behavior modeling.

Acknowledgements

I would like to express my deepest gratitude to all those who have supported me throughout the journey of completing this thesis.

First and foremost, I am profoundly thankful to my supervisors, Prof. Paolo Garza, and Prof. Soroush Korivand for their invaluable guidance, unwavering support, and insightful feedback. Their expertise and encouragement have been instrumental in shaping this research.

On a personal note, I would like to thank my family for their unconditional love and support. To my friends, your understanding and encouragement have been a source of strength throughout this process.

Lastly, I wish to acknowledge God and myself for being a strong team and not giving up.

Thank you all for being part of this journey.

"Ali Abbasi"

Table of Contents

List of Tables

List of Figures

Acronyms

EEG
: Electroencephalography

IRL
: Inverse Reinforcement Learning

DCB
: Dynamic Contextual Belief

PPO
: Proximal Policy Optimization

GAIL
: Generative Adversarial Imitation Learning

IRL-Image
: Inverse Reinforcement Learning for Image Data

IRL-EEG
: Inverse Reinforcement Learning for EEG Data

TFP
: Target Fixation Probability

TFP-AUC
: Target Fixation Probability Area Under Curve

ADHD
: Attention-Deficit/Hyperactivity Disorder

ERP

Event-Related Potential

GAN

Generative Adversarial Network

PSD

Power Spectral Density

FPOG

Fixation Point of Gaze

HR

High Resolution

LR

Low Resolution, Learning Rate

IOR

Inhibition of Return

Coef

Coefficient

Chapter 1

Introduction

The neural overlap between cognitive control and eye movement regulation, including saccadic shifts and fixation, implies a close association between task demands and gaze patterns during the Stroop task. Cognitive load, as modulated by the Stroop task, typically affects gaze behavior, influencing the strategies used to manage competing cognitive demands. These eye movement patterns offer clues about the cognitive load, attention allocation, and executive function, all vital areas of study in both psychological research and potential clinical applications. Studies have shown that Stroop task performance can reflect underlying cognitive mechanisms and identify variations across individual cognitive capacities [1]. Consequently, non-invasive methods like eye-tracking combined with EEG can serve as powerful tools to explore these mechanisms with precision.

Recently, the integration of EEG with eye-tracking has emerged as a prominent area of research, enabling a more nuanced analysis of cognitive processing during attentionally demanding tasks. EEG provides rich data on neural activity, reflecting real-time responses to task demands, while eye-tracking captures the observable outcomes of these processes through gaze behavior. By employing EEG in the Stroop test, it is possible to correlate eye movement patterns directly with neural markers of cognitive load and attentional shifts. However, traditional diagnostic approaches have often used machine learning models that may lack interpretability and flexibility, relying on predefined statistical correlations to infer behavior. Such approaches, while insightful, do not fully capture the adaptive nature of human cognition in real-time.

In addressing these limitations, this project utilizes Inverse Reinforcement Learning (IRL) based on Generative Adversarial Imitation Learning(GAIL) algorithm to model and reconstruct the human gaze behavior during the Stroop test based on human gaze fixations and EEG data. IRL has shown promise in capturing reward-based behavioral patterns across complex tasks by inferring the “reward” functions underlying observed actions [2]. By applying IRL to gaze fixations and EEG data from Stroop task trials, this project aims to infer the underlying cognitive "reward functions" associated with managing interference, processing speed, and attentional control. This approach allows for a dynamic interpretation of gaze and neural patterns, offering insights into the interaction between task demand and cognitive control mechanisms. In summary, the study introduces an EEG-based IRL framework tailored to the Stroop test, integrating gaze data to map cognitive strategies during task execution. This framework aims to enhance our understanding of cognitive control and attentional mechanisms, providing a novel methodological approach to studying gaze and EEG in cognitive neuroscience and gain the ability of reproducing the human gaze during the Stroop test.

1.1 Background and Motivation

The investigation of cognitive processes through the lens of eye movement and brain activity has revealed critical insights into how humans engage with complex tasks, enabling a finer understanding of underlying cognitive mechanisms. Specifically, the Stroop test, a cornerstone in neuropsychological research, challenges participants' cognitive control by requiring them to identify the color of a word while ignoring the word's conflicting meaning. For example, writing the word "Blue" in red color is wrong(incongruent), and writing the word "Yellow" in yellow color is correct(congruent). This test evokes distinct cognitive responses, thereby serving as an ideal basis for studying attention, executive function, and response inhibition. Tracking gaze behavior during such tasks provides valuable data that can quantify these cognitive processes and uncover the mental strategies individuals employ to manage conflict, selective attention, and cognitive load. Besides, there's the potential of using this approach to use in the clinical atmosphere to detect disorders particularly neurodegenerative mental disorders like Alzheimer or even ADHD.

Recent advancements in eye-tracking and EEG technologies have significantly improved the granularity with which these processes can be analyzed. Eye tracking records metrics like fixation duration, saccade patterns, and pupil dilation, all of which offer insights into how attention shifts in response to stimuli. EEG, on the other hand, captures neural correlates of cognitive engagement, providing an avenue for examining brain activity in real-time. These tools, when used together, offer a multidimensional perspective on cognitive functioning during tasks like the Stroop test, where reaction time, attentional shifts, and cognitive load are key components. This project aims to build upon these developments by using inverse reinforcement learning (IRL) to model gaze behavior in the Stroop test, providing an analytical framework for understanding the motivations and decision-making processes underpinning gaze shifts during cognitive tasks. Thus, attempting to reconstruct human gaze by combining gaze scanpaths with EEG data might provide valuable understanding for some mental disorders like Alzheimer or ADHD.

1.2 Problem Statement

In the realm of cognitive neuroscience, understanding the intricate mechanisms that govern human gaze behavior during cognitively demanding tasks remains a formidable challenge. Despite advancements in eye-tracking and EEG technologies, existing analytical models fall short in elucidating the underlying cognitive processes that drive gaze patterns, especially in tasks that require high levels of cognitive control, such as the Stroop test. Traditional machine learning approaches predominantly focus on predicting gaze trajectories based on observable data without delving into the motivations and reward structures that influence such behaviors. These models often lack interpretability and fail to capture the adaptive and goal-directed nature of human visual attention under varying levels of cognitive load.

The Stroop test presents a unique opportunity to investigate these underlying processes due to its inherent cognitive interference and demand for selective attention. Participants must resolve conflicts between the semantic meaning of words and their font colors, engaging executive functions and attentional control mechanisms. However, current models inadequately address how individuals adapt their gaze behavior in response to these cognitive demands. There is a significant gap in our ability to infer the implicit reward functions that guide gaze shifts and fixation patterns during the Stroop task, which is crucial for understanding the cognitive strategies employed to manage interference and allocate attention effectively.

Moreover, while EEG provides a rich source of data reflecting neural activity and cognitive load, integrating this information with gaze behavior analysis has proven to be complex. The dynamic interplay between neural signals and eye movements during tasks involving cognitive interference is not fully understood. Existing methodologies do not effectively leverage EEG data to inform predictions about gaze behavior, thereby missing an opportunity to correlate neural markers of cognitive load with visual attention strategies.

The central problem addressed in this dissertation is the development of a framework that can effectively model and reconstruct human gaze behavior during the Stroop test by integrating EEG data with inverse reinforcement learning (IRL) techniques. Specifically, the challenge lies in decoding the complex interplay between neural activity and gaze patterns to infer the underlying reward-based motivations driving visual attention. By leveraging IRL, the goal is to move beyond mere prediction of gaze trajectories and instead provide insights into the cognitive processes and reward structures that underpin gaze behavior under varying levels of cognitive load.

This problem is significant for several reasons:

1. **Lack of Interpretability in Existing Models**: Traditional machine learning models often act as black boxes, providing little insight into why certain gaze patterns occur. Understanding the 'why' behind gaze behavior is essential for advancing theories of cognitive processing and attention.

2. **Adaptive Nature of Human Gaze**: Human gaze behavior is inherently adaptive, constantly adjusting in response to changes in cognitive load and task demands. Current models do not adequately capture this adaptability, limiting their applicability in real-world scenarios where cognitive demands fluctuate.

3. **Integration of EEG Data**: EEG offers real-time insights into cognitive load and neural activity, but integrating this data into gaze prediction models remains challenging. There is a need for frameworks that can seamlessly combine EEG and eye-tracking data to provide a holistic view of cognitive processes.

4. **Clinical Implications**: Understanding the reward-based motivations behind gaze behavior has implications for diagnosing and treating neurodegenerative disorders and attention-related conditions such as ADHD. Improved models could lead to non-invasive diagnostic tools and interventions that monitor gaze behavior and cognitive load.

5. **Advancement of IRL Techniques**: Applying IRL to cognitive neuroscience represents an innovative approach that can enhance our ability to infer underlying motivations from observed behaviors, providing a more nuanced understanding of human cognition.

Addressing this problem involves several challenges:

- **Data Complexity**: EEG and eye-tracking data are high-dimensional and noisy, requiring sophisticated preprocessing and feature extraction techniques to be useful in modeling.

- **Model Design**: Developing an IRL framework that can handle the complexities of cognitive tasks like the Stroop test necessitates careful consideration of state and action representations, reward functions, and policy learning algorithms.

- **Validation**: Ensuring that the model accurately reflects human gaze behavior and cognitive processes requires rigorous validation against empirical data, including statistical analyses and possibly experimental replication.

In summary, the problem this research tackles is the development of an interpretable and adaptive computational framework that integrates EEG data with IRL to model and reconstruct human gaze behavior during the Stroop test. By inferring the implicit reward functions and cognitive motivations behind gaze patterns, the research aims to provide deeper insights into the cognitive strategies individuals employ under varying levels of cognitive load, with potential applications in both theoretical and clinical domains.

1.3 Objectives and Contributions

The objective of this dissertation is to develop a novel framework that uses IRL to model and decode the gaze behavior of individuals undertaking the Stroop test, with combining gaze and EEG data providing additional depth to this exploration. Specific aims include:

1. **Developing an IRL model tailored to the Stroop task**: This involves constructing an IRL framework that can infer and reconstruct the reward functions underlying gaze behavior during a task characterized by different levels of cognitive load.

2. **Integrating EEG and eye-tracking data to enhance model accuracy**: By synchronizing EEG data with gaze patterns, this project aims to correlate neural activity with inferred gaze motivations, thereby adding a layer of neural validation to the inferred reward functions.

3. **Advancing understanding of attentional control and decision-making strategies**: This study seeks to shade a light on the motivations that guide gaze shifts in the Stroop task, contributing valuable insights into cognitive control mechanisms and attentional dynamics.

Through these objectives, this research contributes an IRL-based methodology for gaze analysis, with applications in cognitive neuroscience, neuropsychological assessments, and potential use in diagnostic settings for conditions involving executive function deficits.

1.4 Structure of the Dissertation

This dissertation is structured as follows:

1. **Chapter 2: Related Work** – This chapter reviews relevant literature on eye tracking in cognitive neuroscience, particularly for neurodegenerative and cognitive disorders. It also explores studies integrating EEG and eye-tracking

data, the application of IRL in behavioral analysis, and prior work on gaze prediction and scanpath analysis.

2. **Chapter 3: Materials and Method** – This chapter describes the study's methodology, including details on participants, experimental design, EEG and eye-tracking data collection, and data preprocessing. A thorough overview of the Stroop test setup and the integration of EEG and eye-tracking data are provided, followed by a description of the IRL framework and neural data analysis methods.

3. **Chapter 4: Scanpath Prediction Framework** – This section presents the design of the state and action models used in the IRL framework, explaining how EEG features guide the gaze modeling process.

4. **Chapter 5: Experiments** – This chapter details the experimental setup, comparing various scanpath prediction methods and evaluating their performance with metrics such as target fixation probability and scanpath sequence scoring.

5. **Chapter 6: Results** – The results chapter presents a detailed analysis of gaze path prediction accuracy, model performance across different cognitive loads, and the validation of inferred rewards with neural data correlations.

6. **Chapter 7: Discussion** – This chapter interprets the results, examining the implications of the findings for cognitive control theories and potential clinical applications, and discusses limitations and future research directions.

7. **Chapter 8: Conclusion** – The dissertation concludes with a summary of key findings, contributions, and the anticipated impact of this research on cognitive neuroscience and neuropsychology.

Chapter 2

Related Work

In the interdisciplinary field examining cognitive processing through gaze and EEG signals, extensive research has explored the integration of eye-tracking and electroencephalography (EEG) to provide insight into cognitive control, attentional shifts, and decision-making under varied cognitive loads. The utility of these technologies is demonstrated across a spectrum of applications, from identifying cognitive load during specific tasks to understanding neurological mechanisms in health and disease. The Stroop test, a classic paradigm in cognitive psychology, is widely recognized for its ability to evoke cognitive conflict and measure executive function. The incorporation of eye-tracking and EEG during the Stroop test represents a novel avenue for decoding gaze patterns and underlying neural mechanisms.

Current literature reflects the growing interest in understanding cognitive control processes by leveraging inverse reinforcement learning (IRL) to infer the "reward" systems guiding eye movement behaviors, and then give the ability of reproducibility to the human gaze during this task. This method has gained traction due to its capability to uncover underlying motivations in observed behavior, transcending the limitations of traditional predictive models which often lack interpretability in various cognitive scenarios. In the context of gaze behavior, IRL can provide a powerful framework for modeling decision-making processes, particularly during tasks like the Stroop test, where participants must frequently override instinctual responses to manage conflicting information.

Studies integrating EEG and eye-tracking for behavioral analysis underscore the distinct neural signatures associated with saccadic shifts, fixation durations, and other gaze metrics in tasks involving varying levels of cognitive demand(in this research the focus is only on the fixations location). However, while these approaches have led to valuable findings, many rely on static correlations and do not fully capture the adaptive, goal-directed nature of human gaze patterns. Integrating IRL with EEG data in the Stroop task may bridge this gap by enabling a dynamic assessment of attention, executive function, and underlying neural processes as individuals navigate task-induced cognitive interference.

2.1 Eye Tracking for Neurodegenerative Disorders

Eye-tracking technology has shown promise as a non-invasive tool for identifying early biomarkers of neurodegenerative disorders, particularly those affecting cognitive control and visual processing[3]. Studies demonstrate that eye movement metrics such as saccadic velocity, fixation duration, and pupil dilation are effective indicators of neurological status and cognitive load, allowing researchers to infer cognitive impairment levels in disorders like Alzheimer's and Parkinson's disease[4]. In recent years, eye tracking has been incorporated into various cognitive assessments, including tasks like reading and arithmetic, to detect abnormal gaze patterns that indicate executive function deficits and attentional shifts associated with neurodegenerative disorders[5].

The Stroop task, with its demands on cognitive control, presents a unique framework for eye-tracking applications. In such tasks, discrepancies in eye movement patterns can reveal deficits in selective attention and response inhibition—symptoms common to neurodegenerative diseases. Research on using eye-tracking for monitoring cognitive function underscores its potential for early diagnosis and tracking disease progression, especially when combined with EEG data to observe neural responses to visual stimuli and cognitive interference[6]. This integration provides a comprehensive toolset for assessing neural function and cognitive control in both clinical and research settings, facilitating detailed analysis of cognitive responses to interference and the ability to suppress automatic responses, which are core aspects of executive function evaluated by the Stroop task.

2.2 Studying Eye Movement and Brain Activities Using EEG

Combining eye tracking with EEG has enriched research on cognitive and neural responses to visual and attentional tasks, offering an in-depth view of how gaze patterns and brain activity align during decision-making and attention regulation[7]. This dual-modality approach allows researchers to correlate gaze shifts with neural markers of cognitive load, providing a real-time assessment of attention allocation and cognitive processes. EEG measures, such as event-related potentials (ERPs), have proven effective in capturing the brain's immediate response to visual stimuli, facilitating a more nuanced understanding of the neural underpinnings of eye movement control, especially in tasks requiring complex cognitive processing like the Stroop task[8].

For cognitive tasks like the Stroop, EEG can capture frontal and parietal activations associated with cognitive interference and response inhibition, while eye-tracking data elucidates the gaze behavior corresponding to these neural events. Research in this area has shown distinct EEG signatures tied to decision-making and cognitive control, particularly in tasks where participants must manage conflicting stimuli, such as color-word incongruence. The joint analysis of EEG and eye-tracking data in cognitive tasks not only enriches the understanding of neural dynamics but also provides diagnostic insights into cognitive function, making this approach valuable for studying both healthy cognition and cognitive impairments.

2.3 Eye Movement Data Analysis

The analysis of eye movement data has evolved significantly, with advanced computational methods enhancing the extraction and interpretation of gaze metrics across various cognitive tasks. Machine learning techniques, ranging from basic classifiers to sophisticated deep learning architectures, are increasingly utilized to identify and predict patterns in eye movement data, focusing on metrics such as fixation duration, saccadic speed, and scanpath regularities[9]. In the context of cognitive assessments like the Stroop task, eye movement data analysis can offer insights into how cognitive interference impacts gaze behavior, revealing adaptations in attention and visual processing strategies.

Despite progress, traditional machine learning approaches often lack flexibility in modeling the adaptive nature of human gaze, particularly under the cognitive demands posed by tasks like the Stroop. Analyzing scanpaths and gaze patterns during Stroop tasks requires algorithms that can handle variability across individuals and cognitive states. These limitations highlight the need for dynamic and context-sensitive methods, such as inverse reinforcement learning (IRL), which is more adept at capturing the evolving reward-based strategies underpinning gaze shifts in response to cognitive interference and attentional demands.

2.4 IRL for Human Behavior Data Analysis

Inverse Reinforcement Learning (IRL) is emerging as a powerful tool for decoding complex human behaviors, allowing researchers to infer the reward structures driving observed actions in various settings, from robotics to cognitive science[10]. By applying IRL to human gaze data, researchers can identify the implicit “rewards” or motivations behind gaze behaviors, facilitating a deeper understanding of decision-making and attentional strategies. This methodology is particularly relevant for studying tasks that require frequent shifts in cognitive strategy, as in the Stroop test, where participants must override automatic responses to achieve task objectives. Also the combination of IRL with GAN has been used for replicating and understanding human behaviors[11].

IRL's application to EEG and eye-tracking data enables researchers to construct a model that interprets gaze behavior as a product of both task demand and cognitive control, making it suitable for exploring reward-based adaptations in the Stroop task. For example, by associating EEG markers of cognitive load with gaze trajectories, IRL frameworks can dynamically infer the cognitive "reward functions" that guide attentional shifts, providing insights into how individuals manage interference and attentional demands. This approach is especially valuable in cognitive neuroscience, where understanding the interaction between cognitive goals and visual attention patterns has implications for clinical diagnosis and intervention.

2.5 Prior Work on Gaze Prediction and Behavioral Scanpath Analysis

Gaze prediction and scanpath analysis have become critical areas of study in understanding how cognitive processes shape visual exploration patterns. Scanpath analysis provides a window into task-related visual strategies, with research increasingly focusing on computational models that predict gaze trajectories based on underlying cognitive processes [12]. This approach is highly applicable in the Stroop task, where gaze paths can indicate how individuals allocate attention to resolve cognitive conflicts.

Prior studies have utilized both traditional machine learning and IRL frameworks to analyze and predict gaze behavior, often focusing on visual attention models that simulate human scanpaths in tasks of varying complexity. In tasks with high cognitive load, such as the Stroop, incorporating EEG data enhances the predictive accuracy of these models by providing neural context to gaze behavior. This enables a deeper analysis of how cognitive control and reward dynamics drive gaze patterns, shedding light on the ways individuals manage attention and response inhibition. The integration of EEG with gaze prediction models thus represents a step forward in capturing the interplay between neural processes and observable behaviors, particularly in challenging cognitive tasks. Fig 2.1

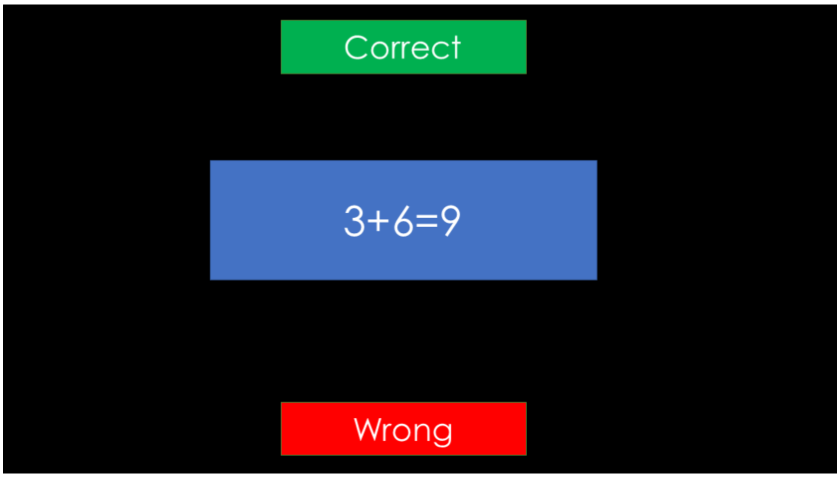

Figure 2.1: Prior work on arithmetic tasks

Chapter 3

Materials and Methods

This section outlines the methodological framework of the study, detailing the participant selection, data gathering procedures, and data preprocessing techniques adapted from established approaches in EEG and eye-tracking-based cognitive load assessments. The study follows a design informed by previous research on decoding cognitive behaviors through multimodal data integration, using eye-tracking and EEG to interpret gaze and brain activities.

3.1 Participants

11 adult participants were initially recruited for this study, comprising a mix of graduate and undergraduate students at the University of Alabama. All individuals had normal or corrected vision, ensuring minimal visual interference, except for one participant who required eyeglasses; due to corrupted data attributed to visual artifacts, their data was excluded from further analysis. After excluding this participant, data from 10 individuals (eight males and two females) were retained. The mean age of these participants was 29.11 years, with a standard deviation of 3.62 years. All participants were informed of the study's aims, methodology, and any potential risks, consistent with ethical standards for research involving human subjects. The study design received prior approval from the Institutional Review Board (IRB) at the University of Alabama.

3.2 Data Gathering

Data collection was conducted in a controlled laboratory setting, where participants completed a set of 20 Stroop trials, while their neural and ocular responses were recorded using EEG and eye-tracking devices. The Stroop task required participants to identify the color of presented words, which were either congruent (color and word matched) or incongruent (color and word did not match), thus creating conditions for assessing cognitive load and selective attention. To collect the EEG data the 64-channel EEG g.tec Nautilus Pro integrating with an EEG cap has been used, and for tracking the subjects' gaze, GP3 eye tracking device and GazePoint software has been used.

During each trial, slides were presented for 5 seconds, followed by a 5-second inter-stimulus interval, allowing participants time to reset visually and cognitively. Each participant completed 20 tasks (combination of congruent and incongruent tasks). This experimental design allowed a balanced analysis of both cognitive load conditions. Eye-tracking was facilitated by a GazePoint device, capturing pupil diameter variations and fixation metrics at a frequency of 60 Hz. EEG signals were recorded using a 64-channel g.tec Nautilus Pro EEG system, with a sampling rate of 250 Hz. This setup enabled high-resolution data collection necessary for subsequent analysis of cognitive load through EEG and eye-tracking synchronization. Fig. 3.1

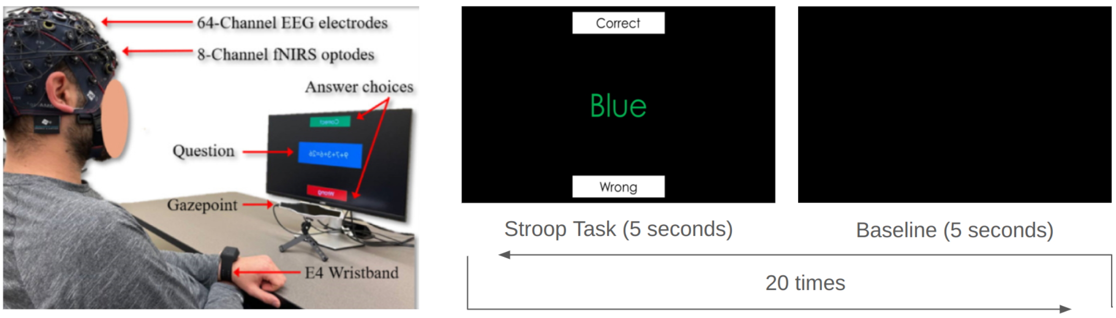

Figure 3.1: Stroop task process

3.3 Data Preprocessing

The initial phase of data preprocessing involved synchronization across the multimodal data streams, aligning EEG and eye-tracking data using timestamps recorded at the onset of the experiment. The data window extracted for analysis extended from 5 seconds before the start to 5 seconds after the experiment's conclusion, thus encompassing all relevant cognitive activity periods.

EEG Data Preprocessing

EEG data preprocessing was conducted using Brainstorm software, following standard procedures to minimize artifacts and enhance signal clarity. A high-pass filter (0.2 Hz) and a low-pass filter (32 Hz) were applied to maintain frequency ranges relevant to cognitive processing while eliminating low-frequency drift and high-frequency noise, including muscle artifacts. Using power spectral density analysis, bad channels were identified and removed from the dataset, while blink and heartbeat artifacts were suppressed using signal projectors based on the statistical characteristics of these physiological noise sources. Remaining EEG data were re-referenced to the instantaneous average of all active channels, ensuring consistent baseline signals across trials. In the next step, the solution used was calculating top-3 important features that yielded a 3-member set of numbers for each task for each user separately. Fig. 3.2

Figure 3.2: G.Tec Nautlilus Pro EEG System

Eye-Tracking Data Preprocessing

Eye-tracking data, specifically the position on the screen, was used as a primary indicator of cognitive load. Using the csv file provided by GazePoint there were the ability of accessing relative location of fixation on the screen ("FPOGX" and "FPOGY" which are scaled by 1), time of the fixation ("FPOGS"), the duration of the fixation ("FPOGD") and other valuable data. Among all of these data we should convert the relative location of the fixation to the exact location. Then convert to find the corresponding patch from the 32x20 patches of each image (640 patches at all and the size of each path was 16x16 pixels) to be able to work with the IRL environment efficiently. Fig. 3.3

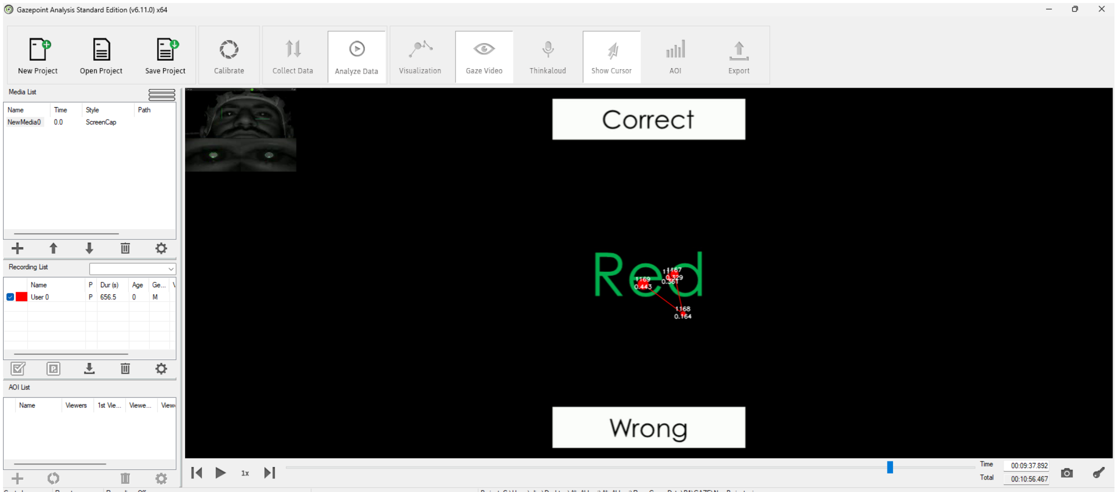

Figure 3.3: GazePoint Analyzer Software

Feature Extraction

EEG Feature Extraction

The EEG preprocessing employs a combination of high-pass and low-pass filters to isolate the frequency bands most relevant to cognitive processing (0.2–32 Hz). The key features are extracted using:

1. **Power Spectral Density (PSD)**:PSD analysis helps identify frequency-based activity and noise removal by analyzing power across various frequency bands.

2. **Artifact Rejection**:The data undergoes artifact rejection, including the suppression of blink and heartbeat artifacts via signal projectors to focus on cognitive signal quality.

3. **Re-referencing**:Channels are re-referenced against the mean signal of the active channels to ensure consistent baseline signals across trials.

Eye-Tracking Feature Extraction

Eye-tracking feature extraction in the study includes:

1. **Fixation Location and Duration**:The gaze coordinates (FPOGX, FPOGY) are converted to exact screen locations. The fixations are mapped to one of 640 grid patches for further analysis.

3.4 Overview of Experimental Design and Approach

This section provides a detailed overview of the experimental design, including the setup of the Stroop test and the integration of electroencephalography (EEG) and eye-tracking with GazePoint technology. By combining EEG and eye-tracking data, this study aims to investigate the cognitive load and attentional dynamics associated with Stroop task performance, thereby offering insights into neural and behavioral responses under conditions of cognitive interference.

3.4.1 STROOP Test Setup

The Stroop test is a widely recognized task in cognitive neuroscience, known for its utility in measuring executive functions such as selective attention, cognitive flexibility, and response inhibition. In this study, the Stroop test is designed to involve congruent and incongruent trials to induce varying levels of cognitive load. Each participant is presented with a color word displayed in a colored font, where the task requires identifying the font color while disregarding the semantic meaning of the word. Congruent trials present the word in a matching color (e.g., "Red" displayed in red font), whereas incongruent trials display the word in a non-matching color (e.g., "Red" displayed in blue font), thus creating conditions that require increased attentional control to suppress automatic reading responses.

Participants in this study consist of 11 adults; however, due to the eye-tracking setup requirements, one participant was excluded due to visual impairments that could affect gaze data quality. Consequently, data from 10 participants were retained for analysis. Each participant completed a total of 20 tasks that contained congruent and incongruent tasks, yielding comprehensive data for both congruent and incongruent conditions. This setup aims to elicit cognitive interference, allowing for a robust analysis of attentional and cognitive load metrics through eye-tracking and EEG.

3.4.2 Integration of EEG and Eye-Tracking with GazePoint

The integration of EEG and eye-tracking technologies enables a multifaceted analysis of cognitive load and gaze behavior during the Stroop test, capitalizing on the temporal precision of EEG and the spatial tracking capabilities of GazePoint. Eye-tracking data, captured at a frequency of 60 Hz, provides real-time metrics on pupil dilation, fixation duration, and saccadic movements—features indicative of the participant's cognitive engagement and attention allocation. GazePoint technology is calibrated to ensure accurate tracking, while it is able to record a lot of metrics like pupil diameter, fixation duration and others, the only ocular

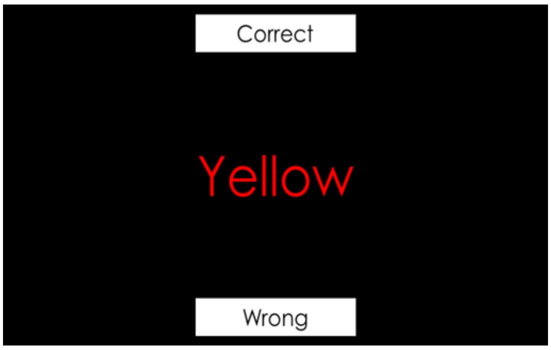

Figure 3.4: Incongruent Task

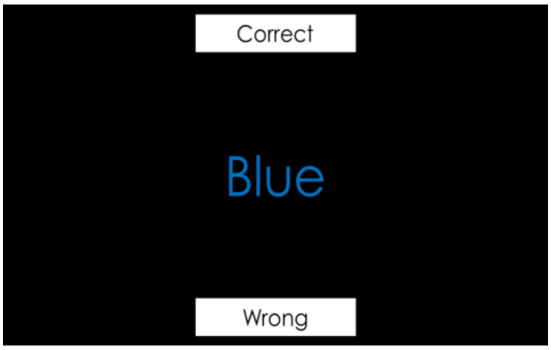

Figure 3.5: Congruent Task

value that employed in this research was fixation location that firstly was defined in a relative location, and afterward it changed to height and width pixel and then patch number.

EEG data acquisition utilizes a 64-channel g.tec Nautilus Pro system sampled at 250 Hz, providing high-resolution insights into neural activity during task performance. The EEG signal is preprocessed following standard protocols, including a 0.2–32 Hz band-pass filter to retain frequencies relevant to cognitive processing while removing noise and artifacts such as muscle movement and heartbeat interference. Data alignment between EEG and eye-tracking streams is performed using a synchronized timestamp, allowing for precise correlation between gaze patterns and neural responses. As previously explained after removing artifacts and noises, the top-3 important features have been chosen as follows: mean, standard deviation, and the number of peaks. These features were sampled and calculated over each distinct task of each subject.

3.5 Dataset Collection and Preprocessing

This section outlines the methods employed for dataset collection and preprocessing in a study combining EEG and eye-tracking data to analyze gaze behavior and cognitive load during Stroop tasks. We structured this approach based on protocols aligned with prior work from Gong et al. (2024)[13] and Zhu et al. (2022)[14], enhancing it with tailored experimental conditions for Stroop tasks. The methods span participant selection, EEG and eye-tracking data acquisition, and preprocessing steps, ensuring that the datasets are robust for analysis through Inverse Reinforcement Learning (IRL).

3.5.1 Participants and Experimental Conditions

Eleven participants were initially recruited for the study, comprising graduate and undergraduate students with normal or corrected vision. Due to specific visual requirements, data from one participant who wore glasses were excluded, resulting in a final sample of ten participants (eight males and two females) aged 29.11 ± 3.62 years. This participant cohort aligns with studies from Gong et al. (2024)[13], which utilized samples of similar demographics to control for variability in cognitive load responses.

Each participant performed 20 Stroop tests per session, yielding data across both congruent (color and word align) and incongruent (color and word conflict) task conditions. This setup allowed for evaluating cognitive interference and attentional control across two levels of task complexity, a design inspired by cognitive load studies such as those from Zhu et al. (2022)[14], which examined responses across tasks of varying cognitive demands. Each task was displayed for 5 seconds with a

5-second inter-stimulus interval, providing participants with a rest period between trials to minimize cognitive carryover effects. This protocol aims to balance cognitive load, providing a robust foundation for analyzing both high and low-interference tasks.

3.5.2 EEG Data Acquisition and Preprocessing

EEG data acquisition was conducted using a 64-channel g.tec Nautilus Pro EEG system with a 250 Hz sampling rate, as applied in similar neurophysiological studies by Gong et al. (2024)[13] and Zhu et al. (2022)[14]. EEG signals were recorded to capture neural responses to both congruent and incongruent Stroop stimuli, focusing on the frontal and parietal regions associated with attentional control and interference resolution.

The preprocessing pipeline, consistent with methods outlined in Zhu et al. (2022)[14], involved:

1. High-pass filtering at 0.2 Hz and low-pass filtering at 32 Hz to eliminate muscle artifacts and physiological noise.

2. Power spectral density analysis to remove channels with excessive noise.

3. Artifact rejection processes to suppress blink and heartbeat noise through independent component analysis.

After these preprocessing steps, each task segment was re-referenced to the average of all channels, establishing a uniform baseline across participants. Following Gong et al. (2024)[13], key EEG features, including mean and peak activity within task windows, were extracted for each task to facilitate a nuanced analysis of cognitive load associated with Stroop conditions.

3.5.3 Eye-Tracking Data Acquisition and Processing

Eye-tracking data were collected using the GP3 eye-tracking device, recorded at 60 Hz to capture fixation duration and location. Following the methods of Gong et al. (2024)[13], gaze coordinates (FPOGX, FPOGY) were recorded, and the data were processed for further analysis.

The fixation data were mapped onto a 640-grid matrix, representing screen positions as 32x20 patches of 16x16 pixels each. Fixations were then converted to specific patch numbers to synchronize with EEG data, aiding in the alignment with IRL-based gaze modeling. This mapping approach ensures that fixation points are accurately represented, enabling scanpath analysis and subsequent integration with EEG data. The Dynamic Contextual Belief (DCB) map is a critical feature within the model, representing the spatial focus of gaze by mapping fixated locations within a defined foveal radius. This DCB map dynamically updates after each fixation to capture areas of cognitive engagement, reflecting changes in visual attention over time and providing a structured map for gaze modeling.

3.5.4 Scanpath and Dataset Structuring

Following Zhu et al. (2022)[14] and Gong et al. (2024)[13], scanpaths for each Stroop trial were created by aggregating fixation sequences aligned with EEG data. Data were organized into time-segmented blocks, capturing EEG and eye-tracking metrics for both congruent and incongruent trials. This structuring provides a basis for the IRL framework, enabling dynamic modeling of gaze shifts under cognitive load and facilitating the reconstruction of gaze behaviors.

The dataset structure includes:

1. **Eye-tracking vectors**: Fixation coordinates

2. **EEG feature vectors**: Encapsulating mean, peak, and standard deviation for each task.

3. **Trial metadata**: Task type

4. **DCBs**: Low-res and High-res Dynamic Contextual Belief(DCB) maps for each task.

3.6 Inverse Reinforcement Learning(IRL) Framework

The Inverse Reinforcement Learning (IRL) framework developed in this study provides an environment (based on Yang et al. (2020)[12] project) that leverages EEG and eye-tracking data to decode and reconstruct gaze behavior during Stroop task trials. By defining and training an IRL-based system, this framework uses a multi-layered approach—specifically integrating Generative Adversarial Imitation Learning (GAIL) algorithm for behavioral replication, Proximal Policy Optimization (PPO) for policy refinement, and Dynamic Contextual Belief (DCB) for state representation. Together, these components contribute to modeling and predicting gaze trajectories as they respond to neural and task-specific cues.

Within this IRL framework, the environment is central, structured to represent the visual field in distinct low- and high-resolution regions, based on where the gaze is focused. This segmentation supports efficient processing by refining high-resolution details only in focal areas, which are derived from participant fixation data. Observations within this environment are designed to capture both state information from fixation points and EEG signals, creating a comprehensive dataset that reflects cognitive and visual processing in response to Stroop task demands.

The training process incorporates task-specific constraints and metrics to ensure alignment with human behavioral patterns observed in Stroop trials. These metrics include Target Fixation Probability (TFP) at various temporal stages, probability mismatch between model-predicted and human eye-tracking data, and cumulative fixations on target locations. Each metric is critical for evaluating the model's ability to replicate human-like gaze paths and adjust dynamically based on EEG signals.

By iteratively optimizing the generator (policy) and discriminator (reward estimator), this IRL framework progresses through a series of policy adjustments, where PPO is employed to refine fixation probabilities and action selection, closely mirroring EEG-guided gaze behaviors. Additionally, GAIL-driven adversarial training enables the model to discriminate between synthetic (model-generated) and real (human) gaze paths, using this feedback to enhance the fidelity of gaze prediction. Through these methods, the IRL framework achieves a robust simulation of attention-driven gaze behaviors under varying cognitive load conditions, as exemplified in the Stroop task.

3.6.1 GAIL: Generative Adversarial Imitation Learning

Generative Adversarial Imitation Learning (GAIL) serves as a core component of this framework, driving the replication of human-like gaze patterns by learning from real fixation sequences collected during Stroop task trials. Within the GAIL model, a generator-discriminator setup is used to mimic observed human behavior, where the generator synthesizes gaze paths and the discriminator evaluates their authenticity by comparing them against real human data. This adversarial structure enables the model to refine its gaze trajectory predictions iteratively, ensuring that generated scanpaths increasingly resemble actual human gaze paths.

The GAIL training process here uses human scanpath data as a reference, with real fixation sequences segmented based on task and congruence level, allowing the model to learn from distinct conditions of cognitive load inherent in the Stroop task. The generator model synthesizes gaze sequences by selecting fixation points within a patch-based environment, with each patch representing a discrete visual field location. In each training iteration, the generator produces a trajectory, which is then assessed by the discriminator in terms of its "realness" (i.e., similarity to human behavior). This evaluation guides the generator's updates, using PPO-based policy optimization to refine fixation probabilities, thereby helping it produce increasingly accurate gaze sequences that adapt to EEG-informed cognitive cues.

The discriminator, on the other hand, calculates the probability of each generated gaze sequence being "real" or human-like. It does so by analyzing state-action pairs, which encapsulate fixation points, EEG-derived cognitive states, and task conditions. Through the inclusion of EEG data, the discriminator is better informed about the cognitive state driving each gaze pattern, enhancing its ability to distinguish authentic behaviors from synthetic ones. This EEG integration thus aligns the GAIL model more closely with the nuanced cognitive demands observed during the Stroop task, reinforcing its predictive accuracy for gaze behavior across varying task conditions.

Training the generator with GAIL also includes specific metrics, such as the probability mismatch and Target Fixation Probability (TFP) at key stages, which assess how well the synthetic gaze matches human data. These metrics are essential for refining the generator's ability to focus on task-relevant areas of the visual field, adapting gaze based on task dynamics. By iteratively optimizing these metrics, the GAIL model effectively learns to generate gaze paths that are not only statistically similar to human data but also responsive to the cognitive cues embedded in EEG signals, providing a nuanced replication of attention and visual exploration under cognitive load.

3.6.2 PPO: Proximal Policy Optimization for Policy Learning

Proximal Policy Optimization (PPO) is employed within this framework to optimize the generator's policy, enabling it to model gaze trajectories that respond adaptively to real-time cognitive and environmental cues derived from EEG data and task conditions. PPO is particularly well-suited for this setup due to its stability and efficiency in high-dimensional action spaces, such as those found in gaze prediction, where the model must select fixation points within a grid-like environment representing the visual field.

PPO enhances the generator's policy by balancing the exploration of new fixation strategies with the exploitation of previously learned patterns. During training, PPO constraints policy updates within a specified range (the "clip" parameter), preventing overly large changes in action probabilities between iterations. This approach ensures that the model does not deviate too drastically from effective fixation strategies, thereby improving its convergence toward realistic gaze sequences that align with EEG-guided cognitive cues.

In each iteration, PPO calculates the advantage estimates, representing the benefit of taking specific fixation actions over baseline options. These advantage estimates guide the selection of subsequent fixation points, informed by the EEG data reflecting the participant's cognitive state. PPO then adjusts the policy based on the gradient of these advantage estimates, refining the likelihood of fixations that align with areas of cognitive interest as indicated by EEG signals. This iterative process allows the generator to learn fixation paths that dynamically adapt to shifts in attention driven by Stroop task demands.

Key metrics, such as Target Fixation Probability (TFP) and probability mismatch, are used to evaluate and fine-tune the policy at specific intervals. TFP, for instance, assesses the model's accuracy in directing gaze to relevant task regions within early steps of each gaze sequence. By optimizing for high TFP values, PPO encourages the generator to prioritize task-relevant areas that align with EEG-based cognitive load indicators, ultimately producing gaze paths that reflect both task demands and cognitive states.

Overall, PPO contributes to this framework by ensuring that the generator's policy is both responsive to EEG-informed attentional shifts and stable across training epochs. This stability is critical for capturing the nuanced dynamics of human gaze behavior under cognitive load, enabling the model to generate gaze paths that mirror the decision-making and attentional strategies observed in real participants during the Stroop task.

3.6.3 Dynamic Contextual Belief (DCB) in State-Action Modeling

In the context of state and action modeling, Dynamic Contextual Belief (DCB) provides a method to structure state representations by leveraging high-level contextual cues that change dynamically as the agent interacts with the environment. DCB integrates information from multiple perspectives to generate a comprehensive view of both the static context and the evolving understanding of the agent's environment during task completion. This approach enables more effective learning and adaptation in reinforcement learning models, particularly in applications like visual search.

Components of DCB

DCB (Fig. 3.7)includes three main components, each essential for building an effective state representation:

1. **Fovea**: This component emulates the high-resolution central area of the human visual field, capturing detailed information around a fixation point. It models how visual information is accumulated, enabling the agent to focus attention on specific image regions iteratively. Fig. 3.6

2. **Contextual Beliefs**: These beliefs represent the agent's high-level knowledge of object locations and spatial relationships within the environment. This component guides attention toward task-relevant areas, even before they are directly observed. For example, when searching for a TV(Or in the current case ”wrong” or “correct” buttons), DCB might prioritize areas of the scene

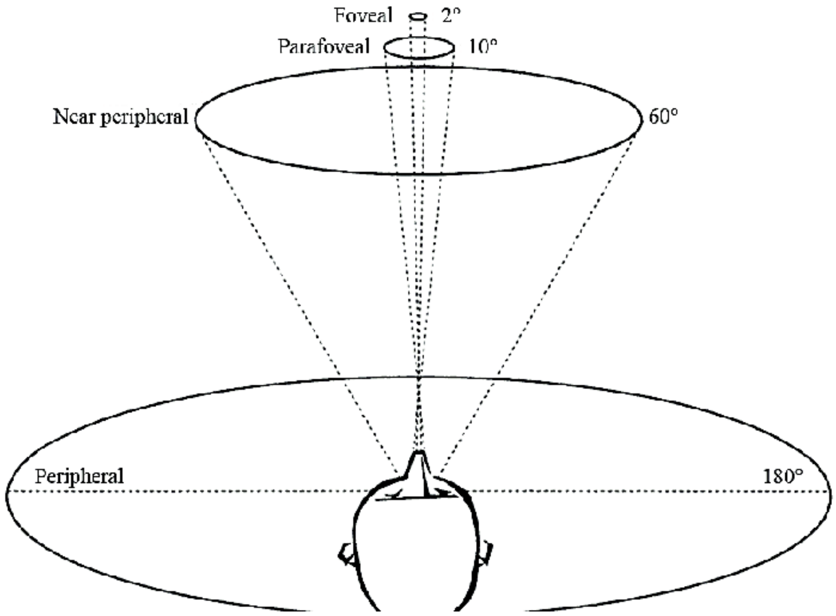

Figure 3.6: Foveal Vision

likely to contain it, such as walls or specific regions known to house the target. In this project, Meta Detectron2 framework is used.

3. **Dynamic Updating**: As the agent gathers information over successive interactions, the state representation is refined. This dynamic update process ensures that new information from each action (such as fixating on a new region) is integrated, updating both the contextual beliefs and the agent's understanding of the environment.

Implementation in Visual Search

In visual search, DCB enables more efficient fixation patterns by using a layered approach to represent both known and discovered areas. The initial state is constructed from contextual beliefs at a low resolution, representing peripheral visual inputs. As the agent selects fixation points, the relevant high-resolution information is incorporated into the belief state, enhancing the agent's knowledge base for subsequent actions.

1. **State Initialization**: The model initializes with a low-resolution contextual belief map that represents the broader scene.

2. **Belief Updating**: With each fixation, a new belief state is generated by combining high-resolution details from the fovea with existing low-resolution

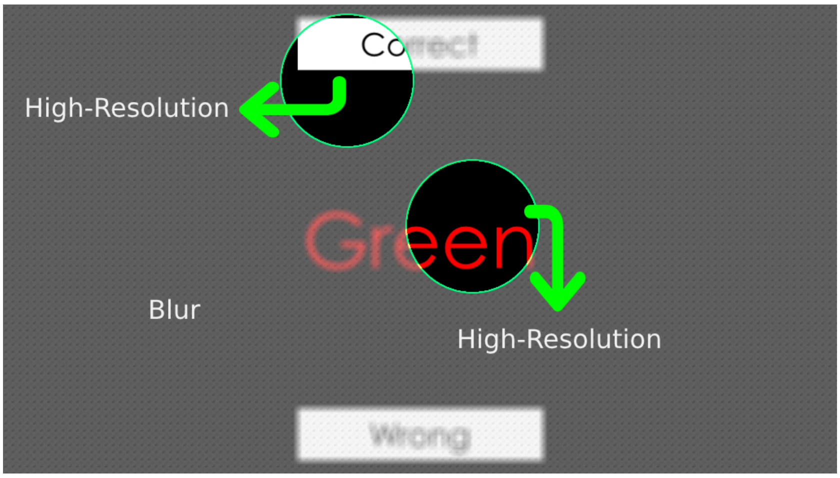

Figure 3.7: DCBs Mask Concept

beliefs. The belief state is adjusted using a mask specific to each fixation location, with high-resolution details updating only within the fixation region while the remaining area retains the lower-resolution context.

3. **Task-Specific Augmentation**: DCB can incorporate task-specific cues as part of the belief state, enhancing the state representation to reflect task priorities dynamically. For instance, when tasked with finding a specific object category, the state includes cues that prioritize relevant scene areas, making search more efficient.

Role of DCB in Reinforcement Learning

In reinforcement learning applications, DCB contributes to a richer state representation that enhances decision-making. By dynamically incorporating new information with each action, DCB refines the policy, improving the agent's ability to identify reward-optimizing actions based on evolving contextual cues.

3.7 Network and Model Architecture

Through an integration of GAIL and IRL, the architecture is designed to process visual data, human fixations, and EEG inputs to reproduce gaze behavior that reflects task-specific demands and neural responses. The overarching framework is structured to support a layered, context-sensitive model that adapts gaze predictions based on EEG-informed insights, offering a nuanced understanding of visual attention under cognitive load.

Central to this framework is the interaction between the generator and discriminator models, trained in a Generative Adversarial Imitation Learning (GAIL) setup to ensure generated gaze trajectories closely resemble authentic human behavior. The generator, or policy network, leverages Proximal Policy Optimization (PPO) to dynamically update gaze probabilities, informed by both fixation data and EEG signals, which enrich the model's ability to adapt to cognitive load and task complexity. This integration of EEG data—captured and processed in both training and validation—adds a significant layer of contextual understanding, allowing gaze patterns to respond adaptively to neural indicators of cognitive engagement.

The training pipeline is built around a robust dataset preparation and training loop that includes:

1. **Data Preprocessing and Storage**: EEG data is systematically associated with specific fixation points and processed alongside gaze trajectories. This step enables synchronized learning from EEG patterns linked to visual attention during each single task.

2. **Dynamic Belief Representation (DCB)**: Low- and high-resolution Dynamic Contextual Belief (DCB) maps are used within the environment, enabling selective attention to critical visual regions while maintaining computational efficiency.

3. **Training Process**: Both models—generator and discriminator—are iteratively refined through PPO and adversarial learning in GAIL, adjusting action probabilities based on task requirements. The training also includes advanced metrics such as Target Fixation Probability (TFP) and probability mismatch, ensuring that predicted gaze paths align with observed human data.

The architecture's modularity and data-centric approach to gaze and EEG integration set a foundational model for decoding gaze in the Stroop task, leveraging IRL to gain insights into the neural correlates of cognitive control and attentional behavior.

3.7.1 Discriminator Network Structure

The discriminator network is designed to evaluate and distinguish between authentic human gaze paths and those generated by the policy network, thereby refining the accuracy of the model's gaze trajectory predictions. This discriminator functions within the Generative Adversarial Imitation Learning (GAIL) setup, where it assigns rewards to generated gaze paths based on their similarity to real human scanpaths, integrating contextual and neural cues from the EEG data associated with each fixation.

The discriminator is a conditional model tailored for the Stroop task environment, built with convolutional layers that incorporate task-specific information through one-hot encoded task vectors. This task vector is appended to each layer, ensuring that the model learns task-relevant features in relation to gaze behavior:

1. **Input Layer:** The input consists of feature maps with dimensions `(batch_size, channels, height, width)` where:
   - Channels include belief maps and task-related information.
   - Task vectors are represented in a one-hot encoding format and modulate feature maps in each layer, preserving task-specific distinctions.
2. **Convolutional Layers:**
   (a) **Conv1:** The initial convolution layer accepts the concatenated input of belief maps and task vectors, producing 128 feature maps. This layer includes:
   - Filter size: 3×3
   - Padding: 1 (to maintain spatial dimensions)
   - Output channels: 128

   (b) **Conv2:** The output from Conv1 is concatenated with the task vector and passed through a second convolutional layer.
   - Filter size: 3×3
   - Padding: 1
   - Output channels: 64

   (c) **Conv3:** Similar to previous layers, task information is appended again, and features are convolved.
   - Filter size: 3×3
   - Padding: 1
   - Output channels: 32

   (d) **Conv4:** This final convolutional layer reduces the feature maps to a single channel.
   - Filter size: 1×1
   - Output channels: 1 (produces a reward map)
3. **Pooling Layers:**
   - A MaxPool layer is optionally applied after the first two convolutional layers, which is activated based on input spatial dimensions (e.g., when height = 80).

4. **Task Modulation Mechanism:**

    - Each layer leverages task modulation by concatenating the one-hot encoded task vector with feature maps, forming a multi-channel input. This modulation ensures that the discriminator learns to identify task-specific gaze patterns and differentiate them from generic gaze behavior.

Training and Loss Calculation

The discriminator is trained to output a probability score that reflects the likelihood of a gaze path being authentic. It calculates a binary cross-entropy loss between real and generated gaze paths, using the following metrics for gradient adjustments:

- **Real Loss**: Probability score for actual human gaze paths.
- **Fake Loss**: Probability score for generated gaze paths, with added regularization from a gradient penalty to enforce smooth decision boundaries.

The network's training procedure incorporates gradient penalty for stability, particularly when training with small datasets or high-dimensional data:

- **Milestones for learning rate adjustment**: 5

This architecture, with task-conditional modulation and convolutional layers, forms a discriminator capable of discerning between EEG-guided, human-authentic gaze trajectories and those generated by the policy network. Its design emphasizes contextual relevance by incorporating task vectors at every layer, improving the discriminator's sensitivity to task-dependent gaze behaviors.

3.7.2 Policy Network Structure

The policy network serves as the generator within the GAIL framework, tasked with producing gaze trajectories that mimic human scanpaths in the Stroop task. This network is structured to integrate both visual and EEG data, enabling it to model gaze paths that adapt to the cognitive demands of the task. The Proximal Policy Optimization (PPO) algorithm is employed to update the network's action probabilities in response to real-time observations, EEG signals, and task-specific details.

Network Architecture

The policy network is a conditional model with dedicated components for actor (action selection) and critic (value estimation) pathways, allowing it to predict gaze actions and estimate their expected rewards. EEG features are integrated into the network to guide gaze decisions based on cognitive load signals, enhancing the model's alignment with real human behavior.

1. **Input Layer:**
   - **Input dimensions:** `(batch_size, channels, height, width)`,where:
     - Channels include low- and high-resolution belief maps and task-specific one-hot vectors.
     - EEG data, processed separately through a fully connected layer, is later combined with other features in the critic pathway.
     - Task vectors are appended as one-hot encodings to modulate feature maps in each layer, maintaining task relevance throughout the network.
2. **Feature Encoding Layer:**
   - **Feature Encoder (Conv):** The initial convolutional layer processes the concatenated input (belief maps + task vector), producing 128 feature maps.
     - Filter size: 5×5
     - Padding: 2 (to preserve spatial dimensions)
     - Output channels: 128
3. **Actor Pathway (Action Selection):**
   (a) **Actor Conv1:** The first convolution in the actor pathway receives the output of the feature encoder with appended task vectors.
      - Filter size: 3×3
      - Padding: 1
      - Output channels: 64

   (b) **Actor Conv2:** Another convolutional layer that continues feature refinement.
      - Filter size: 3×3
      - Padding: 1
      - Output channels: 32

   (c) **Actor Output (Conv3):** Produces the final action logits, reshaped to a single-channel output for action probabilities.
      - Filter size: 1×1
      - Output channels: 1

(d) **Softmax Activation:** Logits are converted into action probabilities across available gaze actions using a softmax function, ensuring the output reflects a probability distribution over gaze locations.

4. **Critic Pathway (Value Estimation):**

(a) **Critic Conv0:** Initial convolution that feeds into the critic pathway, followed by a max-pooling layer.

- Filter size: 3×3
- Output channels: 128

(b) **Critic Conv1:** Further feature refinement for value estimation.

- Filter size: 3×3
- Output channels: 256

(c) **Fully Connected Layer (Critic2):** Following the convolutional layers, a fully connected layer reduces the feature map to 64 units.

- Input size: 256 (or as determined by the previous layer's flattening)
- Output size: 64

5. **EEG Processing Layer:**

- **EEG Fully Connected Layer:** A fully connected layer specifically processes EEG features.
  - Input size: 3 (assuming three critical EEG features)
  - Output size: 32
- This EEG-derived output is concatenated with features from the critic pathway, enhancing value estimation with real-time cognitive state indicators.

6. **Final Critic Layer:**

- **Critic Output Layer:** The concatenated EEG and task features are passed through a final fully connected layer to output the state value estimate.
  - Input size: $64 + 32$ (concatenated from critic pathway and EEG features)
  - Output size: 1

Training and Hyperparameter Details

The policy network's training is driven by the PPO algorithm, which balances stability and exploration in gaze path predictions:

- Clipping Parameter: Ensures stable updates by limiting the changes in action probability across iterations. 0.2
- Learning Rate: 1e-05
- Batch Size and Epochs: 2

The policy network leverages EEG data and task-specific modulation to dynamically adapt gaze predictions to cognitive demands. Its structure integrates attention-guiding features while maintaining spatial relevance, ensuring that gaze actions align with the Stroop task's attention requirements. Through the combination of actor-critic architecture and EEG-informed adaptation, this model successfully generates gaze paths that reflect realistic human-like decision-making under cognitive load.

3.7.3 Training Process and Hyperparameter Tuning

The training process for the gaze prediction model is designed to iteratively refine the generator (policy network) and discriminator within a GAIL framework, utilizing PPO to optimize gaze trajectories based on EEG and visual data. This detailed pipeline ensures that the model learns to generate human-like gaze paths that are both task-specific and sensitive to cognitive demands. The following subsections provide a granular breakdown of the steps, processes, and configurations used in training.

1. **Data Loading and Preprocessing:**
   - The dataset is divided into training and validation sets, containing gaze trajectories annotated with EEG data for each fixation point.
   - Input images are segmented into 32×20 patches (`patch_num`) to cover the entire 512×320 image (`im_w` and `im_h`). Each patch represents a potential fixation region, forming a grid of 640 gaze points. This segmentation allows the model to predict gaze locations at a fine-grained level.
   - Each gaze path consists of a maximum of 6 fixations (`max_traj_length`), restricting trajectories to concise, task-relevant paths.
2. **Batch Processing:**
   - Training is conducted with mini-batches of 4 images (`batch_size`), where each batch includes both EEG and gaze data for joint processing.

- EEG data is preprocessed and linked to each gaze trajectory, providing additional cues for fixation decisions based on cognitive load.

3. **GAIL Framework:**
    - **Policy Network (Generator):** The policy network generates potential gaze trajectories in response to the input data and is updated through PPO. It samples fixation points from a probability distribution over all patches, influenced by task demands and EEG indicators.
    - **Discriminator Network:** The discriminator evaluates the authenticity of each generated gaze trajectory, comparing them against real human scanpaths. It assigns rewards to trajectories that closely resemble human behavior, thus guiding the policy network to refine its predictions iteratively.

4. **Action Collection and State-Action Pairs:**
    - For each batch, the policy network samples actions (fixations) based on the current state representation (visual and EEG data). These sampled actions form state-action pairs representing the model's gaze choices under task-specific conditions.
    - **Trajectory Generation:** The generator samples a complete trajectory of fixations for each image in the batch, constrained to a maximum of 6 fixations per path. This step simulates a sequence of gaze shifts, allowing the model to approximate human scanning behavior.

5. **Policy Network Update with PPO:**
    - **Advantage Estimation:** Advantages are computed using Generalized Advantage Estimation (GAE) with a smoothing factor τ of 0.96. This method calculates the expected return for each action, adjusting for both immediate rewards and long-term value.
    - **Reward Discounting:** The cumulative reward is discounted at a rate of 0.9 (`gamma`), favoring immediate fixations while capturing long-term visual goals. This discount factor encourages the policy to focus on immediate task-relevant regions while maintaining awareness of the broader task context.
    - **Clipping and Gradient Updates:** The PPO algorithm limits updates to the policy network by clipping policy changes to a maximum of 0.2 (`clip_param`), ensuring stable learning and preventing drastic shifts in gaze predictions across iterations. After each mini-batch, the policy is updated by backpropagating the gradients computed from the advantage-weighted loss function.

6. **Discriminator Network Update with GAIL:**

    - **Real and Fake Data Discrimination:** The discriminator network receives both real human trajectories and generated trajectories from the policy network. It assigns higher rewards to paths that align with human patterns, guiding the policy toward more realistic gaze behaviors.
    - **Gradient Penalty:** To improve stability, the discriminator applies a gradient penalty with a coefficient λ (0.15), enforcing smooth decision boundaries between real and generated trajectories. This penalty reduces overfitting and ensures that the discriminator generalizes well to both authentic and synthetic gaze paths.
    - **Learning Rate Adjustment:** The learning rate for the discriminator is adjusted at milestone epochs (`gail_milestones`: [5]), enhancing training stability and allowing for gradual improvements over time.

7. **Environment Interaction and Reward Calculation:**

    - **Dynamic Contextual Belief (DCB):** The environment's DCB maps modulate focus between high- and low-resolution regions based on the agent's current fixation. This dynamic adjustment is guided by the foveal radius (`fovea_radius`: 2) and inhibition of return (IOR) size (`IOR_size`: 1), preventing repeated fixation on previously viewed areas and promoting exploration of new regions.
    - **Stop-On-Target Criterion:** The training process incorporates a Stop-On-Target (SOT) criterion (`stop_criteria`: "SOT") that halts fixation sampling once the gaze reaches the target. This criterion aligns the model's behavior with human tendencies to cease search efforts upon locating relevant stimuli.

8. **Checkpointing and Evaluation:**

    - **Checkpointing:** Model checkpoints are saved after every single epoch (`checkpoint_every`: 1), allowing incremental progress tracking and model state recovery. Only one checkpoint is retained (`max_checkpoints`: 20) to manage storage while keeping the 20 latest model states and access previous models after early stopping.
    - **Evaluation Frequency:** Every 5 steps (`evaluate_every`: 5), the model is evaluated on validation data using metrics such as Target Fixation Probability (TFP) and probability mismatch, which gauge the model's ability to replicate human-like gaze paths and track alignment with observed behaviors.

Algorithm 1 Inverse Reinforcement Learning (IRL) Framework

Require: Human scanpaths S_{human}, bounding boxes B, configuration C
Ensure: Trained policy π_θ

```
Initialize generator π_θ and discriminator D_φ;
Load and preprocess datasets S_train and S_valid;
Set hyperparameters (learning rates, batch size, max steps, etc.);
for each epoch do
    for each batch B_img in S_train do                         ▷ Generate trajectories
        Reset environment with batch B_img;
        while not done do
            Sample action a_t ~ π_θ(s_t);
            Execute action and collect (s_t, a_t, r_t, s_{t+1});
        end while
                                                               ▷ Update discriminator
        Sample fake and real scanpaths;
        Compute loss L_D = −E[log D_φ(S_human)] − E[log(1 − D_φ(S_gen))];
        Update φ using gradient descent;
                                                               ▷ Update generator
        Compute advantages A(s_t, a_t) using GAE;
        Compute PPO loss L_π;
        Update θ using gradient descent;
                                                               ▷ Evaluate policy
        if step mod eval_interval = 0 then
            Generate validation scanpaths S_gen;
            Compute evaluation metrics (TFP, Multimatch, etc.);
        end if
    end for
end for
```

This structured training process, combining PPO and GAIL within a task-specific framework, enables the model to iteratively improve its gaze predictions by leveraging both visual cues and EEG data. Each component—from reward calculations to gradient penalties and checkpointing—contributes to building a robust, accurate gaze prediction model that closely mimics human gaze behavior in the Stroop task.

Chapter 4

Scanpath Prediction Framework

This section presents a comprehensive framework for predicting human gaze paths within the Stroop task, guided by EEG signals and informed by inverse reinforcement learning (IRL) principles. The framework builds upon advanced reinforcement learning strategies, combining Proximal Policy Optimization (PPO) and Generative Adversarial Imitation Learning (GAIL) to enable a dynamic model capable of capturing and predicting attention-driven scanpaths under cognitive load. This approach integrates neural markers from EEG data, providing real-time insights into participants' cognitive states and refining the gaze predictions to align closely with human-like behaviors.

The framework is designed to train a policy network (generator) and a discriminator, each tailored to recognize and replicate gaze patterns within a high- and low-resolution environment. The model employs Dynamic Contextual Belief (DCB) maps to structure the visual field into regions of varying importance, allowing it to adaptively prioritize areas based on the participant's cognitive goals, as indicated by EEG cues. By using these belief maps, the framework effectively mimics human visual attention, focusing on areas likely to capture gaze during complex Stroop tasks.

During training, the policy network learns to generate sequences of fixation points informed by EEG data, structured through a custom environment that simulates the participant's visual experience. The EEG data is preprocessed and integrated with the model, shaping the predicted scanpaths based on real-time cognitive feedback. Through iterative optimization, PPO refines the network's policy by balancing exploration of novel gaze paths and exploitation of learned, task-aligned patterns, adjusting fixation probabilities to optimize cognitive engagement indicators.

The discriminator network, in contrast, operates by distinguishing generated scanpaths from human-generated paths, providing feedback that the generator leverages to better replicate human-like gaze behaviors. By dynamically updating based on EEG cues, the model's reward structure enables it to mirror the cognitive adjustments individuals make during the Stroop task, capturing not only fixation points but the nuanced strategies driving gaze behavior under cognitive interference.

Through this framework, the model achieves robust prediction of gaze patterns that mirror human attentional shifts, producing scanpaths that align with observed cognitive states and the demands of the Stroop task. This EEG-driven IRL approach thus provides a refined tool for gaze behavior modeling, advancing applications in cognitive neuroscience and clinical diagnostics where attention and response inhibition are critical markers.

4.1 State-Action Modeling

State and action modeling form the foundation of this gaze prediction framework, with states representing the visual and cognitive contexts of gaze behavior and actions representing fixation choices within these states. This framework captures a dynamic interplay between the participant's cognitive state (inferred through EEG data) and the visual environment (modeled as a spatial grid of potential fixation points). By leveraging Dynamic Contextual Belief (DCB) maps, the framework structures the visual field into discrete low- and high-resolution patches, dynamically updating based on the participant's fixation history and neural responses.

State Representation

In this framework, the state encapsulates both the visual features within a scene and cognitive markers associated with each task condition. Each state includes:

- **DCB Map Layers**: High-resolution representations are reserved for fixation points, while peripheral regions are rendered in lower resolution. This approach mimics human vision, where the fovea captures detailed information while the periphery maintains a lower resolution. These maps adjust based on fixation history, allowing the model to retain contextual awareness of previously viewed regions.

- **EEG-Informed Cognitive Cues**: EEG data, processed to extract cognitive load indicators, is integrated with the visual state, providing a real-time measure of the participant's attentional engagement and cognitive load. This EEG-driven approach enables the model to adjust its fixation choices in response to dynamic cognitive shifts, a feature crucial in tasks like the Stroop test that involve cognitive interference.

The state is therefore a multi-layered representation combining spatial attention with cognitive context. By integrating EEG data, the framework enhances its sensitivity to cognitive demands, enabling it to differentiate between high and low-load conditions and to prioritize fixation points accordingly.

Action Modeling

In this framework, actions are defined as discrete fixation choices on a grid-like structure, where each grid cell represents a potential fixation region within the DCB map. The choice of fixation points is guided by the reward-based motivations inferred through inverse reinforcement learning. By applying EEG-derived cognitive cues, the framework identifies areas of interest that align with cognitive engagement markers, effectively predicting fixation locations in a manner that reflects human attentional priorities during the Stroop task.

The environment module provides a central mechanism for action modeling by:

1. **Masking Previously Fixated Regions**: To prevent repetitive fixations on recently observed areas, the environment applies inhibition of return (IOR), dynamically updating action availability based on past fixations.

2. **Dynamic Task-Dependent Reward Calculation**: The environment generates rewards based on task alignment and visual context, using EEG-informed state updates to identify areas of high cognitive relevance. This approach allows actions to adapt based on task-specific demands, with the model prioritizing regions associated with higher cognitive load during the Stroop task.

4.2 EEG Feature-Guided Modeling

In this framework, the EEG Feature-Guided Modeling approach focuses on leveraging EEG signals exclusively within the policy network to guide gaze predictions. This component is crucial, as it aligns gaze actions with the participant's cognitive state, captured through EEG data recorded during the Stroop task. By incorporating EEG features into the policy network, the model dynamically adjusts fixation choices based on real-time cognitive load, effectively enhancing the alignment between gaze behavior and neural responses.

EEG Data Processing and Integration in the Policy Network

The EEG data serves as a rich source of information about the participant's cognitive load, which influences attention and visual exploration. In this framework:

- **Feature Extraction and Preprocessing**: EEG data is processed to extract features that reflect cognitive states, including indicators of attentional engagement and cognitive load. The extracted features focus on relevant frequency bands known to correlate with task-related cognitive processing.
- **Policy Network Integration**: Within the policy network, the EEG features are incorporated to shape action selection probabilities. This integration is implemented through a fully connected layer that processes EEG data and combines it with visual and task-specific cues. By incorporating EEG features, the policy network gains sensitivity to shifts in cognitive load, allowing it to prioritize fixation locations that align with attentional demands during high-interference Stroop trials.

EEG-Informed Action Selection

The policy network's EEG integration enables the framework to adaptively select fixation points that reflect cognitive engagement in real-time:

- **Dynamic Adjustment of Gaze Behavior**: EEG signals are mapped directly into the policy network's decision-making process, allowing fixation probabilities to fluctuate based on EEG-informed cognitive cues. For example, when EEG data indicates heightened cognitive load, the policy network may increase fixation probability on task-relevant areas, emulating human focus on critical information under cognitive strain.
- **Enhanced Reward Sensitivity**: Although the EEG data is not directly incorporated into the reward function, the EEG-informed policy network adjusts action selection in a way that indirectly affects the reward outcomes. By prioritizing regions associated with cognitive load, the network aligns gaze paths with the Stroop task's demands, optimizing for realistic fixation sequences without altering the discriminator's learning process.

Training Process and EEG Integration in Policy Optimization

In the training phase, EEG data informs the policy network, guiding it to produce gaze paths that align with observed human behavior under cognitive load:

- **PPO-Based Policy Updates**: The Proximal Policy Optimization (PPO) algorithm is employed to update the policy network, utilizing advantage estimates that indirectly benefit from EEG-informed action choices. By adjusting action probabilities based on EEG features, the policy network refines its gaze path predictions, increasing the accuracy of fixation patterns that mirror human attention shifts in the Stroop task.

- **EEG-Driven Realism in Generated Paths**: The EEG-guided policy enhances the realism of generated gaze paths by ensuring that action selection reflects real-time cognitive states. This alignment with neural data results in a model that can more accurately predict gaze patterns that resemble human behavior, offering insights into how cognitive load and attention impact visual exploration during the Stroop test.

In summary, EEG Feature-Guided Modeling in this framework is achieved by embedding EEG data directly within the policy network, allowing adaptation of gaze predictions based on cognitive state. This targeted integration supports the generation of realistic gaze paths that respond dynamically to the Stroop task's cognitive demands, providing a powerful tool for studying gaze behavior through the lens of EEG-informed cognitive processes.

4.3 Reward and Policy Learning

The Reward and Policy Learning component is the core of the IRL-based gaze prediction framework, where the model learns to replicate human-like gaze patterns in response to task demands and cognitive cues. This section utilizes both reward-based mechanisms and EEG-guided policy updates to enable a dynamic model capable of accurately predicting gaze behavior under the Stroop task's cognitive load. By employing a combination of Proximal Policy Optimization (PPO) and Generative Adversarial Imitation Learning (GAIL), the framework refines fixation predictions through an iterative learning process that balances action selection with realistic behavioral modeling.

Reward Structure and Discriminator Role

The reward structure in this framework is designed to reinforce fixation choices that align closely with human gaze patterns, utilizing a discriminator network to assess the quality of generated gaze paths:

- **Discriminator Network as a Reward Estimator**: The discriminator operates within the GAIL framework, distinguishing between real human gaze paths and those generated by the policy network. It assigns higher rewards to fixation sequences that resemble authentic human behavior. By treating gaze trajectories as state-action pairs, the discriminator's feedback encourages the policy network to adapt its fixation patterns to match human-like exploration behaviors during the Stroop task.

- **Implicit Reward Mechanism**: Unlike traditional IRL approaches that define explicit rewards based on predefined features, this framework utilizes the discriminator's feedback as an implicit reward signal. The discriminator's assessment incentivizes the generator (policy network) to produce gaze paths that align with the cognitive demands of the Stroop task, focusing on fixation points that accurately reflect attentional strategies.

EEG-Guided Policy Network and PPO Optimization

The policy network leverages EEG data to refine fixation choices, with PPO providing an efficient framework for updating action probabilities in response to cognitive state changes:

- **Policy Network Integration with EEG Data**: The policy network, guided by EEG-informed state representations, adjusts fixation probabilities to reflect cognitive load indicators. By dynamically adapting to EEG signals, the policy network is able to prioritize gaze locations that align with task-relevant areas, capturing the shifts in attention typical of Stroop task performance.

- **Proximal Policy Optimization (PPO) for Stability and Efficiency**: PPO is applied to optimize the policy network's action probabilities within a stable update range, ensuring smooth adjustments in fixation decisions across training iterations. PPO's clipping mechanism restricts drastic changes in policy updates, allowing the model to refine its gaze predictions incrementally. This incremental approach ensures that fixation probabilities evolve in a manner consistent with both EEG-informed cognitive cues and the human-like gaze patterns encouraged by the discriminator.

Training Process and Reward-Policy Interaction

The training process balances the EEG-informed policy updates and reward signals from the discriminator, iterating through steps that refine gaze prediction accuracy:

- **Training Loop with GAIL and PPO**: The training loop iterates between generating gaze trajectories through the policy network and evaluating them using the discriminator. With GAIL's adversarial structure, the discriminator provides feedback on the "realness" of generated paths, guiding the policy network to adapt its predictions. Each PPO update step leverages advantage estimates, calculated based on the discriminator's reward, to refine policy parameters. This back-and-forth refinement process ensures that the model's gaze predictions evolve to reflect realistic gaze paths in response to EEG-informed cognitive states.

- **Cumulative Rewards and Advantage Estimation**: Advantage estimation, critical in PPO, computes the expected return for each fixation action, integrating both immediate and longer-term rewards. Although EEG data does not directly influence the reward structure, the EEG-guided fixation choices indirectly shape advantage values, reinforcing gaze paths that are cognitively aligned with the Stroop task's demands. By estimating advantages with EEG-informed actions, the policy network refines fixation choices in a manner that optimally reflects cognitive engagement.

Outcome: Near Realistic and Cognitively Aligned Gaze Predictions

Through the combination of reward and policy learning, this framework produces gaze predictions that mirror human attentional patterns under cognitive load. The reward signals from the discriminator incentivize human-like fixation sequences, while the EEG-guided policy network ensures that these predictions are responsive to real-time cognitive states. This integration of GAIL and PPO with EEG-informed action selection results in a gaze prediction model capable of dynamically adapting to cognitive demands, offering an advanced tool for studying the interaction between visual attention and cognitive processes.

Chapter 5

Experiments

This section delineates the experimental framework, which leverages a sophisticated integration of EEG signals, eye-tracking data, and IRL mechanisms to model and reconstruct human gaze behavior in the Stroop task. The experimental setup involves a dynamic and adaptive IRL environment built around GAIL and PPO methodologies, tailored to predict and analyze gaze patterns in response to task-specific cognitive load and reward structures.

The codebase implements several notable techniques to ensure accurate, adaptive gaze modeling within this environment:

1. **EEG-Driven State Modeling and Action Selection**: A primary focus of the experiment is the incorporation of EEG data to drive both state representation and action selection within the model. Through preprocessing, EEG signals are extracted and fed directly into the IRL framework, specifically modulating the generator's (policy network's) action probabilities. This EEG integration ensures that gaze predictions are not only responsive to visual context but are also dynamically adapted based on real-time cognitive states, which are pivotal in the Stroop task.

2. **Dynamic Contextual Belief (DCB) Maps and Foveal Masking**: The IRL environment is designed with DCB maps that dynamically adjust visual resolution around each fixation point, mimicking human visual processing. These maps include both low-resolution peripheral and high-resolution central areas, where the high-resolution region shifts according to the latest fixation, allowing the model to accumulate scene information similarly to human attention mechanisms. The DCB maps are further enhanced through a foveal masking approach, where fixation history dictates inhibition of return to previously visited regions, promoting naturalistic gaze patterns by encouraging exploration of novel areas within the visual field.

3. **Generative Adversarial Imitation Learning (GAIL) and Proximal Policy Optimization (PPO)**: The experimental structure utilizes a dual-model setup, combining GAIL's generator-discriminator structure with PPO-based optimization. GAIL enables adversarial learning by having the discriminator assess the realism of generated gaze paths against actual human data, providing reward feedback that drives the generator to refine its fixation choices iteratively. PPO, employed within the policy network, ensures stable learning through clipped policy updates, adapting fixation decisions in line with cognitive engagement cues from EEG data. This hybrid approach balances exploratory behavior and exploitation of learned gaze strategies, aligning generated paths closely with human data under Stroop task conditions.

4. **Adaptation to Task-Specific Rewards and Human-Like Gaze Patterns**: The experimental environment applies a flexible reward structure that reflects task-specific demands, essential for tasks involving cognitive control like the Stroop. Task-based rewards guide gaze patterns toward areas that yield higher cognitive relevance as per EEG indicators, helping the model mimic attentional shifts characteristic of Stroop tasks. The PPO's advantage estimation further aligns fixation choices with regions of interest by dynamically recalculating benefits of actions in light of EEG-driven state updates, thereby refining the model's focus on critical task regions.

5. **Evaluation Pipeline for Gaze Prediction Accuracy**: The experimental framework includes an evaluation mechanism that computes metrics such as Target Fixation Probability, Probability Mismatch, and MultiMatch Sequence Score. These metrics serve to quantify the alignment between model-generated scanpaths and actual human data. Additionally, the experimental setup evaluates the EEG-guided policy network's ability to adapt to changes in cognitive load, allowing a nuanced comparison between gaze paths generated with and without EEG input.

Through these methodological components, the experiment provides a robust testbed for analyzing gaze behavior under cognitive load and demonstrates the effectiveness of EEG-integrated IRL frameworks in reconstructing realistic gaze patterns in cognitively demanding tasks. The results from this experimental framework are intended to yield insights into the interaction between cognitive control, reward-driven attention mechanisms, and visual processing under task-induced cognitive demands.

5.1 Compare Scanpath Model

In this subsection, we present a comparative analysis of various scanpath models evaluated within the Stroop task environment. The experimental setup leverages both EEG-enhanced and non-EEG baselines to investigate the impact of cognitive state information on scanpath prediction accuracy and realism.

The comparison framework centers on evaluating models' ability to replicate human scanpaths across key metrics. By implementing multiple configurations—such as EEG-guided policies, purely visual models, and reward-modulated IRL frameworks—the experiment assesses each model's competency in capturing human-like gaze patterns under the Stroop task's cognitive demands.

1. **EEG-Enhanced IRL Model vs. Non-EEG Models**:

    - **EEG-Enhanced IRL Model**: The EEG-informed model leverages EEG signals as direct input into the IRL framework, specifically influencing state representations and fixation probability distributions within the PPO-driven policy network. This approach allows the model to adapt gaze predictions in real-time based on participants' cognitive load, reflecting shifts in attention associated with Stroop task interference.
    - **Non-EEG IRL Models**: To evaluate the contribution of EEG data, we compare the EEG-enhanced model with non-EEG counterparts that rely solely on visual context and reward structures derived from GAIL. These models do not have access to real-time cognitive cues, relying instead on static gaze behavior rules within the IRL framework. By contrasting these approaches, we measure the added value of EEG integration for realistic gaze behavior replication.

2. **Scanpath Sequence Fidelity**:

    - **Temporal Sequence Analysis**: For each model, gaze sequence fidelity is assessed by comparing the temporal order of fixations to human scanpaths using metrics like MultiMatch. This comparison enables insights into how well each model reproduces the step-by-step progression of human gaze as it adapts to task demands and visual stimuli. The EEG-enhanced model is hypothesized to better capture human temporal fixation patterns, especially under high cognitive load.
    - **Probability Distribution Matching**: The probability of fixating on key regions (Target Fixation Probability, TFP) is another core metric, calculated at various stages in the scanpath sequence (steps 1, 3, and 6). These results reveal whether each model directs attention toward

task-relevant areas consistently with human patterns, particularly under incongruent (cognitively challenging) Stroop trials.

3. **Reward Sensitivity and Policy Adaptation**:

   - Each model's policy network adaptation is examined to understand how reward sensitivity shapes scanpath generation under different cognitive loads. The EEG-enhanced model incorporates cognitive states into its reward structure, potentially yielding fixation sequences that dynamically adjust to the Stroop task's interference conditions. In contrast, non-EEG models adhere to a more static reward interpretation, likely resulting in less adaptive scanpaths. This comparison provides insight into how EEG integration fosters policy refinement within the PPO framework, emphasizing areas of cognitive engagement aligned with EEG markers.]

4. **Action Selection and Inhibition of Return (IOR)**:

   - The scanpath models also differ in how they handle the inhibition of return (IOR) mechanism, a core component for promoting human-like exploration by discouraging repeated fixations. The EEG-informed model utilizes cognitive load signals to further modulate IOR, adapting action selections to emphasize task-relevant yet previously unexplored regions. Non-EEG models, while incorporating IOR, lack real-time cognitive inputs, potentially leading to less optimal and more repetitive fixation patterns. This aspect of comparison illuminates the role of cognitive state information in maintaining naturalistic gaze behavior.

5. **Performance Evaluation and Metrics Summary**:

   - A summary of quantitative comparisons across Target Fixation Probability (TFP), Sequence Score, Probability Mismatch, and MultiMatch metrics is provided to highlight each model's strengths and limitations in mirroring human gaze patterns. This summary underscores the utility of EEG-informed gaze modeling for high cognitive-load tasks, validating the EEG-enhanced model's advantage in predicting nuanced attention shifts.

5.2 Implementation Details

5.2.1 Extracting DCBs

The implementation of Dynamic Contextual Belief (DCB) within our Inverse Reinforcement Learning framework necessitated the computation of contextual belief maps for each slide in the Stroop test. These belief maps form a critical component of the state representation, capturing both high-resolution focal details and low-resolution peripheral information. To generate these maps, we developed a comprehensive approach that involved training a custom object detection model using Detectron2 and constructing feature maps based on the detected objects. This section details the processes involved in training the object detection model and computing the high-resolution (HR) and low-resolution (LR) feature maps integral to the DCB representation.

A. Training Detectron2 for Custom Object Detection

To accurately model human visual attention and contextual beliefs within the Stroop task environment, it was essential to detect and classify specific objects present in the slides, such as buttons indicating "correct" or "wrong" responses and words displayed in various colors and congruence conditions. For this purpose, we employed Detectron2, a state-of-the-art object detection and segmentation framework developed by Facebook AI Research. Detectron2 provides a modular and flexible platform for training custom object detection models using various architectures and datasets.

Annotation of Stroop Test Images

We began by collecting a dataset of Stroop test slides, comprising images that presented words in different color-word congruency configurations. Due to the controlled nature of the experiment, we initially had a limited set of 11 distinct images. To prepare these images for training the object detection model, we used Roboflow, an online tool that facilitates image annotation and dataset management.

Using Roboflow's annotation interface, we meticulously labeled each image by drawing bounding boxes around the objects of interest and assigning them to one of 13 predefined classes. These classes were carefully selected to represent all possible combinations of word content and font color relevant to the Stroop task, as well as the response buttons. The classes included:[class-correct, class-wrong, word-blueinblue, word-blueingreen, word-blueinyellow, word-greeninblue, word-greeningreen, word-greeninred, word-redingreen, word-redinred, word-redinyellow, word-yellowinred, word-yellowinyellow]

Each class represents a specific combination of the word's semantic content and its displayed color, capturing both congruent and incongruent conditions critical to the Stroop effect. The "class-correct" and "class-wrong" labels correspond to the response buttons that participants interact with during the task. Fig. 5.1

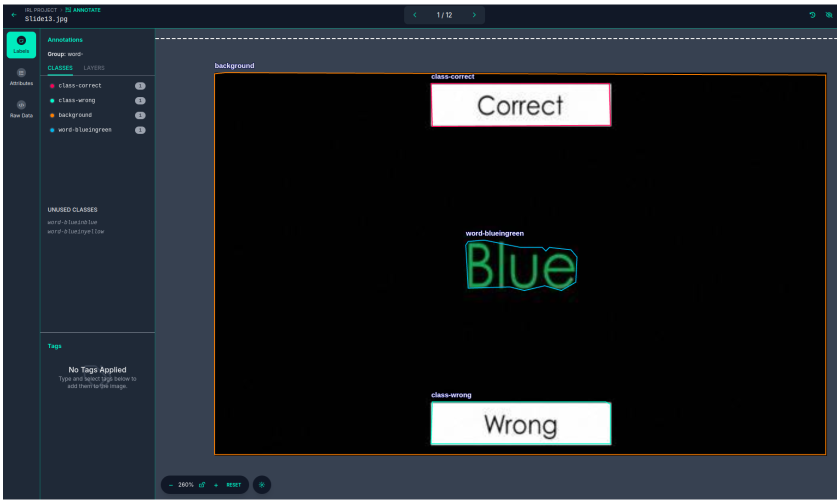

Figure 5.1: Roboflow Annotation System

Conversion to COCO Panoptic Format

After completing the annotations, the dataset was initially in COCO (Common Objects in Context) format, which is suitable for object detection tasks. However, for panoptic segmentation—which combines both instance segmentation and semantic segmentation—we needed to convert the annotations to COCO Panoptic format. This format provides a unified representation of both "stuff" and "thing" classes, allowing for comprehensive scene understanding, however in this case ignoring the “stuff” class was possible because there were only a single “stuff” object that was the background.

The conversion process involved generating a panoptic segmentation map for each image, where each pixel is assigned a class label and, for "thing" classes, an instance ID. We utilized tools provided by the Detectron2 framework to facilitate this conversion, ensuring compatibility with the panoptic segmentation training pipelines.

Data Augmentation to Address Limited Dataset Size

Given the small number of original images (11 in total), training a robust object detection model posed a significant challenge due to the risk of overfitting and poor generalization. To mitigate this issue, we implemented data augmentation techniques focused on cropping. By randomly cropping regions of the original images and treating them as new samples, we effectively increased the diversity of the training data without introducing new content.

This cropping augmentation was carefully managed to maintain the integrity of the annotated objects, ensuring that cropped regions still contained meaningful instances of the classes of interest. The augmented dataset provided a more substantial foundation for training the model, enhancing its ability to detect objects under various spatial configurations and scales.

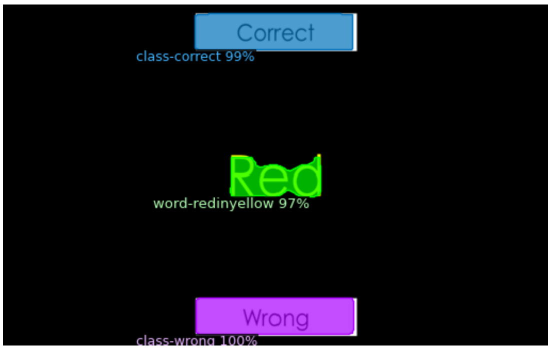

Figure 5.2: HR Image Segmentation

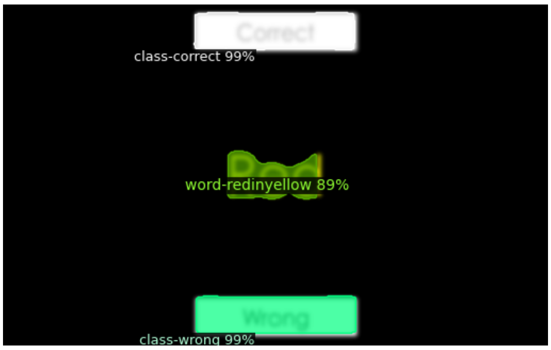

Figure 5.3: LR Image Segmentation

Training the Detectron2 Model

With the augmented dataset prepared and annotations in COCO Panoptic format, we proceeded to train a custom panoptic segmentation model using Detectron2's panoptic_fpn_R_50_3x architecture. This model combines a Feature Pyramid Network (FPN) with a ResNet-50 backbone, providing a balance between accuracy and computational efficiency.

The training process involved the following steps:

1. Configuration Setup: We configured the training parameters, including learning rate schedules, batch sizes, and augmentation settings, to suit our dataset's characteristics. Hyperparameters were tuned to optimize performance given the limited data.

2. Model Initialization: The model was initialized with weights pre-trained on the COCO dataset, allowing it to leverage learned features from general object categories and accelerate convergence.

3. Training Loop: The training loop iterated over the augmented dataset, adjusting the model weights to minimize the loss functions associated with panoptic segmentation. We monitored training metrics such as loss values and segmentation accuracy to assess progress and prevent overfitting.

4. Validation: Although our dataset was small, we set aside a portion for validation to evaluate the model's performance on unseen data. This step was crucial for ensuring that the model generalized well beyond the training samples.

After training, the model demonstrated satisfactory performance in detecting and segmenting the specified classes within the Stroop test images. The model's outputs included segmentation maps that delineated the locations and extents of each object class within the images. Fig. 5.2 Fig. 5.3

B. Computing Feature Maps for Dynamic Contextual Belief

With the trained Detectron2 model, we proceeded to compute the feature maps necessary for constructing the Dynamic Contextual Belief representations used in our IRL framework. The goal was to generate high-resolution (HR) and low-resolution (LR) belief maps for each image, capturing detailed and contextual information about object presence and locations.

Generation of High-Resolution Feature Maps

The high-resolution feature maps were derived directly from the original images using the trained model. For each image:

1. **Object Detection and Segmentation**: The model processed the image to detect and segment instances of the 13 defined classes, producing a panoptic segmentation map where each pixel was assigned a class label.

2. **One-Hot Encoding**: The segmentation map was transformed into a one-hot encoding format across the spatial grid of the image. Specifically, the image was divided into a grid of 32×20 patches, corresponding to the "patch_num": [32, 20] configuration used in our environment. Each patch, therefore, represented a region of 16×16 pixels (since the image dimensions are 512×320 pixels).

3. **Patch-Level Class Representation**: For each patch, we determined the presence of each class by inspecting the class labels within the patch area. The result was a one-hot encoded vector of length 13 for each patch, indicating which classes were present in that region.

The high-resolution feature maps thus provided detailed spatial information about object classes at a fine granularity, essential for modeling the agent's focal attention in the DCB framework.

Generation of Low-Resolution Feature Maps

To simulate peripheral vision and contextual awareness, we generated low-resolution feature maps by applying a Gaussian blur to the original images before processing them through the Detectron2 model. The process involved:

1. **Gaussian Blurring**: Each original image was subjected to Gaussian blur, reducing high-frequency details and simulating the lower acuity of peripheral vision.
2. **Object Detection on Blurred Images**: The blurred images were then input into the trained model to perform object detection and segmentation. While the model's accuracy may be reduced on blurred images, it still captured general contextual information about object presence. The goal is to simulate understanding of the object outside of the fovea radius.
3. One-Hot Encoding and Patch Division: Similar to the high-resolution maps, the resulting segmentation maps were divided into the same 32×20 grid of patches, and one-hot encoding was applied to represent class presence in each patch.

The low-resolution feature maps provided a broader, less detailed overview of the scene, capturing contextual cues that guide the agent's attention toward areas of potential interest.

Integration into Dynamic Contextual Belief Maps

Combining the high-resolution and low-resolution feature maps, we constructed the Dynamic Contextual Belief maps for each image:

1. **Initial State Representation**: At the beginning of each episode (i.e., when the agent starts processing an image), the belief state was initialized with the low-resolution feature map, representing the agent's initial contextual understanding of the scene.
2. **Belief Updating with Fixations**: As the agent selected fixation points during the simulation, the high-resolution feature information corresponding to the fixated patches was integrated into the belief state. This process involved:
    - **Masking**: Applying a mask to update only the patches corresponding to the fixation region, replacing the low-resolution data with high-resolution details.
    - **Dynamic Update**: The belief state was dynamically updated after each fixation, progressively refining the agent's understanding of the scene based on accumulated high-resolution information.

3. **One-Hot Encoding Across Patches**: The belief state at any given time thus comprised a combination of low-resolution and high-resolution feature maps, each represented by one-hot encoded vectors for the 13 classes across the 640 patches.

Role in State and Action Modeling

The computed contextual belief maps played a crucial role in the state representation within our IRL framework:

- **State Input for the Policy Network**: The belief maps served as input to the policy network, providing spatial and semantic information that guided the agent's fixation decisions.

- **Attention Mechanism**: By modeling the foveated nature of human vision, the belief maps enabled the agent to focus on areas of high relevance while maintaining contextual awareness of the broader scene.

- **Integration with EEG Data**: In conjunction with EEG-informed cognitive cues, the belief maps allowed the policy network to make fixation decisions that reflected both the visual context and the participant's cognitive state.

Challenges and Considerations

Several challenges arose during the computation of the contextual belief maps:

- **Limited Dataset and Model Generalization**: Despite data augmentation efforts, the limited diversity of the dataset posed challenges for the model's generalization. Care was taken to validate the model's performance and ensure that it reliably detected the objects of interest across the augmented samples.

- **Balance Between Resolution and Computational Efficiency**: The choice of patch size and grid dimensions involved trade-offs between the granularity of spatial information and computational demands. The 32×20 grid provided a suitable balance, offering sufficient detail without imposing excessive computational overhead.

- **Accuracy of Low-Resolution Detection**: Applying the model to blurred images for low-resolution feature maps introduced some inaccuracies due to the degradation of visual details. However, this approach effectively simulated peripheral vision and contributed valuable contextual information to the belief state.

5.2.2 IRL Framework

This subsection delves into the technical structure and hyperparameter settings for implementing the scanpath models in our experiments. Here, we outline the workflow, data preprocessing, and training nuances, emphasizing key modules, functionality, and parameter choices that drive the system.

Data Processing Pipeline

1. **Preprocessing and Patch-Based Representation**:

    - The visual input space is segmented into a grid of patches, defined by "patch_num": [32, 20], yielding a 32×20 patch grid for 512×320 images. Each patch measures 16×16 pixels, a size configured under "patch_size": [16, 16], effectively balancing granularity with computational efficiency. Therefore it yields 640 patches for each image.
    - The initial fixation for each image is marked within this patch grid to initialize the agent's attention. EEG data, where available, is also linked to each image for EEG-guided models, providing a continuous cognitive state signal that informs gaze distribution.

2. **Fixation Data Augmentation**:

    - Human scanpaths, categorized by task types (e.g., congruent vs. incongruent Stroop conditions), are processed into fixation trajectories capped at "max_traj_length": 6. This constraint ensures that the model predicts a bounded sequence of gaze shifts, maintaining consistency with human data.
    - An inhibition of return (IOR) mechanism is applied to discourage repeated fixations within a predefined area. The IOR radius is set by "IOR_size": 1, encouraging exploration of new regions within the visual space by penalizing revisits to previously attended patches.

3. **EEG Data Integration**:

    - For models incorporating cognitive states, EEG data is preprocessed and stored within the dataset for each trial. This data is later accessed in the policy network, influencing fixation predictions in EEG-enabled conditions. The EEG features are processed using a dense layer, concatenated with visual features to refine policy output.

Training and Model Architecture

1. **Policy and Discriminator Architecture**:
    - The IRL framework uses a Generative Adversarial Imitation Learning (GAIL) setup, where a generator (policy) and discriminator are trained simultaneously. The discriminator (LHF_Discriminator_Cond) differentiates between human and model-generated scanpaths, while the policy (LHF_Policy_Cond_Small) adapts its scanpath predictions to emulate human behavior.
    - The policy network is augmented with an EEG input layer that processes EEG data through a dense layer. The final EEG-enhanced feature is concatenated with visual features, enabling adaptive responses to cognitive states.

2. **Training Parameters and Hyperparameters**:
    - Training Loop Configuration: The Trainer class orchestrates the training loop, setting up logging, checkpoints, and policy evaluation. Training is conducted over num_epoch: 150 and num_step: 4, reflecting a limited run in this experimental setup but modifiable for larger training scales.
    - Batch Size and Device: Training uses a batch size of "batch_size": 4 for images and human gaze data, efficiently balancing memory use and training speed on CUDA-enabled devices.
    - Discount Factor and Advantage Estimation: "gamma": 0.9 defines the discount factor for rewards, favoring shorter fixation sequences. Generalized Advantage Estimation (GAE) is used, with "adv_est": "GAE" and "tau": 0.96, stabilizing training by reducing variance in policy gradient estimates.

3. **Proximal Policy Optimization (PPO) Specifics**:
    - Learning Rate and Gradient Clipping: The PPO optimizer's learning rate is set to "lr": 1e-05 with "clip_param": 0.15, controlling the step size and range for policy updates. These values are critical for managing policy shifts while maintaining stable convergence.
    - Value and Entropy Coefficients: The value coefficient, set to "value_coef": 0.5, ensures that the value function's prediction of future rewards has a balanced influence in the loss function. The entropy coefficient "entropy_coef": 0.01 introduces randomness to action selection, fostering exploration in gaze paths.

4. **Discriminator and GAIL Settings**:

- Learning Rate and Scheduler: The discriminator's learning rate ("gail_lr": 5e-05) and decay schedule are defined by "gail_milestones": [5], moderating the learning pace over time.
- Step Checkpoints: The system saves model checkpoints at intervals defined by "checkpoint_every": 1, supporting quick recovery and monitoring during training.

5. **Evaluation and Logging**:

   - Evaluation Frequency: The evaluation frequency, set at "evaluate_every": 5, governs how often the model is assessed on the validation set, balancing training focus with periodic validation. The log root ("log_root": "./assets") designates where results, losses, and performance metrics are saved for further analysis.

Chapter 6

Results

In this chapter, we present the results of our study on reconstructing human gaze behavior during the Stroop test using inverse reinforcement learning (IRL). The primary objective was to determine whether integrating electroencephalography (EEG) data with human fixation patterns could enhance the prediction accuracy of gaze scanpaths. Specifically, we aimed to assess if the inclusion of EEG-derived cognitive state information would lead to a significant improvement over models that rely solely on fixation data.

To achieve this, we developed and evaluated three models:

- **Human Performance Metrics**: Serving as a baseline, we analyzed the actual human gaze data collected during the Stroop test to establish a performance benchmark. This involved calculating various metrics directly from the participants' eye-tracking data without any modeling or prediction.

- **IRL-EEG Model**: This model incorporates both human fixation data and EEG signals within the IRL framework. By integrating EEG features into the state representation and policy network, the model aims to capture the underlying cognitive processes influencing gaze behavior, potentially leading to more accurate scanpath predictions.

- **IRL-Image Model**: This model utilizes only the fixation data within the IRL framework, excluding EEG information. It serves to isolate the contribution of visual information alone in predicting gaze behavior, allowing for a direct comparison with the IRL-EEG model to assess the added value of EEG integration.

For each model, we conducted experiments across all task conditions as well as specific subsets, namely congruent and incongruent Stroop trials. The congruent tasks involved color words displayed in matching font colors (e.g., the word "Red" written in red), while the incongruent tasks presented a mismatch between the word meaning and its font color (e.g., "Red" written in blue), thus imposing varying levels of cognitive load.

The results were evaluated using a set of quantitative metrics designed to measure different aspects of gaze prediction accuracy and similarity to human behavior:

- **Target Fixation Probability (TFP)**: Assesses the probability of fixating on task-relevant areas at each step of the scanpath, providing insights into how quickly and effectively the model directs attention to critical regions. In addition TFP-AUC (Area Under the Curve), summarizes the overall performance of TFP across all steps, offering a comprehensive measure of the model's ability to focus on target areas over time.

- **Sequence Score**: Evaluates the similarity between the predicted and actual fixation sequences, reflecting how closely the model replicates the temporal order of human gaze shifts.

- **Average Scanpath Ratio (SP Ratio)**: Compares the length of the predicted scanpath to that of the human scanpath, indicating the efficiency of the model's gaze behavior.

- **Probability Mismatch**: Measures the discrepancy between the predicted fixation probabilities and the actual human fixation distribution, highlighting differences in attentional allocation.

- **MultiMatch Scores**: Comprise a set of metrics—**Shape**, **Direction**, **Length**, and **Position**—that quantitatively assess the similarity between predicted and human scanpaths across various dimensions of spatial and temporal characteristics.

These metrics were computed for each model and task category, enabling a detailed comparison of performance. The findings provide insights into the effectiveness of EEG data integration in gaze prediction and the extent to which human fixation patterns alone can suffice in reconstructing gaze behavior during cognitively demanding tasks like the Stroop test.

In the following subsections, we delve into the specific results for each metric and model, analyzing their implications for our understanding of gaze behavior and the potential effects of incorporating EEG data into predictive models.

6.1 Target Fixation Probability AUC

Definition and Importance

Target Fixation Probability (TFP) assesses the likelihood of the scanpaths by measuring the number of fixations done to reach the answer. The TFP-AUC (Area Under the Curve) aggregates the TFP across all steps, from 1 to 6, providing a measure of how effectively and quickly a model directs attention to task-relevant regions over time. A higher TFP-AUC indicates faster performance in guiding gaze toward targets and with fewer steps could reach the answer.

Results Overview

We computed the TFP-AUC for the Human baseline, the IRL-EEG model, and the IRL-Image model across all tasks(average), as well as for the congruent and incongruent conditions separately. **Average**

- **Human**: TFP-AUC = 5.104
- **IRL-EEG**: TFP-AUC = 5.235
- **IRL-Image**: TFP-AUC = 5.504

Congruent Tasks

- **Human**: TFP-AUC = 5.125
- **IRL-EEG**: TFP-AUC = 4.689
- **IRL-Image**: TFP-AUC = 5.189

Incongruent Tasks

- **Human**: TFP-AUC = 5.083
- **IRL-EEG**: TFP-AUC = 5.782
- **IRL-Image**: TFP-AUC = 5.818

Analysis

- **Human Baseline**: The TFP-AUC for human participants was similar across congruent and incongruent tasks, indicating consistent performance in fixating on target areas regardless of cognitive load. Fig. 6.1

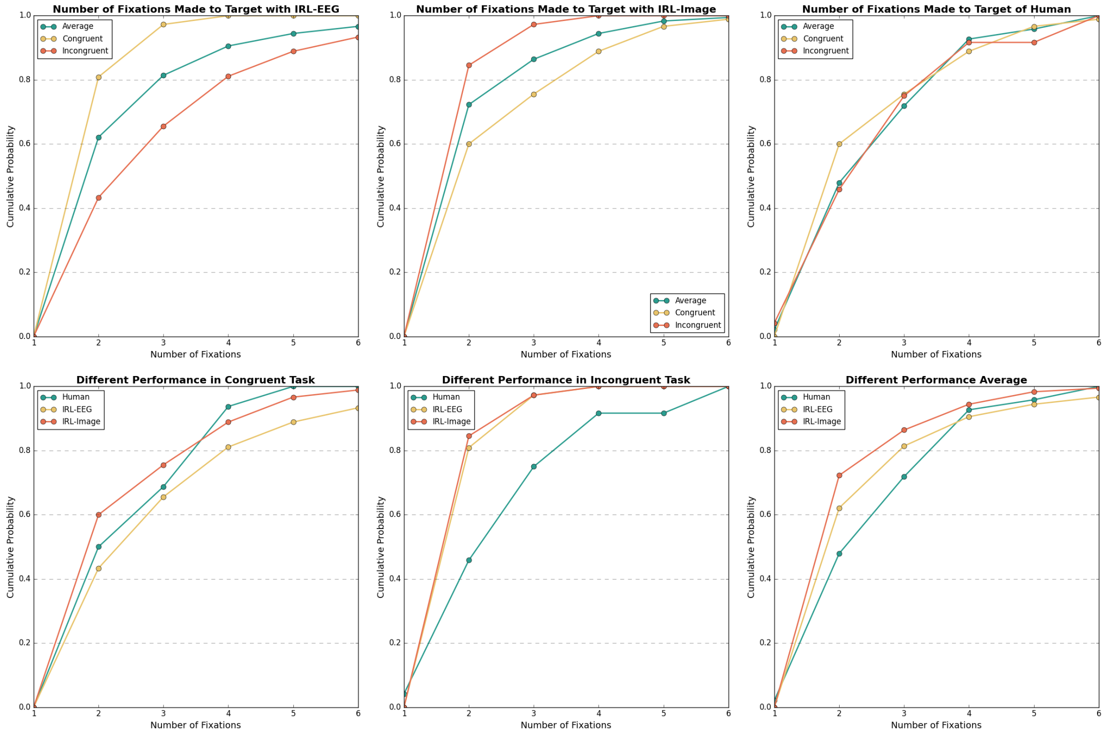

Figure 6.1: Comparison between number of fixation to reach the target

- **IRL-EEG Model**: Showed a higher TFP-AUC in incongruent tasks (5.782) compared to congruent tasks (4.689), suggesting that the model was more effective in directing gaze to target areas under higher cognitive load when EEG data was integrated. Fig. 6.1

- **IRL-Image Model**: Achieved higher TFP-AUC values across all task categories compared to the IRL-EEG model, with the highest TFP-AUC in incongruent tasks (5.818), indicating effective prediction of gaze behavior under cognitive interference using fixation data alone. Fig. 6.1

6.2 Probability Mismatch

Definition and Importance

Probability Mismatch measures the discrepancy between the predicted fixation probabilities generated by the model and the actual fixation distribution observed in human participants. It quantifies how well the model's attention allocation aligns with human behavior, with lower values indicating better alignment. In this comparison there is no option for the human-human comparison, because the mismatch is always zero and is meaningless. Lower probability mismatch means higher similarity.

Results Overview

Probability Mismatch was computed for the IRL-EEG and IRL-Image models across all tasks and separately for congruent and incongruent conditions. **Average**

- **IRL-EEG**: Probability Mismatch = 0.344
- **IRL-Image**: Probability Mismatch = 0.463

Congruent Tasks

- **IRL-EEG**: Probability Mismatch = 0.436
- **IRL-Image**: Probability Mismatch = 0.272

Incongruent Tasks

- **IRL-EEG**: Probability Mismatch = 0.782
- **IRL-Image**: Probability Mismatch = 0.818

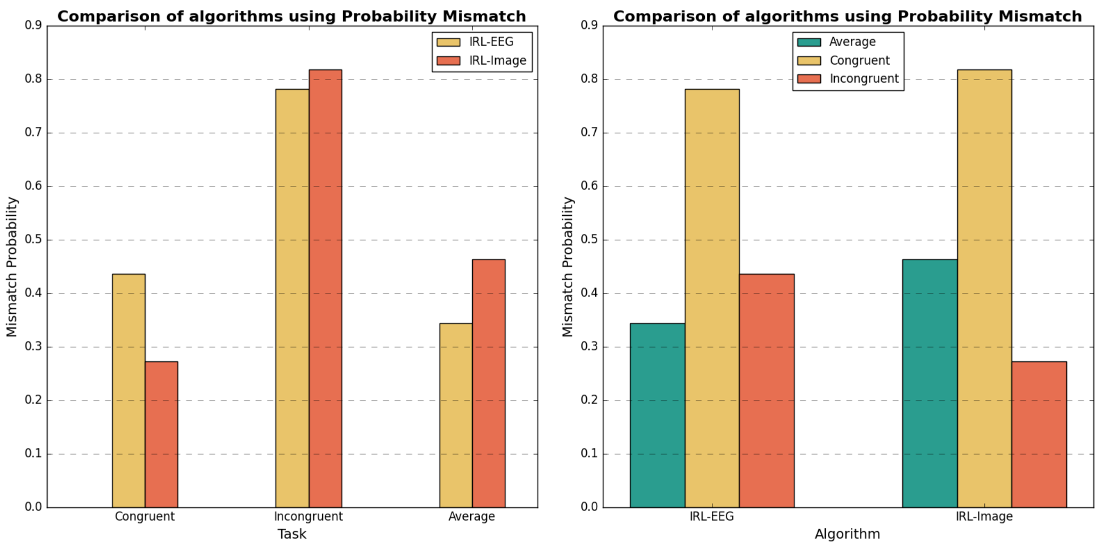

Figure 6.2: Comparison mismatch probability among the algorithms

Analysis

- **All Tasks(Average)**: The IRL-EEG model exhibited a lower Probability Mismatch compared to the IRL-Image model, indicating closer alignment with human fixation probabilities when EEG data was included. Fig. 6.2
- **Congruent Tasks**: The IRL-Image model had a significantly lower Probability Mismatch than the IRL-EEG model, suggesting that fixation data alone provided a better match to human attentional allocation in simpler tasks. Fig. 6.2
- **Incongruent Tasks**: Both models showed higher Probability Mismatch values, reflecting greater difficulty in aligning model predictions with human behavior under higher cognitive load. Fig. 6.2

6.3 Sequence Score

Definition and Importance

The Sequence Score evaluates the similarity between the predicted fixation sequences and the actual sequences observed in human participants. It reflects how closely the model replicates the temporal order and selection of gaze shifts, with higher scores indicating greater similarity.

Results Overview

Sequence Scores were calculated for the Human baseline, IRL-EEG model, and IRL-Image model across all tasks and for congruent and incongruent conditions.
Average

- **Human**: Sequence Score = 0.806
- **IRL-EEG**: Sequence Score = 0.427
- **IRL-Image**: Sequence Score = 0.398

Congruent Tasks

- **Human**: Sequence Score = 0.865
- **IRL-EEG**: Sequence Score = 0.481
- **IRL-Image**: Sequence Score = 0.426

Incongruent Tasks

- **Human**: Sequence Score = 0.767
- **IRL-EEG**: Sequence Score = 0.382
- **IRL-Image**: Sequence Score = 0.375

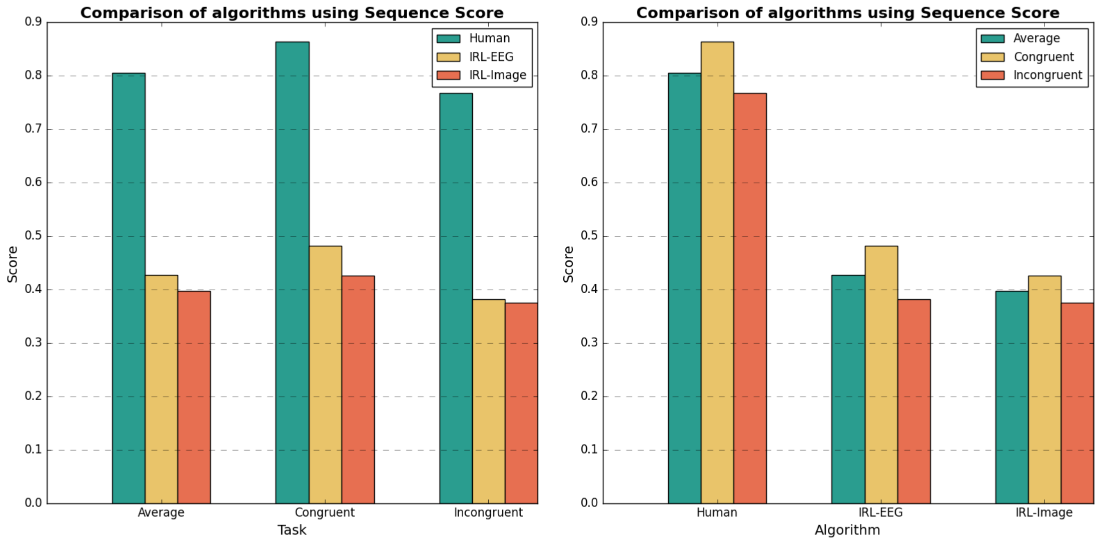

Figure 6.3: Comparison sequence score

Analysis

- **Human Baseline**: High Sequence Scores indicate consistent and coherent fixation sequences during the Stroop task. Fig. 6.3
- **IRL-EEG Model**: Achieved higher Sequence Scores compared to the IRL-Image model across all task categories, suggesting that EEG data improved the temporal replication of human gaze sequences. Fig. 6.3
- **IRL-Image Model**: Showed lower Sequence Scores, indicating less similarity to human fixation sequences, particularly in incongruent tasks. Fig. 6.3

6.4 MultiMatch Analysis

MultiMatch is a comprehensive method for comparing scanpaths by quantifying their similarity across multiple dimensions: Shape, Direction, Length, and Position. Higher scores indicate greater similarity between the predicted and human scanpaths in each dimension.

6.4.1 Shape

Definition and Importance

Shape measures the geometric similarity between scanpaths, considering the overall pattern formed by the sequence of fixations, irrespective of their exact locations.

Results Overview

Average

- **Human**: Shape = 0.928
- **IRL-EEG**: Shape = 0.842
- **IRL-Image**: Shape = 0.847

Congruent Tasks

- **Human**: Shape = 0.944
- **IRL-EEG**: Shape = 0.848
- **IRL-Image**: Shape = 0.854

Incongruent Tasks

- **Human**: Shape = 0.917
- **IRL-EEG**: Shape = 0.837
- **IRL-Image**: Shape = 0.840

Analysis

- Both the IRL-EEG and IRL-Image models achieved similar Shape scores, slightly lower than the human baseline. Fig. 6.4
- The highest Shape scores were observed in congruent tasks for all models. Fig. 6.4

6.4.2 Direction

Definition and Importance

Direction evaluates the similarity in the orientation of saccades between scanpaths, considering the sequence of gaze shifts irrespective of their length or exact position.

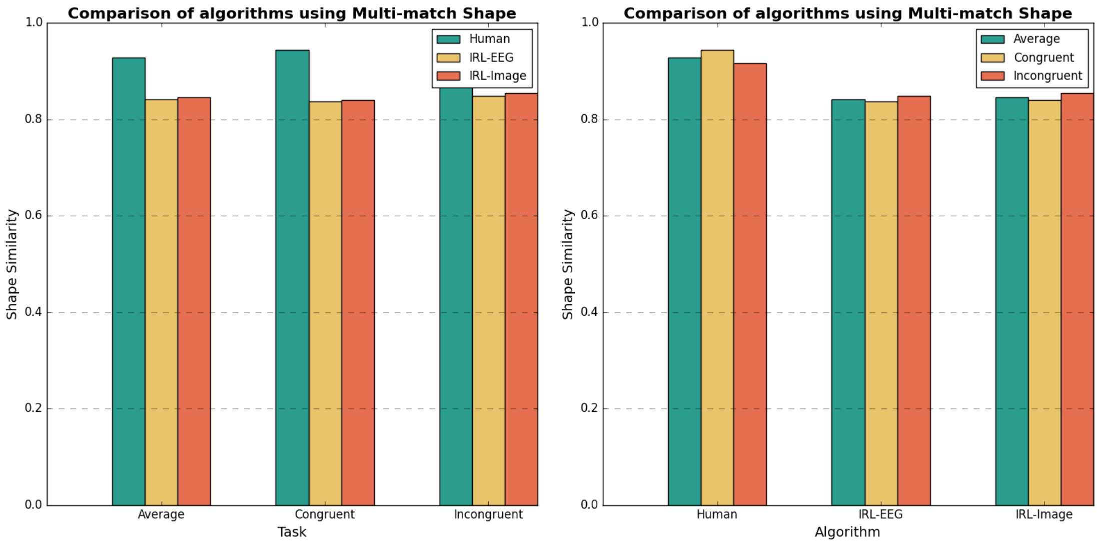

Figure 6.4: Comparison the value of Multi-match Shape

Results Overview

Average

- **Human**: Direction = 0.786
- **IRL-EEG**: Direction = 0.634
- **IRL-Image**: Direction = 0.648

Congruent Tasks

- **Human**: Direction = 0.858
- **IRL-EEG**: Direction = 0.641
- **IRL-Image**: Direction = 0.664

Incongruent Tasks

- **Human**: Direction = 0.738
- **IRL-EEG**: Direction = 0.629
- **IRL-Image**: Direction = 0.635

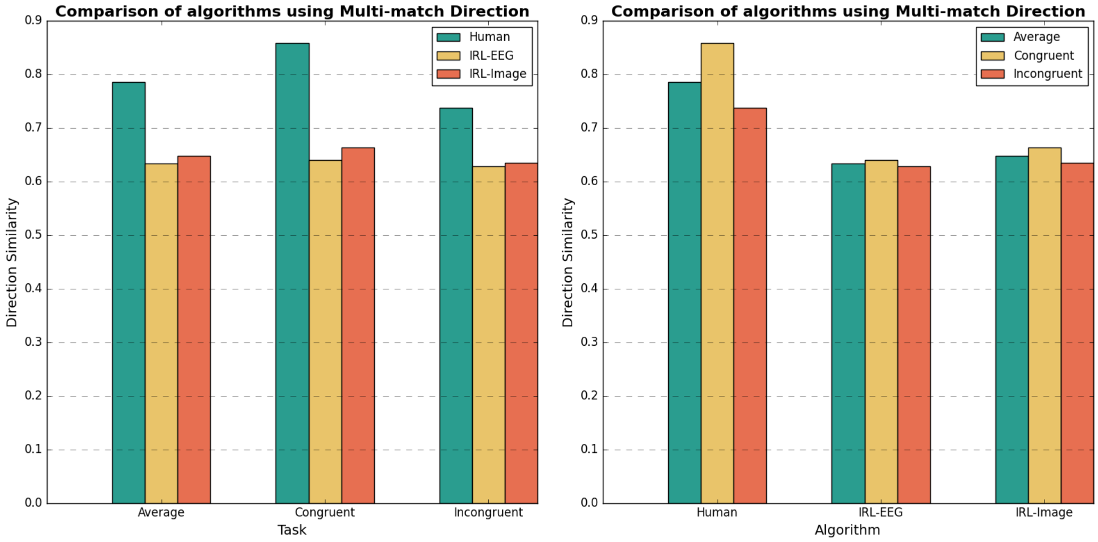

Figure 6.5: Comparison the value of Multi-match Direction

Analysis

- The IRL-Image model slightly outperformed the IRL-EEG model in Direction similarity across all task categories. Fig. 6.5
- Both models showed lower Direction scores compared to the human baseline. Fig. 6.5

6.4.3 Length

Definition and Importance

Length measures the similarity in saccade amplitudes between scanpaths, focusing on the distances covered during gaze shifts.

Results Overview

Average

- **Human**: Length = 0.930
- **IRL-EEG**: Length = 0.839
- **IRL-Image**: Length = 0.839

Congruent Tasks

- **Human**: Length = 0.958
- **IRL-EEG**: Length = 0.870
- **IRL-Image**: Length = 0.870

Incongruent Tasks

- **Human**: Length = 0.911
- **IRL-EEG**: Length = 0.814
- **IRL-Image**: Length = 0.814

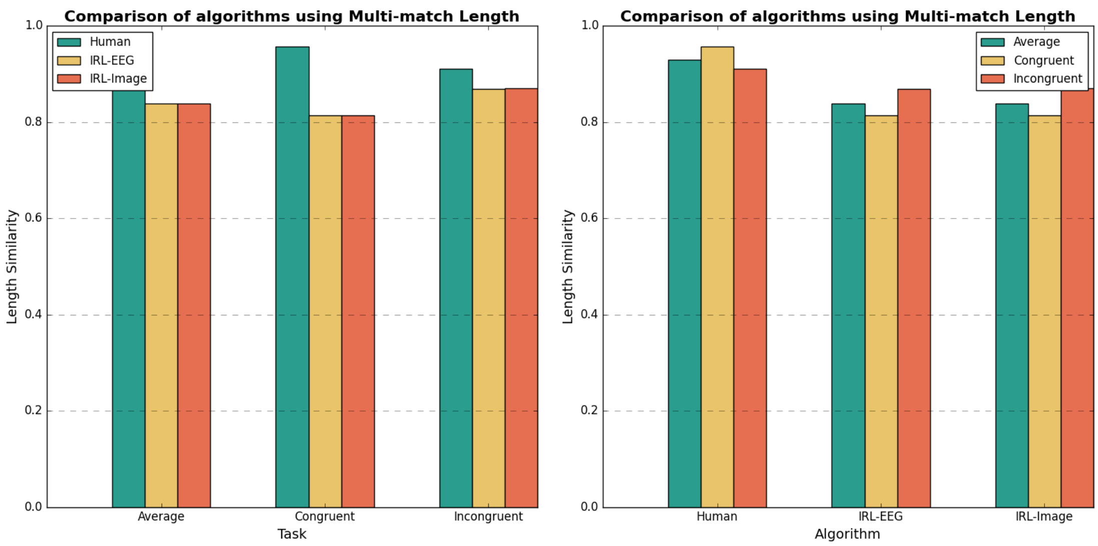

Figure 6.6: Comparison the value of Multi-match Length

Analysis

- Both models achieved identical Length scores across all task categories, slightly lower than the human baseline. Fig. 6.6
- The highest Length scores were in congruent tasks. Fig. 6.6
-

6.4.4 Position

Definition and Importance

Position assesses the spatial similarity between scanpaths, considering the exact locations of fixations.

Results Overview

Average

- **Human**: Position = 0.927
- **IRL-EEG**: Position = 0.831
- **IRL-Image**: Position = 0.840

Congruent Tasks

- **Human**: Position = 0.935
- **IRL-EEG**: Position = 0.806
- **IRL-Image**: Position = 0.819

Incongruent Tasks

- **Human**: Position = 0.922
- **IRL-EEG**: Position = 0.851
- **IRL-Image**: Position = 0.857

Analysis

- Both models achieved identical Length scores across all task categories, slightly lower than the human baseline. Fig. 6.7
- The highest Length scores were in congruent tasks. Fig. 6.7
-

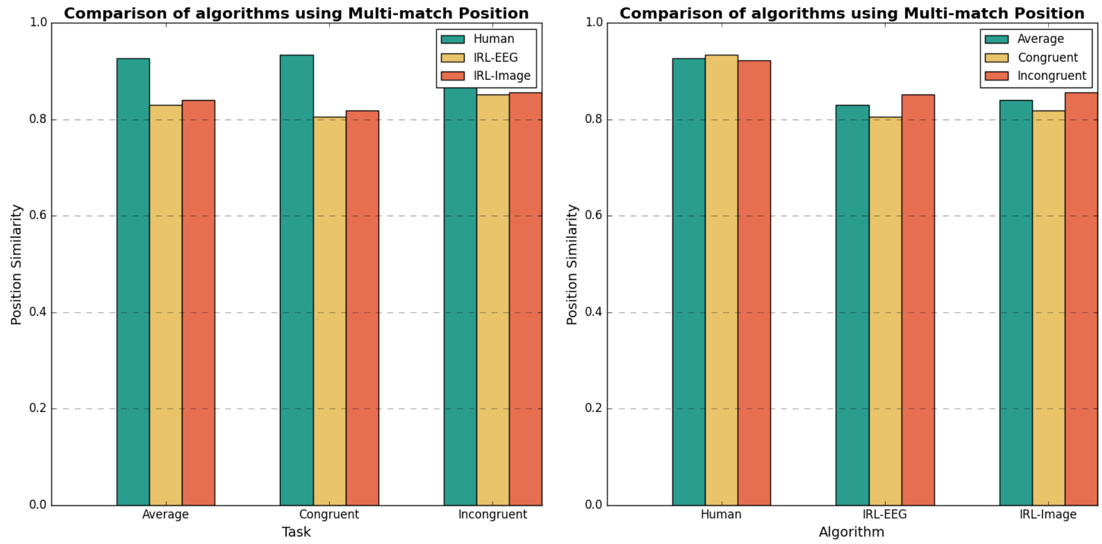

Figure 6.7: Comparison the value of Multi-match Position

6.5 Scanpath Ratio

Definition and Importance

The Scanpath Ratio (SP Ratio) compares the length of the predicted scanpath to the length of the human scanpath. It provides an indication of the efficiency and realism of the model's gaze behavior, with a value closer to 1 indicating a closer match to human scanpath length.

Results Overview

Sequence Scores were calculated for the Human baseline, IRL-EEG model, and IRL-Image model across all tasks and for congruent and incongruent conditions.
Average

- **Human**: Average SP Ratio = 0.721
- **IRL-EEG**: Average SP Ratio = 0.687
- **IRL-Image**: Average SP Ratio = 0.743

Congruent Tasks

- **Human**: Average SP Ratio = 0.664
- **IRL-EEG**: Average SP Ratio = 0.503

- **IRL-Image**: Average SP Ratio = 0.598

Incongruent Tasks

- **Human**: Average SP Ratio = 0.760
- **IRL-EEG**: Average SP Ratio = 0.832
- **IRL-Image**: Average SP Ratio = 0.860

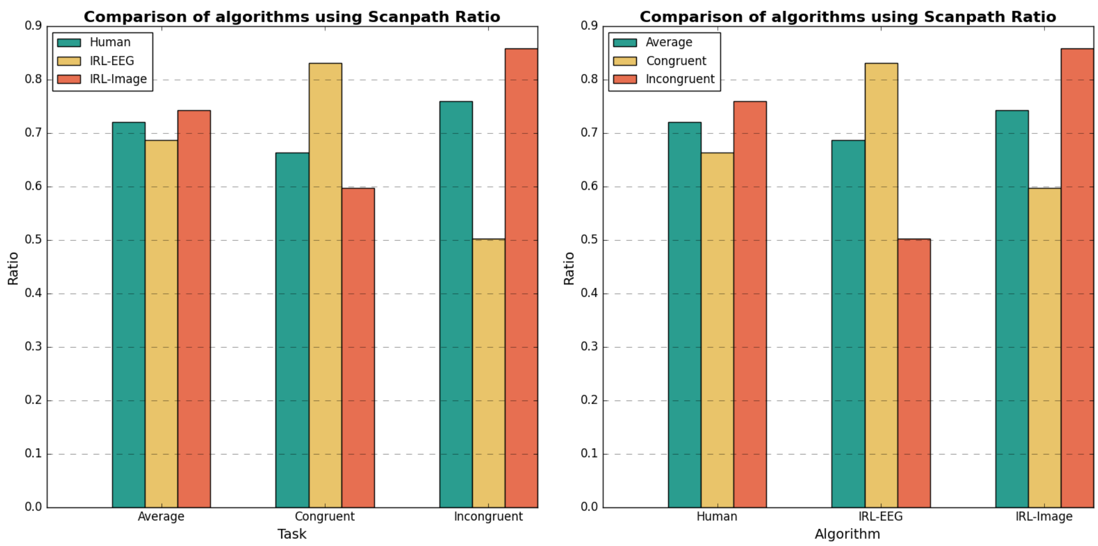

Figure 6.8: Comparison sequence score

Analysis

- **Average**: The IRL-Image model had a Scanpath Ratio closer to the human baseline, indicating better overall efficiency in scanpath length.Fig. 6.8
- **Congruent Tasks**: Both models had lower Scanpath Ratios compared to the human baseline, with the IRL-EEG model showing a notably lower value.Fig. 6.8
- **Incongruent Tasks**: Both models exceeded the human Scanpath Ratio, indicating longer scanpaths under higher cognitive load.Fig. 6.8

		TFP-AUC	Probability Mismatch	Sequence Score	SP Ratio	Multimatch			
						Shape	Direction	Length	Position
Human	Average	5.10	-	0.8062	0.7206	0.9277	0.7861	0.9297	0.9273
	Congruent	5.12	-	0.8645	0.6637	0.9441	0.8583	0.9577	0.9348
	Incongruent	5.08	-	0.7673	0.7601	0.9169	0.7381	0.9109	0.9223
IRL-EEG	Average	5.23	0.3436	0.4266	0.6869	0.8423	0.6344	0.8390	0.8307
	Congruent	4.68	0.4361	0.4814	0.5030	0.8482	0.6410	0.8696	0.8056
	Incongruent	5.78	0.7818	0.3818	0.8322	0.8374	0.6290	0.8141	0.8513
IRL-Image	Average	5.50	0.4632	0.3979	0.7425	0.8465	0.6480	0.8393	0.8398
	Congruent	5.18	0.2722	0.4259	0.5979	0.8542	0.6639	0.8704	0.8192
	Incongruent	5.81	0.8181	0.3749	0.8596	0.8402	0.6350	0.8138	0.8567

Table 6.1: Summary of results of different metrics over algorithms and tasks

Chapter 7

Discussion

The present study explored the efficacy of integrating electroencephalography (EEG) data with eye-tracking metrics within an inverse reinforcement learning (IRL) framework to reconstruct and predict human gaze behavior during the Stroop test. By comparing models that utilized both fixation data and EEG signals (IRL-EEG) against those relying solely on fixation data (IRL-Image) and the human baseline, this research aimed to discern the added value of neural data in enhancing gaze prediction accuracy under varying cognitive loads.

7.1 Interpretation of Key Findings

The analysis revealed that integrating EEG data with fixation patterns did not result in substantial improvements in gaze prediction accuracy compared to models utilizing fixation data alone. Both the IRL-EEG and IRL-Image models demonstrated comparable performance across most metrics, indicating that the inclusion of EEG-derived cognitive state information did not significantly enhance the prediction of gaze scanpaths during the Stroop test.

- **Target Fixation Probability Area Under the Curve (TFP-AUC)**: Both models showed similar TFP-AUC values across all tasks, with the IRL-Image model slightly outperforming the IRL-EEG model in incongruent tasks (IRL-Image: 5.818 vs. IRL-EEG: 5.782). This suggests that fixation data alone is sufficiently effective in directing gaze toward task-relevant areas, even under high cognitive load conditions.

- **Probability Mismatch**: The IRL-EEG model exhibited a slightly lower Probability Mismatch (0.344) compared to the IRL-Image model (0.463) across all tasks. However, this difference was not substantial enough to indicate a significant advantage of EEG integration.

- **Sequence Score**: Both models achieved similar Sequence Scores (IRL-EEG: 0.427 vs. IRL-Image: 0.398), indicating that the temporal order of gaze shifts was equally well replicated by both approaches.

- **MultiMatch Analysis**: The IRL-Image model marginally outperformed the IRL-EEG model in Direction and Position similarities, while the IRL-EEG model showed a slight edge in capturing the Sequence Score. Overall, the differences were minimal, reinforcing the conclusion that EEG integration does not markedly enhance gaze prediction.

- **Scanpath Ratio (SP Ratio)**: Both models presented comparable SP Ratios, with the IRL-Image model slightly closer to the human baseline (SP Ratio: IRL-Image = 0.743 vs. IRL-EEG = 0.687). This indicates that the efficiency and realism of the scanpath lengths were similarly maintained by both models.

Overall, the findings suggest that while EEG data can provide additional cognitive state information, its integration within the IRL framework did not lead to significant improvements in predicting human gaze behavior during the Stroop test compared to using fixation data alone.

7.2 Implications of Findings

Enhancing Gaze Prediction Models

The comparable performance of the IRL-EEG and IRL-Image models underscores that fixation data alone may be sufficient for accurate gaze prediction in tasks like the Stroop test. This implies that the added complexity of integrating EEG data may not be necessary for certain applications, potentially simplifying model architecture and reducing computational overhead without compromising prediction accuracy. During the Stroop test, because there are some limited allowed slides, the model would quickly find out the prediction only by using fixations and somehow it ignores EEG data because fixations (DCBs) have become wild cards in predicting the next fixations and it can proceed without EEG data.

Clinical and Diagnostic Applications

Given that EEG integration did not substantially improve gaze prediction, the use of fixation data alone could be a more efficient approach for clinical diagnostics and monitoring. Eye-tracking metrics, without the need for simultaneous EEG recording, offer a less intrusive and more cost-effective method for assessing cognitive functions and detecting neurodegenerative or attention-related disorders.

Cognitive Neuroscience Research

From a cognitive neuroscience perspective, the findings suggest that while EEG data provides valuable insights into neural activity, its role in enhancing behavioral predictions like gaze scanpaths may be limited in certain contexts. Future research might explore other ways to leverage neural data or investigate different cognitive tasks where EEG integration could have a more pronounced impact on behavioral modeling.

7.3 Limitations

Despite the insights gained, several limitations must be acknowledged:

Nature of Stroop Test

In the Stroop test under NIH standards, there's a strict restriction that prevents the use of arbitrary colors. This could result in very difficult to manage overfitting during the training process. This could lead the model to rapidly detect the answer only by seeing the features of the words, so it somehow ignores the EEG data. In other words, we have to use train DCBs inside the test set, in this case finding a solution to prevent the model from learning too much from the fixations and relying on the EEG data more.

Way of Extracting DCBs

Potentially there's another way of extracting DCB and object annotation that could prevent the model from easily overfitting on the objects that have been seen before.

Sample Size and Diversity

The study's sample size, comprising ten participants, is relatively small and may limit the generalizability of the results. Future studies with larger and more diverse populations are necessary to validate the observed trends and ensure that the models perform consistently across different demographic and cognitive profiles. Even in the context of data augmentation, because of simulating the gaze foveal, using some of the transforms that blur or downsampling were restricted because they became strongly accurate in detecting the images in Low-res mode, that it could prevent us from simulating objects out of the foveal correctly.

EEG Data Integration Complexity

Integrating EEG data into the IRL framework introduces additional layers of complexity, including challenges in effective feature extraction and synchronization with eye-tracking data. The current implementation may not fully exploit the temporal richness of EEG signals, potentially constraining the model's capacity to capture nuanced cognitive states.

7.4 Future Work

Building upon the current study's findings, future research could explore the following avenues:

Changing the Way of Extracting DCBs

The most important effect could be done by using another way of extracting DCBs, particularly like annotating objects as different standards may change the results. It could prevent the model from detecting the answer quickly as soon as it sees the object feature map in belief maps.

Expanding the Dataset

Increasing the number of participants and incorporating a more diverse demographic sample would enhance the robustness and generalizability of the models. A larger dataset would also facilitate the application of more sophisticated EEG feature extraction techniques, potentially uncovering deeper insights into the neural underpinnings of gaze behavior.

Refining EEG Feature Extraction

Employing advanced EEG analysis methods, such as machine learning-based feature extraction or time-frequency analysis, could improve the integration of neural data into the IRL framework. Capturing more granular aspects of cognitive states may enhance the model's ability to replicate complex gaze patterns.

Diverse Cognitive Tasks

Extending the framework to include a variety of cognitive tasks beyond the Stroop test would test the model's versatility and uncover task-specific dynamics in gaze behavior. This would also allow for a broader evaluation of the benefits and limitations of EEG integration across different cognitive contexts.

Refining Model Architecture

Further refinement of the IRL framework, potentially incorporating multi-modal deep learning architectures, could enhance the synergy between EEG and eye-tracking data. Exploring alternative reinforcement learning algorithms beyond Proximal Policy Optimization (PPO) may also yield improvements in model stability and performance.

Chapter 8

Conclusion

This dissertation aimed to investigate the potential benefits of integrating electroencephalography (EEG) data with eye-tracking metrics within an inverse reinforcement learning (IRL) framework to reconstruct and predict human gaze behavior during the Stroop test. The primary objective was to determine whether EEG-derived cognitive state information could enhance the accuracy of gaze scanpath predictions, particularly under varying levels of cognitive load.

Through the development and evaluation of two primary models—the IRL-EEG model, which incorporated both fixation data and EEG signals, and the IRL-Image model, which utilized only fixation data—the study sought to discern the added value of neural data in gaze prediction. Comprehensive assessments using metrics such as Target Fixation Probability Area Under the Curve (TFP-AUC), Probability Mismatch, Sequence Score, MultiMatch Analysis, and Scanpath Ratio (SP Ratio) revealed that both models performed comparably across most metrics. Notably, the integration of EEG data did not result in significant improvements in gaze prediction accuracy compared to the model relying solely on fixation data. Both the IRL-EEG and IRL-Image models exhibited similar performances, indicating that fixation patterns alone possess substantial predictive power in directing gaze towards task-relevant areas, even under high cognitive load conditions like incongruent Stroop trials.

The findings suggest that the added complexity of incorporating EEG data may not be necessary for certain applications, particularly those involving tasks with limited and highly structured stimuli, such as the Stroop test. This has important implications for the design of gaze prediction models, suggesting that simpler models using only eye-tracking data can achieve comparable accuracy without the additional computational and logistical burdens associated with EEG integration.

From a clinical and diagnostic perspective, the results indicate that eye-tracking metrics alone could serve as a more efficient and cost-effective tool for assessing cognitive functions and detecting neurodegenerative or attention-related disorders.

This non-intrusive method offers significant advantages in terms of ease of implementation and participant comfort, making it a viable option for widespread clinical use.

However, the study is not without limitations. The relatively small and homogeneous sample size of ten participants may limit the generalizability of the results. Additionally, the specific nature of the Stroop test, with its stringent color-word combinations, may have facilitated overfitting, allowing the models to rely heavily on fixation patterns while marginalizing the contribution of EEG data. Future research should address these limitations by expanding the participant pool, exploring alternative methods for extracting and integrating Dynamic Contextual Belief (DCB) maps, and testing the framework across a broader range of cognitive tasks to fully understand the potential benefits and constraints of multimodal data integration in gaze behavior modeling.

In conclusion, while the theoretical promise of EEG integration for enhancing gaze prediction exists, this study demonstrates that, in the context of the Stroop test, fixation data alone is sufficient for accurate gaze behavior reconstruction. These insights contribute to the ongoing discourse in cognitive neuroscience and machine learning, highlighting the importance of evaluating the practical utility of multimodal data integration and encouraging further exploration into optimizing gaze prediction frameworks for diverse cognitive applications.

Bibliography

[1] Philipp Kanske and Sonja A. Kotz. «Positive emotion speeds up conflict processing: ERP responses in an auditory Simon task». In: *Biological Psychology* 87.1 (2011), pp. 122–127. ISSN: 0301-0511. DOI: https://doi.org/10.1016/j.biopsycho.2011.02.018. URL: https://www.sciencedirect.com/science/article/pii/S0301051111000664 (cit. on p. 1).

[2] Andrew Y Ng, Stuart Russell, et al. «Algorithms for inverse reinforcement learning.» In: *Icml*. Vol. 1. 2. 2000, p. 2 (cit. on p. 1).

[3] Tim J Anderson and Michael R MacAskill. «Eye movements in patients with neurodegenerative disorders». In: *Nature Reviews Neurology* 9.2 (2013), pp. 74–85 (cit. on p. 8).

[4] Megan Rose Readman, Megan Polden, Melissa Chloe Gibbs, Lettie Wareing, and Trevor J Crawford. «The potential of naturalistic eye movement tasks in the diagnosis of Alzheimer's disease: A review». en. In: *Brain Sci.* 11.11 (Nov. 2021), p. 1503 (cit. on p. 8).

[5] A.P.A. Bueno, J.R. Sato, and M. Hornberger. «Eye tracking – The overlooked method to measure cognition in neurodegeneration?» In: *Neuropsychologia* 133 (2019), p. 107191. ISSN: 0028-3932. DOI: https://doi.org/10.1016/j.neuropsychologia.2019.107191. URL: https://www.sciencedirect.com/science/article/pii/S0028393219302349 (cit. on p. 8).

[6] Sandra Weintraub et al. «Cognition assessment using the NIH Toolbox». In: *Neurology* 80.11_supplement_3 (2013), S54–S64. DOI: 10.1212/WNL.0b013e3182872ded. eprint: https://www.neurology.org/doi/pdf/10.1212/WNL.0b013e3182872ded. URL: https://www.neurology.org/doi/abs/10.1212/WNL.0b013e3182872ded (cit. on p. 8).

[7] Malihe Moghadami, Sahar Moghimi, Ali Moghimi, Gholam Reza Malekzadeh, and Javad Salehi Fadardi. «The Investigation of Simultaneous EEG and Eye Tracking Characteristics During Fixation Task in Mild Alzheimer's Disease». In: *Clinical EEG and Neuroscience* 52.3 (2021). PMID: 32539459, pp. 211–220. DOI: 10.1177/1550059420932752. eprint: https://doi.org/10.1177/

1550059420932752. URL: https://doi.org/10.1177/1550059420932752 (cit. on p. 8).

[8] Sinead Gaubert et al. «EEG evidence of compensatory mechanisms in preclinical Alzheimer's disease». In: *Brain* 142.7 (June 2019), pp. 2096–2112. ISSN: 0006-8950. DOI: 10.1093/brain/awz150. eprint: https://academic.oup.com/brain/article-pdf/142/7/2096/28881336/awz150.pdf. URL: https://doi.org/10.1093/brain/awz150 (cit. on p. 9).

[9] S Akshay, Y J Megha, and Chethan Babu Shetty. «Machine Learning Algorithm to Identify Eye Movement Metrics using Raw Eye Tracking Data». In: *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*. 2020, pp. 949–955. DOI: 10.1109/ICSSIT48917.2020.9214290 (cit. on p. 9).

[10] Saurabh Arora and Prashant Doshi. «A survey of inverse reinforcement learning: Challenges, methods and progress». In: *Artificial Intelligence* 297 (2021), p. 103500. ISSN: 0004-3702. DOI: https://doi.org/10.1016/j.artint.2021.103500. URL: https://www.sciencedirect.com/science/article/pii/S0004370221000515 (cit. on p. 9).

[11] Bing Du, Xiaomu Cheng, Yiping Duan, and Huansheng Ning. «fMRI Brain Decoding and Its Applications in Brain–Computer Interface: A Survey». In: *Brain Sciences* 12.2 (2022). ISSN: 2076-3425. DOI: 10.3390/brainsci12020228. URL: https://www.mdpi.com/2076-3425/12/2/228 (cit. on p. 10).

[12] Zhibo Yang, Lihan Huang, Yupei Chen, Zijun Wei, Seoyoung Ahn, Gregory Zelinsky, Dimitris Samaras, and Minh Hoai. «Predicting goal-directed human attention using inverse reinforcement learning». In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020, pp. 193–202 (cit. on pp. 10, 20).

[13] Jiaqi Gong, Shengting Cao, Soroush Korivand, and Nader Jalili. «Reconstructing human gaze behavior from EEG using inverse reinforcement learning». In: *Smart Health* 32 (2024), p. 100480 (cit. on pp. 18–20).

[14] Xishi Zhu, Soroush Korivand, Kittson Hamill, Nader Jalili, and Jiaqi Gong. «A comprehensive decoding of cognitive load». In: *Smart Health* 26 (2022), p. 100336 (cit. on pp. 18–20).