

Early Detection of Human Decision-Making in Concealed Object Visual Searching Tasks: An EEG-BiLSTM Study

Xuan-The Tran¹, Thomas (Tien-Thong) Do¹ and Chin-Teng Lin¹ *Fellow, IEEE*

Abstract—Detecting concealed objects presents a significant challenge for human and artificial intelligent systems. Detecting concealed objects task necessitates a high level of human attention and cognitive effort to complete the task successfully. Thus, in this study, we use concealed objects as stimuli for our decision-making experimental paradigms to quantify participants’ decision-making performance. We applied a deep learning model, Bi-directional Long Short Term Memory (BiLSTM), to predict the participant’s decision accuracy by using their electroencephalogram (EEG) signals as input. The classifier model demonstrated high accuracy, reaching 96.1% with an epoching time range of 500 ms following the stimulus event onset. The results revealed that the parietal-occipital brain region provides highly informative information for the classifier in the concealed visual searching tasks. Furthermore, the neural mechanism underlying the concealed visual-searching and decision-making process was explained by analyzing serial EEG components. The findings of this study could contribute to the development of a fault alert system, which has the potential to improve human decision-making performance.

I. INTRODUCTION

Decision-making is a ubiquitous aspect of daily life and is crucial for human survival. The decision-making process comprises three primary stages: decision-making preparation, decision-making itself, and decision-making evaluation [1]. Early decision-making prediction means we predict the human decision in the decision-making preparation stage. Early decision-making is also a critical component of a fault alert system, which aims to identify the possibility of human error before a decision is made. Task complexity and individual variability both play a role in the decision-making process. Neurocognitive research has made strides in quantifying the decision-making process through behavioural parameters such as reaction time and decision accuracy. Hierarchical Drift Diffusion Models (HDDM) [2] are commonly employed in the analysis of decision-making behaviour. However, the HDDM model is primarily descriptive and lacks explanatory power regarding the underlying mechanisms of the decision-making process. Additionally, the HDDM model is typically utilized in the context of discrimination-based decision-making tasks.

Computer vision techniques divide the objects of detection tasks into two categories: salient and concealed objects. Salient objects, which are transparent and easy to detect, contrast with concealed objects which pose a greater challenge for human observers and machine learning algorithms.

¹X.-T. T., T. D. and C.-T. L. are with the University of Technology Sydney, Faculty of Engineering and Information Technology, Australian Artificial Intelligence Institute, GrapheneX-UTS Human-centric AI Centre, 15 Broadway, Ultimo, New South Wales 2007, Australia.

Concealed objects are often similar to the image background in terms of colour, texture, and shape. The most popular use case of concealed object detection is in medical endoscopic imaging, such as polyp segmentation and lung infection segmentation [3]. Concealed object detection can also be used for tasks such as detecting locusts to prevent invasions and identifying rare species.

A number of decision-making studies have employed visual searching tasks as a means of investigation. The Patrol [4] paradigms utilize short videos in which participants are required to detect whether a soldier is wearing a helmet or a cap. The Realistic Search paradigm [5] employs multiple objects images that contain penguins and possibly a polar bear, and the participants’ task is detecting the presence of a polar bear. The Suspect Detection experimental paradigm [6] requires participants to detect human appearance in the stimulus images. All of these experimental paradigms involve discrimination tasks that use binary stimuli, where the target either appears or does not appear in videos or images. Participants are also required to make binary decisions (yes or no, target or non-target). However, the use of binary stimuli can lead to the decision-making results being influenced by the randomness of participants’ responses, with a chance level of 1/2, particularly if participants are not fully engaged in the task. Our paradigm offers six options corresponding to the six possible positions of an object within the image, with a chance level of 1/6. As a result, the likelihood of correct answers by chance is lower ($1/6 < 1/2$).

Neuroimaging techniques, such as electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and electrocorticography (ECoG), have the potential to provide insight to explain the decision-making process in the human brain [7]. These methods are able to capture brain activity while participants perform decision-making tasks. Furthermore, advanced data-driven techniques can be applied to neuroimaging data to classify or predict human decisions prior to their actual realization. Therefore, in this research, we employed concealed objects as stimuli in our experimental paradigm to elicit such brain activity.

The objective of this study is to propose a novel visual searching experimental paradigm for decision-making research. The paradigm involves participants detecting concealed objects and making decisions regarding the object’s location within the images. This approach aims to minimize the number of arbitrary decisions and fortunate trials. In addition, using a Bidirectional Long Short Term Memory model with various EEG temporal epoch lengths allows early prediction of the participant’s correct or incorrect decisions.

II. METHODS

A. Participants

This study recruited ten healthy participants (1 female, aged 20-38 years, mean \pm SD: 25.1 ± 4.2 ; 2 left-handed) who all had normal or corrected-to-normal vision. All participants provided informed consent, which was reviewed and approved by the ethical committee of the University of Technology Sydney, Australia (Approval Grant Number: UTS HREC REF NO. ETH22-7038).

Participants took part in four block tests using the same experimental paradigm with different object targets. Each block test comprised 50 trials and took 10 minutes to complete. Before the formal tests, participants completed a practice test comprising 10 trials to familiarize themselves with the task. Participants were given 5 minutes of break time between each test. The entire experiment lasted 2.5 hours, and participants received AUD 60 as compensation for their time.

B. Image dataset and experimental paradigm

Two hundred fifty images were chosen from the publicly available concealed image dataset COD10K [3]. These images contain only one animal object that is challenging to detect. All images were resized to 1000 x 600 pixels in order to enhance visual clarity during the experimental task and to minimize head movement during the object-searching process.

The experimental paradigm is illustrated in Figure 1. Each trial began with the hint, which is the animal species, for 2 seconds, followed by a 1-second fixation period before the image containing the hinted object was presented for 3 seconds. The thin grid lines divided the image into six equally sized areas, and participants were instructed to locate and indicate which area the object was located in. Another fixation period of 1 second was presented before participants had 2 seconds to indicate their decision by pressing one of six keyboard buttons (1, 2, 3, 4, 5 and 6). The correct position of the object in the image was subsequently highlighted for 2 seconds before a 2-second resting period; after that, the next trial began. Each participant completed 200 trials, with the entire testing session lasting 1 hour.

C. EEG data recording and pre-processing

The EEG data were recorded with Neuroscan Synamps 2 amplifier and 64-channels Quik-Cap (Compumedics, Australia). The impedance in all channels was maintained below 5k Ω . The recorded EEG data sampling rate is 1000Hz. A 29-inch screen was used, and the distance between the screen and the participant's chair was 40 cm. The image is shown in the centre of the screen, and participants did not need to move their heads when searching for the image's object. The EEGLAB toolbox v14.1.2 [8] was used for pre-processing recorded EEG data (adapted from [9]). The EEG data was first down-sampling to 250 Hz, a high-pass filter at 1 Hz and a line-noise removal was applied. EEG data were average referenced, and Adaptive Mixed Independent

Component Analysis (AMICA) was applied to remove the artifacts during the data recording.

Stimulus events were recorded during the presentation of the images and were subsequently employed to extract two hundred stimulus-based epochs. These epochs were then averaged (using the average method) for two conditions: correct and incorrect responses. Figure 2 compares the event-related potential (ERP) of the incorrect and correct evoked responses. The epoching and evoking process was implemented using the Python MNE toolbox v1.3.0 [10].

In order to investigate the effect of the epoching time range on decision-making classification, three stimulus-based epochs with time ranges of 125 ms, 250 ms, and 500 ms after the stimulus event onset were extracted. Since the parietal-occipital brain area processes the early visual information, we extract another dataset based on the parietal-occipital EEG channels (OZ, O1, O2, POZ, PO3, PO4, PZ, P1, P2) from these data segments. Finally, both parietal-occipital and whole brain (64-channels) datasets are used as the input for the classifier model separately.

D. Classification

A Bidirectional Long Short Term Memory (BiLSTM) model was utilized to classify correct and incorrect decision-making. Prior to this, the Standard Scaler method was applied to transform the EEG features. For each participant, 80% of the EEG dataset was randomly assigned as the training set, with the remaining 20% used for testing. Ridge regression (L2 regularization) was applied to the first Dense layer, and two Dropout layers (with a 30% rate) were used after each of the Bidirectional layers. Models were trained using the Adam optimizer and a learning rate of 0.0005. Early stopping was implemented with a patience of 15 epochs.

III. RESULTS AND DISCUSSION

As depicted in Figure 2, the comparison of ERP evoked responses in the OZ and POZ channels revealed a similar sequence of components (N170, P2, P3) in the EEG signal. However, differences in temporal and amplitude characteristics were observed. The N170 EEG component, which is known to be related to error-related potentials (ErrP) [11], reflects the moment when the participant has already detected an object and is considering whether or not it is the hinted object. The P2 EEG component has been theorized to reflect the final stage of the decision-making process, indicating the point at which sufficient action-related information has been gathered before a response is emitted [12]. In this experimental paradigm, the P2 EEG component may reflect the point at which the participant makes a decision regarding the location of the object in the image (positions 1 to 6). Finally, The P3 EEG component, which is related to decision confidence [13], reflects the point at which the participant reconsiders the object's position that they have made before.

Tables 1 and 2 present the classification results of all 10 participants, using three different epoch time ranges for all 64 channels and for 9 channels of the parietal-occipital region. Three epoch durations (0 - 125) ms, (0 - 250) ms,

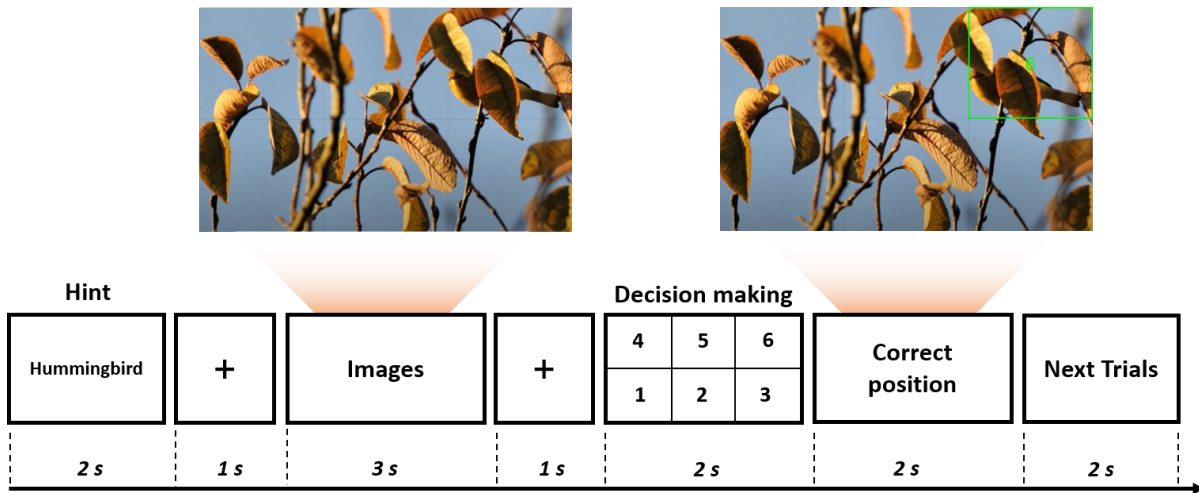


Fig. 1. During a trial, participants were given a hint about an animal object they needed to find in the image. In this figure, the hinted object is a hummingbird. Participants then searched for the hinted object, and made a decision about its location by pressing a button on the keyboard, numbered 1 to 6. The correct location of the object was revealed allowing participants to evaluate their responses. For example, the correct location of the hummingbird was number 6; thus the green bounding box appears around that location.

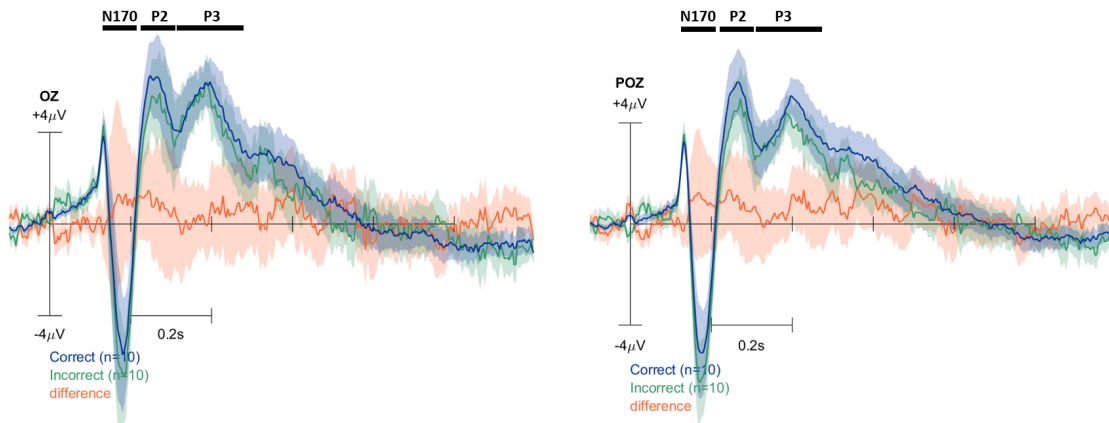


Fig. 2. The evoked responses in the OZ and POZ channels during correct and incorrect trials revealed sequences of similar components (N170, P2, P3) in the EEG signal. The N170 EEG component, which is known to be related to error-related potentials (ErrP), reflects the moment when the participant has already detected an object and is considering whether or not it is the hinted object. The P2 EEG component reflects the point at which the participant makes a decision regarding the location of the object in the image (positions 1 to 6). The P3 EEG component reflects the point at which the participant reconsiders the object's position that they have made before.

and (0-500) ms were selected as inputs for the Bidirectional Long Short Term Memory models based on the analysis of evoked comparisons. The period of (0 - 125) ms corresponds to the point at which participants are actively searching for the hinted object in the images, and it is likely that they have not yet made any decisions about the object. The period of (0 - 250) ms includes the N170 EEG component, which reflects the moment when participants are considering whether the detected object is the hinted object or not. This period may have some impact on the final decision of participants regarding the location of the observed object. By the time period of (0 - 500) ms, the three components (N170, P2, and P3) have already finished. Thus, the participant's final decision is almost completed at this point, and the classification result of the detected object is fully reported.

The mean classification accuracy when using all 64 chan-

nels (89.4 ± 1.9) is higher than when using only 9 parietal-occipital channels (81.0 ± 2.0). In both channel configurations, the mean classification accuracy increases as the epoching time duration increases. When using 64 channels as input for the classifier model, the highest classification accuracy was 96.1% (participant SS07) for the epoch time range (0 - 500) ms, while the highest accuracy for the shortest epoch time range (0 - 125) ms was 88.6% (participant 10). Similarly, when using 9 parietal-occipital channels as the classification dataset, the highest accuracy was 86.3% with an epoch time range (0 - 500) ms (participant 10). The highest classification results were achieved by participants SS07 (64 channels configuration) and SS10 (9 parietal-occipital region channels configuration). The classification results were more consistent when using 64 channels as opposed to 9 parietal-occipital region channels, and when

the epoch time range was higher. The classification results support the explanation of the effect of the three EEG components (N170, P2, and P3) on the participants' final decision-making.

TABLE I

COMPARISON OF CLASSIFICATION RESULTS FOR ALL PARTICIPANTS WITH VARIOUS EPOCHS TIME RANGE, USING ALL 64 CHANNELS

Participants	Classifiers	Accuracy (%)			
		[0-125] ^a ms	[0-250] ^a ms	[0-500] ^a ms	mean
SS01	BiLSTM	83.5	85.5	90.5	86.5
SS02	BiLSTM	84.1	87.3	92.2	87.9
SS03	BiLSTM	87.3	90.0	95.4	90.9
SS04	BiLSTM	86.4	89.6	93.7	89.9
SS05	BiLSTM	87.8	89.4	94.6	90.6
SS06	BiLSTM	86.0	88.5	92.0	88.8
SS07	BiLSTM	88.2	91.6	96.1	92.0
SS08	BiLSTM	85.1	88.1	91.2	88.1
SS08	BiLSTM	85.4	87.6	90.1	87.7
SS10	BiLSTM	88.6	91.5	95.2	91.8
mean		86.2	88.9	93.1	89.4
std		1.7	1.9	2.2	1.9

^a: The stimulus-based epochs' time range

TABLE II

COMPARISON OF CLASSIFICATION RESULTS FOR ALL PARTICIPANTS WITH VARIOUS EPOCHS TIME RANGE, USING 9 CHANNELS OF THE PARIETAL-OCCIPITAL REGION

Participants	Classifiers	Accuracy (%)			
		[0-125] ^a ms	[0-250] ^a ms	[0-500] ^a ms	mean
SS01	BiLSTM	74.5	79.1	81.4	78.3
SS02	BiLSTM	78.1	80.8	82.7	80.5
SS03	BiLSTM	77.3	82.0	84.1	81.1
SS04	BiLSTM	79.4	82.6	83.8	81.9
SS05	BiLSTM	78.5	81.4	84.6	81.5
SS06	BiLSTM	76.6	81.5	83.6	80.6
SS07	BiLSTM	79.9	84.6	85.1	83.2
SS08	BiLSTM	73.4	78.1	81.2	77.6
SS08	BiLSTM	78.1	82.6	84.1	81.6
SS10	BiLSTM	80.8	84.8	86.3	84.0
mean		77.7	81.8	83.7	81.0
std		2.3	2.1	1.6	2.0

^a: The stimulus-based epochs' time range

IV. CONCLUSIONS

In this study, we investigated the feasibility of participants' performance prediction in concealed object visual searching tasks using electroencephalography (EEG) signals. We employed a Bi-directional Long Short Term Memory (BiLSTM) classification model and used three different epoching time ranges (0 - 125 ms, 0 - 250 ms, and 0 - 500 ms) as input datasets. The classification was performed using both all 64 EEG channels and 9 parietal-occipital region channels. The results indicated that the mean classification accuracy was 89.4% when using all 64 channels and 81.0% when using only the 9 parietal-occipital region channels, suggesting the potential for early detection of human performance in challenging visual searching tasks. This can be beneficial in scenarios where people work in teams with other people or

artificial intelligence systems. For future work, we will use different classification models with more epoch time ranges and EEG channels configuration to increase the accuracy of decision-making classification in the shortest time and with the lowest number of EEG channels.

ACKNOWLEDGMENT

This work was supported in part by the Australian Research Council (ARC) under discovery grant DP210101093 and DP220100803, and the UTS Human-Centric AI Centre funding sponsored by GrapheneX (2023-2031). Research was also sponsored in part by the Australia Defence Innovation Hub under Contract No. P18-650825, Australian Cooperative Research Centres Projects (CRC-P) Round 11 CRCPXI000007, US Office of Naval Research Global under Cooperative Agreement Number ONRG - NICOP - N62909-19-1-2058, and AFOSR – DST Australian Autonomy Initiative agreement ID10134. We also thank the NSW Defence Innovation Network and NSW State Government of Australia for financial support in part of this research through grant DINPP2019 S1-03/09 and PP21-22.03.02.

REFERENCES

- [1] Y. Si, X. Wu, F. Li, L. Zhang, K. Duan, P. Li, L. Song, Y. Jiang, T. Zhang, Y. Zhang *et al.*, "Different decision-making responses occupy different brain networks for information processing: a study based on eeg and tms," *Cerebral Cortex*, vol. 29, no. 10, pp. 4119–4129, 2019.
- [2] T. V. Wiecki, I. Sofer, and M. J. Frank, "Hddm: Hierarchical bayesian estimation of the drift-diffusion model in python," *Frontiers in neuroinformatics*, p. 14, 2013.
- [3] D.-P. Fan, G.-P. Ji, M.-M. Cheng, and L. Shao, "Concealed object detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.
- [4] S. Bhattacharyya, D. Valeriani, C. Cinel, L. Citi, and R. Poli, "Target detection in video feeds with selected dyads and groups assisted by collaborative brain-computer interfaces," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2019, pp. 159–162.
- [5] D. Valeriani, C. Cinel, and R. Poli, "Group augmentation in realistic visual-search decisions via a hybrid brain-computer interface," *Scientific reports*, vol. 7, no. 1, p. 7772, 2017.
- [6] D. Valeriani and R. Poli, "Cyborg groups enhance face recognition in crowded environments," *PLoS one*, vol. 14, no. 3, p. e0212935, 2019.
- [7] C.-T. Lin and T.-T. N. Do, "Direct-sense brain-computer interfaces and wearable computers," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 298–312, 2020.
- [8] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis," *Journal of neuroscience methods*, vol. 134, no. 1, pp. 9–21, 2004.
- [9] T.-T. N. Do, C.-T. Lin, and K. Gramann, "Human brain dynamics in active spatial navigation," *Scientific Reports*, vol. 11, no. 1, pp. 1–12, 2021.
- [10] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, L. Parkkonen, and M. S. Hämäläinen, "MNE software for processing MEG and EEG data," *NeuroImage*, vol. 86, pp. 446–460, 2014.
- [11] A. Selimbeyoglu, Y. Keskin-Ergen, and T. Demiralp, "What if you are not sure? electroencephalographic correlates of subjective confidence level about a decision," *Clinical Neurophysiology*, vol. 123, no. 6, pp. 1158–1167, 2012.
- [12] M. Berchicci, D. Spinelli, and F. Di Russo, "New insights into old waves. matching stimulus-and response-locked erps on the same time-window," *Biological psychology*, vol. 117, pp. 202–215, 2016.
- [13] Á. Darriba and F. Waszak, "Predictions through evidence accumulation over time," *Scientific reports*, vol. 8, no. 1, pp. 1–15, 2018.