Distinction of the object recognition and object identification in the brain-computer interfaces applications

Daniel Leong, Tien-Thong Nguyen Do and *Chin-Teng Lin, *Fellow, IEEE*

***Abstract*— Object recognition is a complex cognitive process in which information is integrated and processed by various brain regions. Previous studies have shown that both the visual and temporal cortices are active during object recognition and identification. However, although object recognition and object identification are similar, these processes are considered distinct functions in the brain. Despite this, the differentiation between object recognition and identification has yet to be clearly defined for use in brain-computer interface (BCI) applications. This research aims to utilize neural features related to object recognition and identification and classify these features to differentiate between the two processes. The results demonstrate that several classifiers achieved high levels of accuracy, with the XGBoost classifier using a Linear Booster achieving the highest accuracy at 96% and a F1 score of 0.97. This ability to distinguish between object recognition and identification can be a beneficial aspect of a BCI object recognition system as it could help determine the intended target object for a user.**

Index Terms — Brain-Computer Interfaces (BCIs); Object Identification; Object Recognition; Electroencephalogram (EEG); Functional Connectivity; Classification

I. INTRODUCTION

Humans have to deal with much changing visual information in everyday life. In order to navigate our surroundings, our eyes must instantly identify and process what we see. Despite this, humans are efficient at finding the object they are looking for in even the most cluttered and complex of scenes, with a high accuracy level. While research has been done to understand how the brain processes large amounts of visual information, we still do not understand the process entirely. However, new technology is now being used to help us better understand the brain's visual system.

Researchers have been trying to find a way to measure the brain's ability to recognize and identify objects by decoding various brain signals. One measurement that has gained popularity is functional connectivity, which looks at the different patterns of connections within the brain, also known as brain network properties [1]. By studying these properties, researchers hope to gain a deeper understanding of how the brain processes visual information. Studies have found a link between object recognition and the integration of brain networks, which is a key aspect of functional connectivity. For example, it has been observed that lack of visual input during early development can lead to decreased connections between certain brain areas involved in visual processing and other regions, such as the somatosensory, motor and multisensory cortices. [2]

Most research on functional connectivity has relied on functional magnetic resonance imaging (fMRI) due to its high spatial resolution, allowing for detailed observations of changes in brain regions. However, electroencephalogram (EEG) offers a better temporal resolution and can gather information over shorter time intervals, making it useful for studying specific events such as object recognition. Recent studies have used EEG-based functional connectivity to investigate brain regions related to object recognition. For example, the study by Rizkallah et al. [3] looked at the brain network while participants recognized visual images regarding their meaningfulness. Additionally, EEG-based functional connectivity has been used in BCI research to classify different object categories, as demonstrated in a study by Tafreshi, Daliri & Ghodousi [4], where they used functional and effective connectivity features extracted from EEG signals to classify 12 different object categories.

Object recognition and object identification are similar processes, but object identification is considered to be a distinct process that engages various regions of the brain in processing information [5]. The distinction between object recognition and identification is largely dependent on the quantity of objects presented to people. In instances where a single object is presented, people will utilize object recognition, whereas when required to differentiate between multiple objects, object identification is employed. Despite advancements in the development of brain-computer interface (BCI) systems for distinguishing between object recognition and object identification, these systems remain largely inapplicable for practical use. This is due to the inability of the systems to discern the intent of the user when targeting an object, whether it be for the purpose of object recognition or object identification [6].

In most studies using EEG for object recognition, the experimental design commonly involves the participant pressing a button upon the presentation of a target stimulus. However, a few studies have modified the experimental design to detect object identification within the brain instead. In this study, our experimental design integrates the task of object recognition with object identification in order to distinguish between them. We aimed to use neural features

[1]D. L., T.-T. N. D. and C.-T. L. are with 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.

related to object recognition and identification and classify these features to differentiate between object recognition and object identification. To achieve this goal, various machine learning classifiers were evaluated to identify the one that yielded the best results. The objective was to identify dependable features that could be applied in a BCI system to enhance the system's comprehension of the user's intentions.

II. Methodology

A. Participant and Data Recording

The present study involved 25 participants, all of whom were aged 32.5 ± 10.4 years and had normal or corrected-to-normal vision. Each participant completed a total of 600 trials. The experiment was conducted at the Computation Intelligence and Brain-Computer Interface (CIBCI) Centre located in University of Technology Sydney (UTS). Prior to commencing the experiment, participants were provided with instructions and asked to sign an informed consent form. The study received ethical approval from the University of Technology Sydney under ethics ID ETH20-5519.

Brain activity was recorded utilizing a 64-channel EEG system manufactured by Neuroscan, Compumedics, Australia. This medical-grade device is known for its high-density EEG recordings and high accuracy and has been widely used in previous studies in the field of neuroscience and neurodiagnostic. The EEG electrode placement was consistent with the extended 10-20 international system, and data were referenced to an electrode located closest to the standard position FCZ. The contact impedance of the electrodes was maintained below 5 kΩ, and the EEG recordings were digitally sampled at a rate of 1000 Hz.

B. Experimental Design

In the course of our experiment, participants were presented with two tasks: object recognition and object identification. The object recognition task involved displaying to the participants an image selected at random from the Caltech-256 object category dataset [7]. The images were drawn from four distinct categories: animals, flowers, food, and vehicles, and each category included five different types of objects, with ten images per object. In total, 200 images were utilized during the experiment (see Figure 1). For each trial, participants were shown a picture of their target object for one second, following which they were prompted with a question inquiring as to whether the image belonged to the respective category. For instance, if the target image was of a bear, the participant would be asked, "Is this an animal?" and given two seconds to respond. This question was designed to assess the participant's ability to accurately recognize the target image.

The object identification task was administered next. Four images were randomly selected from the image dataset, with at least one of the images being chosen from the same category and subtype as the target image from the previous recognition task. The four images were presented in four different directions (up, down, left, and right), and the participants were instructed to press the keyboard button corresponding to the image that they felt was most closely related to the target image within three seconds. It should be noted that, due to the random selection of images, it was possible for there to be choices of the same category and subtype. Each trial lasted a total of six seconds: one second for the target image to appear, two seconds for the question, and three seconds for the presentation of the four images. Prior to the appearance of the target image, a fixation cross was displayed for 300 milliseconds, serving as the trial's baseline. An example trial is illustrated in Figure 1.

C. EEG Analysis

The EEG signals were processed utilizing EEGLAB v14.1.2 [8] (and adapted from [9]). The raw EEG data were filtered using a 1 Hz high-pass and a 50 Hz low-pass finite impulse response (FIR) filter and subsequently down-sampled to a rate of 250 Hz. Subsequently, channels deemed as noisy were removed, and the data was re-referenced to the average. Adaptive mixed independent component analysis (AMICA) was then applied to the re-referenced data in order to decompose it into maximally independent components (ICs), which are statistically independent sources of variance in the EEG. Independent components (ICs) associated with eye movement, muscle activity, and other noise were eliminated through the use of the IClabel toolbox [10]. After these bad components were removed, the epochs were extracted. An epoch began 300 ms prior to the onset of the target image (i.e. event onset) and ended 5 seconds after the target image appeared, thus covering the entire duration of the trial. Bad epochs were identified by examining their data values and removed if they contained data values outside a specified standard deviation threshold (threshold = 150 uV).

D. Functional Connectivity Feature Extraction and Classification

The extracted epochs was divided into two segments: the object recognition segment, which began when the target image was presented, and the object identification segment, which began when the four images were presented.

In order to address the inverse problem in EEG, a technique known as distributed source localization was employed. Additionally, functional connectivity was estimated utilizing the Brainstorm toolbox [11]. The distributed brain source localization activity was inferred from the data epochs, and then functional connectivity was determined from the estimated source activity. The data epochs were initially aligned with an MRI template [12], and the locations of the EEG sensors were mapped to the same anatomical landmarks. The openMEEG [13] process was subsequently utilized to compute the lead field of the cortical mesh. A noise covariance matrix was also calculated using the fixation cross period of the trial. The EEG data were then projected onto an anatomical framework consisting of 68 cortical regions as identified by re-segmenting the Desikan-Killiany

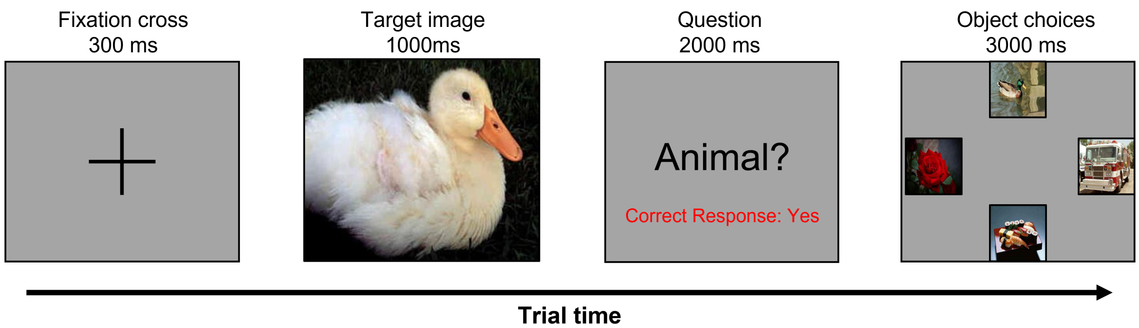

Fig. 1. The experimental design consisted of the following procedures: Each trial began with the fixation cross for 300 milliseconds, followed by the display of the target image for one second. Participants were then prompted to identify the object depicted in the target image and given two seconds to provide a response. Following this, four additional images were presented, one of which was chosen from the same category and subtype as the target image. Participants were given three seconds to select the image that most closely resembled their target object by pressing a button corresponding to the direction of the chosen image (up, down, left, and right). The figure illustrates an example of a trial, including the correct response to the question (shown in red text) and the layout of the object choices.

Cortical Atlas using FreeSurfer [14]. The standardized low-resolution brain electromagnetic tomography (sLORETA) [15] and dipole method were then applied to reconstruct the regional time series from the 68 brain regions. The amplitude envelope correlation (eq. 1) (AEC) was then used to estimate functional connectivity in the regions across six frequency bands: delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), low gamma (30-60 Hz), and high gamma(60-90 Hz). However, only the top 10% of the AEC values were selected for further analysis of network properties, with the remaining values set to zero.

$$\tilde{x}(t) = x(t) + j\mathcal{H}\left\{x(t)\right\} = a_{\tilde{x}}(t)e^{j\phi_{\tilde{x}}(t)} \tag{1}$$

The complex time series $\tilde{x}(t)$ is a representation of the original data time series, $x(t)$. The real component of $\tilde{x}(t)$ is equivalent to the original time series $x(t)$, while the imaginary component is the Hilbert transform of $\tilde{x}(t)$, denoted as $\mathcal{H}\left\{x(t)\right\}$. The instantaneous amplitude (or envelope) of the original time series $x(t)$ is represented by the module $a_{\tilde{x}}(t)$, and the instantaneous phase of $x(t)$ is represented by the variable $\phi_{\tilde{x}}(t)$.

The network properties of the brain were calculated using the chosen AEC connectivity matrix, with a focus on network integration and segregation as measurements for the experimental task. These measurements provide insight into the global information exchange among brain regions and the capacity for specialized processing within densely interconnected groups of regions. The Brain Connectivity Toolbox [16] was utilized to calculate the network properties, including the clustering coefficient, local efficiency, modularity, and participation coefficient. These measures, specifically the clustering coefficient, local efficiency and modularity, provide an assessment of network segregation, while the participation coefficient measures network integration. Since object recognition and object identification tasks involve various brain regions in processing the information, the network properties selected for the classification were extracted from different brain networks, including the frontoparietal network, somatosensory network, default mode network (DMN) and visual network. The network properties were calculated for each participant's object recognition and object identification tasks, and a statistical test was conducted to compare the results (assessed via the Friedman test, followed by a post-hoc pairwise Wilcoxon signed-rank test). However, only the network properties that showed significance from the statistical test ($p < 0.05$) were selected as the feature for classification analysis.

A binary (object recognition/ object identification) classification analysis was conducted utilizing several commonly employed classifier models: Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), K-Nearest Neighbors (K-NN), Random Forest, Decision Tree, Logistic Regression, AdaBoost, Gradient Boosting, Multi-Layer Perceptron, Naive Bayes, XGBoost with Tree Booster, and XGBoost with Linear Booster. These classifier models were obtained from the scikit-learn library [17] for the Python programming language, and the default parameters were utilized for all models.

III. Results and Discussion

The analysis of object recognition and identification was conducted utilizing binary classification techniques and the network properties from all 25 participants. The performance of the classification was evaluated using the 10-fold cross-validation method. The results of the classification analysis are presented in Table I. Among the various classifiers tested, K-NN, decision tree, and gradient boosting exhibited the lowest accuracy above 60%. Some classifiers, such as SVM, Logistic Regression, and XGBoost with Tree Booster, demonstrated an accuracy above 70%. Several classifiers achieved an accuracy above 80%, including LDA, Random Forest, AdaBoost, Multi-Layer Perceptron, and Naive Bayes. The highest classification accuracy was obtained from the XGBoost classifier with Linear Booster, which recorded a result of 96% and a F1 score of 0.97.

The distinction between object recognition and object identification in classification results can potentially help develop a better BCI system. Utilizing this distinction within a BCI system would allow for implementing a classification algorithm based on the user's objective. For instance, if the user intends to have the BCI system recognize a specific object, the system would initiate an object recognition algorithm and provide feedback to the user. Conversely, if the user wishes to select an object within their environment, the BCI system would initiate an object identification algorithm and then select the desired object. Incorporating such a feature in a BCI system could significantly aid individuals with disabilities in selecting and communicating information about objects with others.

Although the binary classification analysis has provided a promising result in distinguishing between object recognition and object identification, several factors could be applied to improve the result. As mentioned in the methodology, the classifiers utilized in this study used the default parameters, which were not optimized for the selected features. Therefore, algorithm tuning could help us to achieve a better classification outcome. Additionally, ensemble learning techniques, combining multiple individual models, could also improve the entire dataset's performance.

TABLE I
BINARY CLASSIFICATION RESULTS FROM DIFFERENT CLASSIFIERS

Classifier	Accuracy	F1-Score
SVM	76%	0.79
LDA	86%	0.81
K-NN	66%	0.64
Random Forest	84%	0.82
Decision Tree	60%	0.65
Logistic Regression	78%	0.79
AdaBoost	80%	0.82
Gradient Boosting	60%	0.59
Multi-Layer Perceptron	86%	0.86
Naive Baye	80%	0.81
XGBoost with Tree Booster	74%	0.74
XGBoost with Linear Booster	96%	0.97

IV. CONCLUSION

In this study, we selected several network properties related to object recognition and object identification and utilized some commonly used classifiers to distinguish them. The result showed that three classifiers achieved an accuracy of 60% or higher, while another three reached 70% or higher. Six classifiers had an accuracy of 80% or more, with the XGBoost classifier using a Linear Booster achieving the top result at 96% accuracy with high F1 score. This feature of distinguish between object recognition and object identification could be a valuable aspect of a BCI object recognition system, as it would aid in determining a user's intended target object.

V. 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] J. M. Shine, "Neuromodulatory influences on integration and segregation in the brain," *Trends in cognitive sciences*, vol. 23, no. 7, pp. 572–583, 2019.

[2] Y. Liu, C. Yu, M. Liang, J. Li, L. Tian, Y. Zhou, W. Qin, K. Li, and T. Jiang, "Whole brain functional connectivity in the early blind," *Brain*, vol. 130, no. 8, pp. 2085–2096, 2007.

[3] J. Rizkallah, P. Benquet, A. Kabbara, O. Dufor, F. Wendling, and M. Hassan, "Dynamic reshaping of functional brain networks during visual object recognition," *Journal of neural engineering*, vol. 15, no. 5, p. 056022, 2018.

[4] T. F. Tafreshi, M. R. Daliri, and M. Ghodousi, "Functional and effective connectivity based features of EEG signals for object recognition," *Cognitive neurodynamics*, vol. 13, no. 6, pp. 555–566, 2019.

[5] N. K. Logothetis and D. L. Sheinberg, "Visual object recognition," *Annual review of neuroscience*, vol. 19, no. 1, pp. 577–621, 1996.

[6] 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.

[7] G. Griffin, A. Holub, and P. Perona, "Caltech-256 object category dataset," 2007.

[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, T.-P. Jung, and C.-T. Lin, "Retrosplenial Segregation Reflects the Navigation Load During Ambulatory Movement," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 488–496, 2021.

[10] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, "Iclabel: An automated electroencephalographic independent component classifier, dataset, and website," *NeuroImage*, vol. 198, pp. 181–197, 2019.

[11] F. Tadel, S. Baillet, J. C. Mosher, D. Pantazis, and R. M. Leahy, "Brainstorm: a user-friendly application for MEG/EEG analysis," *Computational intelligence and neuroscience*, vol. 2011, 2011.

[12] V. S. Fonov, A. C. Evans, R. C. McKinstry, C. R. Almli, and D. Collins, "Unbiased nonlinear average age-appropriate brain templates from birth to adulthood," *NeuroImage*, no. 47, p. S102, 2009.

[13] A. Gramfort, T. Papadopoulo, E. Olivi, and M. Clerc, "OpenMEEG: opensource software for quasistatic bioelectromagnetics," *Biomedical engineering online*, vol. 9, no. 1, pp. 1–20, 2010.

[14] B. Fischl, "Freesurfer," *Neuroimage*, vol. 62, no. 2, pp. 774–781, 2012.

[15] R. D. Pascual-Marqui *et al.*, "Standardized low-resolution brain electromagnetic tomography (sloreta): technical details," *Methods Find Exp Clin Pharmacol*, vol. 24, no. Suppl D, pp. 5–12, 2002.

[16] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010.

[17] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.