Classifying Electroencephalogram (EEG) Signals via Brain Activity Mapping to Distinguish Identified vs Unidentified Information

Adhi Dharma Wibawa[1*], Siti Dwi Suryani[2], and Stralen Pratasik[3]
[1]Department of Medical Technology, Faculty of Medicine and Health,
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia 60111
[1–3]Department of Electrical Engineering, Faculty of Intelligent Electrical and Informatics
Technology, Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia 60111
Email: [1]adhiosa@te.its.ac.id, [2]suryanisitidwi@gmail.com, [3]7022222003@mhs.its.ac.id

***Abstract*—Conventional lie detectors have often been questioned for their accuracy and reliability, which can lead to wrongful accusations. These inaccurate results may compromise legal decisions, threaten national security, or hinder the justice system. Electroencephalogram (EEG) is a technique used to record electrical activity in the brain, which has become a major focus for researchers, especially in the development of lie detection systems. Therefore, the research aims to explore complex patterns in brain activity that play an important role in distinguishing identified and unidentified information by using brain activity mapping as a novel approach. The required data are taken from channels T3, T4, T5, T6, O1, and O2 related to human memory. A total of 30 participants are involved in the research, where their brain activity is analyzed in the Alpha, Beta, and Gamma subbands. Brain activity visualization parameters are based on energy wavelet feature extraction values. The visualization results for each participant in the three subbands are then classified using the Naïve Bayes algorithm with a Gaussian distribution approach. The results of the machine learning method achieve 72% accuracy, with test scenarios using 80% training data and 20% testing data. The research introduces brain heat mapping as an innovative visualization technique to interpret EEG-based deception detection better, offering a more intuitive and explainable approach compared to traditional feature-based methods. The findings contribute to a deeper understanding of brain function and provide a foundation for improving the effectiveness and reliability of EEG-based lie detection in investigative contexts.**

***Index Terms*—Electroencephalogram (EEG) Signals, Brain Activity, Identified Information, Unidentified Information**

Received: Nov. 07, 2024; received in revised form: Feb. 26, 2025; accepted: Feb. 26, 2025; available online: April 28, 2025.
*Corresponding Author

I. Introduction

INTERROGATIONS are carried out by authorities to reveal hidden information in security, investigative, criminal, and social contexts [1]. In the interrogation process, the polygraph is a tool that has long been used to monitor a person's physical reactions when lying. However, its accuracy and reliability are often questioned [2]. In the modern context, the Electroencephalogram (EEG) has become a promising research focus [3–6]. EEG records electrical activity in the brain and can provide insight into a person's honesty. Integrating EEG with lie detector technology can increase accuracy in detecting lies by identifying patterns of brain activity associated with lying. The information revealed in lie detection can be done by distinguishing information that is and is not identified by the brain so that research on EEG continues to be developed to analyze patterns of brain activity recorded via EEG [7–14].

In analyzing brain activity patterns, feature extraction is done to extract information hidden in the EEG signal [15–19]. Furthermore, many researchers are using machine learning to distinguish between normal brain activity patterns and design scenarios that trigger brain reactions. For example, previous research [20] has classified EEG signals based on known and unknown information using machine learning resulting in good accuracy. Furthermore, another previous research [21] uses a combination of feature extraction and tries several machine learning algorithms to determine differences in brain activity when given familiar and unfamiliar image stimuli. It is followed by another previous research [22] which uses energy wavelet feature extraction and Shannon entropy to analyze differences in brain activity when faced with familiar and unfamiliar objects. The energy wavelet values in the research show quite a prominent difference in these two conditions.

The research has integrated the concept of EEG visualization to compare brain activity mapping, including previous research [23] which calculates the similarity of brain activity images in groups of EEG data when the eyes are closed and the eyes are open. It then classifies and produces quite good results. Other studies have also shown the benefits of using brain visualization in identifying certain neurological conditions through non-invasive methods [24–26]. The feature extraction results are then integrated into brain activity mapping, which can now be applied to developing a more sophisticated lie detection system. Brain activity mapping identifies brain areas involved in the lying process, such as the prefrontal cortex and anterior cingulate cortex, which are active when someone is processing false or contradictory information. Activity in these areas indicates cognitive conflict, self-control, and decision-making related to lying behavior [27].

After the integration of brain activity mapping, the classification process is also an important method for distinguishing the characteristics of EEG signal patterns from others. Many studies use several classification algorithms to classify data from EEG signals. Some of the classification algorithms used include Support Vector Machine (SVM), K-Nearest Neighbors (KNN) [28], Logistic Regression, Convolutional Neural Network (CNN), Random Forest [14, 28–30]. Another classification method, namely the Decision Tree, is used by [31] in the development of a lie detector system by analyzing changes in pupil diameter and eye movements. The use of classification algorithms also succeeds in classifying face recognition carried out [32] by trying several algorithms including Long Short-Term Memory (LSTM), Hierarchical Long Short-Term Memory (H-LSTM), Recurrent Neural Network (RNN), Deep CNN, and SVM. Another effective classification method is Naïve Bayes, which can be used to classify images [33–36], EEG signals [37–41], and various other types of data with the simple assumption that each feature is independent of the other.

Based on several previous studies by combining information from EEG and brain activity mapping, the research aims to explore more subtle and complex patterns in brain activity related to distinguishing identified information from unidentified information using energy wavelet feature extraction so that it can be used as a reference in developing a lie detection system. Next, to evaluate the effectiveness of the method used, the researchers apply Naïve Bayes machine learning algorithm, as it is capable of classifying EEG signal data and produces relatively good accuracy [42, 43]. Several similar studies have not explored the proposed scenario, brain activity mapping feature extraction method, and classifier algorithm, thereby creating a research gap in the research.

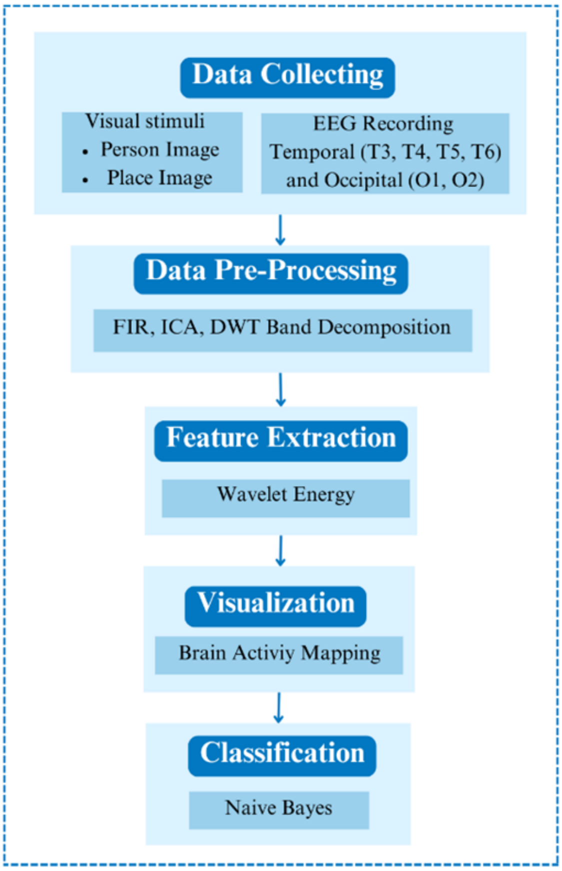
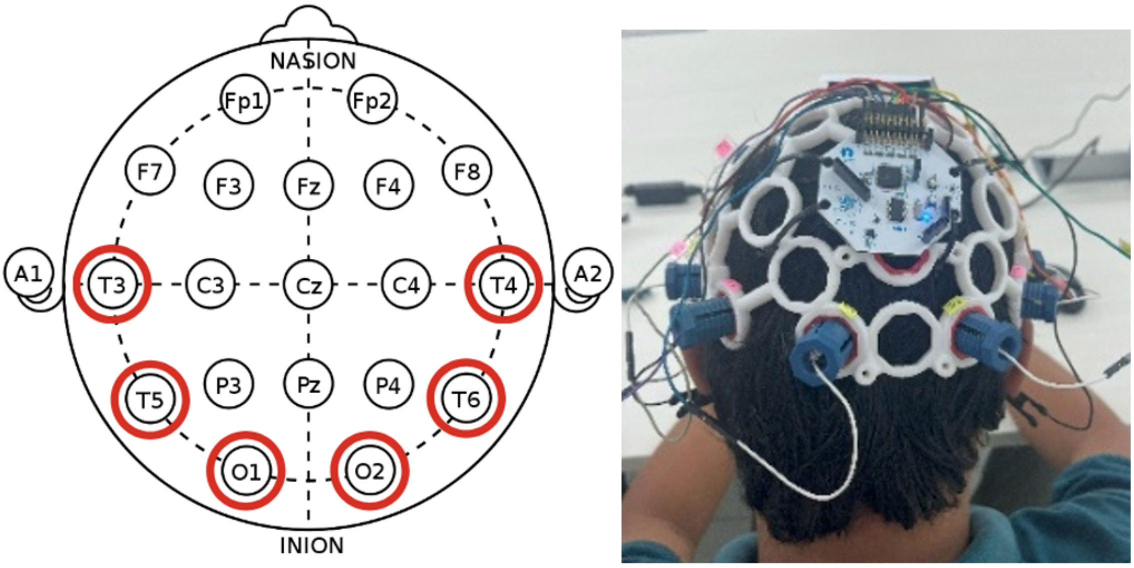


Fig. 1. Research methodology. It has Finite Impulse Response (FIR), Independent Component Analysis (ICA), and Discrete Wavelet Transform (DWT).

II. Research Method

In this stage, the methods and data used in the research are described. Figure 1 presents the five stages, starting from data collection, data pre-processing, feature extraction, visualization, and classification. The details are explained in the following sub-chapters.

Cite this article as: A. D. Wibawa, S. D. Suryani, and S. Pratasik, “Classifying Electroencephalogram (EEG) Signals via Brain Activity Mapping to Distinguish Identified vs Unidentified Information”, CommIT Journal 19(1), 101–114, 2025.



(a) (b)

Fig. 2. (a) Electrode placement and (b) Electroencephalogram (EEG) recording.

A. Data Collection

The research uses an OpenBCI product using the 10-20 system. This system is a method of placing electrodes that refer to specific locations on the scalp, and it is used to measure brain electrical activity according to international standards. Figure 2 shows that six channels are applied. The four channels used are the temporal area on the side of the head consisting of the right temporal lobe (T4 and T6) and the left temporal lobe (T3 and T5). The other two channels are the occipital area, which is at the back of the head, namely the left occipital lobe (O1) and the right occipital lobe (O2).

The selected EEG channels related to memory in the human brain provide important insights into visual processing, color perception, object recognition, and other visual processes that occur in the brain [20–22, 44]. In the EEG recording process, the stimuli used are images of people and places identified and not identified by the respondents. The respondents in the research are 30 healthy students, consisting of 16 females and 14 males, with an average age of 19 years. Before the recording, respondents are asked to fill out a health questionnaire and provide informed consent. To ensure familiarity with the stimuli, respondents are selected based on their ability to identify the given stimuli. In the first recording stage, respondents are presented with images of two people and two places that they can identify. In this recording, a scenario is created using a quiz-like system. If the respondent can identify the stimulus, they will click the “Yes” button, with each image being displayed for seven seconds to answer the question. The stimulus system created is designed in such a way that the respondent minimizes movement except for mouse movements “Yes” or “No”. After the first recording stage, the respondent recovers for one minute. Next, the second stage of recording is carried out, respondents are presented with two pictures of people and two pictures of places that they cannot identify.

B. Data Preprocessing and Feature Extraction

Data preprocessing is a crucial step in preparing raw EEG data for further analysis. Since the EEG signal is often affected by noise, eye blinks, and muscle activity, it is essential to clean and filter the data to ensure accuracy [5, 41, 45, 46]. During the research, the analysis applies a Finite Impulse Response (FIR) filter to eliminate Direct Current (DC) offset and extract the desired frequency signal. This filter process combines band pass and FIR filters to remove noise and unwanted frequency components while preserving relevant frequency components. Next, artifacts, such as eye blinks, from the EEG signal are removed using the Independent Component Analysis (ICA) method. ICA is a method in statistics that separates mixed signals

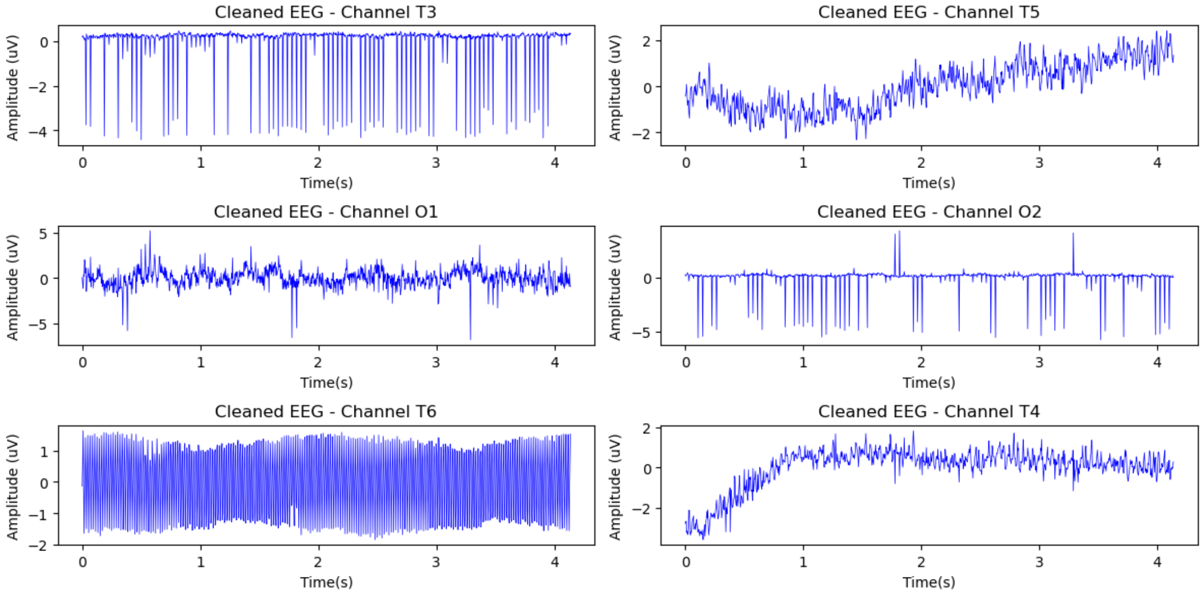


Fig. 3. Cleaned Electroencephalogram (EEG) data.

into independent signals [47–50]. After obtaining clean data from the ICA process, amplitude filtering is applied. If the amplitude value exceeds 100, it is limited to 100. Figure 3 provides an example of clean EEG data from each channel.

The next process is band decomposition using Discrete Wavelet Transform (DWT). The EEG device has a sampling frequency of 256 Hz, which is then reduced by a factor of 2 to 128 Hz through band decomposition using DWT. This reduction is carried out for several purposes, including reducing the computational load and eliminating high-frequency noise without sacrificing important information. The use of DWT is suitable for non-stationary signals which has advantages compared to all types of spectral analysis [51–55]. With DWT, every non-static time domain signal $x(t)$ can be solved in Eq. (1) [56, 57]. It has $\gamma(t)$ as DWT of $x(t)$, a as scale parameter, b as shift parameter, and $\psi(t)$ as mother wavelet.

$$\gamma(t) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{2^a}} \psi\left(\frac{t - b \times 2^a}{2^2}\right) dt. \quad (1)$$

The approximate and detailed coefficients are calculated using Daubechies mother wavelet (db4). The db4 mother wavelet is selected because it effectively captures detailed signals by balancing time and frequency resolution. This mother wavelet has orthogonal and compact support properties, ideal for analyzing non-stationary signals because it can detect frequency changes efficiently [58]. Band decomposition using DWT is illustrated in Fig. 4.

The research focuses on the Alpha, Beta, and Gamma sub-bands. These bands are related to frontal and occipital brain activity as well as emotional and mental reactions. After getting the coefficient values representing each subband, the energy value is calculated. Following decomposition, the wavelet sub-band energy is computed as follows in Eq. (2). It has N as number of samples in the signal X_i.

$$E = \sum_{i=1}^{N} |X[i]|^2. \quad (2)$$

C. Visualization

After getting the feature values of each subband, a heatmap is plotted to visualize these features. The visualization is generated using Python visualization with MNE (Magnetoencephalography (MEG) and EEG data analysis) library. The following is the stages of the process that are conducted.

- Import Libraries: The script first imports the necessary libraries, including NumPy for numerical computing, matplotlib for data visualization, MNE for EEG data analysis, and Operating System (OS) for file system manipulation.
- Data: EEG data is represented as `data_list`, which has several lists (each representing data from one measurement session). Each sublist contains a series of amplitude values from several EEG channels.
- Channel Information: `channel_names` is a list containing the names of the EEG channels used in the measurement. It is used to create informa-

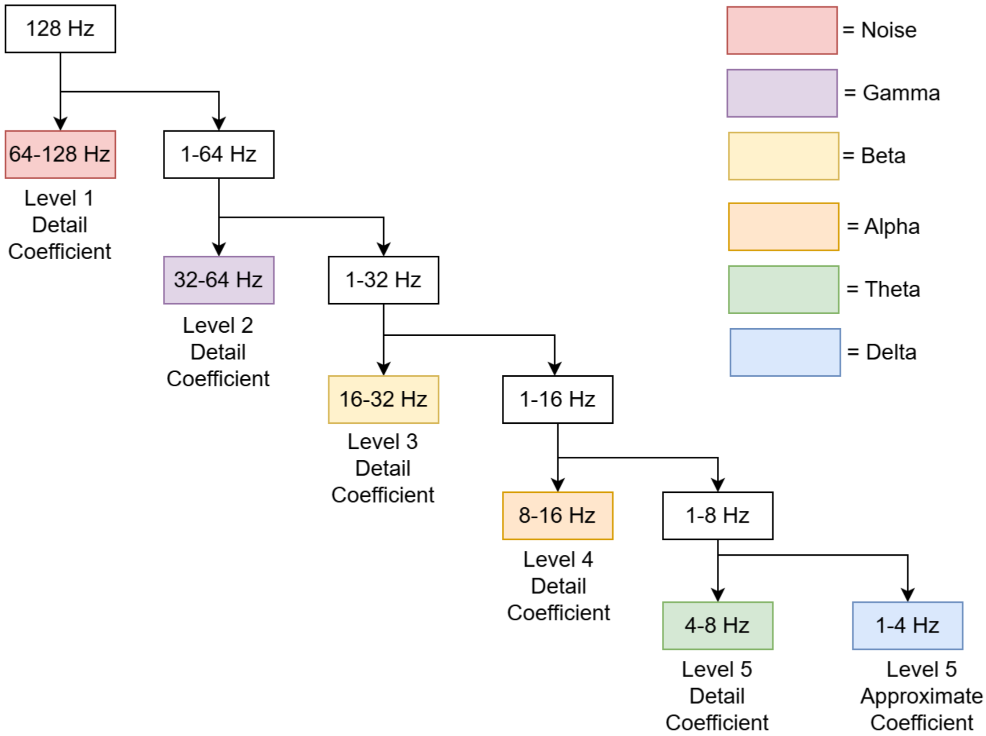


Fig. 4. Implementation of 5-level Discrete Wavelet Transform (DWT) decomposition.

tion objects that store metadata about EEG data, such as channel names, sample frequencies, and channel types.

- Montage: The montage (the physical arrangement of the electrodes on the head) is determined using the `make_standard_montage` function of MNE. These montages are then organized into information objects.
- Visualization: Iteration is performed through each measurement session in the `data_list`. For each session, a topomap image (amplitude spatial distribution map) for the given data are created using `mne.viz.plot_topomap`. The channel names are placed on top of the plot using the coordinates of the `info` object. Each image is saved in the appropriate folder.

D. Classification

After plotting the heatmap, a classification method is applied to identify patterns in the visualization. The Naïve Bayes method with a Gaussian approach is used as the classification method, based on Bayes' theorem, which assumes that each feature is independent of the others. This method is suitable for cases where the data features are continuous and are assumed to come from a Gaussian (normal) distribution. It works by assuming that the feature values in each class are drawn from a Gaussian distribution, with a different mean and variance for each class [59].

In the research, the classifier is trained using an 80:20 train-test split, where feature values extracted through wavelet-based analysis are assumed to follow a Gaussian distribution with class-specific means and variances. The formula for this calculation is shown in Eq. (3). It has x as a continuous feature vector, C_k as class, and the posterior probability for the class C_k. It shows $P(C_k)$ as prior probability for class C_k, μ_{ki} as the average of the i-th feature in the class C_k, and σ_{ki}^2 as variance of the i-th feature in the class C_k.

$$P(X = x_i \mid C = C_k) = \frac{1}{\sqrt{2\pi\sigma_{ki}^2}} \exp\left(-\frac{(x_i - \mu_{ki})^2}{2\sigma_{ki}^2}\right) \tag{3}$$

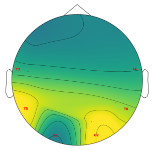
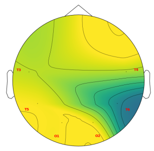
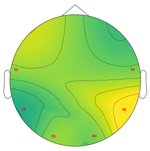
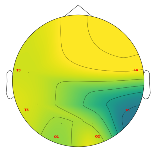
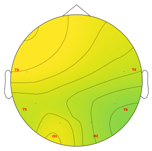
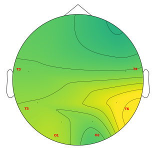
Confusion matrix is an important evaluation tool in the context of classification in machine learning that displays model performance with detailed predictions for each target class. It shows the number of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), providing insight into how well the model differentiates target classes with pinpoint accuracy as shown in Eqs. (4)–(6). It is a key instrument in assessing model reliability and identifying areas that can be improved to increase accuracy.

Cite this article as: A. D. Wibawa, S. D. Suryani, and S. Pratasik, “Classifying Electroencephalogram (EEG) Signals via Brain Activity Mapping to Distinguish Identified vs Unidentified Information”, CommIT Journal 19(1), 101–114, 2025.

TABLE I
MEAN WAVELET ENERGY VALUES.

Subband	Channel	Classification	
		Identified	Unidentified
Alpha	T3	125.3	229.2
	T4	122.0	224.0
	T5	126.1	223.8
	T6	130.3	216.0
	O1	130.2	232.0
	O2	126.7	222.9
Beta	T3	251.4	449.7
	T4	251.6	457.8
	T5	250.8	444.5
	T6	255.5	433.8
	O1	252.6	461.8
	O2	241.4	450.5
Gamma	T3	519.6	941.9
	T4	520.9	945.9
	T5	516.4	924.7
	T6	515.9	995.1
	O1	526.2	937.0
	O2	523.7	962.0

TABLE II
VISUALIZATION RESULTS FROM RESPONDENT 13.

Identified	Unidentified
Alpha Subband	
	
Beta Subband	
	
Gamma Subband	
	

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{4}$$

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{5}$$

$$\text{Recall} = \frac{TP}{TP + FN}. \tag{6}$$

III. RESULTS AND DISCUSSION

A. *Result*

As previously explained, the research uses values from energy features obtained from the band decomposition process using DWT. The results of extracting the average value of the energy feature for 30 respondents for 6 channels each for the Alpha, Beta, and Gamma subbands are shown in Table I.

In Table I, the average wavelet energy values in the unidentified condition tend to be higher than the identified condition in all channels and subbands. The value of the unidentified condition is almost twice the value of the identified condition. This increase in wavelet energy is seen consistently in various EEG channels and in all frequency subbands of the wavelet decomposition results. The results strengthen the indication that the signal characteristics in this condition have a significantly different pattern compared to the identified condition.

Next, the researchers interpret the distribution in a boxplot to compare the distribution of several data groups (Fig. 5). The boxes in the boxplot represent the Interquartile Range (IQR), namely the range between the first quartile (Q1) and the third quartile (Q3). This covers the middle 50% of the data. Figure 5 shows that the distribution values for several data groups in the unidentified condition are higher when compared to the values in the identified condition. These distribution values are grouped based on each subband and channel, combining data from 30 participants. In the condition when participants are given a stimulus that they can identify, the lowest energy wavelet value in the Alpha subband is 43.6 and the highest value is 288.4. While the lowest value in the Beta subband is 99 and the highest value is 536. In the Gamma subband, the lowest value is 238.5 and the highest value is 1,110.2. The result shows that when participants are given a stimulus that they cannot identify, the energy wavelet values in all subbands tends to be higher and has more variable compared to conditions where the stimulus can be identified. This increase in wavelet energy values may reflect increased cognitive activity or brain effort in trying to identify an unidentified stimulus. Additionally, greater variability in the unidentified condition also suggests a wider range of responses among participants, which can result from individual differences in cognitive and perceptual processes.

Next, the extracted feature values from each respondent is visualized into a heatmap map. Table II shows the results of heatmap visualization from respondent 13 as an example in the Alpha, Beta, and Gamma subbands in the identified and unidentified information conditions. In the research, each respondent produces three heatmap images of Alpha, Beta, and Gamma subbands. So, in the image classification test using the Naïve Bayes algorithm, the total image data consists of 180 images (90 for identified labels, and 90 for

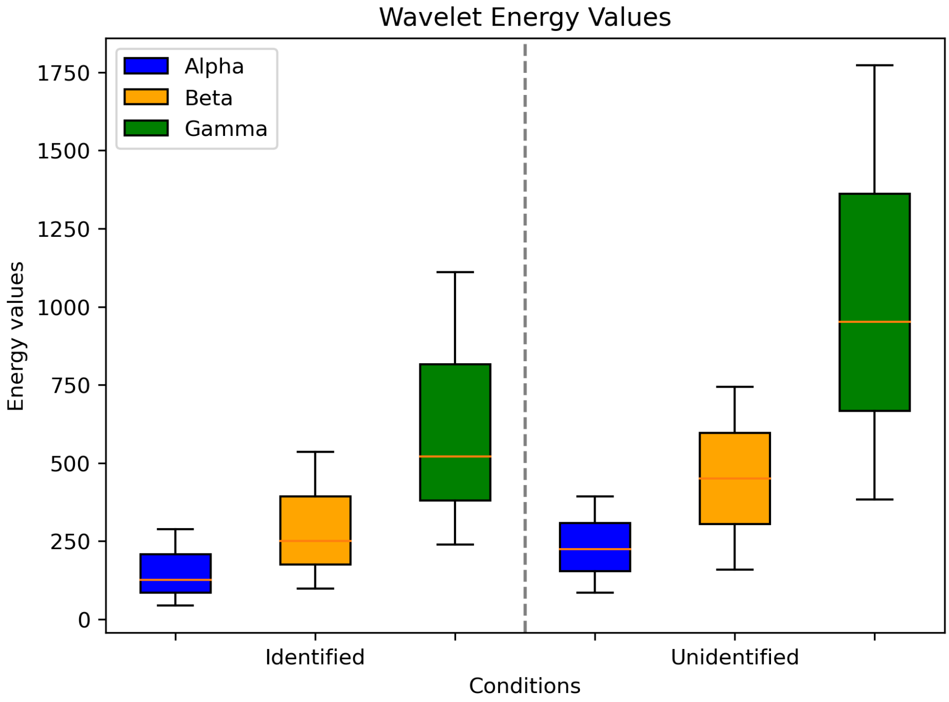


Fig. 5. Distribution of wavelet energy values in the Alpha, Beta, and Gamma band frequencies for identified and unidentified conditions.

unidentified labels).

In Table II within the Alpha subband, the identified condition shows high energy concentrations in frontal and parietal areas, indicating memory and attention-related activity, while the unidentified condition shows a more uniform distribution of energy, reflecting broader cognitive effort. In the Beta subband, the identified condition displays intense activity in frontal and central areas, indicating a large allocation of cognitive resources to information processing, while the unidentified condition has a more homogeneous energy distribution, indicating additional effort to understand the unidentified stimulus. In the Gamma subband, the identified condition shows high brain activity across areas, especially temporal and parietal, associated with complex information processing, whereas the unidentified condition has a more even distribution of energy, indicating a diffuse cognitive effort to recognize the unidentified stimulus.

Table III shows the performance of the Naïve Bayes algorithm using the Gaussian distribution approach with a scenario of 80% training data and 20% testing data. The algorithm achieves an accuracy of 72%. It demonstrates its effectiveness in classifying identified and unidentified conditions based on wavelet energy features extracted from the heatmaps.

Next, the researchers also test other classification algorithms, including CNN, SVM, Random Forest, and KNN as classifiers. However, the results obtained from these algorithms are lower than Naïve Bayes. Therefore, the researchers focus on the best results found as the core of the discussion. Table IV presents a comparison of the performance of various classification algorithms that have been tested in the research.

A classification system must measure its performance using a confusion matrix. Confusion matrix is a table that records the results of classification work.

TABLE III
EVALUATION METRICS VALUES.

Evaluation Metrics	Percentage (%)
Accuracy	72
Precision	75
Recall	66

TABLE IV
PERFORMANCE COMPARISON OF DIFFERENT CLASSIFICATION ALGORITHMS.

Algorithm	Accuracy (%)
Naïve Bayes	72.0
Convolutional Neural Network (CNN)	69.0
Support Vector Machine (SVM)	69.0
Random Forest	63.8
K-Nearest Neighbors (KNN)	61.0

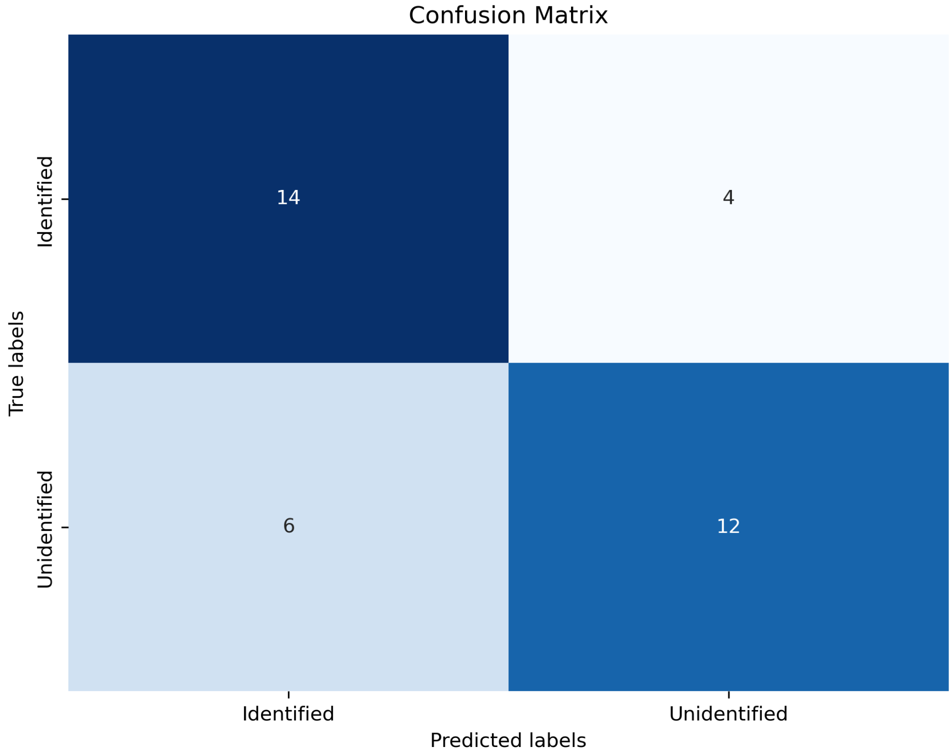


Fig. 6. Confusion matrix of Naïve Bayes model with a Gaussian distribution approach

Figure 6 displays the confusion matrix from Naïve Bayes with a Gaussian distribution approach with a positive value of 14. It means that from the testing data of 18 heatmap maps on brain conditions, 14 correct values are identified, and 4 predicted label values are incorrect. The data testing test value in the identified condition produces fewer FP values when compared to the data testing test value in the unidentified condition. In the unidentified condition, the false value is only 6, and the heatmap incorrect value is 12. In simpler terms of the 18 images that should have been unidentified, the model incorrectly classifies 6 images as identified and succeeds in classifying 12 images correctly.

The FP cases occur when participants do not recognize an image, but the model incorrectly classifies it as “identified”. It may occur due to similar brain activity patterns, even though the participant has no prior experience with the stimulus. Meanwhile, the FN cases happen when participants recognize an image, but the model incorrectly classifies it as “unidentified”. This result may be caused by weak EEG responses or noise interference during recording. Overall, this confusion matrix helps to understand the model’s performance in separating two conditions based on heatmap data, showing that despite some classification errors, the model performs quite well in differentiating between identified and unidentified conditions.

B. Discussion

In general, Table II shows quite good results in distinguishing the recognition of identified and unidentified information using energy wavelet feature values in EEG data. The condition of the brain signal when recognizing unidentified information tends to have a high value. The result is also in line with research [20] where respondent who does not know the information tends to have a higher value using Mean, Mean Absolute Value (MAV), and Standard Deviation (STD) value feature extraction.

The results of the interpreted energy wavelet values show an increase in cognitive activity. When participants attempt to identify an unidentified stimulus, their brains may engage in more complex cognitive processes and make greater efforts to decipher the information received. According to [60], increased cognitive load can lead to increased brain activity as measured through various brain imaging techniques, including wavelet analysis. The research finds that more complex and challenging tasks required more cognitive resources, which are reflected in increased brain wave activity. In addition to increased cognitive activity, greater variability in the unidentified stimulus

Cite this article as: A. D. Wibawa, S. D. Suryani, and S. Pratasik, “Classifying Electroencephalogram (EEG) Signals via Brain Activity Mapping to Distinguish Identified vs Unidentified Information”, CommIT Journal 19(1), 101–114, 2025.

TABLE V
COMPARISON OF ELECTROENCEPHALOGRAPHY (EEG) SIGNAL CLASSIFICATION STUDIES BASED ON FEATURES, ALGORITHMS, AND ACCURACY.

Author	Features	Algorithm	Accuracy
[20]	Mean, MAV, STD	KNN	87.00%
	MAV/STD	Naïve Bayes	50.00%
[14]	P300	CNN	84.44%
[30]	Fisher’s LDA, ApEn	KNN	85.00%
[44]	PSD	Random Forest	95.40%
[9]	CWT, Hjorth parameters	SVM	84.37%
[10]	CWT, Hjorth parameters	LDA	72.20%
[1]	FC Features: C, L, Degree	SVM, KNN	90.58%

Note: MAV= Mean Absolute Value, STD= Standard Deviation, P300= Positive peak at approximately 300 milliseconds, LDA= Linear Discriminant Analysis, PSD= Power Spectral Density, CWT= Continuous Wavelet Transform, FC Features= Functional Connectivity Features, C= Complexity, L= Lateralization, KNN= K-Nearest Neighbors, CNN= Convolutional Neural Network, ApEn= Approximate Entropy, and SVM= Support Vector Machine.

condition suggests a wider range of responses among participants. It can be due to individual differences in cognitive and perceptual strategies. For example, some participants may have a better ability to process ambiguous or unclear information, while others may experience greater difficulty, resulting in greater variability in the data. This variability in cognitive abilities can lead to significant differences in brain responses to unidentifiable stimuli [61]. Previous research [62] has used wavelet analysis to identify patterns of brain activity during complex cognitive tasks and found that wavelet analysis is very effective for detecting dynamic changes in brain activity associated with cognitive load.

Furthermore, the use of the MNE library visualization in Python provides deep insight into the significant differences in accuracy in distinguishing identified and unidentified information. In this context, the Gaussian Naïve Bayes algorithm stands out as a method capable of processing EEG data effectively, especially when applied to wavelet energy features. The visualization provided by MNE makes it possible to see directly how this algorithm manages to separate relevant information and produce quite high accuracy values in classification. It shows great potential in applications in the field of neuroimaging data analysis. The result is supported by previous research [63] which recognizes that visualization of heatmap using the MNE Python library helps to understand and validate classification results in a neuroimaging context.

The researchers have compared the research with other studies on EEG. Table V presents a comparison of several studies related to EEG signal classification based on the features used, classification algorithms, and the accuracy achieved. Each study adopts a different approach to feature extraction and model selection to improve classification accuracy. Some studies use statistical-based methods such as MAV, STD, and Power Spectral Density (PSD), while others rely on more complex techniques such as Continuous Wavelet Transform (CWT), Hjorth parameters, and Approximate Entropy (ApEn).

The ability of the Gaussian Naive Bayes algorithm to differentiate between identified and unidentified information, especially in the context of wavelet energy features, shows great potential for further development in this field. This model produces quite good values because just one feature is enough to produce 72% accuracy based on Table III. The value of the findings produces quite high accuracy when compared to previous research [20] which only obtains 50% accuracy on one feature with the same scenario. Additionally, compared to other studies (Table V) that utilize multiple features and more complex algorithms, such as Random Forest with PSD (95.40%) or KNN with Mean, MAV, and STD (87%), this approach with Gaussian Naive Bayes demonstrates that a simpler model with minimal feature extraction can still yield competitive results. This finding gives quite good results because if the researchers use more features, the computing becomes heavier, and the price also becomes more expensive. Moreover, this approach introduces a novel brain heat map visualization, which is absent in previous studies, providing a more intuitive interpretation of brain activity patterns. So, with the research, researchers and other practitioners can optimize the classification process and gain a better understanding of brain activity, paving the way for significant advances in the understanding of brain function and health-related development that can be used as a reference for lie detection in the field of investigation.

Compared to traditional polygraph tests, EEG-based lie detection offers more reliable information. Polygraphs rely on physiological responses such as heart rate and blood pressure, which individuals can consciously manipulate. In contrast, EEG measures brain activity, which is much harder to control intentionally, making it a more robust tool for detecting deception. Beyond lie detection, this method has potential applications in various fields. In forensic and criminal investigations, EEG-based classification can assist in identifying suspects and verifying testimonies. In national security, it may help to detect individuals attempting to conceal critical information. Additionally, in cognitive psychology and neuroscience, it can contribute to research on how humans recognize and recall information.

IV. Conclusion

The research proposes the Naïve Bayes method with a Gaussian distribution approach to classify EEG signals according to brain conditions when they can or cannot identify information using brain activities mapping. The channels are T3, T4, T5, T6, O1, and O2 which are then analyzed in the Alpha, Beta, and Gamma subbands. The energy wavelet feature extraction value is used as a visualization parameter for brain activity mapping. Next, the visualization image is classified using the Gaussian Naïve Bayes algorithm. It produces an accuracy of 72% with a scenario of 80% training data and 20% testing data. These results demonstrate a fairly good understanding of brain function and provide strong support for the development of lie detection in investigative contexts. Compared to conventional polygraph methods, which rely on physiological responses that can be consciously controlled, EEG-based lie detection offers a more objective and reliable approach by analyzing neural activity that is difficult to manipulate. The research introduces brain heat mapping as an innovative visualization technique, enhancing interpretability in distinguishing identified and unidentified information. These findings highlight its potential applications in forensic investigations, national security, and cognitive neuroscience research.

The research has limitations in the small sample size and the limited variation of brain activity trigger scenarios. For further research, it is suggested to increase the sample size, carry out more varied scenarios, and try a combination of feature extraction to increase accuracy in distinguishing brain signal conditions to identify or unidentify information via EEG signals. Hence, it can be used as a lie detection development system. The bigger sample size and more diverse participants can increase the generalizability of the findings. The research also suggests performing more varied scenarios that can trigger different brain activities to obtain higher difference scores in the recognition of identifiable and unidentifiable objects in the brain.

Author Contribution

Conceived and designed the analysis, A. D. W; Collected the data, S. D. S. and S. P.; Contributed data or analysis tools, A. D. W., S. D. S., and S. P.; Performed feature extraction and brain activity mapping using machine learning techniques for EEG signal analysis, A. D. W.; Wrote the paper, A. D. W., S. D. S., and S. P.; Validated the results by comparing heatmaps from different scenarios to improve model accuracy, S. D. S.; and Adapted machine learning algorithms to improve accuracy, S. P.

Data Availability

The participants of the research did not give written consent for their data to be shared publicly. So, due to the sensitive nature of the research, supporting data are not available.

References

[1] W. Chang, H. Wang, G. Yan, and C. Liu, "An EEG based familiar and unfamiliar person identification and classification system using feature extraction and directed functional brain network," *Expert Systems with Applications*, vol. 158, 2020.

[2] Y. Z. Abd El Gawad, M. I. Youssef, and T. M. Nasser, "Medical system based on thermal optical system and neural network," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 4, pp. 3796–3804, 2023.

[3] J. Immanuel, A. Joshua, and S. T. George, "A study on using blink parameters from EEG data for lie detection," in *2018 International Conference on Computer Communication and Informatics (ICCCI)*. Coimbatore, India: IEEE, Jan. 4–6, 2018, pp. 1–5.

[4] N. A. Sayel, B. M. Sabbar, and S. Albermany, "Real time control system for wheel chair of disabled people using EEG signal," in *2022 4th International Conference on Advanced Science and Engineering (ICOASE)*. Zakho, Iraq: IEEE, Sep. 21–22, 2022, pp. 71–76.

[5] Y. Pamungkas, A. D. Wibawa, and Y. Rais, "Classification of emotions (positive-negative) based on EEG statistical features using RNN, LSTM, and Bi-LSTM algorithms," in *2022 2nd International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)*. Jakarta, Indonesia: IEEE, Sep. 22–23, 2022, pp. 275–280.

[6] Y. Guo, M. Wang, T. Zheng, Y. Li, P. Wang, and X. Qin, "NAO robot limb control method based on motor imagery EEG," in *2020 International Symposium on Computer, Consumer and Control*

(IS3C). Taichung City, Taiwan: IEEE, Nov. 13–16, 2020, pp. 521–524.

[7] A. Turnip, M. F. Amri, M. A. Suhendra, and D. E. Kusumandari, "Lie detection based EEG-P300 signal classified by ANFIS method," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, vol. 9, no. 1-5, pp. 107–110, 2017.

[8] A. I. Simbolon, A. Turnip, J. Hutahaean, Y. Siagian, and N. Irawati, "An experiment of lie detection based EEG-P300 classified by SVM algorithm," in *2015 International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology (ICACOMIT)*. Bandung, Indonesia: IEEE, Oct. 29–30, 2015, pp. 68–71.

[9] P. V. Ravindran and A. P. Vinod, "Name familiarity detection using EEG-based brain computer interface," in *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)*. Kochi, India: IEEE, Oct. 17–20, 2019, pp. 2305–2309.

[10] K. G. Smitha, A. P. Vinod, and K. Mahesh, "Voice familiarity detection using EEG-based brain-computer interface," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. Budapest, Hungary: IEEE, Oct. 9–12, 2016, pp. 001 626–001 631.

[11] R. Alazrai, R. Homoud, H. Alwanni, and M. I. Daoud, "EEG-based emotion recognition using quadratic time-frequency distribution," *Sensors*, vol. 18, no. 8, pp. 1–32, 2018.

[12] K. K. Koganti, S. Kameswari, and V. J. Naveen, "Artifical intelligance based real time lie detector using eye gaze pattern," in *2022 International Conference on Electronic Systems and Intelligent Computing (ICESIC)*. Chennai, India: IEEE, April 22–23, 2022, pp. 235–240.

[13] C. H. Ma, H. Y. Chang, H. C. Lee, Y. F. Yu, H. S. Tien, Y. H. Lin, M. Y. Liu, Y. L. Lin, H. M. Ma, K. F. Lin, and H. W. Lieh, "The psychological and physiological effects of integrated cognitive-behavioral and biofeedback therapy on panic disorder: A randomized controlled trial," *Journal of the Formosan Medical Association*, vol. 122, no. 12, pp. 1305–1312, 2023.

[14] N. Baghel, D. Singh, M. K. Dutta, R. Burget, and V. Myska, "Truth identification from EEG signal by using convolution neural network: Lie detection," in *2020 43rd International Conference on Telecommunications and Signal Processing (TSP)*. Milan, Italy: IEEE, July 7–9, 2020, pp. 550–553.

[15] R. Holker and S. Susan, "Quantitative EEG feature selection by MajorityVoting for alcohol use disorder detection," in *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*. Athens, Greece: IEEE, July 27–30, 2021, pp. 1–4.

[16] O. K. Cura, S. K. Atli, S. Y. Sen, and A. Akan, "Classification of ADHD by using multiple feature maps of EEG signals and deep feature extraction," in *2023 31st European Signal Processing Conference (EUSIPCO)*. Helsinki, Finland: IEEE, Sep. 4–8, 2023, pp. 1065–1069.

[17] X. Xu, F. Wei, Z. Zhu, J. Liu, X. Wu, X. Xu, F. Wei, Z. Zhu, J. Liu, and X. Wu, "EEG feature selection using orthogonal regression: Application to emotion recognition," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Barcelona, Spain: IEEE, May 4–8, 2020, pp. 1239–1243.

[18] M. Murtazina and T. Avdeenko, "Feature selection techniques analysis for identification of cognitive and resting states based on EEG data," in *2022 VIII International Conference on Information Technology and Nanotechnology (ITNT)*. Samara, Russian Federation: IEEE, May 23–27, 2022, pp. 1–4.

[19] M. Huang, L. Ye, J. Chen, R. Fu, and C. Zhou, "Feature representation for meditation state classification in EEG signal," in *2021 11th International Conference on Information Technology in Medicine and Education (ITME)*. Fujian, China: IEEE, Nov. 19–21, 2021, pp. 267–270.

[20] A. Farizal, A. D. Wibawa, Y. Pamungkas, M. Pratiwi, and A. Mas, "Classifying known/unknown information in the brain using Electroencephalography (EEG) signal analysis," in *2022 11th Electrical Power, Electronics, Communications, Controls and Informatics Seminar (EECCIS)*. Malang, Indonesia: IEEE, Aug. 23–25, 2022, pp. 362–367.

[21] Y. Pamungkas, "Data analytics on EEG signal features to distinguish familiar and unfamiliar information in human brain memory," in *2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*. Jakarta, Indonesia: IEEE, Feb. 16, 2023, pp. 369–374.

[22] S. D. Suryani, A. D. Wibawa, and D. P. Wulandari, "EEG analysis of familiar and unfamiliar objects using wavelet energy and Shannon entropy," in *2024 16th International Conference on Knowledge and Smart Technology (KST)*. Krabi, Thailand: IEEE, Feb. 28–March 2, 2024, pp. 226–

231.

[23] A. V. Garenskaya, M. A. Bakaev, and O. M. Razumnikova, "Telling minds apart: Classification of EEG signals based on comparison of brain activity maps," in *2021 XV International Scientific-Technical Conference on Actual Problems of Electronic Instrument Engineering (APEIE)*. Novosibirsk, Russian Federation: IEEE, Nov. 19–21, 2021, pp. 459–464.

[24] M. G. Puxeddu, M. Petti, and L. Astolfi, "Multilayer analysis of multi-frequency brain networks as a new tool to study EEG topological organization," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. Mexico: IEEE, Nov. 1–5, 2021, pp. 924–927.

[25] P. Volf, M. Stehlik, P. Kutilek, G. Kloudova, K. Rusnakova, S. Kozlova, M. Braunova, J. Hejda, V. Krivanek, and R. Doskocil, "Brain electrical activity mapping in military pilots during simulator trainings," in *2019 International Conference on Military Technologies (ICMT)*. Brno, Czech Republic: IEEE, May 30–31, 2019, pp. 1–6.

[26] A. Abdalmalak, G. Laforge, L. C. M. Yip, D. Milej, L. E. Gonzalez-Lara, U. Anazodo, A. M. Owen, and K. S. Lawrence, "Shining light on the human brain: An optical BCI for communicating with patients with brain injuries," in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. ON, Canada: IEEE, Oct. 11–14, 2020, pp. 502–507.

[27] N. Abe, J. Okuda, M. Suzuki, H. Sasaki, T. Matsuda, E. Mori, M. Tsukada, and T. Fujii, "Neural correlates of true memory, false memory, and deception," *Cerebral Cortex*, vol. 18, no. 12, pp. 2811–2819, 2008.

[28] D. Risqiwati, A. D. Wibawa, E. S. Pane, E. M. Yuniarno, W. R. Islamiyah, and M. H. Purnomo, "Effective relax acquisition: A novel approach to classify relaxed state in alpha band EEG-based transformation," *Brain Informatics*, vol. 11, no. 1, pp. 1–14, 2024.

[29] P. D. Purnamasari and T. W. Junika, "Frequency-based EEG human concentration detection system methods with SVM classification," in *2019 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*. Banda Aceh, Indonesia: IEEE, Aug. 22–24, 2019, pp. 29–34.

[30] I. Lakshan, L. Wickramasinghe, S. Disala, S. Chandrasegar, and P. S. Haddela, "Real time deception detection for criminal investigation," in *2019 National Information Technology Conference (NITC)*. Colombo, Sri Lanka: IEEE, Oct. 8–10, 2019, pp. 90–96.

[31] Z. Labibah, M. Nasrun, and C. Setianingsih, "Lie detector with the analysis of the change of diameter pupil and the eye movement use method Gabor wavelet transform and decision tree," in *2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS)*. Bali, Indonesia: IEEE, Nov. 1–3, 2018, pp. 214–220.

[32] L. Ghosh, S. Ghosh, A. Konar, P. Rakshit, and A. K. Nagar, "Decoding of EEG signals using deep long short-term memory network in face recognition task," in *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*. Bangalore, India: IEEE, Nov. 18–21, 2018, pp. 477–483.

[33] R. Ismayanti, F. Alameka, D. Mirwansyah, N. W. W. Sari, A. Rahim, and Riyayatsyah, "Identification of pests on black orchid plants using naïve bayes method based on leaf image texture," in *2022 International Conference of Science and Information Technology in Smart Administration (ICSINTESA)*. Bali, Indonesia: IEEE, Nov. 10–12, 2022, pp. 167–172.

[34] P. C. Kumar and B. T. Geetha, "Efficient removal of real time rain streaks from a image using novel Naive Bayes (NB) compare over Linear Regression (LR) with improved accuracy," in *2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*. Chennai, India: IEEE, May 25–26, 2023, pp. 1–6.

[35] D. Divyamary, S. Gopika, S. Pradeeba, and M. Bhuvaneswari, "Brain tumor detection from MRI images using naive classifier," in *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*. Coimbatore, India: IEEE, March 6–7, 2020, pp. 620–622.

[36] C. A. Sari, I. P. Sari, E. H. Rachmawanto, E. Proborini, R. R. Ali, and I. Rizqa, "Papaya fruit type classification using LBP features extraction and naive bayes classifier," in *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*. Semarang, Indonesia: IEEE, Sep. 19–20, 2020, pp. 28–33.

[37] A. Setiawan, M. A. S. Noor, A. Trisanto, G. F. Yohanes, D. D. Rumani, B. M. Wibawa, N. Rohadi, and A. Turnip, "Pilot anxiety detection through brain signal using naïve bayes method," in *2024 IEEE International Conference on Ar-*

tificial Intelligence and Mechatronics Systems (AIMS). Bandung, Indonesia: IEEE, Feb. 21–23, 2024, pp. 1–6.

[38] J. S. Akshitha, C. Y. Sai, D. Mishra, A. Sinha, and B. B. S. Sai, “Detection of alcoholism based on EEG signals using machine learning,” in *2024 International Conference on Integrated Circuits, Communication, and Computing Systems (ICIC3S)*, vol. 1. Una, India: IEEE, June 8–9, 2024, pp. 1–6.

[39] A. Turnip, E. F. Athalla, D. D. Rumani, A. Setiawan, D. E. Kusumandari, and N. S. Syafei, “Level of pilot fear on maneuvering based brain wave using naïve bayes method,” in *2024 IEEE International Conference on Artificial Intelligence and Mechatronics Systems (AIMS)*. Bandung, Indonesia: IEEE, Feb. 21–23, 2024, pp. 1–6.

[40] N. Memon, S. S. H. Zaidi, and M. Suleman, “The prognosis of epilepsy with naive bayes classifier on FPGA using HDL coder,” in *2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*. Langkawi Island, Malaysia: IEEE, March 1–3, 2021, pp. 18–23.

[41] A. D. Wibawa, B. S. Y. Mohammad, M. A. K. Fata, F. A. Nuraini, A. Prasetyo, and Y. Pamungkas, “Comparison of EEG-based biometrics system using naive bayes, neural network, and support vector machine,” in *2022 International Conference on Electrical and Information Technology (IEIT)*. Malang, Indonesia: IEEE, Sep. 15–16, 2022, pp. 408–413.

[42] M. A. Mawalid, A. Z. Khoirunnisa, M. H. Purnomo, and A. D. Wibawa, “Classification of EEG signal for detecting cybersickness through time domain feature extraction using naïve bayes,” in *2018 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM)*. Surabaya, Indonesia: IEEE, Nov. 26–27, 2018, pp. 29–34.

[43] N. Y. Oktavia, A. D. Wibawa, E. S. Pane, and M. H. Purnomo, “Human emotion classification based on EEG signals using naïve bayes method,” in *2019 International Seminar on Application for Technology of Information and Communication (iSemantic)*. Semarang, Indonesia: IEEE, Sep. 21–22, 2019, pp. 319–324.

[44] A. Farizal, A. D. Wibawa, D. P. Wulandari, and Y. Pamungkas, “Investigation of human brain waves (EEG) to recognize familiar and unfamiliar objects based on power spectral density features,” in *2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)*. Surabaya, Indonesia: IEEE, July 26–27, 2023, pp. 77–82.

[45] Y. Rais, A. D. Wibawa, and D. P. Wulandari, “An EEG pattern depicting the attention of customers while viewing video advertisements,” in *2024 7th International Conference on Informatics and Computational Sciences (ICICoS)*. Semarang, Indonesia: IEEE, July 17–18, 2024, pp. 353–358.

[46] B. G. Pratama, A. D. Wibawa, D. P. Wulandari, and S. Pratasik, “Neuromarketing study of purchase decisions using advertising videos based on EEG signal analysis,” in *2024 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*. Bali, Indonesia: IEEE, July 4–6, 2024, pp. 315–320.

[47] A. M. Judith, S. B. Priya, and R. K. Mahendran, “Artifact removal from EEG signals using regenerative multi-dimensional singular value decomposition and independent component analysis,” *Biomedical Signal Processing and Control*, vol. 74, 2022.

[48] S. Pratasik, A. D. Wibawa, and D. P. Wulandari, “Coherence analysis of EEG signal in happy and sad emotions during visual stimulation,” in *2023 IEEE International Biomedical Instrumentation and Technology Conference (IBITeC)*. Yogyakarta, Indonesia: IEEE, Nov. 9–10, 2023, pp. 171–176.

[49] N. Fatih, A. D. Wibawa, M. H. Purnomo, and A. Mas, “Comparative analysis of EEG-based emotion recognition between male and female participants using Hjorth parameter,” in *2023 International Electronics Symposium (IES)*. Denpasar, Indonesia: IEEE, Aug. 8–10, 2023, pp. 179–185.

[50] J. M. Suhendro, A. D. Wibawa, and A. Mas, “Analysis of attraction response on product packaging based on EEG signal:(A preliminary study of neuromarketing on packaging),” in *2023 International Electronics Symposium (IES)*. Denpasar, Indonesia: IEEE, Aug. 8–10, 2023, pp. 479–485.

[51] M. Hajian and M. H. Moradi, “Quantification of depression disorder using EEG signal,” in *2017 24th National and 2nd International Iranian Conference on Biomedical Engineering (ICBME)*. Tehran, Iran: IEEE, Nov. 30–Dec. 1, 2017, pp. 1–5.

[52] N. Shweta, M. Moinuddin, S. Suma, K. S. Raju, S. Patil, and S. Patil, “DWT approach based on analysis of seizures in EEG signal,” in *2023 International Conference on Integrated Intelligence*

and Communication Systems (ICIICS). Kalaburagi, India: IEEE, Nov. 24–25, 2023, pp. 1–5.

[53] Masitoh, Suprijanto, and V. Nadhira, "DWT decomposition of EEG signal and source localization using ICA-eLORETA method for basic hand motor activity," in *2021 International Conference on Instrumentation, Control, and Automation (ICA)*. Bandung, Indonesia: IEEE, Aug. 25–27, 2021, pp. 127–132.

[54] M. Bekbalanova, A. Zhunis, and Z. Duisebekov, "Epileptic seizure prediction in EEG signals using EMD and DWT," in *2019 15th International Conference on Electronics, Computer and Computation (ICECCO)*. Abuja, Nigeria: IEEE, Dec. 10–12, 2019, pp. 1–4.

[55] M. Shahbakhti, M. Beiramvand, M. Nazari, A. Broniec-Wójcik, P. Augustyniak, A. S. Rodrigues, M. Wierzchon, and V. Marozas, "VME-DWT: An efficient algorithm for detection and elimination of eye blink from short segments of single EEG channel," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 408–417, 2021.

[56] M. R. Islam and M. Ahmad, "Wavelet analysis based classification of emotion from EEG signal," in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*. Cox's Bazar, Bangladesh: IEEE, Feb. 7–9, 2019, pp. 1–6.

[57] R. Matei and D. Matei, "Multiresolution spectral analysis of epileptic EEG signals using various wavelet types," in *2019 E-Health and Bioengineering Conference (EHB)*. Iasi, Romania: IEEE, Nov. 21–23, 2019, pp. 1–4.

[58] P. Pavithara, C. Kalaivanan, P. Ponmurugan, V. L. Jothi, K. Karthik, and K. K. Kumar, "Implementation of EEG signal decomposition and feature extraction through efficient wavelet transforms," in *2024 International Conference on Communication, Computing and Internet of Things (IC3IoT)*. Chennai, India: IEEE, April 17–18, 2024, pp. 1–6.

[59] L. He and B. Liu, "Motor imagery EEG signals analysis based on Bayesian network with Gaussian distribution," in *Intelligent Computing in Bioinformatics (ICIC 2014)*. Taiyuan, China: Springer, Aug. 3–6, 2019, pp. 241–247.

[60] O. Jensen and C. D. Tesche, "Frontal theta activity in humans increases with memory load in a working memory task," *European Journal of Neuroscience*, vol. 15, no. 8, pp. 1395–1399, 2002.

[61] M. J. Kane, A. R. A. Conway, D. Z. Hambrick, and R. W. Engle, "Variation in working memory capacity as variation in executive attention and control," in *Variation in working memory*. Oxford University Press, 2007, pp. 21–46.

[62] A. Grinsted, J. C. Moore, and S. Jevrejeva, "Application of the cross wavelet transform and wavelet coherence to geophysical time series," *Nonlinear Processes in Geophysics*, vol. 11, no. 5/6, pp. 561–566, 2004.

[63] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. Hamalainen, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroinformatics*, vol. 7, pp. 1–13, 2013.