Investigating the Effect of Sustainable Product Packaging on Consumer Behavior Using EEG Signals

Eray Ozcan Kilic
Dept. of Electrical and Electronics Eng.
Izmir University of Economics
Izmir, TURKEY
eray.kilic@std.izmirekonomi.edu.tr

Burak Akbugday and Aydin Akan
Dept. of Electrical and Electronics Eng.
Izmir University of Economics
Izmir, TURKEY
{burak.akbugday, akan.aydin}@ieu.edu.tr

Riza Sadikzade
School of Applied Management Sci.
Izmir University of Economics
Izmir, TURKEY
riza.sadikzade@ieu.edu.tr

***Abstract*—This study examines the emotional impact of sustainable versus conventional product packaging on consumers using electroencephalography (EEG). Brain signals were recorded from 56 participants while they viewed images of a wood fiber bottle and a glass bottle. A 32-channel EEG system was used at a 256 Hz sampling rate. After bandpass filtering and independent component analysis, the cleaned signals were segmented into 4-second windows. From each segment, 896 features were extracted across four categories: time-domain, frequency-domain, nonlinear, and entropy-based measures. The ReliefF algorithm selected the 30 most informative features, and a Subspace K-Nearest Neighbor ensemble classifier achieved 94.4% test accuracy for the glass condition and 90.7% for the wood fiber condition in four-class valence–arousal classification. Shapley value analysis showed that nonlinear features, especially fuzzy entropy and Hurst exponent, were more effective in emotion discrimination than classical band power features. The two packaging conditions also activated different brain regions, suggesting that packaging material produces measurable differences in neural processing.**

***Keywords*—Sustainability, emotion recognition, EEG, machine learning, neuromarketing, valence–arousal.**

I. Introduction

Packaging is the first point of contact between a product and a consumer. Since purchasing decisions are often made at the shelf within seconds, the visual and material properties of packaging play a key role in shaping consumer preferences [1]. With growing environmental awareness, sustainable packaging has become a central topic in both research and industry. However, consumer attitudes toward eco-friendly packaging are typically measured through surveys and focus groups, which are prone to social desirability bias and fail to capture unconscious emotional reactions [2].

Neuromarketing applies neuroscience tools to marketing research to understand consumer behavior [3]. Among these tools, electroencephalography (EEG) is the most widely used due to its millisecond-level temporal resolution and low cost [4]. EEG-derived features such as band powers, entropy measures, and fractal dynamics can identify neural signatures of different emotional states [5]. Russell's circumplex model [6], which describes emotions along valence and arousal dimensions, has been widely adopted in EEG-based emotion recognition [7]. Fig. 1 shows how valence and arousal values are placed in a two-dimensional coordinate system. However, most existing studies use music clips or film excerpts as stimuli [8], and research on packaging-evoked emotional responses through EEG remains very limited [9].

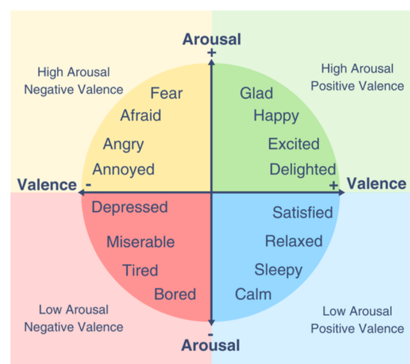


Fig. 1: Valence–Arousal Model.

This study investigates the emotional effect of sustainable packaging on consumers using EEG. Two beverage packaging types were compared: a wood fiber bottle and a conventional glass bottle. EEG data from 56 participants were acquired and processed, extracted 896 features were reduced to 30 via ReliefF for classification via Subspace k-Nearest Neighbors (KNN) algorithm. The main contributions of the study are: (1) applying EEG-based emotion recognition to sustainable packaging; (2) demonstrating that nonlinear features outperform classical band power features in emotion discrimination; and (3) showing that different packaging materials activate distinct regional brain networks.

*This study was supported by Izmir University of Economics, Scientific Research Projects Coordination Unit. Project number: 2022-07, and by TUBITAK 2209-A grant no: 1919B012425267.

II. Materials and Methods

A. Participants

Sixty healthy, right-handed volunteers under the age of 25 participated in the study. The sample was balanced in terms of gender, with 30 female and 30 male participants. All participants were university students with no known neurological or psychiatric conditions. Before the experiment, each participant was informed about the purpose and procedure of the study, and written informed consent was obtained. The study protocol was approved by the Science and Engineering Research Ethics Committee of Izmir University of Economics, with number 101658, dated 10 July 2025.

During the data quality control stage, four participants were excluded from the analysis due to poor signal quality, missing channels, or recording duration inconsistencies. As a result, the final dataset consisted of EEG recordings from 56 participants.

B. Stimuli and Experimental Protocol

Two beverage packaging images were used as visual stimuli in the experiment. The first was a bottle made from wood fiber, representing a sustainable packaging material. The second was a conventional glass bottle. Both images were presented on a green background and did not contain any brand names or labels, in order to prevent brand-related associations from influencing emotional responses.

Fig. 2: Samples of fiber bottle (right hand-side) and conventional glass bottle (left hand-side) used as visual stimuli.

The experiments were conducted in a controlled laboratory environment at Izmir University of Economics. After the EEG cap was placed and electrode impedances were verified, a resting-state baseline recording was taken. Each stimulus image was then displayed on a screen for 45 seconds, with a 15-second black screen inserted between the two stimuli. To control for order effects, a counterbalancing design was applied: half of the participants (Group A, $n = 30$) saw the wood fiber packaging first, followed by the glass bottle, while the remaining half (Group B, $n = 30$) viewed the stimuli in the reverse order.

Immediately after stimulus presentation, participants completed a bipolar self-report questionnaire for each packaging condition. The valence subscale consisted of six items (e.g., unhappy–happy, hopeless–hopeful, dissatisfied–satisfied) and the arousal subscale consisted of five items (e.g., calm–excited, sleepy–alert, relaxed–stimulated). Each item was rated on a five-point scale ranging from -2 to $+2$. For each participant, valence and arousal scores were computed by averaging the respective subscale items.

C. EEG Recording System

Brain electrical activity was recorded using the BitBrain Versatile 32-channel EEG system. The electrodes were positioned according to the international 10–20 placement system, covering frontal, fronto-central, central, temporal, parietal, and occipital regions of the scalp given in Table I. The sampling rate was set to 256 Hz. During recording, electrode impedances were kept within acceptable limits, and electrodes with high impedance were corrected by reapplying conductive gel. Data were recorded using BitBrain's SennsLab software and exported to Python for preprocessing.

TABLE I: EEG channels and their regional distribution.

Region	Channels
Frontal	Fp1, Fp2, Fpz, AF3, AF4, F3, F4, F7, F8, Fz
Fronto-central	FC1, FC2, FC5, FC6
Central	C3, C4, Cz
Temporal	T7, T8
Parietal	CP1, CP2, CP5, CP6, P3, P4, P7, P8, Pz
Occipital	O1, O2, POz

D. Signal Preprocessing

The raw EEG signals underwent a two-stage frequency filtering process. First, a fourth-order Butterworth bandpass filter with a passband of 0.5–45 Hz was applied to remove low-frequency baseline drift and high-frequency noise components. Second, a 50 Hz notch filter was used to suppress power line interference from the electrical grid. After these filtering steps, the signals were restricted to the frequency bands commonly used in emotion recognition research: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–45 Hz).

Following the frequency filtering stage, independent component analysis (ICA) was applied to further reduce artifacts caused by eye blinks, muscle movements, and environmental noise. The FastICA algorithm was used to decompose the 32-channel signal into statistically independent components. Three automatic criteria were employed for artifact component identification: (i) components showing high correlation with frontal channels were flagged as eye movement artifacts; (ii) components with kurtosis values exceeding a predefined threshold were classified as muscle artifacts; and (iii) components with abnormally high variance were identified as environmental noise. On average, seven components per participant were detected as artifacts and removed. The signal was then reconstructed from the remaining components.

E. Segmentation

After artifact removal, the cleaned EEG signals were divided into fixed-length segments suitable for feature extraction. A sliding window approach was used with a window length of 4 seconds (1024 samples at 256 Hz) and 50% overlap between consecutive windows (512-sample step size). This configuration yielded 22 segments from each 45-second recording, and a total of 1232 segments per packaging condition across all 56 participants. The combined dataset contained 2464 segments (1232 for wood fiber and 1232 for glass).

The choice of a 4-second window length was based on the trade-off between two competing factors: shorter windows provide insufficient frequency resolution for low-frequency bands such as delta, while longer windows risk violating the stationarity assumption that underlies spectral analysis methods [10]. The 50% overlap was applied to increase the number of available samples and to reduce boundary effects between adjacent segments.

F. Feature Extraction

From each segment, 28 features were computed for each of the 32 EEG channels, resulting in a 896-dimensional feature vector per segment (28×32). The features were organized into four categories:

Time-domain features (9 per channel): Hjorth parameters (activity, mobility, and complexity), which provide a compact statistical description of the signal waveform [11]; mean amplitude; skewness; kurtosis; root mean square (RMS); mean absolute deviation (MAD); and signal energy.

Frequency-domain features (10 per channel): Absolute and relative power for each of the five sub-bands (delta, theta, alpha, beta, and gamma), computed using Welch's power spectral density (PSD) estimation method. Relative power values were obtained by dividing each band's power by the total power, which helps normalize across participants with different overall signal amplitudes [5].

Nonlinear features (5 per channel): The largest Lyapunov exponent (LLE), Katz fractal dimension, Higuchi fractal dimension, Hurst exponent, and detrended fluctuation analysis (DFA). The Hurst exponent captures long-range temporal correlations in the signal, while DFA measures the scaling behavior of the signal after removing trends. These features are particularly effective in differentiating between emotional states through their sensitivity to neural complexity profiles.

Entropy features (4 per channel): Approximate entropy (ApEn), permutation entropy (PE), fuzzy entropy (FuzzyEn), and spectral entropy (SE). Fuzzy entropy is a more stable generalization of approximate entropy and is more robust to noise. Permutation entropy evaluates the ordinal pattern structure of the signal and is computationally efficient.

Table II provides a summary of the feature categories and their total dimensions.

No missing values (NaN) were detected in the feature matrix, confirming that the preprocessing workflow operated correctly across all recordings and channels.

TABLE II: Feature categories and dimensions.

Category	Per Ch.	Channels	Total
Time-domain	9	32	288
Frequency-domain	10	32	320
Nonlinear	5	32	160
Entropy	4	32	128
Total	**28**	**32**	**896**

G. Labeling

After feature extraction, each segment needed to be assigned to an emotional class. The labeling was based on the valence and arousal scores obtained from the post-experiment questionnaire. For each participant, the valence score was computed as the mean of the six valence items, and the arousal score as the mean of the five arousal items.

To define the four emotion classes, a condition-specific median split method was applied. The median values of valence and arousal were calculated separately for each packaging condition (wood fiber and glass). Scores above the median were labeled as "high" and scores below the median were labeled as "low." The intersection of these two binary dimensions produced four quadrants that are aligned with Russell's circumplex model of affect [6]: high valence–high arousal (HVHA) which corresponds to high excitement, high valence–low arousal (HVLA) corresponding to calmness, low valence–high arousal (LVHA) corresponding to discomfort, and low valence–low arousal (LVLA) corresponding to sadness.

The glass condition had positive median values for both valence (+1.167) and arousal (+0.600), indicating that participants generally responded to the glass bottle with more positive and activated emotions. In contrast, the wood fiber condition had near-zero and slightly negative medians (valence: -0.167, arousal: -0.200), suggesting a more neutral emotional response. No severe class imbalance was observed, with the largest class ratio remaining within 54%/46%.

H. Feature Selection and Classification

The 896-dimensional feature space poses challenges for classification in terms of both computational cost and the curse of dimensionality. To address this, the ReliefF algorithm was used for feature selection. ReliefF evaluates each feature by measuring how well it distinguishes between nearest neighbors from the same class and from different classes [12]. It is widely used in EEG-based classification because of its ability to handle multi-class problems and to capture feature interactions [13]. Feature selection was performed separately for each packaging condition, based on the assumption that different stimulus types may activate different neural and feature profiles. For each condition, the top 30 features were selected as a trade-off between cost and model accuracy as given in Table III. The configuration with 30 features is selected as it achieves a near-optimal balance between predictive performance and computational efficiency, maintaining high F1-scores while substantially reducing memory usage and inference latency compared to higher-dimensional settings. Although configurations with 100–200 features yield marginally higher accuracy,

TABLE III: Feature reduction efficiency comparison for wood fiber and glass conditions for test sets. Chosen feature set count is given in bold.

Features	Wood fiber					Glass					Reduction
	Acc.(%)	F1(%)	Train(s)	Predict(s)	Mem(KB)	Acc.(%)	F1(%)	Train(s)	Predict(s)	Mem(KB)	
10	75.1	70.7	1.211	0.0484	2530.7	82.3	78.9	1.161	0.0429	2532.4	89.6×
20	84.3	81.4	1.034	0.0407	2843.1	90.0	88.9	1.342	0.0777	2845.0	44.8×
30	**87.1**	**84.3**	**0.949**	**0.0487**	**3155.4**	**94.8**	**94.4**	**1.193**	**0.0573**	**3157.6**	**29.9×**
50	90.7	90.1	0.907	0.0476	3780.1	96.1	95.7	1.166	0.0441	3782.7	17.9×
100	92.3	90.8	0.896	0.0572	4405.9	94.7	94.4	1.215	0.0476	4409.0	9.0×
200	95.0	94.1	0.721	0.0762	4965.3	96.2	96.0	1.025	0.0686	4969.3	4.5×
500	82.7	76.3	0.763	0.0751	7236.5	93.9	92.8	0.956	0.0926	7242.4	1.8×
896	86.5	83.5	0.790	0.0741	9228.8	86.0	82.3	1.060	0.0892	9236.3	1.0×

the associated increases in memory footprint and computational cost make them less suitable for resource-constrained environments where efficiency and scalability are critical.

The selected features were then used to train classification models in the latest version of MATLAB's Classification Learner application. Multiple classifier families were tested, including decision trees, support vector machines (SVM), k-nearest neighbors (KNN), artificial neural networks, naive Bayes, discriminant analysis, and several ensemble methods. Among all tested models, the Subspace KNN ensemble consistently achieved the best performance in both conditions.

Subspace KNN is an ensemble method that trains multiple KNN classifiers, each operating on a randomly selected subset (subspace) of the feature space. The individual predictions are then combined through a voting mechanism. This approach reduces sensitivity to noise and provides more stable generalization compared to a single KNN classifier, especially in moderately high-dimensional feature spaces.

All models were evaluated using 10-fold cross-validation on the training set, and a separate held-out test set (approximately 29% of the data) was used to verify that the models generalize to unseen data. To interpret the model decisions, Shapley value analysis was applied [14]. Shapley values, originally derived from cooperative game theory, quantify the marginal contribution of each feature to the model output. By computing Shapley summary plots for each emotion class, it is possible to identify not only which features are most important overall, but also how their high or low values affect the prediction of specific emotional states.

III. Results and Discussion

A. Classification Results

The Subspace KNN ensemble classifier was evaluated on both packaging conditions. Table IV presents the class-level test set performance for the glass and wood fiber conditions.

The glass condition achieved 94.4% test accuracy with a macro F1 of 93.8%, while the wood fiber condition reached 90.7% accuracy with a macro F1 of 88.4%. The validation-to-test accuracy gap was less than 1 percentage point in both conditions, confirming that the models did not overfit. The glass condition produced only 20 misclassifications compared to 33 for the wood fiber condition, a 40% reduction in total errors. This difference suggests that the glass stimulus evoked stronger and more distinguishable emotional responses. In both conditions, HVLA was the weakest class due to its smaller sample size, while LVLA in the glass condition achieved near-perfect performance (F1 = 97.3%).

TABLE IV: Test set performance for both packaging.

Condition	Class	Prec.	Rec.	F1	N
Glass	HVHA	94.5	93.9	94.2	165
	HVLA	89.6	93.5	91.5	46
	LVHA	94.1	90.6	92.3	53
	LVLA	96.8	97.8	97.3	92
	Macro	**93.7**	**94.0**	**93.8**	**356**
Wood fiber	HVHA	87.7	98.8	92.9	145
	HVLA	96.0	72.7	82.8	33
	LVHA	90.5	82.6	86.4	46
	LVLA	93.7	89.4	91.5	132
	Macro	**92.0**	**85.8**	**88.4**	**356**

B. Model Complexity Comparison

To support the choice of the Subspace KNN ensemble, 16 classifiers from different families were compared using 10-fold cross-validation on the 30 features selected by ReliefF. Table V shows the accuracy, macro F1 score, prediction time, and model size for each classifier under both packaging conditions.

TABLE V: Model performance and complexity comparison for test set for both packaging conditions (30 ReliefF features, 10-fold CV).

Model	Glass		Wood fiber		Pred (s)	Size (KB)
	Acc	F1	Acc	F1		
Subspace KNN	**94.5**	**94.0**	**90.5**	**89.3**	0.0290	3157.6
Bagged Trees	92.4	91.8	88.8	88.1	0.0190	1307.6
Cubic SVM	89.9	88.6	85.6	84.3	0.0039	178.8
Wide NN (100)	88.6	87.6	84.7	83.2	0.0008	33.4
Medium NN (25)	88.1	86.7	83.0	81.1	0.0007	12.9
Bilayered NN	86.3	84.8	82.6	80.3	0.0007	15.3
Boosted Trees	86.0	84.0	80.0	75.3	0.0152	218.4
Fine Tree	83.7	80.1	79.6	77.8	0.0006	31.0
Narrow NN (10)	82.8	80.5	76.3	73.4	0.0006	8.8
Quadratic SVM	80.3	77.4	74.0	69.8	0.0036	310.3

Subspace KNN achieved the highest accuracy and F1 score in both conditions, scoring 2.1 and 1.7 percentage points above the second-best model (Bagged Trees) in the glass and wood fiber conditions, respectively. Although it has a larger model size (3157.6 KB) than simpler classifiers, its prediction time stays under 30 ms per observation, giving a throughput of over 34,000 observations per second. This shows that the model can be used in real-time settings despite being an ensemble. The neural network models (Wide NN, Medium NN, Bilayered NN) did not perform better than the ensemble methods, and linear models such as Linear Discriminant and Linear SVM gave much lower results. This supports the idea that the emotion classification task is nonlinear in nature.

C. Feature Importance

Shapley value analysis given in Fig. 3 showed that nonlinear features dominated the top-ranked contributions in both conditions. Fuzzy entropy and Hurst exponent consistently appeared among the most important features across all four emotion classes. Among frequency-domain features, only the relative delta power of the Fp1 channel showed a consistent contribution in both conditions. This suggests that the complexity structure and temporal correlation dynamics of EEG signals carry stronger emotion-related information than classical spectral power measures.

From a regional perspective, the wood fiber condition concentrated its discriminative features in frontal and temporal channels (Fp1, Fp2, Fpz, F3, F7, FC2, T7), while the glass condition showed a broader spatial distribution, recruiting parieto-occipital (POz) and centro-parietal (CP5) regions alongside frontal areas. This pattern is consistent with the idea that the glass bottle triggered a more complex cognitive evaluation involving visual and conceptual processing, while the wood fiber packaging was processed through a faster, more automatic frontal-temporal pathway [15]. A notable condition-specific finding was the high Shapley contribution of the DFA feature from the Fp1 channel in the glass condition across all four classes, which was absent in the wood fiber results.

D. Comparison with Literature

Table VI compares the results of this study with recent EEG-based four-class valence–arousal classification studies published after 2020.

TABLE VI: Comparison with V-A classification studies.

Study	Data	N	Method	Acc.
Galvão et al. [5]	DEAP	32	KNN	84.4%
Cui et al. [16]	DEAP	32	CNN-BiLSTM	89.3%
Liu et al. [17]	DEAP	32	FCAN-XGB	90.8%
Vempati & Sharma [18]	DEAP	32	Bagging	91.9%
Lin et al. [19]	DEAP	32	DSE-Mixer	91.9%
This study (Glass)	Custom	56	Sub. KNN	**94.4%**
This study (W. fiber)	Custom	56	Sub. KNN	**90.7%**

The glass condition result (94.4%) exceeds all compared studies, while the wood fiber result (90.7%) is comparable to the best-performing methods. It should be noted that all compared studies used the DEAP benchmark dataset with music video stimuli and 32 participants, whereas this study uses a custom dataset with packaging stimuli and a larger participant pool ($N = 56$). Additionally, the Subspace KNN classifier is a lightweight ensemble that requires no graphical processing unit (GPU), unlike the deep learning methods used in several of the compared studies.

E. Computational Performance

To evaluate whether the proposed system can work in practice, runtime measurements were recorded for each processing stage. Table VII shows the average execution time per recording for both the feature extraction and classification stages.

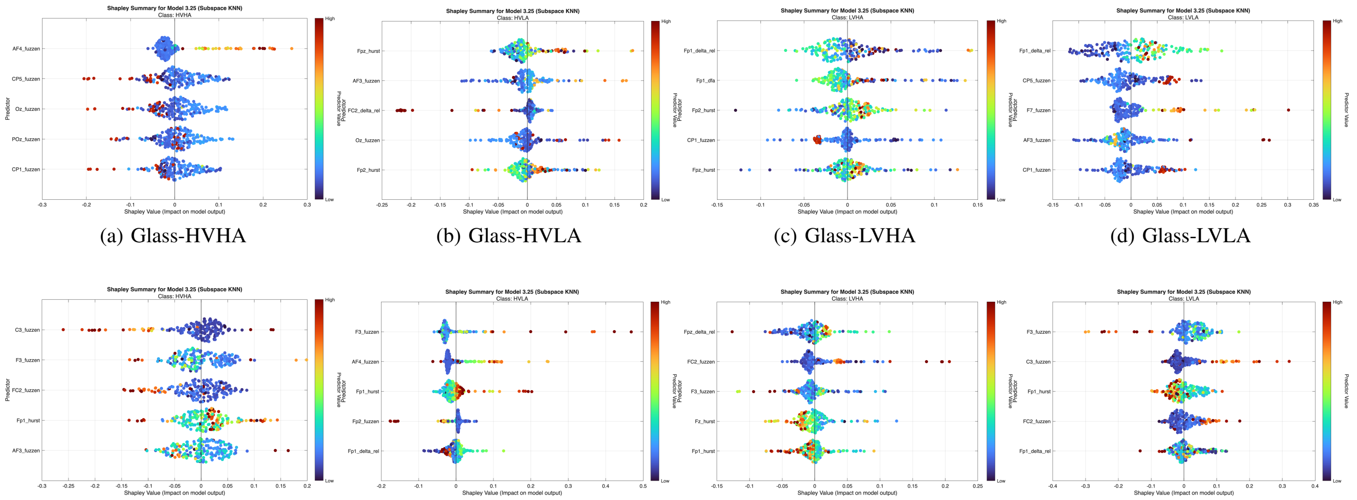
TABLE VII: Empirical runtime analysis of the processing stages.

Stage	Mean (s)	Std (s)	% Total
Feature Extraction			
Data Loading	0.0666	0.0141	0.1
Filtering	0.0119	0.0035	<0.1
ICA	2.1411	0.7406	3.4
Segmentation	0.0002	0.0000	<0.1
Feature Computation	16.3515	7.1088	25.6
Subtotal	*18.5713*		*29.1*
Classification			
ReliefF Selection	44.9135	11.7878	70.4
Subspace KNN Training	0.2493	0.1685	0.4
Prediction (Inference)	0.0690	0.0474	0.1
Subtotal	*45.2318*		*70.9*
Total	**63.8031**		**100.0**

The total processing time per recording was about 63.8 seconds. The ReliefF feature selection step took the most time, making up 70.4% of the total. This is expected because ReliefF calculates pairwise distances across all 896 features for every training sample. Feature computation was the second most time-consuming stage (25.6%), mainly because the nonlinear and entropy features require iterative calculations. On the other hand, Subspace KNN training took only 0.25 seconds and inference required just 0.069 seconds, which corresponds to a per-sample inference time of 0.029 ms and a throughput of about 34,459 observations per second. For the full dataset of 112 recordings (56 participants $\times$ 2 conditions), the total feature extraction time was estimated at 34.7 minutes. The ReliefF step reduced the feature matrix from 17.66 MB to 0.59 MB, a 29.9$\times$ reduction in size, and the compact model occupied 6.31 MB. These results show that the entire workflow, from raw signal to prediction, can run on standard hardware without the need for GPU resources.

IV. Conclusions

This study examined the emotional impact of sustainable vs. conventional packaging using EEG-based emotion recognition. Three main findings emerged from the analysis. First, nonlinear features provided stronger emotion discrimination than classical frequency band powers, suggesting that signal complexity carries important emotional information that is often overlooked in conventional EEG analysis. Second, the two packaging conditions activated different brain networks: the wood fiber condition engaged mainly frontal-temporal areas, while the glass condition additionally recruited parieto-occipital regions, indicating a deeper cognitive evaluation process. Third, the glass stimulus produced more polarized emotional neural responses, resulting in higher classification accuracy. Additionally, a comparison of 10 classifiers showed



(a) Glass-HVHA (b) Glass-HVLA (c) Glass-LVHA (d) Glass-LVLA

(e) Fiber-HVHA (f) Fiber-HVLA (g) Fiber-LVHA (h) Fiber-LVLA

Fig. 3: Class activation maps for two groups under four valence–arousal conditions. The first row corresponds to contribution of best performing 5 features to emotional states under Glass, and second row for Fiber condition.

that the Subspace KNN ensemble achieved the best accuracy in both conditions, outperforming neural networks, SVMs, and tree-based methods. Runtime analysis showed that the full processing workflow takes about 64 seconds per recording on standard hardware and does not require a GPU. The model reached a per-sample inference time of 0.029 ms with a throughput of over 34,000 observations per second, which makes it suitable for real-time use.

The study has some limitations that will be addressed in future studies. The participant sample consisted of university students under 25, which limits the generalization of the findings to other age groups. Only two packaging types were tested, with stimuli restricted to the visual modality. The median split labeling method may also cause some information loss by converting continuous scores into binary categories. Future work will extend this study by including a wider variety of packaging materials and by incorporating tactile stimuli to increase ecological validity. Another limitation of the study is the possible omission of neural habituation.

Overall, these findings demonstrate that packaging material choice produces measurable differences at the brain activity level, providing objective neural evidence that can inform both neuromarketing research and sustainable product design.

References

[1] C. Spence, C. Velasco, and O. Petit, "The consumer neuroscience of packaging," in *Multisensory Packaging*, pp. 319–347, Springer, 2019.

[2] L. Alvino, L. Pavone, A. Abhishta, and H. Robben, "Picking your brains: Where and how neuroscience tools can enhance marketing research," *Frontiers in Neuroscience*, vol. 14, 2020.

[3] J. Wang, A. H. Alsharif, N. Z. Abd Aziz, A. Khraiwish, and N. Z. M. Salleh, "Neuro-insights in marketing research: A PRISMA-based analysis of EEG studies on consumer behavior," *SAGE Open*, vol. 14, no. 4, 2024.

[4] P. Cherubino *et al.*, "Consumer behaviour through the eyes of neurophysiological measures: State-of-the-art and future trends," *Computational Intelligence and Neuroscience*, vol. 2019, 2019.

[5] F. Galvão, S. M. Alarcão, and M. J. Fonseca, "Predicting exact valence and arousal values from EEG," *Sensors*, vol. 21, no. 10, p. 3414, 2021.

[6] J. A. Russell, "A circumplex model of affect," *J. Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.

[7] O. A. Bozbas, B. Akbugday, A. Akan, and S. P. Akbugday, "Classification of emotions under multiple olfactory stimuli using EEG signals and machine learning," in *2025 Medical Technologies Congress (TIPTEKNO)*, Gazi Magusa, Turkiye, pp. 1–5, 2025.

[8] S. Koelstra *et al.*, "DEAP: A database for emotion analysis using physiological signals," *IEEE Trans. Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.

[9] N. Leeuwis, T. van Bommel, and M. Alimardani, "A framework for application of consumer neuroscience in pro-environmental behavior change interventions," *Frontiers in Human Neuroscience*, vol. 16, 2022.

[10] O. Bazgir, Z. Mohammadi, and S. A. H. Habibi, "Emotion recognition with machine learning using EEG signals," in *25th Iranian Conf. Biomedical Engineering (ICBME)*, 2018.

[11] B. Akbugday, A. Akan, S. Pehlivan, and R. Sadighzadeh, "An assessment of linear and nonlinear features for detecting olfactory stimulus in EEG," in *2022 Medical Technologies Congress (TIPTEKNO)*, Antalya, Turkey, pp. 1–4, 2022.

[12] S. P. Akbugday, B. Akbugday, F. Yeganli, A. Akan, and R. Sadikzade, "Decoding olfactory EEG signals using multi-domain features and machine learning," in *2024 Medical Technologies Congress (TIPTEKNO)*, Mugla, Turkiye, pp. 1–4, 2024.

[13] J. Zhang *et al.*, "ReliefF-based EEG sensor selection methods for emotion recognition," *Sensors*, vol. 16, no. 10, p. 1558, 2016.

[14] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, vol. 30, pp. 4766–4777, 2017.

[15] R. J. Davidson, "Anterior cerebral asymmetry and the nature of emotion," *Brain and Cognition*, vol. 20, no. 1, pp. 125–151, 1992.

[16] F. Cui, R. Wang, W. Ding, Y. Chen, and L. Huang, "A novel DE-CNN-BiLSTM multi-fusion model for EEG emotion recognition," *Applied Sciences*, vol. 12, no. 21, p. 11255, 2022.

[17] Y. Liu, G. Fu, J. Liu, and S. Chen, "FCAN–XGBoost: A novel hybrid model for EEG emotion recognition," *Sensors*, vol. 23, no. 12, p. 5680, 2023.

[18] R. Vempati and L. D. Sharma, "EEG rhythm based emotion recognition using multivariate decomposition and ensemble machine learning classifier," *J. Neuroscience Methods*, vol. 390, p. 109845, 2023.

[19] K. Lin, L. Zhang, J. Cai, J. Sun, W. Cui, and G. Liu, "DSE-Mixer: A pure multilayer perceptron network for emotion recognition from EEG feature maps," *J. Neuroscience Methods*, vol. 401, p. 110008, 2024.