Analyzing EEG Patterns in Functional Food Consumption: The Role of PCA in Decision-Making Processes

Mauro Daniel Castillo Pérez[1], Jesús Jaime Moreno Escobar[2], Verónica de Jesús Pérez Franco[3],
Ana Lilia Coria Paéz[4], Oswaldo Morales Matamoros[5]
Escuela Superior de Ingeniería Mecánica y Eléctrica, Zacatenco, Instituto Politécnico Nacional, México[1]
Centro de Investigación en Computación, Instituto Politécnico Nacional, México[2,5]
Unidad Profesional Interdisciplinaria de Ingeniería y Ciencias Sociales y Administrativas, Instituto Politécnico Nacional, México[3]
Escuela Superior de Comercio y Admistración, Tepepan, Instituto Politécnico Nacional, México[4]

***Abstract*—The impact of obesity and diabetes are two central reasons for the high rate of developing cardiovascular diseases in this country, which is largely due to their ultra-processed, diet-rich foods. Supervised Learning for Decision Making: A Case Study of Functional Food Taste Perceptions In this experiment, we trained ordinary consumers to estimate the taste preferences of a unique group from which no ratings were available(11) and established that decision making can be performed through supervision. A deep learning neural network architecture according to the present disclosure is designed to model the decision-making behavior of consumers consuming functional products. The efficiency of the model can be increased upto 1.23% making use of proper values for the rest of the hyperparameters as explained in experiments carried out where we set the optimal configuration so that nurturing it gives the best results.**

***Keywords*—EEG analysis; functional foods; decision-making; deep learning; Principal Component Analysis (PCA)**

I. Introduction

By 2023, decision-making has become essential and complex due to the vast amount of information and uncertainty present today. Additionally, various factors such as the fear of failure can negatively affect decision-making. This also impacts health, specifically obesity and overweight, which are linked to dietary decisions. In Latin America, obesity has increased by 55%, and in Mexico, it affects 76.4% of adults and 35.6% of children between the ages of 5 and 11, [1].

Some of these factors could be:

- Deceptive Advertising: Unhealthy food products are promoted as healthy, making informed decisions difficult (for example, whole grain bread labeled as “light”), [2].
- Social Pressure: The discrediting of people with obesity influences their dietary choices and self-image, sometimes leading them to consume unhealthy foods to cope with stress, [3].

Decision-making affects not only at a personal level but also in governmental and corporate spheres, where incorrect decisions can cause significant damage.

Some examples where a wrong decision could have an impact are:

- Policy: Lack of transparency and corruption can have devastating consequences.
- Environment: Disputes and short-term economic decisions harm the planet.
- Social Justice: Decisions about economic and social policies can influence inequality.
- Education: Decisions about curricula and funding affect the quality of education.

Principal Component Analysis (PCA) and Artificial Intelligence (AI), including Deep Convolutional Neural Networks (DCNN), [4], can help analyze complex data and improve decision-making. These technologies can identify patterns in health data and predict diseases by organizing large datasets and providing more accurate and predictive analyses. In the context of overweight and obesity, they can enhance dietary decision-making by providing a deeper understanding of the factors influencing these conditions. On the other hand, four main deficiencies of PCA can be stated: i) Assumes linear relationships: PCA can only capture linear relationships between variables. If the relationships between the data are nonlinear, PCA will not be able to model them adequately; ii) Loss of interpretability: Although PCA reduces dimensionality, the principal components do not have a clear interpretation in terms of the original variables, which can make the results difficult to interpret; iii) Sensitivity to noise: If the data contains noise, it can influence the computation of the principal components, affecting the results; and iv) Computationally expensive for large datasets: For very large or high-dimensional datasets, the computation of the covariance matrix and its significant vectors can be computationally expensive.

The design of an intelligent system using deep learning networks makes it possible to classify a person’s like/dislike choices with an efficiency of 80%.

This work is based on EEG Analysis, a technique that uses electroencephalographic signals to study brain activity, being useful in the analysis of brain patterns.

The analysis of EEG is important for various reasons, primarily due to its ability to provide information about the brain's electrical activity, allowing the study and understanding of a variety of cognitive, emotional, and physiological processes. It also has several key study reasons, such as:

- Study of brain activity in real time: EEG provides a real-time window into brain activity, allowing the direct observation of changes in brain wave patterns related to different mental states, such as attention, relaxation, sleep, or decision-making [5].
- Study of decision-making: EEG is also used to investigate how the brain makes decisions by detecting neuronal responses to external stimuli [6].
- Interaction between emotions and decisions: EEG is used to measure brain activity, providing information about how emotions influence decision-making [7].

It also incorporates the use of Functional Foods, which are foods that, in addition to their nutritional value, offer additional health benefits, such as disease prevention or improving a person's health. The philosophy of Decision-Making is applied, which is a cognitive process through which a person evaluates options and selects the most appropriate one to make an informed decision. This process is complemented by the tool of Deep Learning, a branch of artificial intelligence based on artificial neural networks that allows analyzing complex data, identifying patterns, and making accurate predictions and the tool Principal Component Analysis (PCA) is a statistical technique that reduces the dimensionality of complex data sets while preserving the most relevant characteristics. The combination of these concepts enables efficient classification in decision-making.

This work consists of a total of five sections, starting with the introduction, which provides the context of the work. Similarly, related work is given in Section II, presenting an analysis of studies directly related to the proposed one. This is followed by the theoretical framework in Section III, which explains the mathematical and theoretical foundations used in the work. The methodology in Section IV explains how the network operates, leading to the experiment in Section V, where the functionality of the proposed method is tested, and finally, the conclusions in Section VI is presented, where the obtained results are discussed.

II. Related Works

A search was conducted in the databases of IEEE Xplore and Hindawi; this research was carried out with the purpose of finding works that have a similar relationship with the approach followed in the preparation of this document. A total of 13 works were found, which relate to the fundamental topics for the development of this study. From these 13 works, 4 works were derived, whose focus provides a greater contribution to the research.

Fig. 1 is a PRISMA flow diagram, which includes the systematic review of citations, where less relevant works to the objective are discarded, and where these citations were searched in the aforementioned search sites. In this way, 13 articles were found, for example, there is one titled Development of AI model to analyze customer behavior by decision making system, written by L. Rong, Y. Ding, M. Wang, M.S. Hossain et al. in [8], which discusses a model that uses artificial intelligence for facial recognition and deep neural networks to analyze customer behavior. It investigates how the consumer's mental process influences the purchasing process. Another example is the paper titled Analysis of Consumer Coffee Brand Preferences Using Brain-Computer Interface and Deep Learning, written by M. Maram, M.A. Khalil, K. George et al. in [9]. This work focuses on the acquisition and analysis of EEG signals using a wireless device, where tools like MATLAB and Python in Google Colab are used to design an interface aimed at identifying coffee brand preferences through brain activity. A different case is the paper titled A Deep Learning Model for Classification of EEG written by S.M. Usman, S.M.A. Shah, O.C. Edo, J. Emakhu et al. in [10], where the study focuses on the use of EEG signals to classify product preferences by observing their packaging. Brain signals from 25 volunteers were monitored while they viewed different packages, using data analysis tools and machine learning models to automatically identify consumer preferences. Finally, there is the paper titled A Survey on Neuromarketing Using EEG Signals, written by V. Khurana, M. Gahalawat, P. Kumar, P.P. Roy, D.P. Dogra, E. Scheme, M. Soleymani et al. in [11], which explains how neuroscience uses tools like EEG. This technique is used to analyze brain activity during the consumer decision-making process. Electrodes are employed to measure brain waves, and analysis software is used to interpret the collected data, thus enabling an understanding of consumer preferences.

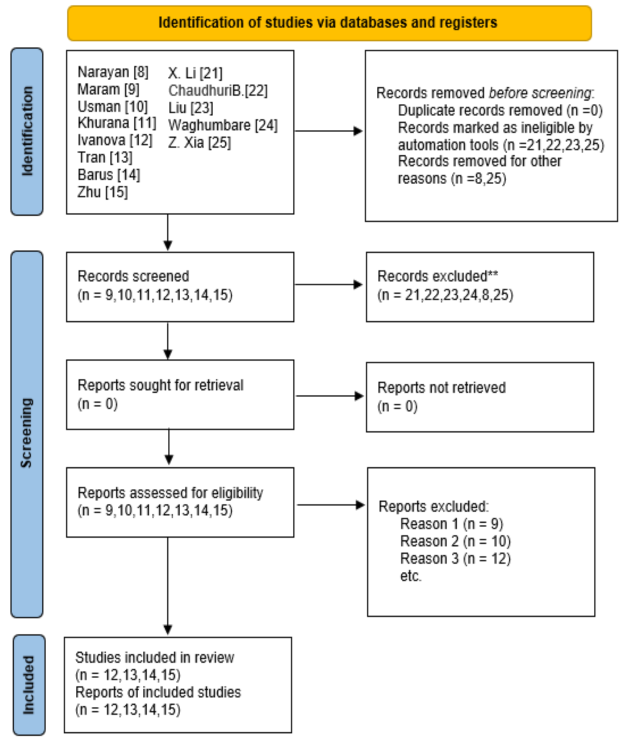

Fig. 1. Identification of related works via research databases such as IEEE Xplore and Hindawi.

Once some relevant works are presented, the following section will present those that are most directly related to this work, fulfilling similar objectives and/or analyses, these are:

i) *Motor Learning and Decision Making in Volatile Environment in Bipolar Disorder*,
ii) *On a Development of Sparse PCA Method for Face Recognition Problem*,
iii)*Prediction Using Support Vector Machine and Logistic Regression Model with Combination of PCA and SMOTE*,
iv)*Continuous Speech Recognition Based on DCNN-LSTM*.

A. Motor Learning and Decision Making in Volatile Environment in Bipolar Disorder

This study by M. Ivanova and M. H. Ruiz in [12] investigates how motor learning and decision-making in a volatile environment affect individuals with bipolar disorder. The hypothesis is that disruptions in learning in volatile environments may play a crucial role in the development and manifestation of bipolar disorder. The study compares patients with bipolar disorder to a control group of healthy individuals, using psychological tests and experimental tasks that simulate conditions of uncertainty. A neural network model is employed to analyze the data and seek patterns that relate motor learning to neurophysiological markers, aiming to improve the diagnosis and treatment of bipolar disorder. The diagram of this study can be seen in Fig. 2.

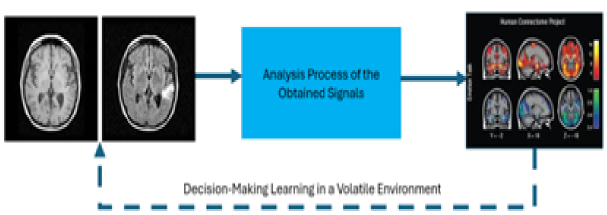

Fig. 2. Model diagram, motor learning and decision making in a volatile environment in bipolar disorder.

B. On a Development of Sparse PCA Method for Face Recognition Problem

This work by L. Tran and colleagues in [13], proposes an innovative semi-supervised learning method based on the unnormalized p-Laplacian graph for speech recognition. The method aims to improve the accuracy and performance of speech recognition through machine learning techniques and signal processing. A semi-supervised learning model is constructed using voice samples, and experiments are conducted to evaluate its effectiveness compared to other methods, achieving significant improvements in accuracy and performance, Fig. 3.

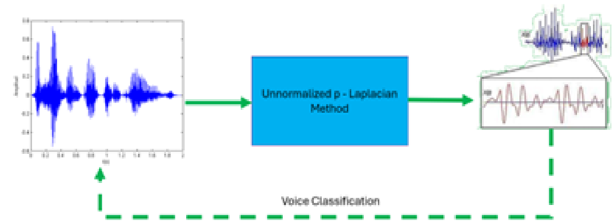

Fig. 3. Model diagram, on a development of sparse PCA method for face recognition problem.

C. Prediction Using Support Vector Machine and Logistic Regression Model with Combination of PCA and SMOTE

This study conducted by O. P. Barus and colleagues in [14], focuses on the use of machine learning for liver disease prediction and facial recognition. In the first study, Logistic Regression (LR) and Support Vector Machine (SVM) algorithms are employed, along with PCA and SMOTE, to improve early and accurate detection of liver diseases. In the second study, advanced versions of sparse PCA are used to improve the accuracy of facial recognition, applying methods such as Proximal Gradient Sparse PCA and Fast Iterative Shrinkage-Thresholding Algorithm Sparse PCA. Fig. 4 shows both studies combing clinical and facial data to achieve accurate and effective predictions.

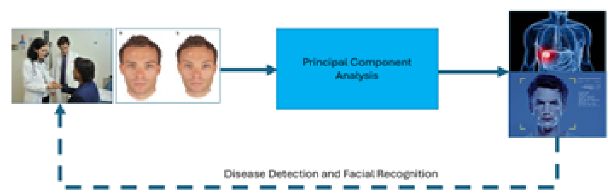

Fig. 4. Model diagram, prediction using support vector machine and logistic regression model with combination of PCA and SMOTE.

D. Continuous Speech Recognition Based on DCNN-LSTM

This work by Y. Zhu and Q. Zeng in [15], compares different acoustic features and network structures in automatic speech recognition. Using Mandarin datasets (THCHS-30 and ST-CMDS), the study extracts acoustic features such as the spectrogram and Mel cepstral coefficient (MFCC). Deep Convolutional Neural Networks (DCNN) and Long Short-Term Memory Neural Networks (LSTM) are evaluated, concluding that the spectrogram is the most effective feature and that LSTM networks significantly improve speech recognition accuracy compared to DCNNs. Fig. 5 depicts the combination of DCNN and LSTM results in a remarkable improvement in speech-to-text conversion.

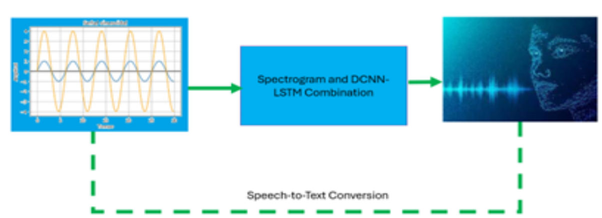

Fig. 5. Model diagram, continuous speech recognition based on DCNN-LSTM.

III. Theoretical Framework

A. Principal Component Analysis PCA

Principal Component Analysis (PCA) is a fundamental statistical technique used in multivariate data analysis. Its objective is to transform a set of P correlated data into a new set of fewer and uncorrelated variables. PCA reduces the complexity of the original data, facilitating its interpretation and analysis, and allows for the identification of hidden patterns and structures within the data. The PCA method takes data described by k variables in an $n \times k$ matrix X, representing n subjects in a k-dimensional space, Eq. (1):

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \tag{1}$$

The new variables, or principal components, are generated through linear combinations of the original variables. The aim is for the first principal component to contain the maximum possible variance, while restricting the magnitude of its elements to avoid excessive variance. This process is repeated to find the subsequent principal components, ensuring that each new component is uncorrelated with the previous ones.

B. Electroencephalogram (EEG)

The Electroencephalogram (EEG) is a study that detects and records the electrical activity of the brain under different conditions, including the basal state and through activation methods such as hyperventilation and photostimulation. The electrical signal collected is amplified and represented in the form of lines that show the activity of various brain areas over time, a representation of this can be seen in the Fig. 6.

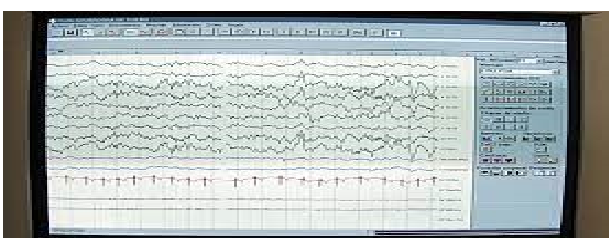

Fig. 6. Brain electrical signal.

The EEG monitors the electrical functioning of the brain and can detect both global alterations and changes in specific areas. It is useful for identifying various lesions such as tumors, hemorrhages, encephalitis, and trauma, Fig. 7.

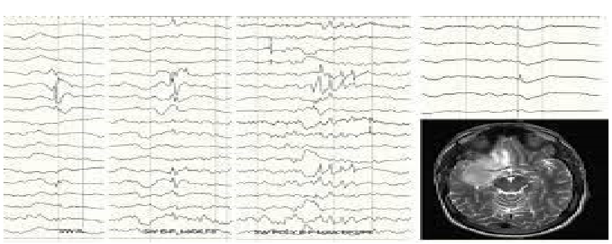

Fig. 7. EEG of a brain tumor.

C. Deep Learning Neural Networks

1) VGG16 Neural Network: The VGG16, Fig. 8, neural network is a deep convolutional network architecture developed by the Visual Geometry Group (VGG) at the University of Oxford. Used in computer vision tasks such as image recognition and classification, VGG16 is characterized by its 16-layer depth, divided into 13 convolutional layers and 3 fully connected layers.

a) Main features:

- Convolution Filters: Uses 3x3 filters with a fixed depth of 64 and 128 in its initial layers, followed by additional layers with more filters.
- Pooling Layers: Employs pooling layers with a 2x2 window to reduce dimensionality and extract prominent features.

The training of VGG16 follows the standard supervised learning approach, using labeled datasets and adjusting the network weights through backpropagation and optimization with stochastic gradient descent (SGD).

This is a complex architecture due to the fact that it consists of 16 training layers, where 13 are convolutional layers, using a 3x3 matrix for feature extraction, and 3 fully connected layers at the end, to reduce the dimensionality it has max pooling layers. At the input, it receives images of 224x224 pixels with three channels (RGB), and in the end, it has a softmax layer that allows for image classification.

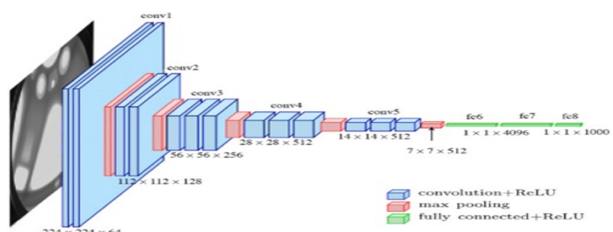

Fig. 8. VGG16 neural network architecture.

2) Inception Neural Network: The Inception neural network was created by Google Brain in 2014 to address the challenges of image classification in large and complex datasets. Its modular structure uses multiple convolutional layers and Inception modules, allowing the network to learn and extract features at different spatial scales within an image.

a) Main features:

- Inception Modules: They combine parallel convolutional operations with different filter sizes to capture information at different scales.
- Efficiency and Accuracy: The modular structure enables more efficient and accurate image classification.
- Text Recognition and Classification:Adapted for text processing tasks such as classification and sentiment analysis.

The training of the Inception network uses large labeled datasets and deep learning techniques, with a focus on optimization through stochastic gradient descent.

This architecture is composed of modules, meaning that instead of using a single convolutional operation, this architecture uses filters of different matrix types, which allows capturing various features at different scales within the same block. These results are combined, resulting in better characterization. Additionally, it uses dimensionality reduction layers to maintain efficiency, which makes it effective in applications with images.

3) AlexNet Neural Network: The AlexNet neural network, Fig. 9, is an architecture that marked a significant shift in the field of deep learning by significantly outperforming other architectures in image classification. Developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, AlexNet consists of 5 convolutional layers and 3 fully connected layers.

a) Main features:

- Pooling Layers: It uses pooling layers to reduce dimensionality and prevent overfitting.
- Non-linear Activation Functions: It employs non-linear activation functions to accelerate training and prevent the vanishing gradient.

The training of AlexNet follows the supervised learning paradigm, using backpropagation and stochastic gradient descent on large datasets such as ImageNet.

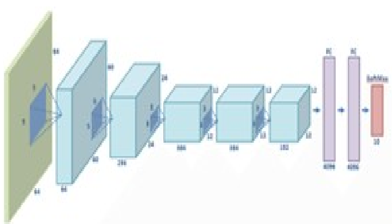

Fig. 9. AlexNet neural network architecture.

AlexNet is composed of a set of layers designed to extract spatial features at different levels. This architecture consists of 5 convolutional layers with large kernels that capture broad spatial patterns, followed by 3 fully connected layers that process and classify the learned features. Similarly, it incorporates max-pooling layers to reduce dimensionality and uses ReLU activation functions. Additionally, it includes mechanisms such as dropout to prevent overfitting during training.

4) ResNet Neural Network: The ResNet neural network was developed by Microsoft Research in 2015 to address the vanishing gradient problem and enable the training of much deeper models without performance degradation.

a) Main features:

- Residual Blocks: They include connections that allow information to flow directly through the layers, making it easier to train deep networks.

ResNet, Fig. 10, training follows the principles of supervised deep learning, using large labeled datasets and optimization via stochastic gradient descent.

The ResNet50 architecture is composed of residual blocks, meaning that instead of learning features directly, the layers

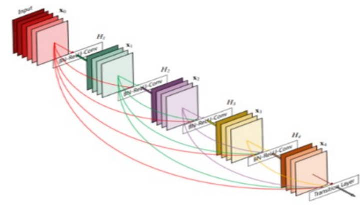

Fig. 10. ResNet neural network architecture.

focus on capturing the differences between the input and the output within each block. This is achieved through shortcut connections, which allow data to pass directly through the network without going through all intermediate layers. This architecture not only facilitates network training but also avoids issues like gradient vanishing. Thanks to this, it is widely used for processing images with high precision.

5) EfficientNet Neural Network: The EfficientNet neural network was developed by Google in 2019 to optimize the balance between accuracy and computational efficiency in image processing.

a) Main features:

- Compound Scaling: It optimizes the width, depth, and resolution of the network in a balanced manner using a method called Compound Scaling.

The training of EfficientNet uses labeled datasets and supervised learning techniques, adjusting the weights of the network through optimization algorithms such as stochastic gradient descent.

The EfficientNet architecture is designed to optimize efficiency in computer vision tasks, focusing on a single aspect. This architecture uses a compound scaling approach to uniformly scale the depth, width, and input resolution dimensions of the model, achieving a more comprehensive characterization of image features. It is based on modules that combine convolutional operations with dimensionality reduction techniques, achieving a balance between accuracy and computational efficiency.

6) You Only Look Once (YOLO) Algorithm: The YOLO algorithm stands out for its ability to detect objects in images quickly and accurately. Developed in 2016, YOLO approaches object detection as a simultaneous classification and bounding box regression problem.

a) Main features:

- Single-Pass Detection: It performs object detection in a single pass through the convolutional neural network.
- Multi-object predictions: It can detect multiple objects in an image, assigning a class probability to each bounding box.
- Real-Time: Its ability to perform real-time object detection makes it suitable for applications requiring rapid response.

The training of YOLO, Fig. 11, involves using labeled datasets and applying supervised learning techniques. The network is trained to recognize patterns and features of objects, adjusting the weights through backpropagation and optimization techniques.

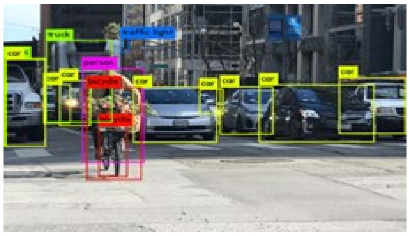

Fig. 11. Example of object detection with YOLO.

The YOLO architecture is designed for real-time object detection through the use of modules organized into blocks, where each block includes convolutional operations of various matrix sizes, allowing it to capture features at different scales within the same structure. These layers consist of max-pooling operations to reduce dimensions and preserve relevant features.

7) Data Augmentation: Data augmentation is a fundamental technique in machine learning and data processing. It involves creating new data samples by applying transformations to existing data while preserving the original information and label.

a) Applications and techniques:

- Image Augmentation: Rotations, flips, crops, changes in lighting, brightness and zoom, Fig. 12.

Fig. 12. Example of data augmentation on an image.

- Text Augmentation: Change of word order, synonyms, punctuation, and sentence structure, Fig. 13.

Fig. 13. Example of data augmentation in text.

- Audio Augmentation: Change of pitch, addition of noise, and variation of speed.

Data augmentation is applied before training the model to increase data variety and improve the model's predictive capability and robustness. There are various libraries in Python such as TensorFlow, Keras, PyTorch, and Augmentor that facilitate the implementation of these techniques.

IV. METHOD

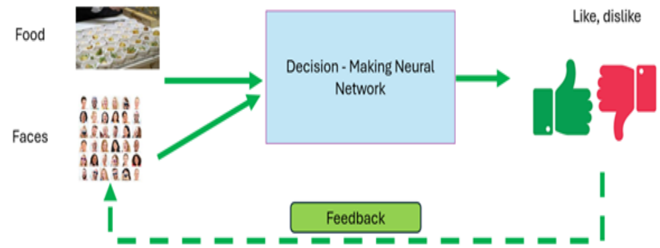

Fig. 14. General model of the system.

The decision-making system, Fig. 14, uses a neural network that analyzes images of food and the faces of consumers. This data enters a model that interprets the images to determine whether the food is considered "like" or "dislike" by the person. The model is self-correcting: if the interpretation does not match the expected outcome, it learns and adjusts its response.

A. Acquisition of the EEG Signal

The EEG signal measures the electrical activity during the excitation of pyramidal neurons in the cerebral cortex. This activity generates an electric and magnetic field, measurable with EEG systems through the skull and scalp, which attenuate the signal and add noise. The electrodes used to capture the EEG signal can be superficial, basal, or surgical. For this research, superficial electrodes are used, placed on the scalp following the International 10-20 Positioning System, based on anatomical points such as the inion, nasion, and lobes of the ears, this configuration does not generate any bias in decision-making.

For signal capture, the ThinkGear TGAM1 is used, which collects neural signals and processes them into usable data, filtering out interference. The NeuroSky ThinkGear module is a non-invasive interface with a 98% confidence level. However, due to noise levels, the device applies a reliable notch filter of 50Hz or 60Hz to reduce such noise. If not adequately filtered, it can affect the values of attention, meditation, and raw EEG data [16]. Additionally, signal loss through the scalp can impact the device's reliability. To address this, the module uses the poor quality code [16], which monitors electrode quality. If the connection with the scalp is suboptimal, this value increases to 200, indicating issues in signal acquisition, such as contact losses due to movement or incorrect placement.

The device is positioned at Fp1, as shown in Fig. 15, which measures electrical activity in the left frontal part of the brain, making it important for neurological diagnostics and scientific studies. Due to the fact that the Fp1 position is more directly associated with the person's emotional activity and attention, it makes it a more relevant indicator for assessing emotional preferences in decision-making. On the other hand, the Fp2 position is not considered suitable for this study, as it is not as directly linked to the person's emotional regulation and attention, which would affect the result [17], therefore the Fp1 position is the most suitable for the study of decision-making.

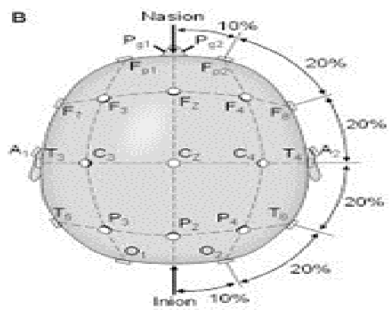

Fig. 15. Fp1 position.

B. The internal Model of the System

Internally, the model uses data reception from the products and the participants' faces to output classified and useful data. This helps to reduce dimensionality using Principal Component Analysis. Once reduced, we move to the next block where we extract data samples to send to the training model block. At the end of this process, the proposed model will indicate "like" or "dislike.", Fig. 16.

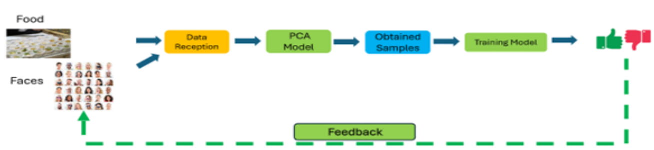

Fig. 16. Internal model of the system.

V. Experimental Results

A. Initial Configuration

The minimum requirements to run the model are as follows:

- Windows 10
- Intel Core i5
- 8.00 GB
- NeuroSky ThinkGear
- Google Colab
- Test Subjects

B. Results

C. Functionality Test

The functionality test is designed to demonstrate the use of the model, where the experiment is divided into faces and products.

1) Functionality Test of Products: As a first step, the file containing the images of the products and faces must be downloaded to train and test the model. In this case, the folder is named flavor.zip, Fig. 17. Additionally, the test image is loaded to verify that the data has been loaded correctly.

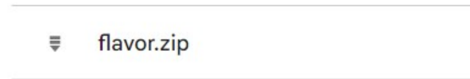

Fig. 17. Samples folder.

Once it is verified that the data has been loaded correctly, we proceed to build and train the neural network, thus providing the following model, which shows the convolutional and pooling layers, one example of this can be see in the Fig. 18.

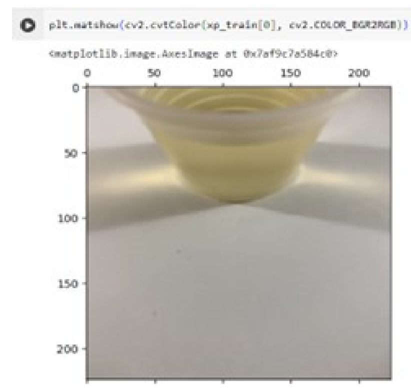

Fig. 18. Example of product samples.

For the case of the product experiment, we have the following inference matrix, which, as can be seen, classifies "like" and "dislike.", Fig. 19.

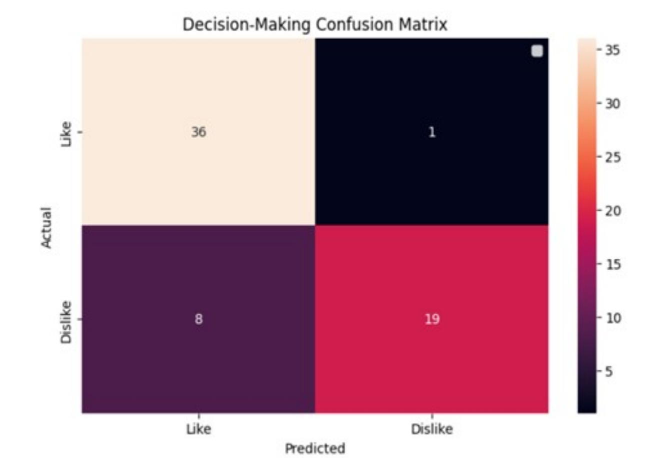

Fig. 19. Product inference matrix.

Similarly, for the case of the face experiment, we have the following inference matrix, which, as can be seen, classifies "like" and "dislike", Fig. 20.

As the final step, the inference matrix is shown with the combination of both experiments, resulting in the following outcome, Fig. 21.

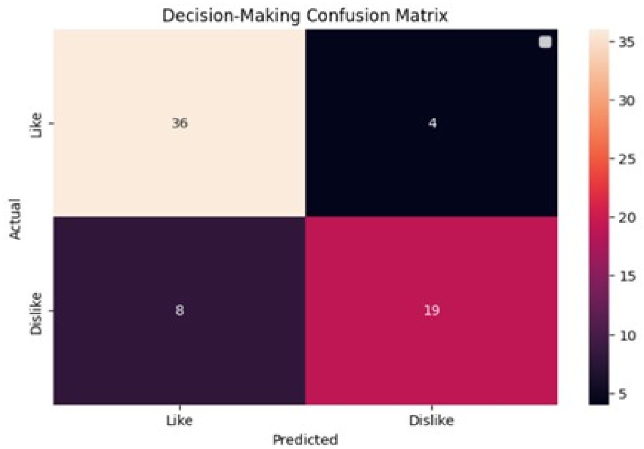

Fig. 20. Face inference matrix.

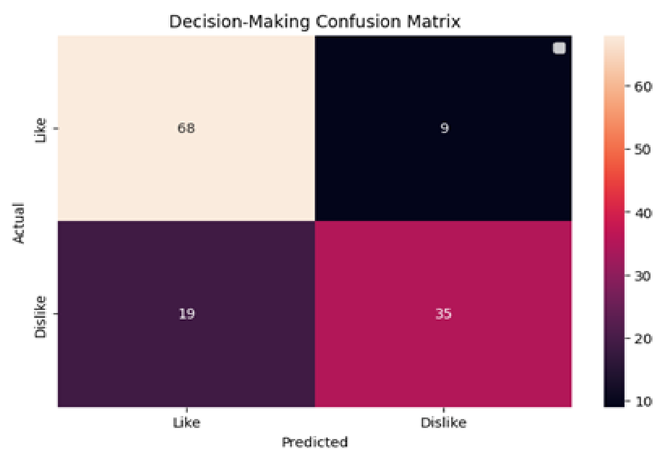

Fig. 21. Inference matrix of both experiments.

D. Experiment

To improve or observe changes in the percentage of effectiveness of the model, a series of experiments were implemented, using 5 variables that are part of the model. These are:

- Activation fuction of the model (A).
- Early Stopping during model training (B).
- Optimizer of the model (C).
- Sample size of images when entering the model (D).
- Amount of training data for the model (E).

The following the Table I, shows the aforementioned variables, along with their high and low values (or zero and one). It is worth noting that combinations of these variables were made to conduct possible tests, both in the Product model and in the Face model, resulting in two experiments.

Since both experiments use the same variables, the following hypothesis can be proposed:

TABLE I. VALUES OF THE VARIABLES TO ANALYZE

Variable	Low (0)	High (1)
Activation Fuction	ReLU	Leaky ReLU
Patience	5	10
Optimizer	ADAM	NADAM
Image Size	48	224
Batch Size	32	64

$$
\begin{array}{llllll}
H_0: & A_i = 0; & B_i = 0; & C_i = 0; & D_i = 0; & E_i = 0; \\
H_1: & A_i \neq 0; & B_i \neq 0; & C_i \neq 0; & D_i \neq 0; & E_i \neq 0;
\end{array}
$$

$$A_j \quad B_j \quad C_j \quad D_j \quad E_j$$

The results of both experiments are shown below

1) Product Experiment: The experiment is conducted using analysis of variance (ANOVA) for multiple factors on the F1 Score, Table II. Several tests and graphs are performed to determine which variables have a statistically significant effect on the F1 Score. It also evaluates the significance of interactions between variables, as long as sufficient data is available. Once the experiment is conducted, combining the variables to measure the outcome, the following results are obtained.

F1 SCORE HACER LO MISMO, duda en la pregunta

TABLE II. VARIANCE ANALYSIS F1 SCORE – TYPE III SUM OF SQUARES VALUES OF PRODUCTS

Source	P-Value
MAIN EFFECTS	
A: Activation Fuction	0.1560
B: Patience	0.0000
C: Optimizer	0.9295
D: Image Size	0.0038
E: Batch Size	0.0004

In the previous table, the results of the variance analysis are shown through a Type III sum of squares. The results indicate that both hypotheses A and C do not contribute to the outcome, and therefore, they are discarded.

$$
\begin{array}{ll}
H_0: & A_i = 0 \\
H_0: & C_i = 0
\end{array}
$$

Once the previous hypotheses are eliminated, the experiment is recalculated, where the following result is obtained, in the Table III.

TABLE III. VARIANCE ANALYSIS F1 SCORE – TYPE III SUM OF SQUARES OF SIGNIFICANT VARIABLES FOR PRODUCTS, SECOND EXPERIMENT

Source	P-Value
MAIN EFFECTS	
A: Patience	0.0000
B: Image Size	0.0038
C: Batch Size	0.0004

As can be seen in the table above, the remaining hypotheses do influence the experiment, which is why they are considered important. Once obtained, through the comparison of means, the value that effectively influences the percentage of effectiveness of the model can be observed, thus having the following results.

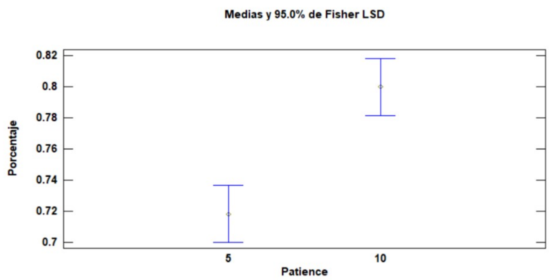

Fig. 22. Fisher LSD mean of the `Patience` variable for products.

This graph, Fig. 22, shows that the `patience` variable has a higher degree of confidence when its value is 10, as its representation is shifted to the right and separated from the other function.

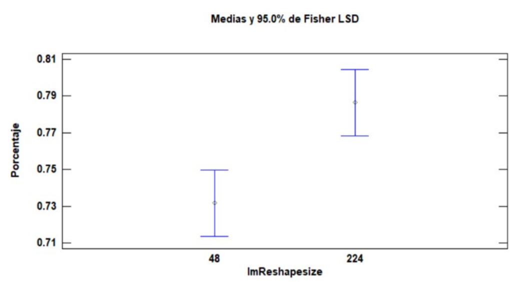

Fig. 23. Fisher LSD mean of the `ImReshapeSize` variable for products.

The previous graph, Fig. 23, shows that the `ImReshapeSize` variable has an effect on the experiment, with an image size of 224 being the most optimal.

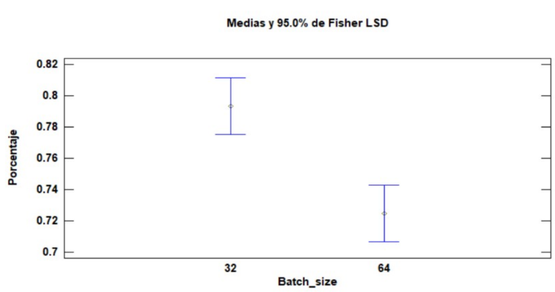

Fig. 24. Fisher LSD mean of the `Batch Size` variable for products.

Fig. 24 shows that the `BatchSize` variable has an effect on the experiment, with a value of 32 being the most optimal. This is because it is separated from the other value and is the first to represent the best option. The model uses a `Batchsize` of 32 and 64. With the applied statistical analysis, it is concluded that the value of 32 influences the model's performance more, particularly in the *Low-Beta*() and *Low-Gamma ()* frequency bands, while the value of 64 decreases the model's performance. As shown in Fig. 24, the value of 32 is above 64.

2) Faces Experiment: For this experiment, the same procedure is carried out as in the previous case, with the same variables and hypotheses.

TABLE IV. VARIANCE ANALYSIS F1 SCORE – TYPE III SUM OF SQUARES FOR FACES

Source	P-Value
MAIN EFFECTS	
A: Activation Fuction	0.0042
B: Patience	0.0019
C: Optimizer	0.4491
D: Image Size	0.8359
E: Batch Size	0.0012

In previous table, Table IV, the results of the analysis of variance are shown, through a Type III sum of squares, the results show that both hypothesis C and D do not contribute to the result, so they are discarded.

$$H_0: \quad C_i = 0$$
$$H_0: \quad D_i = 0$$

Once the previous hypotheses are eliminated, the experiment is recalculated, where the following result is obtained.

TABLE V. VARIANCE ANALYSIS F1 SCORE – TYPE III SUM OF SQUARES OF SIGNIFICANT VARIABLES FOR FACES.

Source	P-Value
MAIN EFFCETS	
A: Activation Fuction	0.0038
B: Patience	0.0017
C: Batch Size	0.0010

As seen in the previous table, Table V, the remaining hypotheses do influence the experiment, making them important. Once the results are obtained through the comparison of means, we can observe which value effectively impacts the model's effectiveness percentage, leading to the following outcomes.

The graph, Fig. 25,indicates that the ReLU `activation function` has a higher degree of confidence, as its representation is shifted to the right and separated from the other function.

The graph, Fig. 26, shows that the value of the `patience` variable has a higher degree of confidence when set to 10. This is because its representation is shifted to the right and separated from the other function.

The graph, shows that the `Batch Size` variable has an effect on the experiment, with a value of 32 being the most optimal, Fig. 27,this is because it is separated from the other value and is the first to represent the best option.

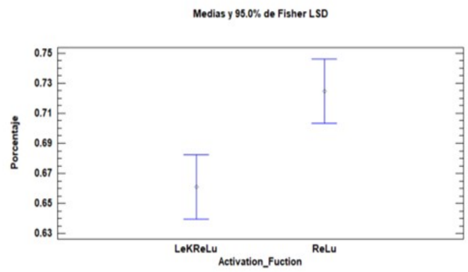

Fig. 25. Fisher LSD mean of the `Activation function` variable for faces.

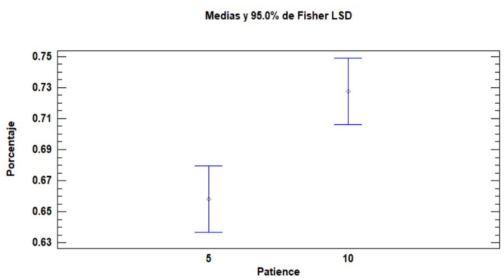

Fig. 26. Fisher LSD mean of the `Patience` variable for faces.

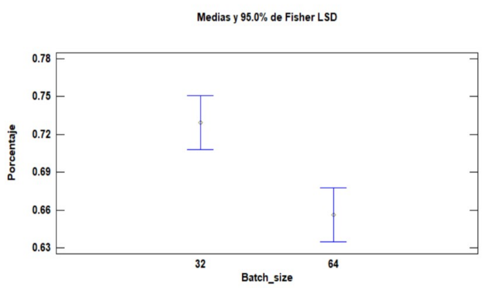

Fig. 27. Fisher LSD mean of the `Batch size` variable for faces.

E. Discussion

The two experiments have different training methods. One tests Products consumed by the subject, and the other tests the subject's Face. Both experiments use the same independent variables, but vary in the dependent variable. Using Statgraphics software, it was determined which variables influence each experiment. The following table, Table VI,shows these variables and their highest confidence values.

TABLE VI. VALUES OF THE VARIABLES FOR EACH EXPERIMENT

Products	Value	Faces	Value
Patience	10	Patience	10
ImReshapesize	224	Activation_Fuction	ReLu
Batch_size	32	Batch_size	32

The table indicates that both experiments share two variables with the same values: `Patience` (10) and `Batch_size` (32). `Patience` determines how many epochs the model can worsen before stopping, preventing overfitting. `Batch_size` is the number of training examples used by the network to update its weights, affecting the training process. In both cases, these values are essential for achieving 95% effectiveness.

Each experiment has a unique variable. In the Products experiment, `ImReshapesize` refers to the size of the image (224) input into the network. In the Faces experiment, `Activation_Fuction` is the mathematical function that determines the output of a neuron, with ReLU being the function used, which returns zero for negative values.

With the adjustment of hyperparameters such as the `ReLU` activation function, the `ADAM` optimizer, a `batchsize` of 32, and a `patience` of 10, while increasing the image size to 224, the model's accuracy was optimized. Additionally, the *Kolmogorov* complexity was reduced, which refers to the concept of algorithmic complexity that measures the amount of information required to describe or generate a dataset [18]. In other words, it is the length of the code used in the shortest program that can produce a given sequence of data as output, streamlining the model and decreasing its computational process [19]. By adjusting the hyperparameters and simplifying the data, the neural network was able to learn more efficiently, reducing the amount of information and the number of input data, thus becoming more agile in analyzing the image and EEG signals.

VI. CONCLUSIONS

In this work, 13 related studies were found concerning the use of decision-making, principal component analysis (PCA), and convolutional neural networks, of which 4 studies were selected for in-depth analysis. These studies were chosen for their connection to the use of convolutional neural networks and the methods they employ, with the main differences being in the decision-making aspect and target audience.

Furthermore, the theoretical foundations of the project were defined, integrating various convolutional neural networks used for different tasks, as well as knowledge about EEG, and lastly, techniques for more optimal data analysis were studied.

Additionally, the processes and subprocesses necessary for developing the decision-making model of whether I like it or not were presented. The three main processes were: i) Data collection through EEG and photographs, ii) Training the neural network with the collected data, and iii) Model functioning.

To verify the model's functionality, a series of experiments were designed with the interaction of 5 variables: a) ImReashepesize, b) Optimizer, c) Patience, d) Activation function, e) Batch size; experimentation was repeated with products and faces. The generated values for F1 Score and their combinations were examined through a Multifactorial Analysis of Variance. This analysis highlights that the two main variables influencing this model are patience and batch size, making the model more efficient.

Moreover, regarding the project's profitability, it is concluded that it is rejected, as the investment made for this project is not recovered, indicating that when not considering inflation, profitability decreases.

Therefore, this work contributes to the design of a neural network focused on decision-making in young people aged 18 to 20, contributing to the food and health sector by indicating whether a person likes what they consume, thus preventing overweight among these individuals.

As future work or derivations of this proposal, the following should be considered:

- Expand the target audience to include older adults and children of different ages.
- Improve the database by increasing its size with the participation of specialists.
- Design a new series of experiments that take into account other variables, such as Dropout, Epochs, etc.
- Implement the model in an IoT (Internet of Things) system.
- Combining Fp1, Fp2, O1, and O2 to create a brain mapping or qEEG analysis, which is an assessment tool used to measure electrical activity in the cerebral cortex. This map is then used to help diagnose mental health conditions by providing a statistical means of evaluating electrical activity in the cortex, [20].
- Expand the context of assistance to diseases such as diabetes or hypertension.

ACKNOWLEDGMENT

The Instituto Politécnico Nacional of Mexico, through the Comisión de Operación y Fomento de Actividades Académicas (COFAA) and SIP-projects No. *20241623* and *20241762*, has provided funding for this work. The research was carried out at the Escuela Superior de Ingeniera Mecánica y Eléctrica, Campus Zacatenco. It should be mentioned that this study is a crucial component of the doctoral dissertation titled *Modelo integral de Neuromarketing para innovar en las empresas del sector alimentario* supported by *Verónica Pérez*, under the supervision of Dr. Ana Coria and Dr. Jaime Moreno. Moreover, this research is also part of the master's thesis titled *Modelo sistémico para explicar el comportamiento de los consumidores durante la toma de decisiones, basado en inteligencia artificial*, guided by Dr. Oswaldo Morales and Dr. Jaime Moreno, and supported by Mauro Castillo.

REFERENCES

[1] C. Oropeza Abúndez, Ed., *Encuesta Nacional de Salud y Nutrición 2018-19: Resultados Nacionales*, primera edición ed. Cuernavaca, Morelos, México: Instituto Nacional de Salud Pública, 2020.

[2] O. C. E. Posible. 5 enganos alimentarios. ¡que no te la cuelen! Accedido el 20 de noviembre de 2024. [Online]. Available: https://www.otroconsumoposible.es/5-enganos-alimentarios-que-no-te-cuelen/

[3] C. Mitchell, "Ops-oms: Sobrepeso afecta a casi la mitad de la población de todos los países de américa latina y el caribe salvo por haití," *Journal of Embedded Systems*, vol. 15, no. 4, pp. 123–145, 2017.

[4] CASADOMO. El mit diseña un sistema de red neuronal de aprendizaje profundo para dispositivos iot. Accedido el 20 de noviembre de 2024. [Online]. Available: https://www.casadomo.com/2020/11/20/mit-disena-sistema-red-neuronal-aprendizaje-profundo-dispositivos-iot

[5] E. Niedermeyer and F. L. da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, 2004.

[6] T. H. Lee and K. S. Lee, "A new method for analyzing the decision-making process using eeg signals," *Neuroscience Letters*, vol. 639, pp. 18–24, 2017.

[7] L. I. Aftanas and S. A. Golocheikine, "Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and were related to self-regulation of emotion," *Neuroscience Letters*, vol. 310, no. 1, pp. 57–60, 2001.

[8] Y. Narayan, M. A. Tripathi, P. P. Singh, D. A. Vidhate, R. Singh, and M. M. S. Rao, "Development of ai model to analyze the customer behavior by decision making system," in *2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS)*, vol. 1, 2023, pp. 1–5.

[9] M. Maram, M. A. Khalil, and K. George, "Analysis of consumer coffee brand preferences using brain-computer interface and deep learning," in *2023 IEEE 7th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 2023, pp. 227–232.

[10] S. M. Usman, S. M. Ali Shah, O. C. Edo, and J. Emakhu, "A deep learning model for classification of eeg signals for neuromarketing," in *2023 International Conference on IT Innovation and Knowledge Discovery (ITIKD)*, 2023, pp. 1–6.

[11] V. Khurana, M. Gahalawat, P. Kumar, P. P. Roy, D. P. Dogra, E. Scheme, and M. Soleymani, "A survey on neuromarketing using eeg signals," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 13, no. 4, pp. 732–749, 2021.

[12] M. Ivanova and M. H. Ruiz, "Motor learning and decision making in volatile environment in bipolar disorder," in *2022 Fourth International Conference Neurotechnologies and Neurointerfaces (CNN)*, Kaliningrad, Russian Federation, 2022, pp. 55–58.

[13] L. Tran, B. Ngo, T. Tran, L. Pham, and A. Mai, "On a development of sparse pca method for face recognition problem," in *2021 International Conference on Advanced Technologies for Communications (ATC)*, Ho Chi Minh City, Vietnam, 2021, pp. 265–269.

[14] O. P. Barus, J. Happy, Jusin, J. J. Pangaribuan, S. Z. H, and F. Nadjar, "Liver disease prediction using support vector machine and logistic regression model with combination of pca and smote," in *2022 1st International Conference on Technology Innovation and Its Applications (ICTIIA)*, Tangerang, Indonesia, 2022, pp. 1–6.

[15] Y. Zhu and Q. Zeng, "Continuous speech recognition based on dcnn-lstm," in *2023 5th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP)*, Chengdu, China, 2023, pp. 1247–1250.

[16] I. NeuroSky, "Neurosky eeg brainwave chip and board TGAM1," http://www.neurosky.com, San Jose, CA, USA, 2010, technical Specifications Document.

[17] D. G. Chávez, "Toma de decisiones guiada por emociones detectadas en señales de electroencefalograma mediante redes convolucionales," Ph.D. dissertation, Universidad Autónoma Metropolitana, 2023.

[18] H. Kabir and N. Garg, "Machine learning enabled orthogonal camera goniometry for accurate and robust contact angle measurements," *Scientific Reports*, vol. 13, p. 1497, 2023. [Online]. Available: https://doi.org/10.1038/s41598-023-28763-1

[19] V. Bolón-Canedo and B. Remeseiro, "Feature selection in image analysis: a survey," *Artificial Intelligence Review*, vol. 53, pp. 2905–2931, 2020. [Online]. Available: https://doi.org/10.1007/s10462-019-09750-3

[20] M. Das and M. Mahadevappa, "Enhancing asd cognitive assessment with p300 classification using emd-based qeeg wavelet features," in *2023 7th International Conference on Computer Applications in Electrical Engineering-Recent Advances (CERA)*, 2023, pp. 1–6.