



# EEG Classification to Food Stimuli in Diverse Weight Groups with Regression Analysis of Eating Behavior Questionnaires

Halil İbrahim Coşar<sup>1\*</sup>, Ferhat Kılıç<sup>2</sup>, Cemil Altın<sup>1</sup>, Nermin Tanık<sup>3</sup>

<sup>1</sup> Department of Electrical and Electronics Engineering, Yozgat Bozok University, Yozgat 66900, Turkey

<sup>2</sup> Department of Mechatronics, Çorum Hitit University, Çorum 19169, Turkey

<sup>3</sup> Department of Neurology, Yozgat Bozok University, Yozgat 66900, Turkey

Corresponding Author Email: [halil.cosar@yobu.edu.tr](mailto:halil.cosar@yobu.edu.tr)

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.410402>

## ABSTRACT

**Received:** 21 February 2024

**Revised:** 22 March 2024

**Accepted:** 11 June 2024

**Available online:** 31 August 2024

### Keywords:

deep learning (DL),  
electroencephalography (EEG)  
classification, overweight, supervised  
tabular meta-learning (SuperTML),  
machine learning (ML), regression

Obesity and overweight are well-documented risk factors for numerous diseases that negatively impact life expectancy and quality of life, including cardiovascular diseases, diabetes, and cancer. Although the effects of weight status on brain function have been extensively studied, the application of machine learning (ML) and deep learning (DL) techniques in this domain remains underexplored. This study aims to address this gap by creating a unique dataset comprising electroencephalography (EEG) data from 19 channels, recorded while participants with varying body mass indices were exposed to visual food cues. The primary objective was to classify the differences in brain signals between normal-weight and overweight/obese individuals using advanced DL methods. To mitigate overfitting and data imbalance, tabular data augmentation was employed. Additionally, the Supervised Tabular Meta-Learning (SuperTML) method was utilized to embed EEG features into images, marking a novel application for this type of data. Classification results indicate that DenseNet-121 achieved the highest accuracy, with a rate of 0.97 at channel T4. Regionally, the temporal area yielded the best average accuracy rates. Furthermore, the study investigated the correlation between EEG data and eating behavior through regression analysis, applying Random Forest, eXtreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and Voting ensemble regression models to the participants' questionnaire responses. A significant relationship between EEG data and the questionnaires was identified, with the LightGBM regressor achieving an  $R^2$  value of 0.966. These findings demonstrate superior performance compared to existing literature in several aspects. This study underscores the potential of DL in enhancing our understanding of the neural mechanisms underlying eating behaviors in individuals with different body weights and provides a robust methodological framework for future research in this field.

## 1. INTRODUCTION

Overweight and obesity are two of the leading causes of health issues influencing individuals of all ages. In 2016, the WHO reported that the number of obese and overweight individuals exceeded 1.9 billion worldwide, with 650 million adults, 340 million adolescents, and 39 million children being obese. In addition, this figure continues to rise [1]. Obesity is an established risk factor for metabolic disorders such as diabetes, cardiovascular disease, cancer, and sleep apnea. Nonetheless, obesity frequently co-occurs with mental illness, particularly affective disorders such as depression, and imposes a massive disease burden owing to the decline in quality of life and social functioning it causes [2, 3]. The Body Mass Index (BMI) is the most commonly used conventional method to measure obesity. However, calculating body fat percentage exclusively using BMI is difficult because of its unreliability, particularly when assessing various age groups. Bioelectric impedance analyzer (BIA), bio-impedance spectroscopy, dilution technique, hydrostatic weighing

method, air displacement plethysmography (ADP), dual energy x-ray absorptiometry (DEXA), ultrasound, CT and MRI modalities are the other most commonly utilized techniques for evaluating body structure in obese subjects [4, 5].

EEG is a non-invasive technique that measures the electrical activity generated by the nerve cells of the cerebral cortex using head-mounted electrodes. EEG signals are especially beneficial for studying how the human brain processes information due to their high temporal resolution [6]. EEG signals are increasingly being combined with ML and DL techniques to perform a variety of projects. Some applications of EEG signals include the following: emotion recognition [7, 8], mental disorders [9, 10], sleep stage scoring [11], apnea detection [12], motor imagery [13], event related potential (ERP) [14, 15].

An online search for food addiction yields millions of search results, indicating a perceived link between food consumption and addiction in everyday language [16]. There are clear behavioral parallels between obesity and traditional substance

abuse. Both involve the persistent, uncontrollable consumption of a substance that has immediate rewarding effects but long-term detrimental effects on the individual's physical and psychosocial well-being [17]. Several research papers in the scientific literature examine the EEG signals of individuals with food-related stimuli; diverse situations were examined as stimuli while these studies were conducted. Some of these include reactions to food photographs, reactions to the aroma of food, reactions based on food sampling, and written food-related phrases. However, eating behavior questionnaires and regression analysis were not used in these studies.

In their study of 66 women (26 overweight or obese, 40 normal), Nijs et al. [18] studied the reactions of individuals to visual food stimuli in the states of hunger and satiety. The study found that in the deprivation condition, P300-related attentional bias to food was significantly higher in normal-weight individuals compared to overweight/obese individuals, and that in the satiety condition, it was more likely to be higher in overweight/obese women than in normal-weight women [18]. The impact of food odours and visuals on 62 participants in three different weight categories (normal, overweight, and obese) was examined in the study by Zsoldos et al. [19]. As a result, it was found that weight status influences the brain activity that supports the unconscious processing of food signals, starting with the first observable P100 peak [19]. Hume et al. studied the EEG responses of normal, overweight, and obese women to food-related visual stimuli. Once more, this study's findings showed that different measurements were made in various weight groups for the P300 and P200 values [20]. In a different study, two groups of 28 participants were formed, with 14 of them being overweight or obese and 14 being normal. The EEG responses of these two groups to images of high-calorie and low-calorie meals were then compared. The processing of foods with fewer calories eventually separated the two weight classes on assessments of neurological activity and was linked to snack food consumption in the overweight/obese group, despite the fact that every individual showed an attentional bias towards calorie-rich foods [21]. The EEG response of 20 participants to 188 distinct food images was examined in the study by Zhao and Xu done in 2019 [22]. As a result, within the same feeling, the EEG data gathered can be distinguished by mood and frequency range. Additionally, the variance in the size of the brain signal within the same frequency band is shown by the difference in standard deviation, which suggests that the band of signals needs to be broken down in order to get characteristics [22]. Nijs et al. [23] evaluated the EEG responses of 40 participants (20 obese, 20 normal weight) to written food-related sentences in another study. Therefore, it was found that during the instinctual stage of information processing, obese participants tended to pay more attention to food-related stimuli than neutral ones. There was a general food-related bias observed in reaction times and P300 scores, which are supposedly indicators of more aware, sustained attention, but there were no significant group differences. Food yearning scores and P200/P300 amplitude biases exhibited a positive link in the normal-weight group; however, in the obese group, there were substantial positive relationships between food-related reaction time bias, food want, and external consumption [23]. In the 2021 investigation by Woltering et al. [24] on 40 participants (18 healthy, 22 obese), EEG recordings were made by exposing the participants to food stimuli. The inverse relationship between BMI and P3 readings was highlighted as a conclusion [24]. In

the study done by Ammar Ali et al. [25], it was determined how a long-term diet affected the EEG responses of overweight and obese people to food-related stimuli, and it was also discovered how closely healthy people's event-related potential data were affected by the diet of the overweight and obese group. Additionally, ANN and SVM were used to categorize the groups, and the results of SVM and ANN differ when the control group and the overweight/obese group after the diet are compared. The ANN has an accuracy of 92,95% for nonfood photos compared to the SVM's 85,45% for food images [25]. In a different study, 24 male participants underwent resting-state functional magnetic resonance imaging (rs-fMRI) to record their hunger and satiety states. These images were then categorized using ML techniques. As a result, their classification accuracy using the support vector machine technique was 81% [26]. Snehalatha and Sangamithirai [5] tried to distinguish between obese and normal instances in thermal pictures. The study built a bespoke DL network and evaluated how well it performed in comparison to the most advanced pre-trained CNN and ML models. In comparison to the pre-trained VGG16 net model, which had an accuracy of 79% and an AUC value of 0.90, the proposed custom CNN network-2 achieved an overall accuracy of 92% and an AUC value of 0.948 [5]. In order to create a ML model for identifying young people at risk of becoming overweight or obese, Singh and Tawfik used data from the UK's Millennium Cohort Study. The Synthetic Minority Oversampling Technique (SMOTE) was used in the study to overcome concerns with low prediction accuracy brought on by data imbalance. To benefit from the prediction accuracy of individual classifier algorithms, an ensemble of classifiers was deployed. The outcomes were positive, with the target class achieving a prediction accuracy of over 90% [27]. In other work using alpha band functional connectivity features generated from EEG data, the study created a unique ML model to identify the brain networks of obese females, attaining an overall classification accuracy of 0.937. According to the research, the obese brain exhibits a disordered network with weak regions in charge of handling information about oneself and the surroundings [28].

The impact of weight status on human brains has been the subject of numerous studies; however, ML and DL algorithms have not been widely used. Within the scope of this study, it is aimed to evaluate the detection of obese and overweight individuals, which are rapidly increasing worldwide, not only by looking at the BMI, but also by examining the EEG signals of the person and comparing them with different eating behavior questionnaires. Collecting unique data is important for the authenticity of the study. In addition, the ML methods that are used for the evaluation of the survey results make this study important since they have not been applied to the surveys conducted in previous studies, which shows the novel contribution of this study. Some of the methods, such as Super TML, are applied for the first time to the EEG dataset, and successful results are obtained. The following is a list of the study's contributions:

- A unique EEG dataset is created within the scope of the study.
- Data augmentation is applied effectively to address overfitting and imbalanced data issues.
- The applicability of the Super TML technique to EEG data by converting table data to images is demonstrated.

- A CNN-based classification is implemented to obtain EEG data using ResNet152-v2, ResNet50-v2, Inception-v3, and DenseNet-121 algorithms.
- The methods used show excellent performance in certain channels up to 0.97 accuracy rate, which is superior to studies in the literature according to the authors knowledge.
- When conducting eating behavior surveys, a strong correlation between EEG data and questionnaire results is shown using regression methods (Random Forest, XGBoost, LightGBM, and Voting Ensemble Regression) with an  $R^2$  value of 0.966. This analysis has not been seen before in the literature.
- It has been discovered which part of the brain is more decisive when you are exposed to visual stimuli in different weight groups.

## 2. MATERIALS AND METHODS

### 2.1 Participant

In this research, EEG data were acquired from 20 male volunteers between the ages of 21 and 41 with a BMI ranging from 21.45 to 39.43kg/m<sup>2</sup> from Yozgat Bozok University, Türkiye. The characteristics of the participants are displayed in Table 1. Participants were excluded from the study if they disclosed having a mental, neurological, or somatic condition or taking any medication within the previous month that could have influenced their eating patterns, body weight, or EEG activity. Each participant conducted three or four trials, with a ten-minute break between each session, for a total of seventy trials.

**Table 1.** Details of participants

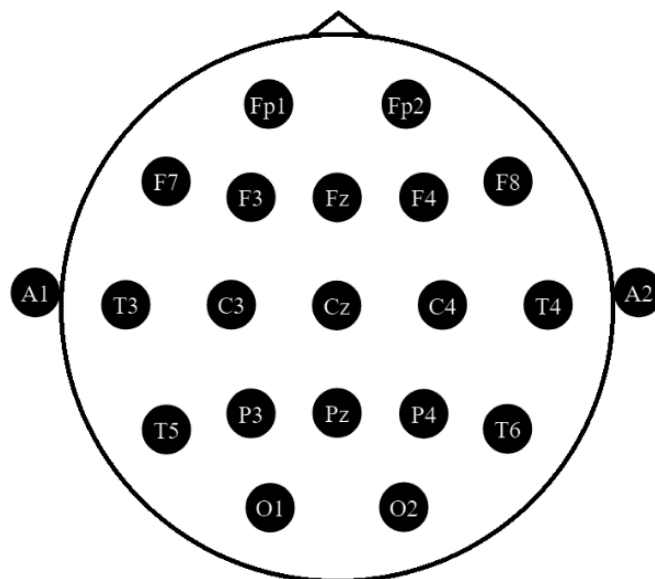
Age	Weight	Height	BMI
35	79	175	25.80
28	66	170	22.84
21	62	170	21.45
22	110	180	33.95
31	72	174	23.78
22	105	193	28.19
35	85	170	29.41
41	73	181	22.28
27	130	186	37.58
23	76	182	22.94
35	118	173	39.43
22	70	170	24.22
26	75	169	26.26
38	105	182	31.70
37	80	176	25.83
33	87	188	24.62
35	90	170	31.14
25	77	183	22.99
35	82	174	27.08
26	86	175	28.08

### 2.2 Data acquisition and filtering

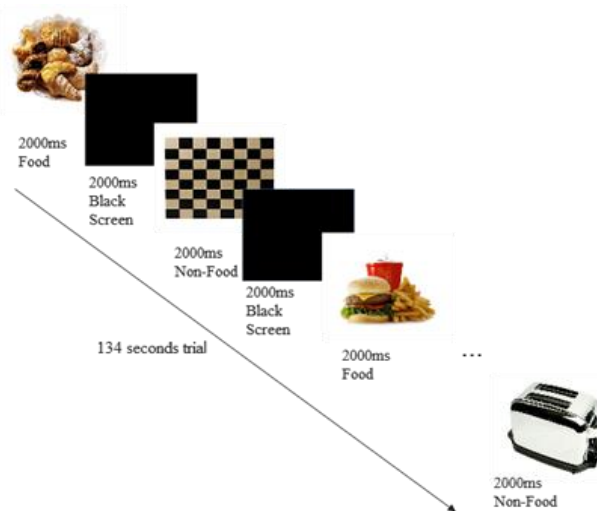
In this study, 19 channels were used with experimental parameters to record EEG signals over the scalp. Using the international 10-20 electrode placement technique, the electrodes were positioned non-invasively. A1 and A2's left-right earlobes were designated as the reference points, while the left eyebrow was designated as the ground. The EEG

electrodes for monopolar placement are listed in Figure 1. Each electrode positioning point is designated by a letter associated with the specific brain lobe or region it is recording from: frontal (F), temporal (T), parietal (P), occipital (O), and central (C). The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of Yozgat Bozok University (protocol code: 2017-KAEK-189\_2022.08.25\_08 and date of approval: 25/08/2022) for studies involving humans.

In this investigation, images from the FoodPics dataset were used as stimuli. In total, 34 photographs (17 food and 17 non-food) were chosen. All images in this dataset are colour photographs with a resolution of 600x450 pixels (96 dpi, sRGB colour format). In terms of viewing distance (up to 80 cm), angle, and fundamental figure-ground composition, images were selected/edited to be relatively uniform. Photos were standardised based on the colour of the background (white) [29]. During the experiment, each participant sat in a chair 50 cm from the computer monitor and was instructed to focus on the images. Each image was shown for two seconds, and an example trial is depicted in Figure 2.



**Figure 1.** Channel names of EEG electrodes and placement



**Figure 2.** A depiction of the experimental setup for the examination of visual stimuli

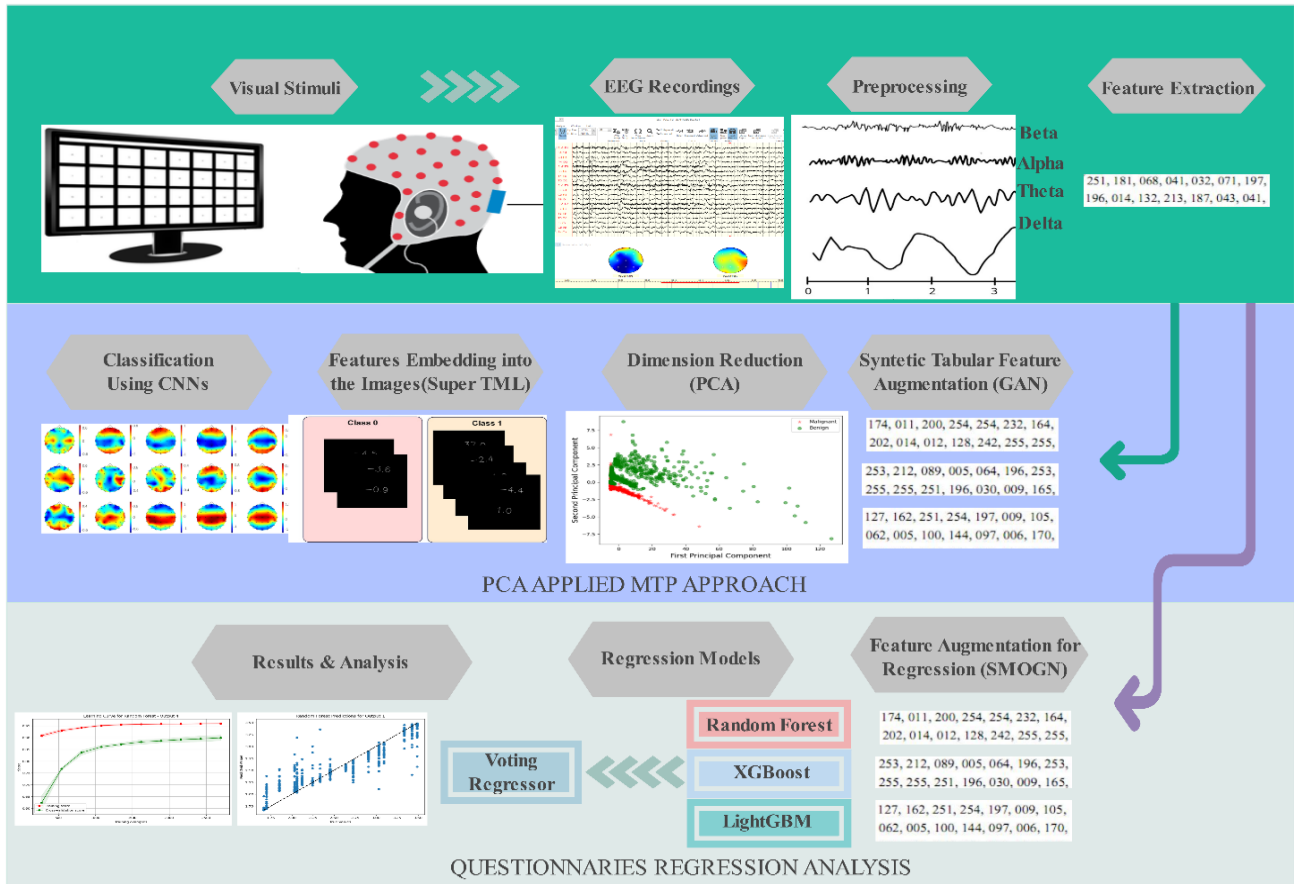


Figure 3. Workflow of the method used

During the research, EEG signals generated in response to visual stimuli were recorded using a Micromed SAM32RFO signal acquisition device and a Medcap electrode cap. The impedance of each channel was maintained below 10 kOhm, and each trial recording lasted 134 seconds. The sampling rate was set to 4096Hz, and the data was filtered at a frequency range of 0.03-70Hz. After the raw EEG data collection procedure was completed, the filtering step was applied. First of all, a 50Hz notch filter was applied to reduce power line noise. Then, a discrete wavelet transform was applied. In this step, a 10<sup>th</sup>-order Butterworth bandpass filter was used to filter the four sub-band frequency components: alpha, beta, theta, and delta. This step helps to remove artifacts, which can affect the proper classification results badly. Due to the superiority of the delta wave result, a classification frequency of 0.03-4Hz was chosen. After the filtering step, data normalization was applied. The data matrix was centered by making the mean value zero. Figure 3 illustrates the fundamental steps of this work. Multi-Task Process (MTP) approach can also be seen in the figure, which is explained in detail in the study [30]. In this method, feature extraction is followed by feature data augmentation. Then, the tabular data is converted into images and given to classification algorithms. In this work, Principle Component Analysis (PCA) is added as an extra step because the size of the data matrix is large, and using PCA, features that do not have much impact can be removed. Additionally, eating behavior questionnaires and EEG results were analyzed in order to find any relation between participant answers to questionnaires and EEG recordings. In this step, three different methods, which are Random Forest, XGBoost, and LightGBM, were used. Finally, using the combination of these methods, a voting regressor was applied.

### 2.3 Feature extraction

The feature extraction stage involves the transformation of unprocessed signal data into a feature vector in order to extract crucial information. This stage is capable of emphasizing the distinguishing characteristics of the signal. Using the Burg method, power spectral density (PSD) was utilized in this investigation for feature extraction. PSD is a valuable signal processing technique for stationary signals, and it is particularly effective for narrowband signals. It is a standard method of signal processing that distributes signal power across frequencies and displays signal energy as a function of frequency [31].

Burg's method generates an estimate of reflection coefficients by minimizing the mean of prediction errors in both forward and backward directions, while also adhering to the Levinson-Durbin recursion. This method resolves closely spaced, low-level sinusoids and estimates transient data records [32]. Following is an explanation of the Burg method:

$$P_{burg}(f) = \frac{\hat{E}_p}{|1 + \sum_{k=0}^p \hat{a}_p(k)e^{-2\pi f}|^2} \quad (1)$$

In this investigation, raw data was collected in segments of size  $(2*67*4096)*19$  for each trial, where 19 represents the number of channels, 4096 the sampling frequency, 67 the number of images, and 2 the duration of each image in seconds. After collecting raw data, responses to 2-second food stimuli for each participant were subtracted. 17 food images multiplied by 2 seconds equals 34 data length, yielding  $34*4096$  data matrices per trial. Finally, the method of feature



extraction was implemented. Using the Burg method, a 34\*129 data matrix was acquired, and for 70 trials, a 2380\*129 data matrix was created.

## 2.4 Tabular data augmentation

In the field of data science, datasets can be divided into structured and unstructured data. Tabular data, the most prevalent form of data in real-world applications, consists of a set of rows and columns containing features. Many difficulties arise when supplying them to deep neural networks, such as lack of locality, absent values, mixed feature types, and, unlike text or images, lack of prior knowledge of the data set's structure [33]. It can be said that there are several reasons for data augmentation [34], such as: limited training data [35], lack of relevant data [36], model overfitting [37] and imbalanced data [38].

There are a total of 2380 data columns in the dataset compiled for this study, of which 1462 represent individuals who are overweight/obese and 918 represent individuals who are normal. In order to eliminate imbalance and overfitting problems, data augmentation is applied. After synthetically increasing the data, the total number reached 2924 (1462 overweight/obese and 1462 normal).

## 2.5 Dimension reduction

PCA is a method for reducing the number of features. It calculates the variance within the data by creating a new set of perpendicular characteristics called principal components (PCs). PCA endeavors to eradicate duplicate dimensions from a given data space. This provides an easy method for ML algorithms to generate reduced data sets [39, 40]. In this investigation, feature-extracted data for each trial was reduced to 2924\*2 dimensions using the Burg method. Figure 4 illustrates the principal components extracted for the data set used in the study.

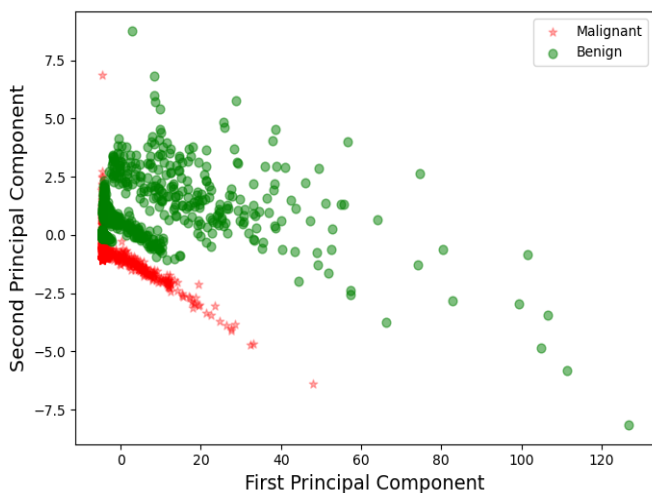


Figure 4. Principal component distribution of used data

## 2.6 Feature embedding into the images

In this stage, the SuperTML method was used in order to convert the tabular data into images. This method is motivated by the resemblance between the challenges encountered in TML and those in text classification tasks [41]. Using this method, all instances and their features were embedded in the

images, resulting in 2924 images. The dimension of the images was chosen as 224\*224 because of the deep model input size. Figure 5 shows the tabular data conversion to images and filing procedures.

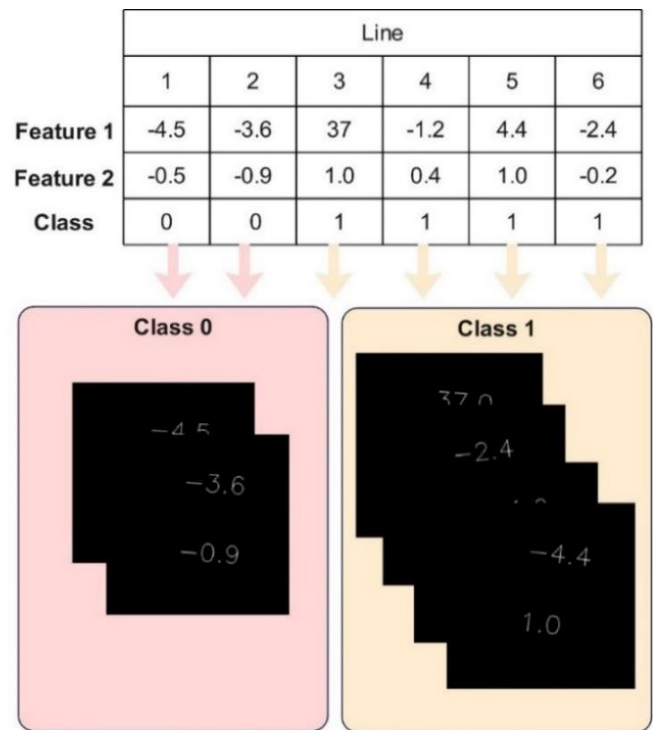


Figure 5. Feature conversion from tabular data to images

After all the pictures were saved in class folders, the data was split into train and test parts with a ratio of 80:20. Then, 30% of the train data is set aside for validation with the Keras image data generator pipeline. As a result, 2339 of 2924 data points were reserved for training, and 702 of them were selected for validation. For the test set, 585 pieces of data were reserved.

## 2.7 Classification using CNN models

CNNs have attained remarkable efficacy in image classification and detection. A CNN identifies an object by searching for basic characteristics, such as edges, lines, and curves, and then constructs additional features with an overall perspective. DL draws its inspiration from traditional neural networks but significantly outperforms its predecessors. Furthermore, DL utilizes transformations and graph technologies concurrently to construct multi-layer learning models [42].

Performance evaluation metrics are crucial evaluation criteria for the ML system used in biomedical signal processing. This study included four assessment metrics: accuracy, precision, recall, and F1-Score. The following are mathematical representations of assessment metrics:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (5)$$

Here, *TP* represents true positive decision, *TN* means true negative, *FP* stands for false positive and *FN* is false negative decision.

In this study, four different CNN models (DenseNet121, Inception-v3, ResNet50-v2 and ResNet152-v2) were used. These models are well known and widely used in the field of biomedical signal processing. One of our aims in using these models is to compare our results with established benchmarks in the literature, which provide a reference point for evaluating the effectiveness of our approach. It is important for the study that each of the models used has a different level of complexity and capacity. Here, ResNet-50 is a lighter model than ResNet-152 or DenseNet-121. Using these models, it is aimed at exploring how model complexity affects performance on a given task. Different architectures can capture different types of features from input images. For example, Inception-v3 is known for its ability to capture fine-grained details due to its use of multiple filter sizes. By using various models, different types of features were utilized, and higher classification results were sought. These models, which have been tested for robustness and generalization using different datasets and architectures, help to see the robustness of our findings on different architectures. These models have been used with high accuracy in various studies in the literature. They are also used both in the processing of the EEG signal and in the processing of the EEG signal converted into an image [43, 44]. The pre-trained CNN models were supplied with ImageNet weights for transfer learning, and the model's top property was set to false. Gaussian noise was added at a rate of 0.7, and dropout was chosen as 0.3 in order to avoid the model becoming overfitted. Additionally, 4096 neurons were added for fully connected layers. CNN models were trained with Keras over the course of 30 epochs. All convolutional layers were defroze, time-based decay was implemented, and the initial learning rate was set to 0.004. The rate of decay is proportional to the rate of learning, the epoch value, and momentum. Momentum applied 0.4 to the Stochastic Gradient Descent (SGD) optimizer, which tends to produce superior outcomes.

## 2.8 Application of eating behavior questionnaires

In the paper, two different eating behavior questionnaires were applied to the participants. These are Turkish versions of the Dutch Eating Behavior Questionnaires (DEBQ) and Three-Factor Eating Questionnaires (TFEQ). Validity and reliability of the Turkish versions of these questionnaires were proven in the researches [45, 46]. While the questionnaires were scored, the Likert scale was used and the results are given in Table 2.

**Table 2.** Questionnaire results

No	Dutch Eating Behavior				
	BMI	Mean	Restraint Eating	Emotional Eating	External Eating
1	25.80	2.45	2.80	1.77	3.00
2	22.84	2.06	1.80	1.54	3.00
3	21.45	2.73	1.60	2.92	3.60
4	33.95	3.03	2.20	3.69	3.00
5	23.78	2.24	2.20	1.54	3.20
6	28.19	3.09	2.20	2.92	4.20
7	29.41	1.70	1.40	1.23	2.60
8	22.28	1.67	1.90	1.08	2.20

9	37.58	3.42	1.80	3.62	4.80
10	22.94	3.48	2.50	3.77	4.10
11	39.43	2.12	1.20	2.15	3.00
12	24.22	2.33	1.60	2.00	3.50
13	26.26	2.33	2.30	1.54	3.40
14	31.70	1.85	2.20	1.08	2.50
15	25.83	3.33	4.00	2.23	4.10
16	24.62	2.06	1.80	1.15	3.50
17	31.14	2.45	2.80	1.23	3.70
18	22.99	2.64	1.70	2.15	4.20
19	27.08	2.91	3.20	2.00	3.80
20	28.08	2.03	1.90	1.08	3.40

No	BMI	Mean	Three-Factor Eating Behavior		
			Uncontrolled Eating	Cognitive Restraint	Emotional Eating
1	25.80	2.67	2.88	2.83	2.00
2	22.84	2.00	2.25	1.67	2.00
3	21.45	2.39	3.00	1.33	3.00
4	33.95	2.50	2.63	1.83	3.67
5	23.78	2.17	2.63	2.17	1.00
6	28.19	2.72	3.13	2.33	2.67
7	29.41	1.83	2.38	1.67	1.00
8	22.28	1.72	1.38	2.50	1.00
9	37.58	2.67	3.25	1.50	3.67
10	22.94	2.78	2.75	2.33	4.00
11	39.43	1.72	1.88	1.50	2.00
12	24.22	2.50	2.63	2.33	2.33
13	26.26	2.00	1.25	3.00	2.00
14	31.70	2.11	2.25	2.33	1.67
15	25.83	2.67	2.88	3.00	1.67
16	24.62	2.00	2.13	2.17	1.33
17	31.14	2.33	2.50	2.33	1.67
18	22.99	2.17	2.63	1.67	1.67
19	27.08	2.39	2.88	2.17	2.00
20	28.08	2.06	1.88	2.33	2.00

After questionnaires were scored, regression analyses were carried out in order to seek out the correlation between the EEG recording of participants and the questionnaire results. In this step, the features extracted from the EEG using the burg method were augmented utilizing the SMOGN algorithm. This technique can be used in order to eliminate imbalance problems in regression models and is explained in detail in the research [47]. A widely used technique for model selection is the method of cross-validation. Essentially, this involves partitioning the data into subsets. One portion is used for fitting each candidate model or method, and the remaining portion is used for evaluating their performance. The model exhibiting the most favorable overall performance is ultimately chosen [48]. In this study, 10-fold cross validation is applied. Then, Random Forest, XGBoost, and LightGBM techniques were applied for regression analysis using the best-scoring EEG channel, which is T4. Finally, the voting regressor, which is combination of these three methods, was applied. Random Forest, XGBoost, and LightGBM, all tree-based ensemble methods, are based on different algorithms and hyperparameters. Using a combination of these models can capture the various patterns and relationships present in the data. The aim of using an ensemble approach is to reduce the weaknesses of individual models and improve their overall performance. These ML-based algorithms can manage complex relationships between predictor variables and target variables, making them suitable for capturing fine details in survey data. Ensemble methods such as Random Forest and boosting algorithms (XGBoost and LightGBM) are less likely to suffer from overfitting than some other ML methods. This is important in regression tasks where the goal is to generalize well to unseen data, ensuring that the predictions of the model

created are reliable. Random Forest, XGBoost, and LightGBM have built-in mechanisms to evaluate feature importance. This feature can be valuable for interpreting the results of your regression analysis. Knowing which characteristics have the most significant impact on eating behavior may provide insight for further research. It is possible to improve the overall performance and robustness of regression analysis by combining the predictions of several models using a voting regressor. The ensemble reduces the risk of making decisions based on the weaknesses of a single model, while exploiting the strengths of each model. In addition, these models are widely used in regression analysis and give very promising results [49, 50].

### 3. RESULTS

#### 3.1 Classification results

For the purpose of determining the most informative EEG channels and brain regions, the unprocessed EEG signals from all 19 channels in 20 subjects have been collected. After preprocessing the collected data, the results were organized

according to five different brain regions. The first region is the frontal, which consists of channels Fp1, Fp2, F7, F3, FZ, F4, and F8. The second part is the temporal, which has 4 channels: T3, T4, T5, and T6. The third region has channels: C4, Cz, and C3, which denote the central region. While the parietal region consists of channels P4, Pz, and P3, the last region is occipital, which indicates O1 and O2.

In Table 3, Frontal lobe including 7 channels results are given. When the classification results are analyzed according to average accuracy, DenseNet-121 achieved the best performance with a rate of 0.859 among the others. On the other hand, ResNet50-v2 has the lowest average accuracy rate of 0.831.

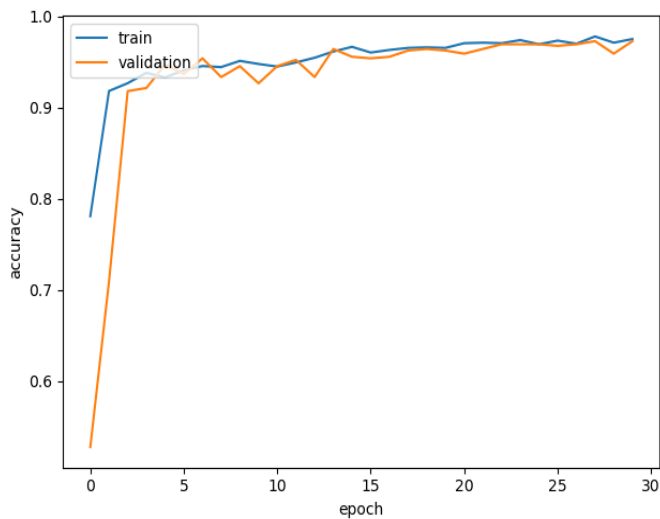
In terms of channel results, it is seen that F7 shows the least successful results, whereas F4 has the highest accuracy rate for all classification algorithms. Inception-v3 and DenseNet-121 reach a rate of 0.96 at this channel.

Figure 6 shows the accuracy, loss, and confusion matrix of the best classification algorithm for the frontal lobe after 30 epochs. Although Inception-v3 and DenseNet-121 have the same accuracy rate, the average accuracy of DenseNet-121 is higher than that of Inception-v3. Therefore, the DenseNet-121 results are illustrated in Figure 6.

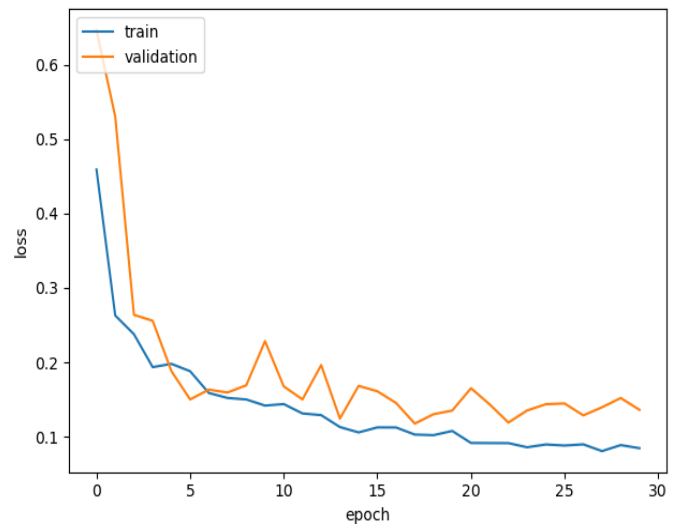
**Table 3.** Frontal region results for used algorithms

	Channel Fp1	Prec	Rec	F1-Score	Channel Fp2	Prec	Rec	F1-Score
<b>ResNet152-v2</b>	Channel Fp1	Prec	Rec	F1-Score	Channel Fp2	Prec	Rec	F1-Score
	Type0	0.98	0.86	0.91	Type0	0.88	0.88	0.88
	Type1	0.88	0.98	0.93	Type1	0.89	0.89	0.89
	Accuracy			0.92	Accuracy			0.88
	Macro avg	0.92	0.92	0.92	Macro avg	0.88	0.88	0.88
<b>ResNet50-v2</b>	Weighted avg	0.92	0.92	0.92	Weighted avg	0.88	0.88	0.88
	Type0	0.86	0.90	0.88	Type0	0.95	0.77	0.85
	Type1	0.90	0.86	0.88	Type1	0.80	0.96	0.87
	Accuracy			0.88	Accuracy			0.86
	Macro avg	0.88	0.88	0.88	Macro avg	0.87	0.87	0.86
<b>Inception-v3</b>	Weighted avg	0.88	0.88	0.88	Weighted avg	0.88	0.86	0.86
	Type0	0.88	0.93	0.90	Type0	0.89	0.89	0.89
	Type1	0.93	0.88	0.90	Type1	0.88	0.88	0.88
	Accuracy			0.90	Accuracy			0.88
	Macro avg	0.90	0.90	0.90	Macro avg	0.88	0.88	0.88
<b>DenseNet-121</b>	Weighted avg	0.90	0.90	0.90	Weighted avg	0.88	0.88	0.88
	Type0	0.94	0.91	0.92	Type0	0.89	0.89	0.89
	Type1	0.92	0.94	0.93	Type1	0.89	0.89	0.89
	Accuracy			0.93	Accuracy			0.89
	Macro avg	0.93	0.93	0.93	Macro avg	0.89	0.89	0.89
	Channel F7	Prec	Rec	F1-Score	Channel F3	Prec	Rec	F1-Score
<b>ResNet152-v2</b>	Type0	0.72	0.78	0.75	Type0	0.88	0.78	0.83
	Type1	0.78	0.70	0.73	Type1	0.83	0.89	0.86
	Accuracy			0.74	Accuracy			0.84
	Macro avg	0.74	0.74	0.74	Macro avg	0.85	0.84	0.84
	Weighted avg	0.75	0.74	0.74	Weighted avg	0.85	0.84	0.84
<b>ResNet50-v2</b>	Type0	0.70	0.86	0.77	Type0	0.95	0.64	0.76
	Type1	0.84	0.66	0.74	Type1	0.74	0.97	0.84
	Accuracy			0.75	Accuracy			0.81
	Macro avg	0.77	0.76	0.75	Macro avg	0.84	0.80	0.80
	Weighted avg	0.77	0.75	0.75	Weighted avg	0.84	0.81	0.80
<b>Inception-v3</b>	Type0	0.77	0.54	0.63	Type0	0.97	0.75	0.85
	Type1	0.67	0.85	0.75	Type1	0.81	0.97	0.89
	Accuracy			0.70	Accuracy			0.87
	Macro avg	0.72	0.70	0.69	Macro avg	0.89	0.86	0.87
	Weighted avg	0.72	0.70	0.69	Weighted avg	0.88	0.87	0.87
<b>DenseNet-121</b>	Type0	0.74	0.76	0.75	Type0	0.90	0.82	0.86
	Type1	0.78	0.76	0.77	Type1	0.84	0.91	0.88
	Accuracy			0.76	Accuracy			0.87
	Macro avg	0.76	0.76	0.76	Macro avg	0.87	0.86	0.87
	Weighted avg	0.76	0.76	0.76	Weighted avg	0.87	0.87	0.87

	Channel Fz	Prec	Rec	F1-Score	Channel F4	Prec	Rec	F1-Score
<b>ResNet152-v2</b>	Type0	0.81	0.83	0.82	Type0	0.98	0.92	0.95
	Type1	0.84	0.83	0.83	Type1	0.92	0.98	0.95
	Accuracy			0.83	Accuracy			0.95
	Macro avg	0.83	0.83	0.83	Macro avg	0.95	0.95	0.95
	Weighted avg	0.83	0.83	0.83	Weighted avg	0.95	0.95	0.95
<b>ResNet50-v2</b>	Type0	0.75	0.78	0.77	Type0	0.98	0.93	0.95
	Type1	0.79	0.76	0.78	Type1	0.93	0.98	0.95
	Accuracy			0.77	Accuracy			0.95
	Macro avg	0.77	0.77	0.77	Macro avg	0.95	0.95	0.95
	Weighted avg	0.77	0.77	0.77	Weighted avg	0.95	0.95	0.95
<b>Inception-v3</b>	Type0	0.82	0.78	0.80	Type0	0.99	0.93	0.96
	Type1	0.81	0.85	0.83	Type1	0.93	0.99	0.96
	Accuracy			0.81	Accuracy			0.96
	Macro avg	0.81	0.81	0.81	Macro avg	0.96	0.96	0.96
	Weighted avg	0.81	0.81	0.81	Weighted avg	0.96	0.96	0.96
<b>DenseNet-121</b>	Type0	0.79	0.79	0.79	Type0	0.98	0.94	0.96
	Type1	0.81	0.81	0.81	Type1	0.94	0.98	0.96
	Accuracy			0.80	Accuracy			0.96
	Macro avg	0.80	0.80	0.80	Macro avg	0.96	0.96	0.96
	Weighted avg	0.80	0.80	0.80	Weighted avg	0.96	0.96	0.96
	Channel F8	Prec	Rec	F1-Score	Average Accuracy			
<b>ResNet152-v2</b>	Type0	0.82	0.86	0.84	0.856			
	Type1	0.84	0.80	0.82				
	Accuracy			0.83				
	Macro avg	0.83	0.83	0.83				
	Weighted avg	0.83	0.83	0.83				
<b>ResNet50-v2</b>	Type0	0.78	0.86	0.82	0.831			
	Type1	0.83	0.74	0.78				
	Accuracy			0.80				
	Macro avg	0.80	0.80	0.80				
	Weighted avg	0.80	0.80	0.80				
<b>Inception-v3</b>	Type0	0.88	0.69	0.77	0.844			
	Type1	0.73	0.90	0.80				
	Accuracy			0.79				
	Macro avg	0.80	0.79	0.79				
	Weighted avg	0.81	0.79	0.79				
<b>DenseNet-121</b>	Type0	0.87	0.72	0.79	0.859			
	Type1	0.74	0.88	0.81				
	Accuracy			0.80				
	Macro avg	0.81	0.80	0.80				
	Weighted avg	0.81	0.80	0.80				

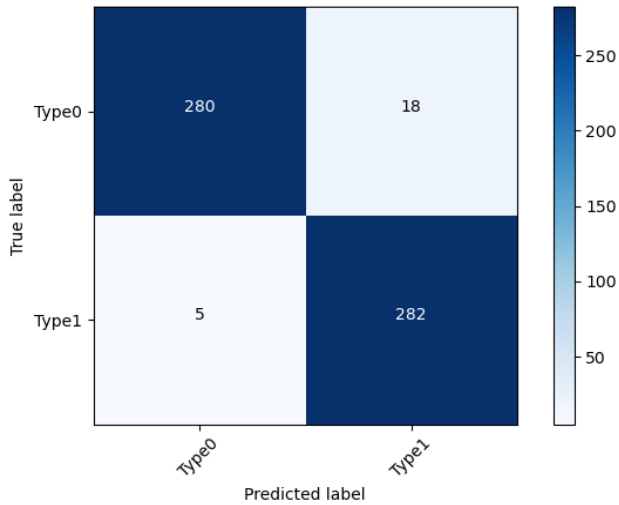


(a)



(b)





(c)

**Figure 6.** DenseNet-121 performance at channel F4: (a) Accuracy graph, (b) Loss graph and (c) Confusion matrix

Temporal lobe channel results are shown in Table 4. In terms of average accuracy rate, DenseNet-121 is the most successful algorithm, with a rate of 0.917. RestNet50-v2 has the lowest score of 0.897 for the temporal lobe, similar to the frontal lobe. When channel performances are analyzed, it can be seen that T4 has the highest score for this region. In addition, DenseNet-121 achieves the best classification performance of 0.97 at this channel. On the other hand, T3 rates are the lowest for the temporal region.

Figure 7 illustrates the best classification algorithm results for the temporal region at channel T4. Similar to the frontal lobe, DenseNet-121 has the highest accuracy score in this region, which is 0.97.

The results of the central region channels can be seen in Table 5. The best average accuracy rate is observed for DenseNet-121, with a rate of 0.827 in the central region, while Inception-v3 has the lowest score, which is 0.803.

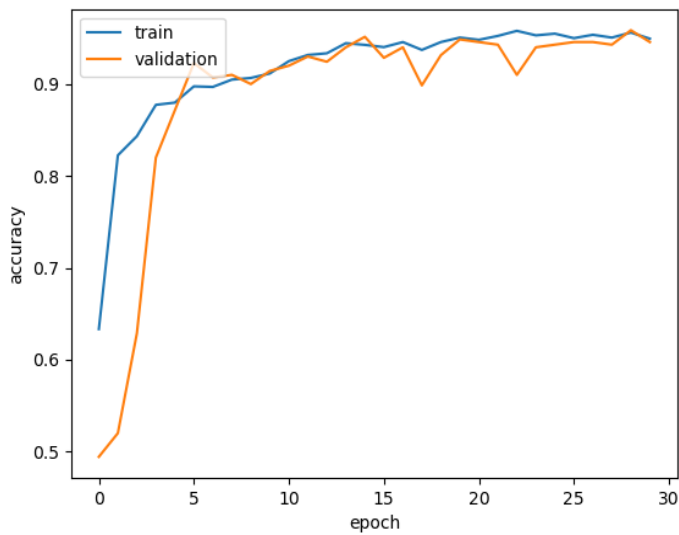
In contrast to the other regions, when the single channel results are analyzed, the highest classification results are obtained with 0.85 in ResNet50-v2, Inception-v3, and DenseNet-121. Additionally, channel C4 rates are higher than other channels in this region.

**Table 4.** Temporal region results for used algorithms

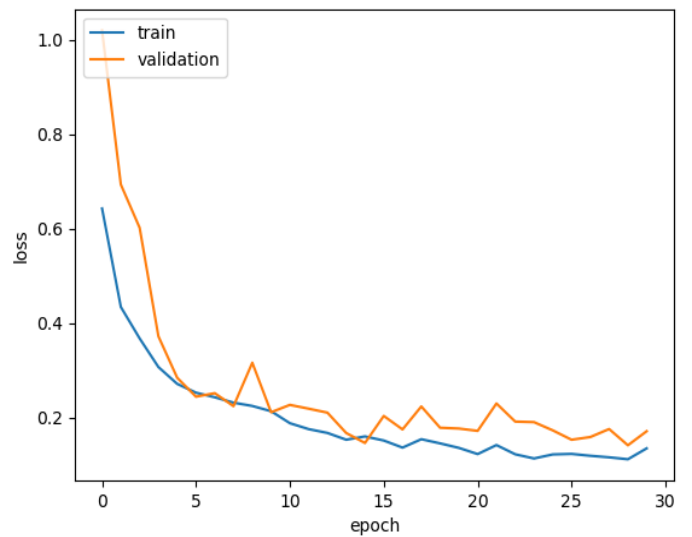
	Channel T3	Prec	Rec	F1-Score	Channel T4	Prec	Rec	F1-Score
<b>ResNet152-v2</b>	Type0	0.86	0.91	0.88	Type0	0.98	0.92	0.95
	Type1	0.91	0.86	0.88	Type1	0.91	0.98	0.95
	Accuracy			0.88	Accuracy			0.95
	Macro avg	0.88	0.88	0.88	Macro avg	0.95	0.95	0.95
	Weighted avg	0.88	0.88	0.88	Weighted avg	0.95	0.95	0.95
<b>ResNet50-v2</b>	Type0	0.92	0.82	0.87	Type0	0.98	0.90	0.93
	Type1	0.83	0.93	0.88	Type1	0.91	0.98	0.94
	Accuracy			0.87	Accuracy			0.94
	Macro avg	0.88	0.87	0.87	Macro avg	0.94	0.94	0.94
	Weighted avg	0.88	0.87	0.87	Weighted avg	0.94	0.94	0.94
<b>Inception-v3</b>	Type0	0.89	0.83	0.86	Type0	0.96	0.92	0.94
	Type1	0.84	0.89	0.86	Type1	0.91	0.95	0.93
	Accuracy			0.86	Accuracy			0.94
	Macro avg	0.86	0.86	0.86	Macro avg	0.93	0.94	0.94
	Weighted avg	0.86	0.86	0.86	Weighted avg	0.94	0.94	0.94
<b>DenseNet-121</b>	Type0	0.89	0.89	0.89	Type0	0.98	0.95	0.97
	Type1	0.89	0.89	0.89	Type1	0.95	0.98	0.96
	Accuracy			0.89	Accuracy			0.97
	Macro avg	0.89	0.89	0.89	Macro avg	0.97	0.97	0.97
	Weighted avg	0.89	0.89	0.89	Weighted avg	0.97	0.97	0.97
	Channel T5	Prec	Rec	F1-Score	Channel T6	Prec	Rec	F1-Score
<b>ResNet152-v2</b>	Type0	0.90	0.91	0.91	Type0	0.88	0.90	0.89
	Type1	0.91	0.91	0.91	Type1	0.90	0.89	0.90
	Accuracy			0.91	Accuracy			0.89
	Macro avg	0.91	0.91	0.91	Macro avg	0.89	0.90	0.89
	Weighted avg	0.91	0.91	0.91	Weighted avg	0.90	0.89	0.89
<b>ResNet50-v2</b>	Type0	0.94	0.91	0.92	Type0	0.90	0.77	0.83
	Type1	0.92	0.94	0.93	Type1	0.81	0.92	0.86
	Accuracy			0.93	Accuracy			0.85
	Macro avg	0.93	0.93	0.93	Macro avg	0.85	0.84	0.85
	Weighted avg	0.93	0.93	0.93	Weighted avg	0.85	0.85	0.85
<b>Inception-v3</b>	Type0	0.96	0.88	0.92	Type0	0.91	0.85	0.88
	Type1	0.88	0.97	0.92	Type1	0.86	0.91	0.88
	Accuracy			0.92	Accuracy			0.88
	Macro avg	0.92	0.92	0.92	Macro avg	0.88	0.88	0.88
	Weighted avg	0.92	0.92	0.92	Weighted avg	0.88	0.88	0.88
<b>DenseNet-121</b>	Type0	0.93	0.90	0.92	Type0	0.86	0.92	0.89
	Type1	0.91	0.94	0.92	Type1	0.92	0.86	0.89
	Accuracy			0.92	Accuracy			0.89
	Macro avg	0.92	0.92	0.92	Macro avg	0.89	0.89	0.89
	Weighted avg	0.92	0.92	0.92	Weighted avg	0.89	0.89	0.89
<b>Avr. Acc.</b>	<b>ResNet152-v2</b>			<b>ResNet50-v2</b>			<b>Inception-v3</b>	<b>DenseNet-121</b>
	0.907			0.897			0.900	0.917

**Table 5.** Central region results for used algorithms

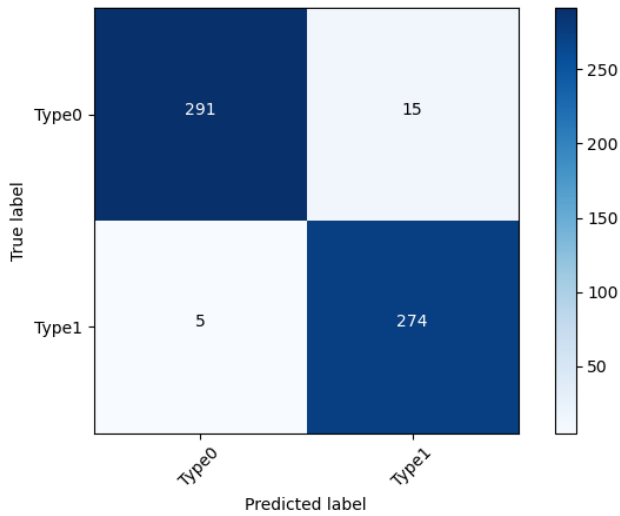
	Channel C3	Prec	Rec	F1-Score	Channel Cz	Prec	Rec	F1-Score
<b>ResNet152-v2</b>	Type0	0.82	0.79	0.81	Type0	0.86	0.68	0.76
	Type1	0.79	0.82	0.81	Type1	0.73	0.89	0.80
	Accuracy			0.81	Accuracy			0.78
	Macro avg	0.81	0.81	0.81	Macro avg	0.79	0.78	0.78
	Weighted avg	0.81	0.81	0.81	Weighted avg	0.80	0.78	0.78
<b>ResNet50-v2</b>	Type0	0.90	0.64	0.75	Type0	0.93	0.65	0.76
	Type1	0.71	0.92	0.81	Type1	0.72	0.95	0.82
	Accuracy			0.78	Accuracy			0.79
	Macro avg	0.81	0.78	0.78	Macro avg	0.82	0.80	0.79
	Weighted avg	0.81	0.78	0.78	Weighted avg	0.83	0.79	0.79
<b>Inception-v3</b>	Type0	0.88	0.77	0.82	Type0	0.80	0.62	0.70
	Type1	0.79	0.89	0.84	Type1	0.68	0.84	0.76
	Accuracy			0.83	Accuracy			0.73
	Macro avg	0.83	0.83	0.83	Macro avg	0.74	0.73	0.73
	Weighted avg	0.83	0.83	0.83	Weighted avg	0.74	0.73	0.73
<b>DenseNet-121</b>	Type0	0.90	0.77	0.83	Type0	0.92	0.65	0.76
	Type1	0.79	0.91	0.85	Type1	0.72	0.94	0.82
	Accuracy			0.84	Accuracy			0.79
	Macro avg	0.85	0.84	0.84	Macro avg	0.82	0.79	0.79
	Weighted avg	0.85	0.84	0.84	Weighted avg	0.82	0.79	0.79
	Channel C4	Prec	Rec	F1-Score	Avr.Acc.			
<b>ResNet152-v2</b>	Type0	0.89	0.76	0.82				
	Type1	0.81	0.92	0.86				
	Accuracy			0.84	0.810			
	Macro avg	0.85	0.84	0.84				
	Weighted avg	0.85	0.84	0.84				
<b>ResNet50-v2</b>	Type0	0.93	0.73	0.82				
	Type1	0.80	0.95	0.87				
	Accuracy			0.85	0.807			
	Macro avg	0.87	0.84	0.84				
	Weighted avg	0.86	0.85	0.85				
<b>Inception-v3</b>	Type0	0.97	0.72	0.83				
	Type1	0.77	0.97	0.86				
	Accuracy			0.85	0.803			
	Macro avg	0.87	0.85	0.84				
	Weighted avg	0.87	0.85	0.84				
<b>DenseNet-121</b>	Type0	0.88	0.79	0.83				
	Type1	0.83	0.91	0.87				
	Accuracy			0.85	0.827			
	Macro avg	0.86	0.85	0.85				
	Weighted avg	0.86	0.85	0.85				



(a)

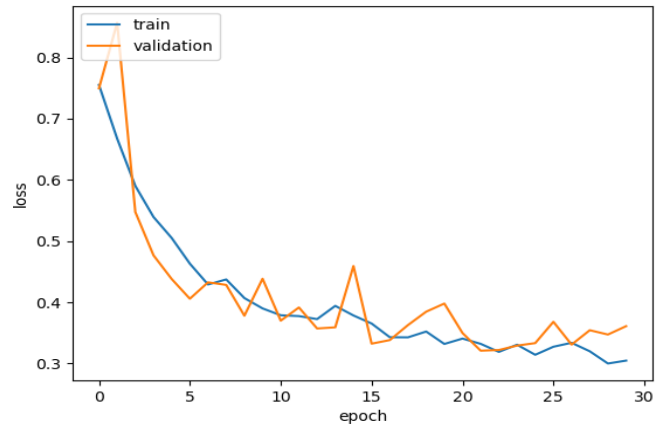


(b)



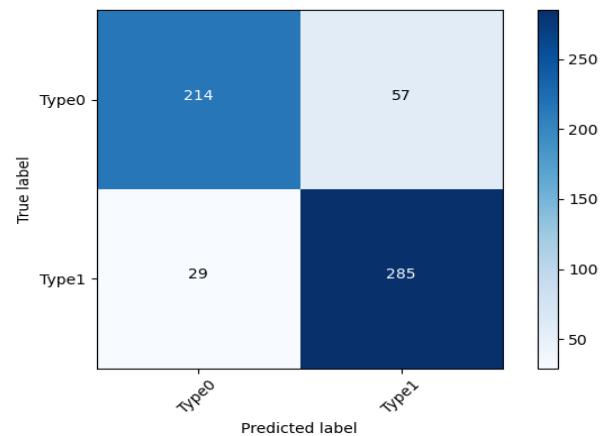
(c)

**Figure 7.** DenseNet-121 performance at channel T4: (a) Accuracy graph, (b) Loss graph and (c) Confusion matrix



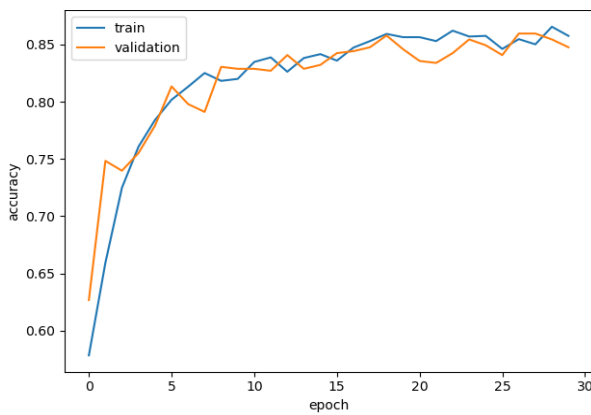
(b)

Although ResNet50-v2, Inception-v3, and DenseNet-121 have the same accuracy rate, which is 0.85, the average accuracy of DenseNet-121 is higher than the others. Therefore, the accuracy, loss, and confusion matrix graphs of DenseNet-121 are illustrated in Figure 8.



(c)

**Figure 8.** DenseNet-121 performance at channel C4: (a) Accuracy graph, (b) Loss graph and (c) Confusion matrix



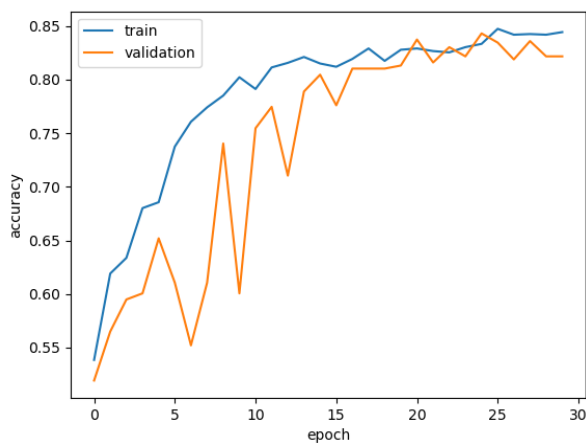
(a)

In Table 6, the classification results of the parietal lobe are shown. In terms of average accuracy rate, Resnet152-v2 and DenseNet-121 have the highest score with a rate of 0.837, while ResNet50-v2 shows the lowest rate of 0.810. Additionally, according to single channel results, Inception-v3 is the most successful algorithm, reaching a 0.86 accuracy rate at channel P3.

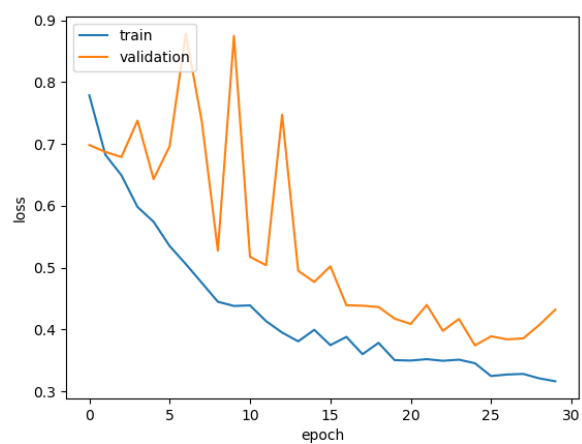
**Table 6.** Parietal region results for used algorithms

	Channel C3	Prec	Rec	F1-Score	Channel Cz	Prec	Rec	F1-Score
<b>ResNet152-v2</b>	Channel P3	Prec	Rec	F1-Score	Channel Pz	Prec	Rec	F1-Score
	Type0	0.94	0.74	0.81	Type0	0.78	0.85	0.81
	Type1	0.80	0.98	0.88	Type1	0.85	0.78	0.81
	Accuracy			0.85	Accuracy			0.81
	Macro avg	0.89	0.84	0.84	Macro avg	0.81	0.81	0.81
<b>ResNet50-v2</b>	Weighted avg	0.88	0.85	0.84	Weighted avg	0.81	0.81	0.81
	Type0	0.85	0.80	0.82	Type0	0.71	0.83	0.76
	Type1	0.82	0.87	0.85	Type1	0.81	0.68	0.74
	Accuracy			0.84	Accuracy			0.75
	Macro avg	0.84	0.83	0.84	Macro avg	0.76	0.76	0.75
<b>Inception-v3</b>	Weighted avg	0.84	0.84	0.84	Weighted avg	0.76	0.75	0.75
	Type0	0.92	0.81	0.86	Type0	0.77	0.72	0.74
	Type1	0.81	0.92	0.88	Type1	0.76	0.80	0.78
	Accuracy			0.86	Accuracy			0.76
	Macro avg	0.87	0.87	0.86	Macro avg	0.76	0.76	0.76
<b>DenseNet-121</b>	Weighted avg	0.87	0.86	0.86	Weighted avg	0.76	0.76	0.76
	Type0	0.91	0.79	0.85	Type0	0.81	0.81	0.81
	Type1	0.79	0.91	0.85	Type1	0.83	0.83	0.83

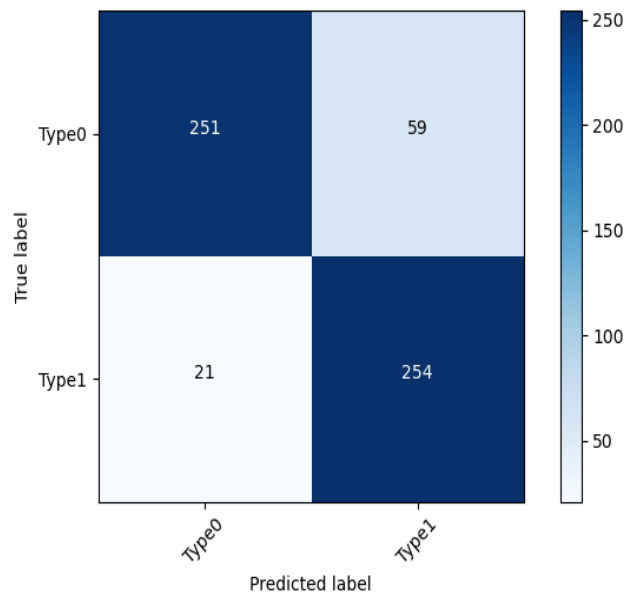
	Accuracy			0.85	Accuracy			0.82
	Macro avg	0.85	0.85	0.85	Macro avg	0.82	0.82	0.82
	<b>Channel P4</b>	<b>Prec</b>	<b>Rec</b>	<b>F1-Score</b>	<b>Avr. Acc.</b>			
<b>ResNet152-v2</b>	Type0	0.96	0.74	0.84	0.837			
	Type1	0.78	0.97	0.87				
	Accuracy			0.85				
	Macro avg	0.87	0.86	0.85				
	Weighted avg	0.87	0.85	0.85				
<b>ResNet50-v2</b>	Type0	0.92	0.76	0.83	0.810			
	Type1	0.78	0.93	0.85				
	Accuracy			0.84				
	Macro avg	0.85	0.84	0.84				
	Weighted avg	0.85	0.84	0.84				
<b>Inception-v3</b>	Type0	0.93	0.73	0.82	0.823			
	Type1	0.80	0.95	0.87				
	Accuracy			0.85				
	Macro avg	0.87	0.84	0.84				
	Weighted avg	0.86	0.85	0.85				
<b>DenseNet-121</b>	Type0	0.96	0.72	0.82	0.837			
	Type1	0.77	0.97	0.85				
	Accuracy			0.84				
	Macro avg	0.86	0.84	0.84				
	Weighted avg	0.87	0.84	0.84				



(a)



(b)



(c)

**Figure 9.** Inception-v3 performance at channel P3: (a) Accuracy graph, (b) Loss graph and (c) Confusion matrix

Figure 9 shows the performance graphs of Inception-v3 at channel P3, which is the best algorithm for this region.

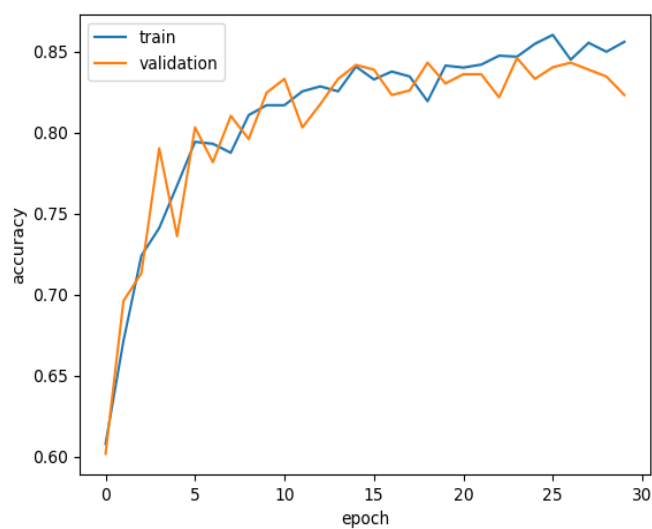
Lastly, the results of the occipital lobe can be seen in Table 7. DenseNet-121 performed the best in both average accuracy and single channel results, with a rate of 0.865 for this region. On the other hand, the lowest accuracy rate is observed in ResNet50-v2.

**Table 7.** Occipital region results for used algorithms

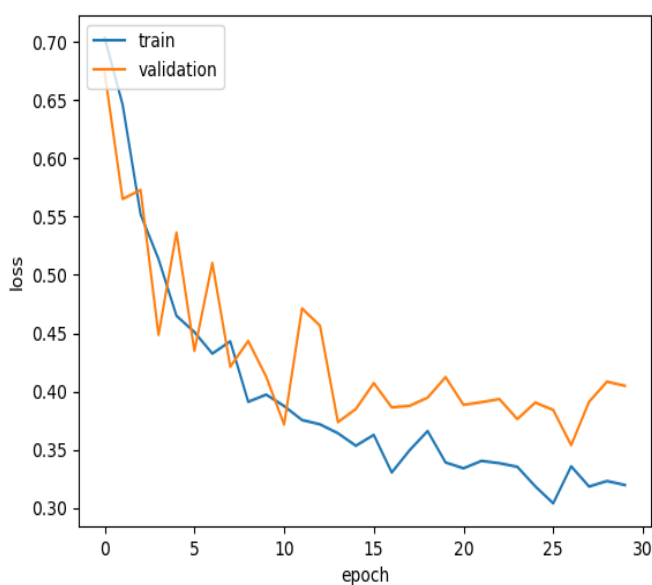
	<b>Channel O1</b>	<b>Prec</b>	<b>Rec</b>	<b>F1-Score</b>	<b>Channel O2</b>	<b>Prec</b>	<b>Rec</b>	<b>F1-Score</b>
<b>ResNet152-v2</b>	Type0	0.96	0.70	0.81	Type0	1.00	0.68	0.81
	Type1	0.76	0.97	0.85	Type1	0.78	1.00	0.88
	Accuracy			0.83	Accuracy			0.85
	Macro avg	0.86	0.83	0.83	Macro avg	0.89	0.84	0.84
	Weighted avg	0.86	0.83	0.83	Weighted avg	0.88	0.85	0.84

<b>ResNet50-v2</b>	Type0	0.74	0.86	0.80	Type0	0.86	0.77	0.81
	Type1	0.83	0.69	0.76	Type1	0.82	0.89	0.85
	Accuracy			0.78	Accuracy			0.84
	Macro avg	0.79	0.78	0.78	Macro avg	0.84	0.83	0.83
	Weighted avg	0.79	0.78	0.78	Weighted avg	0.84	0.84	0.83
<b>Inception-v3</b>	Type0	0.95	0.69	0.80	Type0	0.99	0.67	0.80
	Type1	0.75	0.96	0.84	Type1	0.77	1.00	0.87
	Accuracy			0.82	Accuracy			0.84
	Macro avg	0.85	0.83	0.82	Macro avg	0.88	0.83	0.83
	Weighted avg	0.85	0.82	0.82	Weighted avg	0.88	0.84	0.84
<b>DenseNet-121</b>	Type0	0.94	0.79	0.86	Type0	0.93	0.78	0.85
	Type1	0.83	0.95	0.89	Type1	0.82	0.94	0.88
	Accuracy			0.87	Accuracy			0.86
	Macro avg	0.88	0.87	0.87	Macro avg	0.88	0.86	0.86
	Weighted avg	0.88	0.87	0.87	Weighted avg	0.88	0.86	0.86
<b>Avr. Acc.</b>	<b>ResNet152-v2</b>		<b>ResNet50-v2</b>		<b>Inception-v3</b>		<b>DenseNet-121</b>	
	0.840		0.810		0.830		0.865	

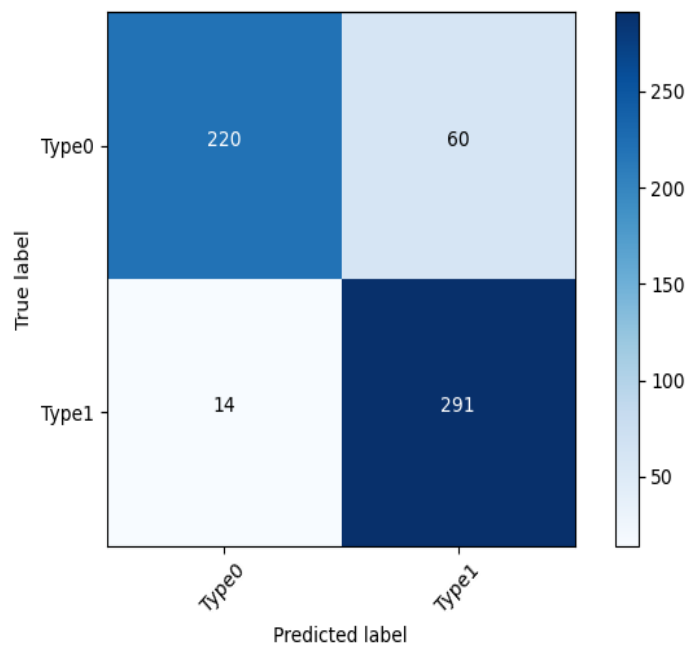
Figure 10 illustrates the accuracy, loss, and confusion matrix graphs of the best classification algorithm, DenseNet-121, for the occipital region.



(a)



(b)



(c)

**Figure 10.** DenseNet-121 performance at channel O2: (a) Accuracy graph, (b) Loss graph and (c) Confusion matrix

### 3.2 Eating behavior questionnaires regression results

In this section, eight different metrics are taken into account under two different questionnaires, as mentioned Table 2. DEBQ has four metrics, which are mean results (output1), restrain eating (output2), emotional eating (output3) and external eating (output4). Likewise, TFEQ metrics are as follows: mean results (output5), uncontrolled eating (output6), cognitive restrain (output7) and emotional eating (output8). For each algorithms output, Median Absolute Error (MedAE), Mean Square Error (MSE), Mean Absolute Error (MAE) and Correlation Square (R2) are given in Figure 11. The performance of the regression models is measured by these metrics. Lower values of MSE, MedAE, and MAE indicate improved performance, while higher values of R2 indicate a more suitable fit of the model to the data.

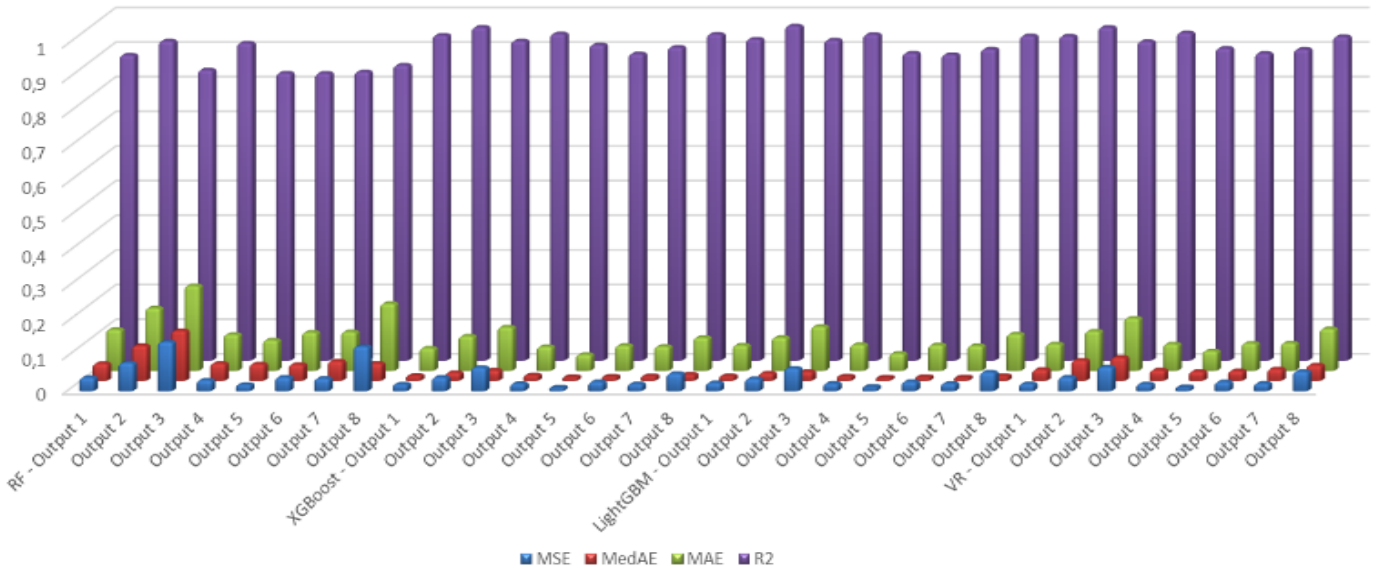


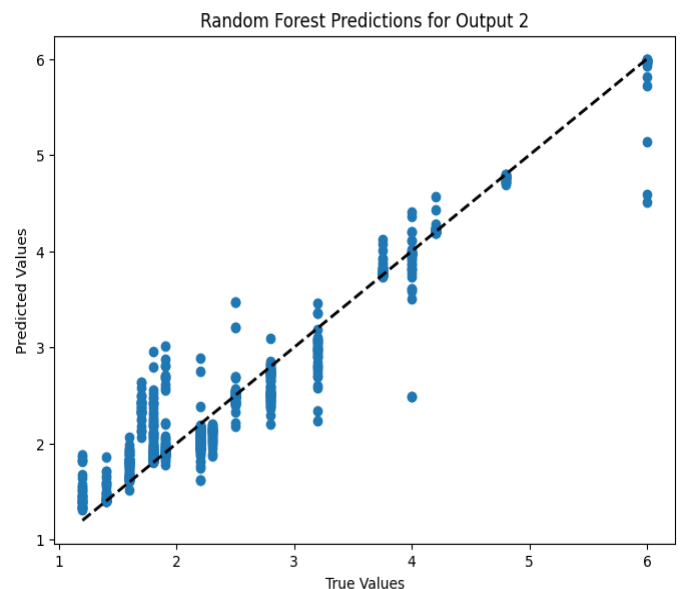
Figure 11. Model results for each questionnaires score

Table 8. Performance metrics of regression models

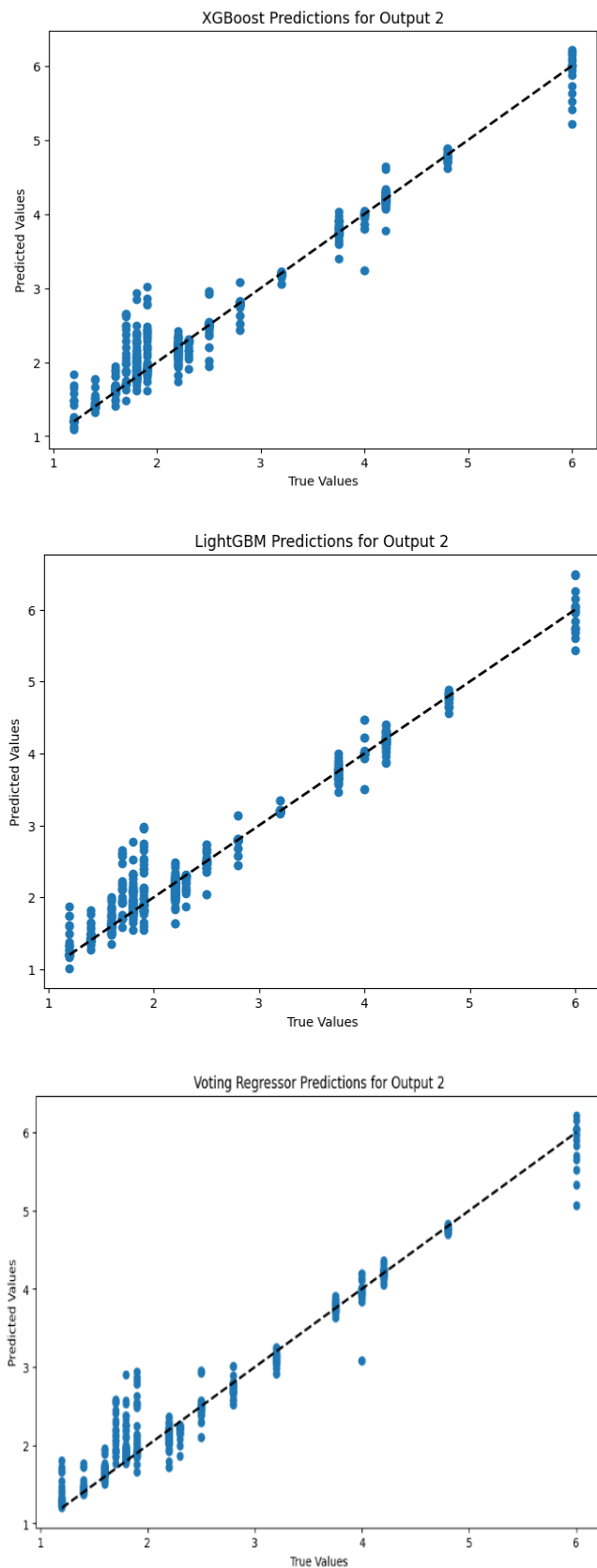
		MSE	MedAE	MAE	R2
Random Forest	Output 1	0.039	0.050	0.119	0.882
	Output 2	0.081	0.102	0.181	0.922
	Output 3	0.142	0.145	0.245	0.839
	Output 4	0.032	0.051	0.105	0.916
	Output 5	0.019	0.049	0.088	0.830
	Output 6	0.040	0.047	0.111	0.829
	Output 7	0.037	0.057	0.112	0.833
	Output 8	0.129	0.051	0.194	0.853
	Mean	0.057	0.054	0.137	0.872
XGBoost	Output 1	0.020	0.015	0.065	0.939
	Output 2	0.040	0.024	0.100	0.962
	Output 3	0.068	0.031	0.126	0.922
	Output 4	0.022	0.017	0.070	0.943
	Output 5	0.010	0.011	0.047	0.911
	Output 6	0.027	0.013	0.073	0.885
	Output 7	0.021	0.014	0.070	0.904
	Output 8	0.051	0.019	0.096	0.942
	Mean	0.031	0.028	0.084	0.902
LightGBM	Output 1	0.024	0.015	0.074	0.928
	Output 2	0.036	0.023	0.096	0.966
	Output 3	0.066	0.027	0.128	0.925
	Output 4	0.022	0.013	0.076	0.941
	Output 5	0.013	0.009	0.050	0.887
	Output 6	0.027	0.011	0.075	0.883
	Output 7	0.022	0.010	0.072	0.899
	Output 8	0.055	0.015	0.106	0.937
	Mean	0.034	0.020	0.101	0.918
Voting Regressor	Output 1	0.021	0.033	0.078	0.936
	Output 2	0.040	0.060	0.114	0.962
	Output 3	0.070	0.068	0.152	0.921
	Output 4	0.021	0.031	0.077	0.946
	Output 5	0.011	0.027	0.057	0.901
	Output 6	0.026	0.029	0.079	0.887
	Output 7	0.022	0.035	0.080	0.898
	Output 8	0.057	0.046	0.121	0.935
	Mean	0.036	0.042	0.106	0.921

Table 8 appears to show the performance metrics of different regression models for various output variables. XGBoost achieved the lowest mean MSE (0.0314), MedAE (0.0276), and MAE (0.0848), indicating superior performance compared to the other models in minimizing prediction errors.

Random Forest and LightGBM also score well. They exhibit slightly higher mean values for MSE, MedAE, and MAE compared to XGBoost. However, the Voting Regressor has the highest mean values for these variables among the models, indicating that it performs slightly worse on average than the others. Random Forest performs well across different outputs, with R2 values ranging from 0.829 to 0.922. XGBoost consistently exhibits strong performance across various outputs, as evidenced by its R-squared values, which range from 0.885 to 0.962. In addition, LightGBM consistently displays impressive performance, as indicated by its R-squared values, which span from 0.883 to 0.966. The Voting Regressor, which amalgamates predictions from three models, exhibits strong performance, as reflected in its R-squared values, which range from 0.887 to 0.962. In general, XGBoost and LightGBM outperform Random Forest on MSE, MedAE, MAE, and R2 across the entire range of outputs. The choice between XGBoost and LightGBM for each output depends on the specific metric. However, overall, XGBoost and Voting Regressor tend to have higher R2 values.

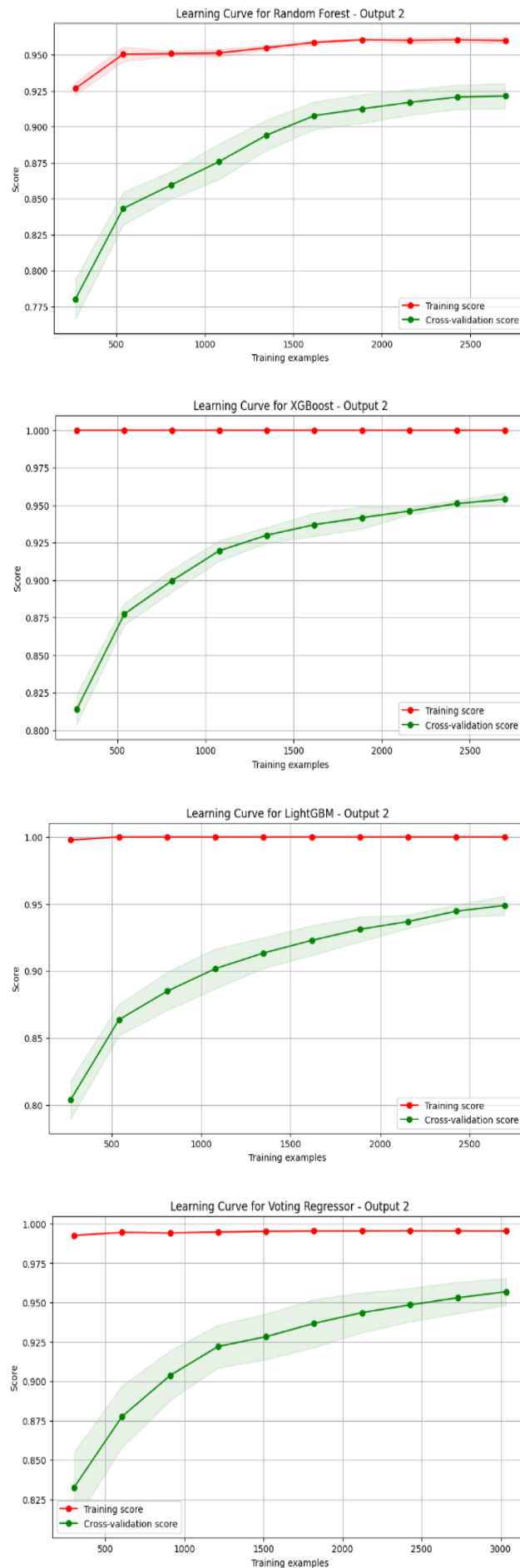






**Figure 12.** Correlation graphs at best results of models

In conclusion, both the individual models (Random Forest, XGBoost, LightGBM) and the ensemble model (Voting Regressor) showcase commendable predictive performance. While output 5 (mean value of TFEQ) has the lowest R2 values, output 2 (restrain eating of DEBQ) reaches the best R2 for each model. The best output correlation graphs are shown in Figure 12.



**Figure 13.** Learning curves for regression models at best results of models



behavior questionnaires are administered to gather relevant data. Firstly, CNN based classification is applied using EEG signals of overweight/obese and normal individuals. Then, regression analysis is done using EEG signals and questionnaire results. At the classification stage, by using tabular data augmentation, overfitting and imbalance problems were eliminated. Additionally, the SuperTML method was used to embed the features into images, which has not been used for EEG data before, and it's observed that this method can be used for EEG signals. Even in the face of the limited quantities, poor qualities, and fluctuating qualities of medical data, training procedures with CNNs were made viable by employing transfer learning approaches by converting the tabular data set into new image representations. The trials show that the suggested method performed better than results from previous work in the literature in most cases. At the regression stage, the SMOGN data augmentation technique, which is more suitable for regression analysis, was applied to eliminate overfitting and imbalance problems. Then, different regression models were applied to see the correlation between questionnaires and EEG data. As a result, it is observed that there is a strong correlation between EEG data and questionnaires.

To conclude, overweight/obese and normal participants are shown to have different brain connectivity patterns. These distinct brain connections can be used to inform the development of neuromodulation therapies to enhance behaviors associated with obesity.

## ACKNOWLEDGMENT

A significant part of this paper includes the Doctorate Thesis data of Halil İbrahim COŞAR.

## REFERENCES

- [1] WHO. "Obesity and overweight fact sheet." <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>.
- [2] Jehan, S., Myers, A.K., Zizi, F., Pandi-Perumal, S.R., Jean-Louis, G., McFarlane, S.I. (2018). Obesity, obstructive sleep apnea and type 2 diabetes mellitus: Epidemiology and pathophysiologic insights. *Sleep Medicine and Disorders: International Journal*, 2(3): 52. <https://doi.org/10.15406/smdij.2018.02.00045>
- [3] Carey, M., Small, H., Yoong, S.L., Boyes, A., Bisquera, A., Sanson-Fisher, R. (2014). Prevalence of comorbid depression and obesity in general practice: A cross-sectional survey. *British Journal of General Practice*, 64(620): e122-e127. <https://doi.org/10.3399/bjgp14X677482>
- [4] Beechy, L., Galpern, J., Petrone, A., Das, S.K. (2012). Assessment tools in obesity-Psychological measures, diet, activity, and body composition. *Physiology & Behavior*, 107(1): 154-171. <https://doi.org/10.1016/j.physbeh.2012.04.013>
- [5] Snekhalatha, U., Sangamithirai, K. (2021). Computer aided diagnosis of obesity based on thermal imaging using various convolutional neural networks. *Biomedical Signal Processing and Control*, 63: 102233. <https://doi.org/10.1016/j.bspc.2020.102233>
- [6] Mathur, N., Gupta, A., Jaswal, S., Verma, R. (2021). Deep learning helps EEG signals predict different stages of visual processing in the human brain. *Biomedical Signal Processing and Control*, 70: 102996. <https://doi.org/10.1016/j.bspc.2021.102996>
- [7] Zhang, X., Huang, D., Li, H., Zhang, Y., Xia, Y., Liu, J. (2023). Self-training maximum classifier discrepancy for EEG emotion recognition. *CAAI Transactions on Intelligence Technology*, 8(4): 1480-1491. <https://doi.org/10.1049/cit2.12174>
- [8] Liu, S., Wang, Z., An, Y., Zhao, J., Zhao, Y., Zhang, Y.D. (2023). EEG emotion recognition based on the attention mechanism and pre-trained convolution capsule network. *Knowledge-Based Systems*, 265: 110372. <https://doi.org/10.1016/j.knosys.2023.110372>
- [9] Göker, H. (2023). Welch spectral analysis and deep learning approach for diagnosing Alzheimer's disease from resting-state EEG recordings. *Traitement du Signal*, 40(1): 257-264. <http://doi.org/10.18280/ts.400125>
- [10] Jayanthi, A.K. (2022). Wavelet scattering transform and deep learning networks based autism spectrum disorder identification using EEG signals. *Traitement du Signal*, 39(6): 2069-2076. <http://doi.org/10.18280/TS.390619>
- [11] Arslan, R., Ulutaş, H., Köksal, A., Bakir, M., Çiftçi, B. (2023). Tree-based machine learning techniques for automated human sleep stage classification. *Traitement du Signal*, 40(4): 1385-1400. <http://doi.org/10.18280/ts.400408>
- [12] Arslan, R.S., Ulutas, H., Köksal, A.S., Bakir, M., Çiftçi, B. (2023). End-to end decision support system for sleep apnea detection and Apnea-Hypopnea Index calculation using hybrid feature vector and Machine learning. *Biocybernetics and Biomedical Engineering*, 43(4): 684-699. <https://doi.org/10.1016/j.bbe.2023.10.002>
- [13] Khademi, Z., Ebrahimi, F., Kordy, H.M. (2022). A transfer learning-based CNN and LSTM hybrid deep learning model to classify motor imagery EEG signals. *Computers in Biology and Medicine*, 143: 105288. <https://doi.org/10.1016/j.compbimed.2022.105288>
- [14] Liu, M., Wu, W., Gu, Z., Yu, Z., Qi, F., Li, Y. (2018). Deep learning based on batch normalization for P300 signal detection. *Neurocomputing*, 275: 288-297. <https://doi.org/10.1016/j.neucom.2017.08.039>
- [15] Waytowich, N., Lawhern, V.J., Garcia, J.O., Cummings, J., Faller, J., Sajda, P., Vettel, J.M. (2018). Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials. *Journal of Neural Engineering*, 15(6): 066031. <https://doi.org/10.1088/1741-2552/aae5d8>
- [16] Nijs, I. (2010). Attentional mechanisms in food craving and overeating: A study of an addiction model of obesity. *Erasmus Universiteit*.
- [17] Volkow, N.D., Wise, R.A. (2005). How can drug addiction help us understand obesity?. *Nature Neuroscience*, 8(5): 555-560. <https://doi.org/10.1038/nn1452>
- [18] Nijs, I.M., Muris, P., Euser, A.S., Franken, I.H. (2010). Differences in attention to food and food intake between overweight/obese and normal-weight females under conditions of hunger and satiety. *Appetite*, 54(2): 243-254. <https://doi.org/10.1016/j.appet.2009.11.004>
- [19] Zsoldos, I., Sinding, C., Godet, A., Chambaron, S. (2021). Do food odors differently influence cerebral activity depending on weight status? An electroencephalography study of implicit olfactory

- priming effects on the processing of food pictures. *Neuroscience*, 460: 130-144. <https://doi.org/10.1016/j.neuroscience.2021.01.015>
- [20] Hume, D.J., Howells, F.M., Rauch, H.L., Kroff, J., Lambert, E.V. (2015). Electrophysiological indices of visual food cue-reactivity. Differences in obese, overweight and normal weight women. *Appetite*, 85: 126-137. <https://doi.org/10.1016/j.appet.2014.11.012>
- [21] Biehl, S.C., Keil, J., Naumann, E., Svaldi, J. (2020). ERP and oscillatory differences in overweight/obese and normal-weight adolescents in response to food stimuli. *Journal of Eating Disorders*, 8: 1-11. <https://doi.org/10.1186/s40337-020-00290-8>
- [22] Zhao, K., Xu, D. (2019). Food image-induced discrete emotion recognition using a single-channel scalp-EEG recording. In 2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Suzhou, China, IEEE, pp. 1-6. <https://doi.org/10.1109/CISP-BMEI48845.2019.8966064>
- [23] Nijs, I.M., Franken, I.H., Muris, P. (2010). Food-related Stroop interference in obese and normal-weight individuals: Behavioral and electrophysiological indices. *Eating Behaviors*, 11(4): 258-265. <https://doi.org/10.1016/j.eatbeh.2010.07.002>
- [24] Woltering, S., Chen, S., Jia, Y. (2021). Neural correlates of attentional bias to food stimuli in obese adolescents. *Brain Topography*, 34: 182-191. <https://doi.org/10.1007/s10548-020-00812-2>
- [25] Ammar Ali, M., Özögür-Akyüz, S., Duru, A.D., Caliskan, M., Demir, C., Bostancı, T., Elsallak, F., Shkokani, M., Dokur, Z., Ölmez, T., Ergün, C., Bebek, N., Yilmaz, G. (2022). Neurological effects of long-term diet on obese and overweight individuals: An electroencephalogram and event-related potential study. *Computational Intelligence*, 38(3): 1163-1182. <https://doi.org/10.1111/coin.12444>
- [26] Al-Zubaidi, A., Mertins, A., Heldmann, M., Jauch-Chara, K., Münte, T.F. (2019). Machine learning based classification of resting-state fMRI features exemplified by metabolic state (hunger/satiety). *Frontiers in Human Neuroscience*, 13: 164. <https://doi.org/10.3389/fnhum.2019.00164>
- [27] Singh, B., Tawfik, H. (2020). Machine learning approach for the early prediction of the risk of overweight and obesity in young people. In *Computational Science-ICCS 2020: 20th International Conference, Amsterdam, The Netherlands, Proceedings, Part IV*. Springer International Publishing. Springer, Cham, 20: 523-535. [https://doi.org/10.1007/978-3-030-50423-6\\_39](https://doi.org/10.1007/978-3-030-50423-6_39)
- [28] Yue, Y., De Ridder, D., Manning, P., Ross, S., Deng, J.D. (2022). Finding neural signatures for obesity using source-localized EEG features. *CoRR*.
- [29] Blechert, J., Lender, A., Polk, S., Busch, N.A., Ohla, K. (2019). Food-pics\_extended – An image database for experimental research on eating and appetite: Additional images, normative ratings and an updated review. *Frontiers in Psychology*, 10(307): 1-9. <https://doi.org/10.3389/fpsyg.2019.00307>
- [30] Kilic, F., Korkmaz, M., Er, O., Altin, C. (2023). A CNN-based novel approach for classification of sacral hiatus with GAN-powered tabular data set. *Elektronika Ir Elektrotechnika*, 29(2): 44-53. <https://doi.org/10.5755/j02.eie.33852>
- [31] Ong, Z.Y., Saidatul, A., Ibrahim, Z. (2018). Power spectral density analysis for human EEG-based biometric identification. In 2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA), Kuching, Malaysia, pp. 1-6. <https://doi.org/10.1109/ICASSDA.2018.8477604>
- [32] Noshadi, S., Abootalebi, V., Sadeghi, M.T., Shahvazian, M.S. (2014). Selection of an efficient feature space for EEG-based mental task discrimination. *Biocybernetics and Biomedical Engineering*, 34(3): 159-168. <https://doi.org/10.1016/j.bbe.2014.03.004>
- [33] Shwartz-Ziv, R., Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion*, 81: 84-90. <https://doi.org/10.1016/j.inffus.2021.11.011>
- [34] Strelcenia, E., Prakoonwit, S. (2023). A survey on GAN techniques for data augmentation to address the imbalanced data issues in credit card fraud detection. *Machine Learning and Knowledge Extraction*, 5(1): 304-329. <https://doi.org/10.3390/make5010019>
- [35] Zhong, Z., Zheng, L., Kang, G., Li, S., Yang, Y. (2020). Random erasing data augmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(07): 13001-13008. <https://doi.org/10.1609/aaai.v34i07.7000>
- [36] Chen, J., Tam, D., Raffel, C., Bansal, M., Yang, D. (2023). An empirical survey of data augmentation for limited data learning in NLP. *Transactions of the Association for Computational Linguistics*, 11: 191-211. [https://doi.org/10.1162/tacl\\_a\\_00542](https://doi.org/10.1162/tacl_a_00542)
- [37] Laddha, S., Kumar, V. (2022). DGCNN: Deep convolutional generative adversarial network based convolutional neural network for diagnosis of COVID-19. *Multimedia Tools and Applications*, 81(22): 31201-31218. <https://doi.org/10.1007/s11042-022-12640-6>
- [38] Burks, R., Islam, K.A., Lu, Y., Li, J. (2019). Data augmentation with generative models for improved malware detection: A comparative study. In 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, pp. 0660-0665. <https://doi.org/10.1109/UEMCON47517.2019.8993085>
- [39] Mahajan, K., Vargantwar, M.R., Rajput, S.M. (2011). Classification of EEG using PCA, ICA and Neural Network. *International Journal of Engineering and Advanced Technology*, 1(1): 80-83.
- [40] Gorur, K., Bozkurt, M.R., Bascil, M.S., Temurtas, F. (2018). Glossokinetic potential based tongue-machine interface for 1-D extraction using neural networks. *Biocybernetics and Biomedical Engineering*, 38(3): 745-759. <https://doi.org/10.1016/j.bbe.2018.06.004>
- [41] Sun, B., Yang, L., Zhang, W., Lin, M., Dong, P., Young, C., Dong, J. (2019). Supertml: Two-dimensional word embedding for the precognition on structured tabular data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, Long Beach, CA, USA, pp. 2973-2981. <https://doi.org/10.1109/CVPRW.2019.00360>
- [42] Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaria, J., Fadhel, M.A., Al-Amidie, M., Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8: 1-74. <https://doi.org/10.1186/s40537-021-00444-8>

- [43] Radhakrishnan, M., Ramamurthy, K., Choudhury, K.K., Won, D., Manoharan, T.A. (2021). Performance analysis of deep learning models for detection of autism spectrum disorder from EEG signals. *Traitement du Signal*, 38(3): 853-863. <https://doi.org/10.18280/ts.380332>
- [44] Bizopoulos, P., Lambrou, G.I., Koutsouris, D. (2019). Signal2image modules in deep neural networks for EEG classification. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, pp. 702-705. <https://doi.org/10.1109/EMBC.2019.8856620>
- [45] Bozan, N., Bas, M., Asci, F.H. (2011). Psychometric properties of Turkish version of Dutch eating behaviour questionnaire (DEBQ). A preliminary results. *Appetite*, 56(3): 564-566. <https://doi.org/10.1016/j.appet.2011.01.025>
- [46] Kirac, D., Kaspar, E., Avcilar, T., Cakir, O., Ulucan, K., Kurtel, H., Deyneli, O., Guney, A. (2015). A new method for investigating eating behaviours related with obesity: Three-factor eating questionnaire. *Journal of Marmara University Institute of Health Sciences*, 5(3): 162-169. <http://doi.org/10.5455/musbed.20150602015512>
- [47] Branco, P., Torgo, L., Ribeiro, R.P. (2017). SMOGN: A pre-processing approach for imbalanced regression. In *Proceedings of the First International Workshop on Learning with Imbalanced Domains: Theory and Applications*. PMLR, pp. 36-50. <https://proceedings.mlr.press/v74/branco17a.html>
- [48] Yang, Y. (2007). Consistency of cross validation for comparing regression procedures. *The Annals Statistics*, 35(6): 2450-2473. <https://doi.org/10.1214/009053607000000514>
- [49] Suenaga, D., Takase, Y., Abe, T., Orita, G., Ando, S. (2023). Prediction accuracy of Random Forest, XGBoost, LightGBM, and artificial neural network for shear resistance of post-installed anchors. *Structures*, 50: 1252-1263. <https://doi.org/10.1016/j.istruc.2023.02.066>
- [50] Guo, X., Gui, X., Xiong, H., Hu, X., Li, Y., Cui, H., Qiu, Y., Ma, C. (2023). Critical role of climate factors for groundwater potential mapping in arid regions: Insights from random forest, XGBoost, and LightGBM algorithms. *Journal of Hydrology*, 621: 129599. <https://doi.org/10.1016/j.jhydrol.2023.129599>
- [51] Sheldon, M.R., Fillyaw, M.J., Thompson, W.D. (1996). The use and interpretation of the Friedman test in the analysis of ordinal-scale data in repeated measures designs. *Physiotherapy Research International*, 1(4): 221-228. <https://doi.org/10.1002/pri.66>
- [52] Hume, D.J., Howells, F.M., Karpul, D., Rauch, H.L., Kroff, J., Lambert, E.V. (2015). Cognitive control over visual food cue saliency is greater in reduced-overweight/obese but not in weight relapsed women: An EEG study. *Eating Behaviors*, 19: 76-80. <https://doi.org/10.1016/j.eatbeh.2015.06.013>
- [53] Bin, N.W., Awang, S.A., Fook, C.Y., Chin, L.C., Ying, O.Z. (2019). A study of informative EEG channel and brain region for typing activity. In *Journal of Physics: Conference Series*. IOP Publishing, 1372(1): 012008. <https://doi.org/10.1088/1742-6596/1372/1/012008>
- [54] Karhunen, L.J., Lappalainen, R.I., Vanninen, E.J., Kuikka, J.T., Uusitupa, M.I. (1997). Regional cerebral blood flow during food exposure in obese and normal-weight women. *Brain: A Journal of Neurology*, 120(9): 1675-1684. <https://doi.org/10.1093/brain/120.9.1675>
- [55] Carbine, K.A., Larson, M.J., Romney, L., Bailey, B.W., Tucker, L.A., Christensen, W.F., LeCheminant, J.D. (2017). Disparity in neural and subjective responses to food images in women with obesity and normal-weight women. *Obesity*, 25(2): 384-390. <https://doi.org/10.1002/oby.21710>