machine learning & knowledge extraction

MDPI

Article

The Impact of Light Conditions on Neural Affect Classification: A Deep Learning Approach

Sophie Zentner, Alberto Barradas Chacon and Selina C. Wriessnegger *

Institute of Neural Engineering, Graz University of Technology, 8010 Graz, Austria; barradaschacon@tugraz.at (A.B.C.)
* Correspondence: s.wriessnegger@tugraz.at

Abstract: Understanding and detecting human emotions is crucial for enhancing mental health, cognitive performance and human–computer interactions. This field in affective computing is relatively unexplored, and gaining knowledge about which external factors impact emotions could enhance communication between users and machines. Furthermore, it could also help us to manage affective disorders or understand affective physiological responses to human spatial and digital environments. The main objective of the current study was to investigate the influence of external stimulation, specifically the influence of different light conditions, on brain activity while observing affect-eliciting pictures and their classification. In this context, a multichannel electroencephalography (EEG) was recorded in 30 participants as they observed images from the Nencki Affective Picture System (NAPS) database in an art-gallery-style Virtual Reality (VR) environment. The elicited affect states were classified into three affect classes within the two-dimensional valence–arousal plane. Valence (positive/negative) and arousal (high/low) values were reported by participants on continuous scales. The experiment was conducted in two experimental conditions: a warm light condition and a cold light condition. Thus, three classification tasks arose with regard to the recorded brain data: classification of an affect state within a warm-light condition, classification of an affect state within a cold light condition, and warm light vs. cold light classification during observation of affect-eliciting images. For all classification tasks, Linear Discriminant Analysis, a Spatial Filter Model, a Convolutional Neural Network, the EEGNet, and the SincNet were compared. The EEGNet architecture performed best in all tasks. It could significantly classify three affect states with 43.12% accuracy under the influence of warm light. Under the influence of cold light, no model could achieve significant results. The classification between visual stimulus with warm light vs. cold light could be classified significantly with 76.65% accuracy from the EEGNet, well above any other machine learning or deep learning model. No significant differences could be detected between affect recognition in different light conditions, but the results point towards the advantage of gradient-based learning methods for data-driven experimental designs for the problem of affect decoding from EEG, providing modern tools for affective computing in digital spaces. Moreover, the ability to discern externally driven affective states through deep learning not only advances our understanding of the human mind but also opens avenues for developing innovative therapeutic interventions and improving human–computer interaction.

Keywords: EEG; machine learning; affect classification; emotion decoding; virtual reality

check for updates

Citation: Zentner, S.; Barradas Chacon, A.; Wriessnegger, S.C. The Impact of Light Conditions on Neural Affect Classification: A Deep Learning Approach. *Mach. Learn. Knowl. Extr.* **2024**, *6*, 199–214. https://doi.org/10.3390/make6010011

Academic Editor: Andreas Holzinger

Received: 7 November 2023
Revised: 16 January 2024
Accepted: 16 January 2024
Published: 18 January 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Many human needs, behaviors, and human performance are related to emotions, as humans are affective beings [1]. Thus, the need for applications that are able to detect human affect states has been growing in recent years so as not only to enhance communication between the user and the machine, but also to treat disorders and achieve marketing goals or sports goals that require specific mental states [1–4]. This can be achieved through a system that performs an action using brain signals as an input while the user does not have voluntary control over it, also called passive brain–computer interface (BCI), using electroencephalography (EEG) [5].

Emotions can be viewed from a categorical or a dimensional perspective. The categorical perspective on emotions was summarized by Ekman, who listed six basic emotions, happiness, sadness, anger, fear, surprise, and disgust, as categories [6]. On the other hand, one of the commonly used dimensional models is Russel's circumplex model, which uses two dimensions, valence and arousal, to capture emotional states [7,8]. Valence measures negative/positive feelings on a continuous scale and arousal is captured from low to high on a continuous scale [7]. Thus, it is possible to span a two-dimensional plane that includes all of the categorical emotions and is quantifiable. This is called affect, which is less subjective and less dependent on the perception of the single words that are intended to describe an emotion [9].

Previous attempts to classify affect from EEG data have involved the application of deep learning models using videos as visual stimuli. For instance, Ramzan et al. [10] classified high arousal versus low arousal and high valence versus low valence with a Fused CNN-LSTM with the mentioned stimuli. While video stimuli offer a more immersive experience than images, it is crucial to note that the rated affect for these videos may not be valid for every moment, but rather reflects the overall impression. This introduces a potential bias when evaluating affect using shorter trials, a common scenario in many studies [9,10].

Affect-eliciting pictures like those in the International Affective Picture System (IAPS) have already been used to classify affect in terms of valence and arousal with machine learning approaches. In these studies, binary classification was performed on one of the two parameters rather than classification on a two-dimensional valence–arousal plane [11,12].

There are a limited number of studies using Virtual Reality (VR) showing affect-eliciting stimuli, although it promises a more immersive and almost realistic experience. One of them classified EEG trials into low or high negativity and low or high positivity using video stimuli. The opposing label names and the fact that the stimuli are videos raise questions about the study design, emphasizing the need for more meaningful labeling [9].

Marin-Morales et al. and Horvat et al. combined the idea of showing images and being in a VR environment by guiding the participant through a virtual 3D art museum. The advantage of this approach is that image databases like NAPS or IAPS are available and just have to be embedded in a VR space. But like all studies that performed affect classification using the circumplex model, the named study also induced geometric splitting of the valence or arousal scale. This might not be reasonable regarding the distribution of ratings from the participant. A more meaningful labeling approach has to be developed. The named study reached 75% accuracy in classifying high or low arousal and 71% accuracy in classifying positive or negative valence, using EEG data and heart-rate variability data. Also, they used a Support Vector Machine for classification, although the literature approaches with Neural Networks performed better than simpler machine learning approaches when comparing similar affect classification tasks with each other [13,14]. Still, simpler machine learning approaches have to be used and tested, because they are less complex and in general take less time to train while resulting in high accuracy [15,16]. Other studies used different VR scenes to elicit emotions and classified them according to a binary system regarding valence and/or arousal, respectively [16–22]. Marin-Morales et al. [23] split the valence arousal plane into four quadrants and achieved successful classification using four classes. These approaches, again, use geometrical splitting of the valence–arousal plane.

The influence of colors on the human EEG signals can be studied using several different visual stimuli. Hosseini et al. [24] presented colored screens, colored words on a white background, and black words on a colored background to participants and analyzed the EEG signals with regard to several band powers. The study could not find any significant differences in specific brain waves between the presented colors.

Hassib et al. [25], on the other hand, were able to discriminate between blue and orange ambient light influence while participants played a drive simulator game and rated their valence and arousal. Based on the EEG data measured, feature extraction was performed and a random forest classifier was utilized to classify positive or negative valence and high or low arousal, and thus blue or orange light condition [24]. Also, experiments in VR were performed, where participants sat in a virtual room with colored walls [22,25]. Differences in brain waves in response to changing wall colors were analyzed and an increase in brain activity linked to emotional processing for blue wall colors was detected [26]. No classification algorithm was applied to see how well color conditions could be distinguished. Schilling et al. investigated the influence of colored lenses on late positive potentials (LPP) in EEG signals while watching images from the IAPS database. According to the study, red-tinted lenses enhanced the LPP during emotionally arousing and neutral picture presentation compared with uncolored lenses, while green-tinted lenses increased LPP as well, but mainly in positive valence picture condition [27].

From the literature, we see that colored lenses influence LPP in participants observing specific emotionally triggering material and colored walls also influence EEG signals [26,27]. To the best of our knowledge, no previous study has investigated the influence of light conditions on brain activity while observing affect-eliciting pictures. Based on that knowledge, the following main research question arises: Is it possible to distinguish brain signals from participants experiencing different light conditions while observing emotional content? To address this, we will further formulate the following hypotheses:

(1) Can we train a network to predict an affect state? To answer this question, two tasks arise: an affect-classification task under the influence of warm light and an affect-classification task under the influence of cold light.
(2) Can we train a network to predict light conditions given an affect state? To answer this question, the following task arises: classification of a warm light vs. a cold light condition.
(3) How well can we classify affect states in one condition vs. another? To answer this question, a statistical analysis of the results of the previous tasks has to be performed.

Section 2 describes the VR environment and the study itself and develops the classification methods and their architecture. This is followed by Section 3, where the results of all classification tasks are listed. Section 4 discusses the results and answers the research questions. Section 5 presents the most important results and describes how affect recognition can be further investigated.

2. Methods

2.1. Participants

A total of 30 healthy participants (15 men, 15 women) took part in the study, with a mean age of 28.80 $\pm$ 9.09. All participants were informed in writing about the aim of the study and the measurement technique, and they signed informed consent prior to the experiment. The study was approved by the local ethics committee (Medical University of Graz) and was in accordance with the ethical standards of the Declaration of Helsinki.

2.2. Experimental Paradigm

Prior to the start of the experiment, participants completed a measurement record detailing demographic features. Subsequently, they were briefed on the experimental task. Initially, participants were instructed to assess images displayed on a standard screen using the Self-Assessment Manikin (SAM) scales for valence and arousal [28]. Following this, participants were seated in a comfortable chair, equipped with EEG, and, lastly, wore VR glasses (HTC Vive) as illustrated in Figure 1. The experiment unfolded within a VR environment resembling an art gallery, where images were positioned on a wall in front of the viewer. Each image was displayed for a duration of 10 s with a 3 s interval in between. During the pause, the VR camera simulated a gradual transition, as if participants were moving to the next image in the gallery. A total of 100 images from the NAPS database were presented in a pre-randomized order, which was consistent for every participant. Following the observation of all images in one light condition, participants had the option to take an individual break without the VR glasses before proceeding to the second light-condition. The order of the light conditions was randomly determined for each participant.

Figure 1. Participant during the study wearing EEG device and VR glasses.

2.3. Affect Elicitation and Labeling

One hundred images from the NAPS database with a rather strong affect score (very high or very low on the valence and arousal scale) were used for affect elicitation. For this purpose, the image ratings with the largest Euclidean distance to the origin of the valence/arousal scale were selected, with the origin of the scale lying at a scoring value of 5. Also, the images were evenly distributed throughout all quarters of the valence arousal scale. Finally, 25 images were selected from each quadrant.

To obtain individual affect ratings, the participants scored all images on continuous valence and arousal scales prior to the VR experiment. The mean participant scores were clustered with the density-based spatial clustering of applications with noise (DBScan) algorithm and the cluster means were calculated; see Figure 2. This clustering resulted in 3 affect classes that will be used subsequently. To individualize the labels for each participant's score, the distance from every single participant score to all cluster means was calculated and the image rating was assigned to the class label with the smallest distance. Following that, we obtained 3 affect classes used as affect-classification labels for each light condition, respectively.

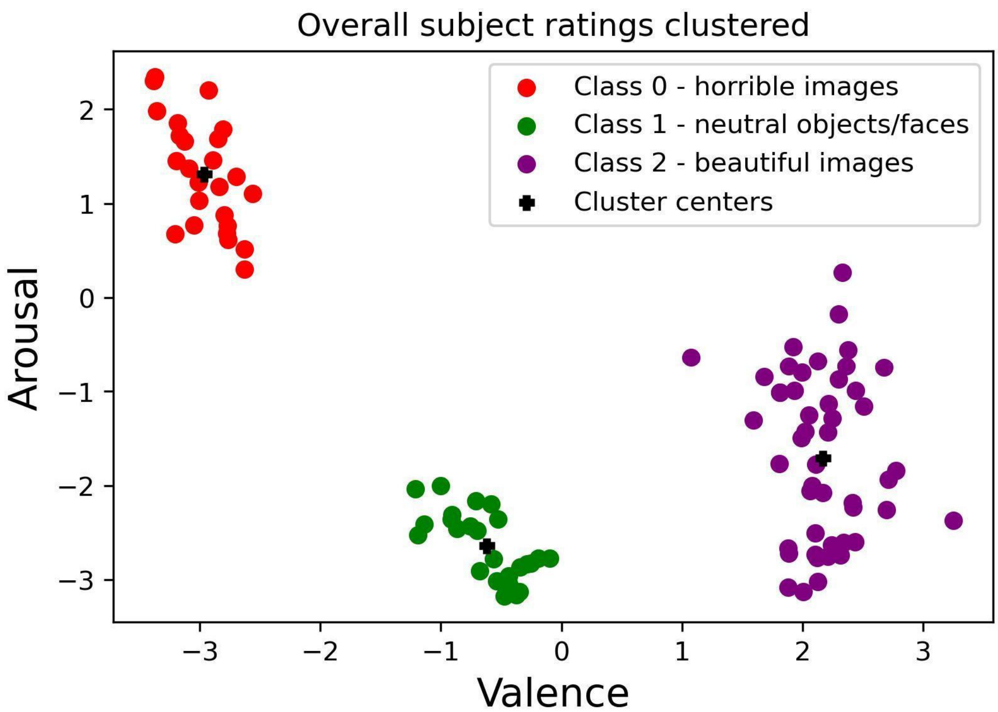

Figure 2. Mean subject ratings for 100 images resulted in three clusters with DBScan algorithm. Black crosses indicate the cluster centers.

The warm vs. cold task yielded two labels, namely warm light and cold light, that were used in the following sections. All images were embedded into a VR environment that was created with UNITY XR (version 2020.3.27f1). The VR environment was created once with warm illumination (Figure 3A) and once with cold illumination, as shown in Figure 3B. Figure 4 shows the procedure of an individual experimental trial.

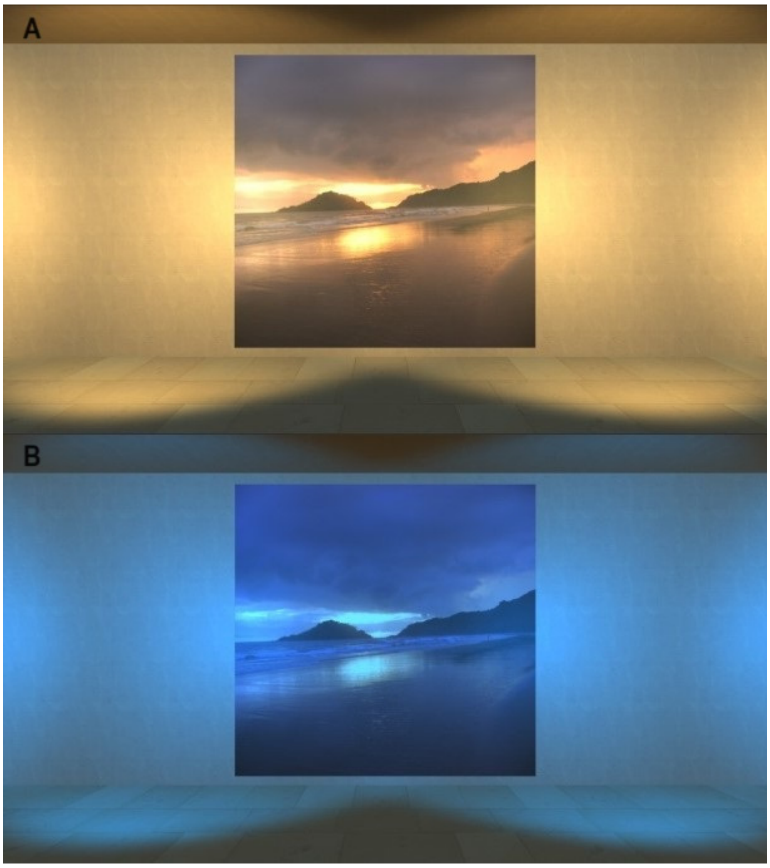

Figure 3. (**A**) VR Environment of one image under warm light condition and (**B**) under cold light condition.

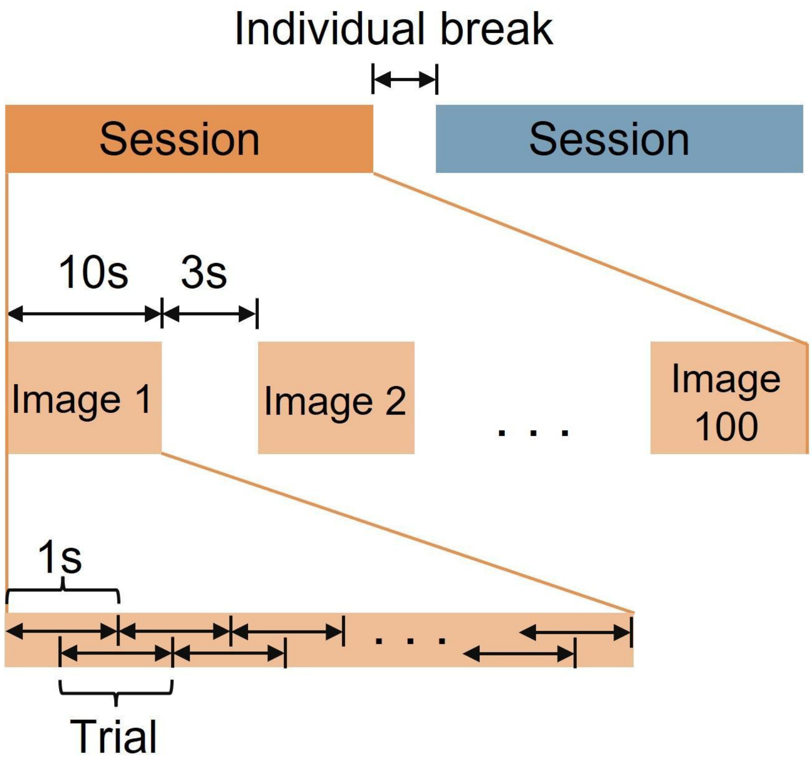

Figure 4. Overview of the experimental paradigm: 2 sessions (1 warm and 1 cold session) with 100 image presentations and 19 extracted trials per image.

2.4. EEG Setup and Data Preprocessing

The EEG was recorded with 34 active electrodes, which were mounted according to the 10–20 international system and connected to a wearable amplifier (LiveAmp: Brain Products GmbH, Gilching, Germany) with a sampling rate of 500 Hz. The ground electrode was placed at the Fpz electrode position and the reference electrode at position FCz. The impedances were kept below 25 kOhm.

First, ICA visual inspection of the signals was performed. The signals were then filtered with a FIR bandpass filter using a Hamming window and cutoff frequencies of 0.2 Hz and 30 Hz. The experimental data were segmented into overlapping windows, each 1 s long and overlapping by 0.5 s. This yielded 19 trials per image per participant with the dimension channels by time before trial rejection (Figure 4). Trial rejection was performed based on the amplitude of the signal, so that approximately eight percent of each participant's trials were rejected.

2.5. Classification Methods

Several approaches were tested, from linear machine learning algorithms to deep neural networks. When performing the affect classification tasks, the EEG trials from each image of each participant, as well as the participant number, were input into the network. The latter was provided to allow the network to learn about inter-subject variability [29] in human brains, which has been demonstrated by Horvat et al. [30], among others. The output would be one of the 3 affect classes introduced in Section 2.3.

When performing the warm vs. cold light-classification task, the EEG trials from each image of each participant, the participant number, and the affect label corresponding to the trial were used as model input. The affect label was also provided so as to allow the model to compare the conditions given the affect, and thus it would be able to extract the features relevant for the light condition and not the affect. The output would be one of the 2 light conditions (warm or cold light).

The model hyperparameters were chosen after performing a random grid search on the affect classification task with the warm light condition using cross-validation with 4 splits and 500 epochs each, while splitting was based on individual trials. The metric chosen for the hyperparameter search was the area under the ROC curve (AUROC). The AUROC might be too general in many cases, as it evaluates all decision thresholds, including unrealistic thresholds—which should not be disregarded, especially in clinical–medical problems [31]. On the other hand, the accuracy, sensitivity, specificity, positive predictive value, and the F1 score are too specific—they are measured against a single threshold, which is also only optimal for some cases but not for others. Carrington et al. [31] described a deep ROC analysis that was intended to balance the aforementioned problems. In our case, we used the AUROC value because it was shown to be a better measure for model evaluation than accuracy, because it can be compared within tasks by considering class imbalances and number of classes [32]. The model features that yielded the highest mean AUROC value were used to train the models for each single task.

Each model performed cross-validation with 4 splits to optimize the computational architecture. The splitting of the data was based on individual trials. In each iteration, three of the four splits were used as training data and one split was used as testing data. This procedure yields four accuracies per model per task, which provides accuracy distributions, respectively. The evaluation method employed to compare the following models involves computing the average accuracy values they generate. Subsequently, we gauged the significance of these results by comparing them to a model trained with random labels. This comparative analysis serves a dual purpose—it not only facilitates the assessment of model performance but also effectively addresses the challenge posed by data imbalance.

(1) Linear discriminant analysis (LDA)

Linear discriminant analysis (LDA) is a linear classification approach using linear decision boundaries [33]. Due to its computational simplicity, it is widely used in BCI implementations. With adequate shrinkage, a large number of variables can be modeled. Moderate accuracies can be achieved in some BCI problems simply by bypassing EEG signal values as variables for parameter fitting in LDA. For this experiment, feature engineering and selection was preferred over direct voltage modeling. To select proper EEG features for this classification method, the following feature engineering pipeline was performed for each task; 1664 features per trial were extracted with the "MNE-features" library. To find the most discriminant features, all features were sorted according to their capability of capturing decision-making knowledge with a decision tree algorithm [34]. Then, the number of most-used discriminant features had to be evaluated. Therefore, it was iterated several times over the number of best features used. The validation accuracy using the LDA model and a validation size of 25 percent of the whole data was observed for every number of best features and for every iteration. The number of features that yielded the best LDA result most often corresponds to the optimal number of features used. For the affect classification task with the warm light condition, 4 of the best features were used, and for the affect classification task with cold light condition, 5 of the best features were used. For the warm vs. cold classification task, 17 of the best features were used.

The LDA model does not take the participation number as input, because it cannot adapt to the participants data as a DNN could.

(2) Spatial Filter Model

The Spatial Filter Model extracts spatial filters for discriminating between classes with a convolutional layer that operates through channels. Spatial filtering with 30 filters (found through hyperparameter search) was performed with a two-dimensional convolutional layer and a kernel size of 34 × 1. This ensured convolution through the channels, i.e., through space. Finally, a linear classification layer was added to classify the spatial filters with a logarithmic softmax activation function. A negative log-likelihood loss function was used to train the model.

(3) CNN Model

Besides two-dimensional convolutional layers, CNNs incorporate non-linearity layers, dropout layers, two-dimensional pooling layers, and fully connected layers [35]. These layers were stacked twice in the stated order until the trial became flattened and a fully connected layer with a non-linear activation function followed that formed the output of the network. The hyperparameter search yielded the best results at 100 filters, 100 nodes, and a dropout rate of 0.0. For the non-linearity layers within the network, ReLu activation functions were used and a softmax activation function was used for the output layer. For training, a negative log-likelihood loss function was used.

(4) EEGNet Model

The EEGNet is a compact convolutional neural network designed for a variety of EEG-based BCI paradigms. It combines depthwise, separable convolutions and thus extracts common EEG features. It possesses the ability to condense a selection of established EEG feature extraction techniques, including the optimal spatial filter bank. One notable benefit of this architecture is its adaptability to a limited dataset during the training process, enabling the generation of distinguishable decoder features [36,37]. A hyperparameter random grid search yielded the best results at a dropout rate of 0.5, 4 spatial filters (34 × 1), 8 temporal filters (1 × 250), and 16 pointwise filters (1 × 16).

(5) SincNet Model

The SincNet, designed by Ranvelli et al. [38], is a CNN that tries to incorporate more meaningful features by using prior knowledge about the data and thus constraining the shape of the filters. The filters consist of Sinc functions that implement bandpass filters. Only the low and the high cut-off frequency have to be learned by the network. Thus, the learned parameters of the Sinc layer might have a physical meaning [38]. The SincNet code was only accessible for one-dimensional audio data, and the code had to be reasonably adjusted to fit two-dimensional EEG data. Also, the range of frequency to be evaluated by the first Sinc layer was changed to 4–125 Hz. The kernel size was chosen to be 125 Hz to catch frequencies of 4 Hz and higher. This model provides an explainability layer that indicates frequency bands. This has been proposed as a method of explaining learned parameters in contexts commonly used in quantitative analysis of EEG, such as frequency bands.

3. Results

Each of the three complementary research questions was approached with different machine learning training procedures or statistical analysis of previous results. Each classification task can be seen as solving complementary problems for the various experimental conditions. The first two tasks are based on using the data of the respective experimental categories, warm or cold, and predicting affect from EEG. The first two tasks were executed to answer the first research question. The third task classifies between experimental conditions, providing a measure of the amount of information about the visual stimulus that can be predicted from EEG, and at the same time providing a baseline for a comparison of models that predict affect from the same data. This task was performed to answer the second research question. Lastly, the model classification accuracies and any changes in it between experimental conditions were compared to answer the third research question. Together, these three experiments can form a data-driven baseline for a functional analysis of the experimental conditions.

3.1. Affect Classification

Figure 5 illustrates the accuracy results for the affect classification task with warm light influence and Figure 6 presents the accuracy results for the affect classification task with cold light influence. The distributions were obtained by training the model four times each. The purple distribution represents training the model with the EEG trials as input and the class labels as output. The green distribution represents the model trained with the same input but shuffled output, and is thus called the "Permutation Test Model". The darker purple and green lines indicate the means of the distributions, respectively. The red line indicates the chance level, without considering class imbalance. This means that each outcome has the same probability, which corresponds to, e.g., 50% per class if two classes are given. The blue line indicates the naive model level, which is the chance level considering the class imbalance. This means that each output comes with a probability according to its occurrence. If, e.g., two output classes exist, where 70% of the samples belong to one class, the naive model level will be 70% for that class. Table 1 shows the corresponding mean accuracies and whether their distributions reached a significant level. Significance was tested with the paired samples t-test between results distribution and permutation test distribution. The significance level was chosen at $p = 0.05$. The CNN and the EEGNet performed significantly above the permutation test model within the warm light condition, although only the EEGNet performed on average above the naive model (see Figure 6). The LDA and the SincNet model performed on the mean worse than the corresponding permutation test model. The CNN and the SincNet show very broad distributions compared with the rest of the models. All the models and their permutation test models, apart from the SincNet and the CNN permutation test model, performed above the theoretical chance level.

Within the cold-light condition, no model performed significantly better than the permutation test model, although the LDA model, EEGNet, and SincNet performed on the mean above the permutation test model (see Figure 6). LDA model results and Spatial Filter model results performed on the mean above the naive model, while only the CNN model performed on the mean below the chance level. The Spatial Filter Model, CNN, and the SincNet permutation test distributions are very broad compared with the rest of the results.

In both Figures 5 and 6, it can be seen that the permutation test for the simpler models (LDA model and Spatial Filter Model) resulted in higher mean accuracies than the permutation test for the models with a higher number of learned parameters (CNN, EEGNet, SincNet).

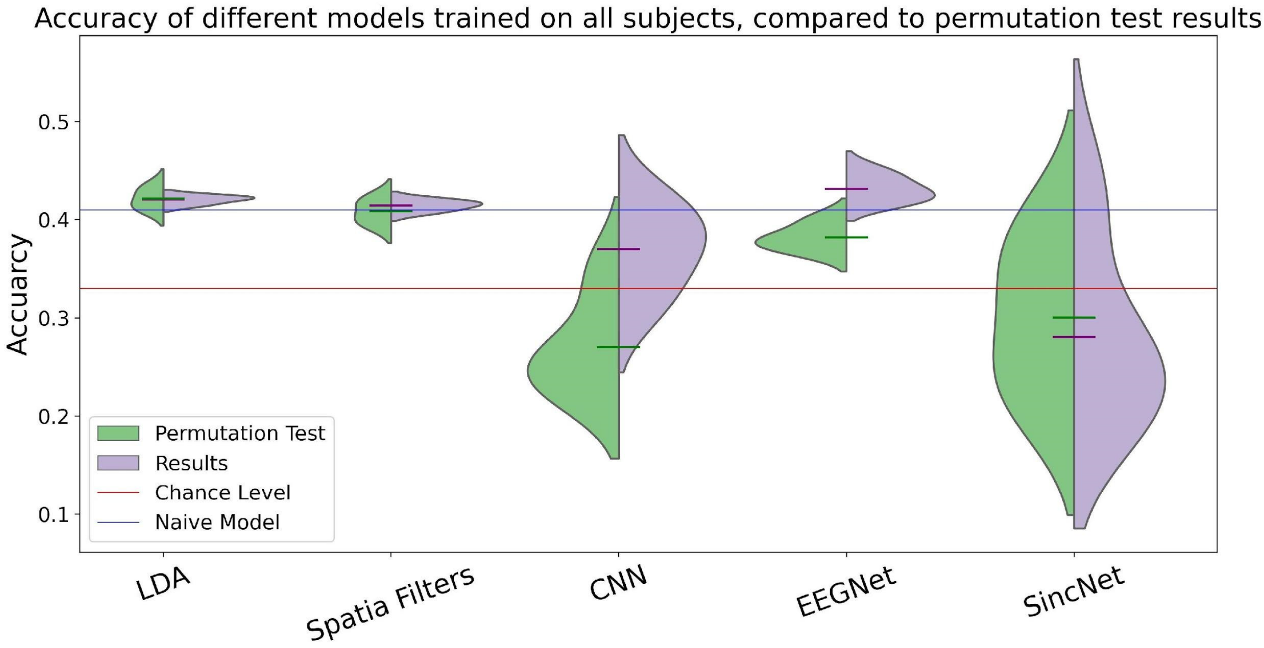

Figure 5. Results for affect-classification task under the influence of warm light (purple) compared with results for the permutation test (green). Models are trained on subsets with trials from all participants.

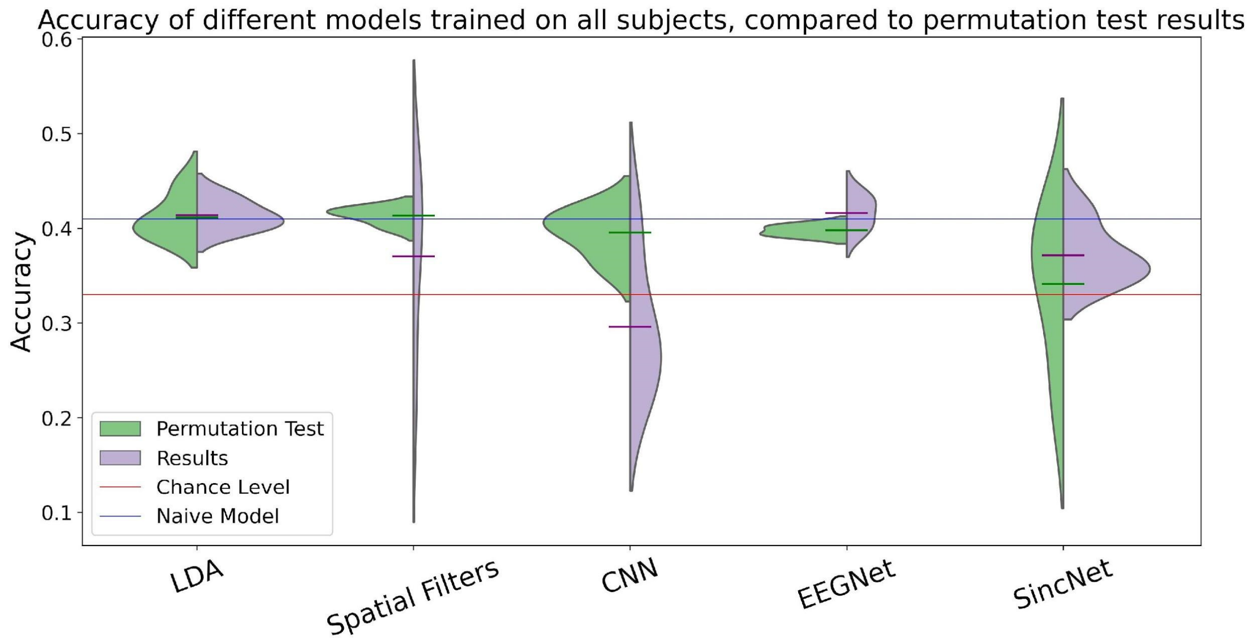

Figure 6. Results for affect-classification task under the influence of cold light (purple) compared with results for the permutation test (green). Models are trained on subsets with trials from all participants.

Table 1. Mean accuracy values for all tasks and all models. Bold values correspond to the distributions that perform significantly better than their permutation test model.

	Emo-Warm	**Emo-Cold**	**Warm vs. Cold**
LDA	0.4213	0.4136	**0.5732**
Spatial Filters	0.4143	0.3705	0.5008
CNN	**0.3698**	0.2958	**0.5567**
EEGNet	**0.4312**	0.4163	**0.7666**
SincNet	0.2804	0.3714	0.5014

3.2. Warm vs. Cold Light Classification

Figure 7 shows accuracy results for the light-classification task. The LDA model, Spatial Filter Model, and EEGNet performed significantly, and all other models performed on the mean, above the permutation test model (see Table 1). The EEGNet shows the best results for this task with a classification accuracy of 76.66% and all models performed on the mean above the naive model. Parts of the permutation test distributions lie at the level of the naive model and the chance level, respectively, for all models. The Spatial Filter model, the CNN, and the EEGNet show the broadest distributions compared with the results of the other models.

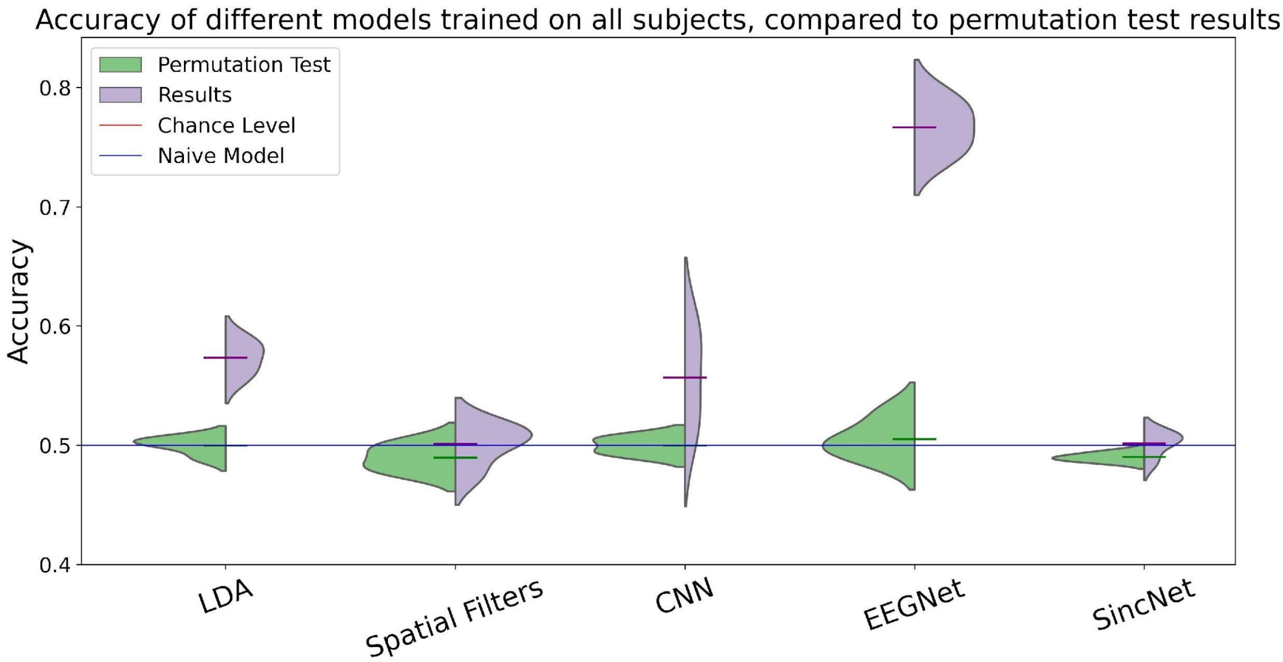

Figure 7. Accuracy results for warm vs. cold task (purple) compared with results for the permutation test (green). Models are trained on subsets with trials from all participants.

3.3. Warm Affect Classification vs. Cold Affect Classification

Figure 8 illustrates accuracy distributions for warm-light affect classification versus cold-light affect classification. No model showed significant differences in affect classification tasks, while classification in warm light resulted in higher mean accuracy levels than classification in cold light in four out of five models. The distribution that yields a higher mean accuracy is narrower than the respective distribution for the opposite task for all models. Only accuracy distributions of the LDA model and the EEGNet performed on the mean above the naive model level for both affect-classification tasks.

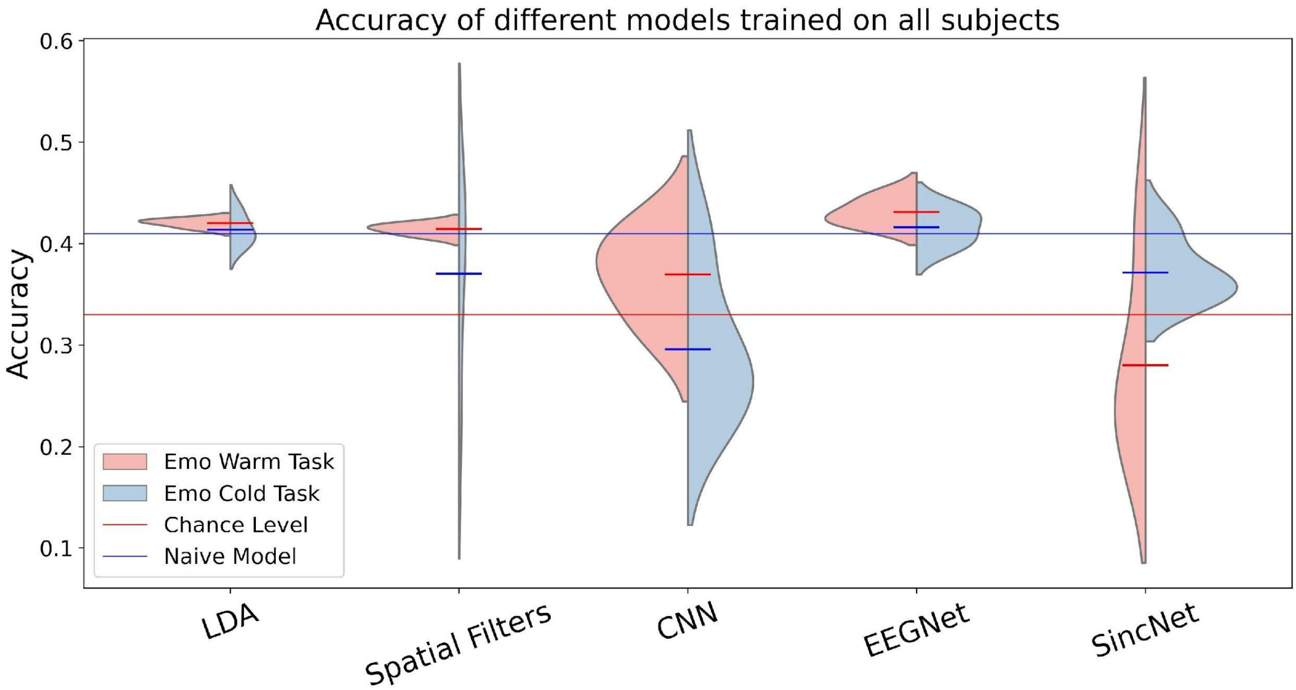

Figure 8. Accuracy distributions for affect-recognition task under the influence of warm (red) and cold (blue) light. Models are trained on subsets with trials from all participants.

4. Discussion

This work investigated, by means of EEG, different affective states under two experimental conditions: a warm light condition and a cold light condition. Three classification tasks arose with regard to the recorded brain data: classification of an affect state within warm light condition, classification of an affect state within cold light condition and warm vs. cold light classification during the observation of affect-eliciting images. In summary, we modeled three complementary classification problems; each of them established a baseline for the next one. In the following section, the results are discussed in relation to the three main research questions.

4.1. Training a Model to Predict Affect States

The LDA model, the Spatial Filter Model, and the EEGNet were able to classify affect states under the influence of warm light on the mean above the naive model, and the CNN (36.98%) and EEGNet (43.12%) were able to classify significantly above the corresponding mean of the permutation test model. As a result, it is possible to train a network to predict an affect state. In the literature, similar affect-classification tasks without light conditions were solved with variants of the EEGNet. They achieved an accuracy of 89.91% for a two-class valence classification and 88.31% for a two-class arousal classification [39]. Qiao et al. [40] reported 99.45% accuracy for a binary valence classification and 99.51% for a binary arousal classification. The named studies performed binary classification on either valence or arousal and are thus not comparable with our study because our study uses three classification labels on the valence–arousal plane. For a four-class affect classification problem, another study reported an accuracy of over 99% for EEGNet [41]. The latter study used video and sound stimulus, which may vary in affect perception over time, and is thus not comparable with our study, which ensures a constant emotion-triggering stimulus.

The LDA and EEGNet demonstrated the capability to classify affective states under the influence of cold light, surpassing both the naive model and the mean of the permutation test model, albeit not in a statistically significant manner. Among the models, EEGNet exhibited the highest performance with a mean accuracy of 41.63%. Consequently, we were unable to establish that training a network to predict affective states under the influence of cold light is feasible. It is important to note, however, that this lack of demonstration does not imply impossibility, as alternative approaches may yield different results. One reason for this could be the lack of concentration induced by the cold light, even though the order of light conditions was randomly assigned. Knez [42] found that short-term memory and problem solving, which is connected to the capacity of concentration, performed better in warm than in cold light.

The permutation test results for the CNN, EEGNet, and SincNet were poorer than the LDA and Spatial Filter Model permutation test results. Notably, the former models were trained using batches, unlike the latter models. Batch training, which involves not seeing all the data and the distribution of class labels simultaneously, might lead the CNN, EEGNet, and SincNet to overlook class imbalances and instead focus on genuine features within the data. Despite the CNN's underperformance relative to the LDA model in the affect classification task under warm-light conditions, it is possible that the CNN learned from the data, whereas the LDA primarily learned from the class imbalance.

In summary, our classification accuracies for affect under cold light conditions were surprisingly low, in contrast to the warmer light conditions. However, the overall findings and the novelty of the paradigm suggest the need for further investigation into this topic, considering the following limitations: Firstly, our study results cannot be directly compared with other studies that primarily use videos as affect-eliciting material, and they may not classify into three affect classes as we did. Secondly, many studies incorporate additional biophysiological signals, such as galvanic skin response (GSR) or heart rate (HR), which can facilitate classification. The intriguing result that subjects perceived affect with lower intensity under cold compared to warm light warrants further exploration, as it could potentially open up a new research field with numerous applications. For instance, one application could involve using cold light in criminal cases to mitigate the impact of images on police officers.

4.2. Training a Model to Predict Light Conditions Based on Affect States

All models demonstrated the ability to classify warm or cold light conditions, yielding average results surpassing both the naive model and the corresponding mean of the permutation test model. Notably, the results of LDA, CNN, and EEGNet significantly outperformed those of the permutation test model. This supports the notion that it is feasible to train a network to predict light conditions based on an affective state. It is worth mentioning that previous studies did not provide classification accuracies or other metrics for result comparison. For instance, Hosseini et al. found no significant difference between colors in EEG signals [24].

These results could be explained by the impact of varying light on eliciting different affective responses or could be due to distinct color processing mechanisms in the brain. The proficiency of the LDA model in this task may be attributed to more distinctly differentiated features compared to the affect classification task. Additionally, non-linear calculations might have been necessary in the latter task to effectively classify the data. The successful classification of light conditions likely stems from the inherent color processing of the light itself.

4.3. Comparing Affective Classification: Evaluating Model Performance across Different Conditions

No significant differences were observed in the classification results under different light conditions. Therefore, we lack evidence to support the notion that affect states are more effectively classifiable in one condition over another, leading us to consider the study's limitations. These findings do not rule out the possibility that the distinctions in classification between light conditions may stem from the mere activation of visual processing areas in the brain. Due to the novelty of our experiment, comparable results were not found in the existing literature. Despite the observed significant differences in each light condition, no distinctions between them were identified. To gain a clearer understanding of color's influences on classification results, future research should incorporate a control condition with a neutral light color. Furthermore, the employed machine learning methods could benefit from enhancements by integrating additional prior knowledge into the models. One potential avenue involves the integration of augmented EEG data, incorporating class-preserving transformations [43]. This approach requires prior domain knowledge to define the transformations to which the network should be invariant. This aligns with the principle of incorporating prior domain knowledge to guide the learning process and improve the overall performance of machine learning models. Additionally, this method addresses the challenge posed by imbalanced data. Balancing classes can be achieved through generating new data, reducing the size of the over-represented class (under-sampling), or applying various statistical methods to compare results without being influenced by this bias.

Another approach that improves classification performance by tackling the class imbalance learning problem might be the use of extreme learning machine (ELM) based on the Bayesian approach, namely a parallel one-class ELM (P-ELM), as proposed by Li et al. [44].

5. Conclusions

This study explores the impact of light on the affective state of individuals viewing emotion-inducing images under two different lighting conditions. The resulting affective states were categorized into three classes within the valence–arousal plane. Consequently, affect classification under two experimental conditions was assessed using four distinct models, including various deep learning models and commonly used linear models in affective brain–computer interface (BCI) research.

For the task of affective recognition, both linear models and the best-performing deep learning models exhibited performance above chance, which was comparable to the naive model. The EEGNet architecture demonstrated superior performance across all tasks. While no significant differences were found between affect recognition in different light conditions, the substantial variation in accuracies achieved by the EEGNet architecture when classifying between experimental conditions underscores the potential of appropriately implemented deep learning models. This points towards the possible refinement of training procedures for deep learning in EEG-based affective recognition.

In the future, the visualization and explanation of learned parameters may offer valuable insights into the neural correlates of experimental conditions. Additionally, the use of Virtual Reality (VR) as a transformative tool in cognitive neuroscience provides an immersive and controlled environment for studying human experiences and affect.

We propose these tools and methods as a possible approach to studying neurophysiology by asking data-driven research questions with deep learning architectures. Being able to classify EEG signals based on properties of the light observed by participants indicates that deep learning architectures can automatically learn features from EEG data. Compared with the accuracy of common modern linear models given selected features, deep learning architectures offer the opportunity for data-driven experiments with the barrier of visualization and interpretation.

Author Contributions: Conceived and designed the experiments: S.Z., A.B.C. and S.C.W. Performed the experiments: S.Z. Analyzed the data: S.Z. and A.B.C. Supervision: S.C.W. and A.B.C. Writing—original draft: S.Z. Writing, review and editing: S.Z. and S.C.W. All authors have read and agreed to the published version of the manuscript.

Funding: Open Access Funding by the Graz University of Technology.

Data Availability Statement: The datasets generated and/or analyzed during the current study are available in the Open Science Framework repository: https://osf.io/gv8k6/ (accessed on 7 September 2023).

Acknowledgments: Thanks to all the study participants for their patience and effort.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bos, D.O. EEG-based emotion recognition. The influence of visual and auditory stimuli. *Comput. Sci.* **2006**, *56*, 1–17.
2. Micoulaud-Franchi, J.A.; Jeunet, C.; Pelissolo, A.; Ros, T. EEG Neurofeedback for Anxiety Disorders and Post-Traumatic Stress Disorders: A Blueprint for a Promising Brain-Based Therapy. *Curr. Psychiatry Rep.* **2021**, *23*, 84. [CrossRef]
3. De Brito, M.A.; Fernandes, J.R.; Esteves, N.S.; Müller, V.T.; Alexandria, D.B.; Pérez, D.I.V.; Slimani, M.; Brito, C.J.; Bragazzi, N.L.; Miarka, B. The effect of neurofeedback on the reaction time and cognitive performance of athletes: A systematic review and meta-analysis. *Front. Hum. Neurosci.* **2022**, *16*, 868450. [PubMed]
4. Lawrence, L.M.; Ciorciari, J.; Kyrios, M. Cognitive processes associated with compulsive buying behaviours and related EEG coherence. *Psychiatry Res. Neuroimaging* **2014**, *221*, 97–103. [CrossRef]
5. Jalilpour, S.; Müller-Putz, G. Toward passive BCI: Asynchronous decoding of neural responses to direction-and angle-specific perturbations during a simulated cockpit scenario. *Sci. Rep.* **2022**, *12*, 6802.
6. Ekman, P. Basic emotions. In *Handbook of Cognition and Emotion*; Dalgleish, T., Power, M.J., Eds.; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2005; pp. 45–60. [CrossRef]
7. Russell, J.A. A circumplex model of affect. *J. Personal. Soc. Psychol.* **1980**, *39*, 1161–1178. [CrossRef]
8. Dadebayev, D.; Goh, W.W.; Tan, E.X. EEG-based emotion recognition: Review of commercial EEG devices and machine learning techniques. *J. King Saud. Univ.-Comput. Inf. Sci.* **2022**, *34*, 4385–4401.
9. Alarcao, S.M.; Fonseca, M.J. Emotions Recognition Using EEG Signals: A Survey. *IEEE Trans. Affect. Comput.* **2019**, *9*, 374–393.
10. Ramzan, M.; Dawn, S. Fused CNN-LSTM deep learning emotion recognition model using electroencephalography signals. *Int. J. Neurosci.* **2023**, *133*, 587–597. [CrossRef]
11. Pinilla, A.; Voigt-Antons, J.-N.; Garcia, J.; Raffe, W.; Möller, S. Real-time affect detection in virtual reality: A technique based on a three-dimensional model of affect and EEG signals. *Front. Virtual Real.* **2023**, *3*, 964754. [CrossRef]
12. Liu, Y.H.; Wu, C.T.; Cheng, W.T.; Hsiao, Y.T.; Chen, P.M.; Teng, J.T. Emotion recognition from single-trial EEG based on kernel Fisher's emotion pattern and imbalanced quasiconformal kernel support vector machine. *Sensors* **2014**, *14*, 13361–13388. [CrossRef]
13. Gonzalez, H.A.; Yoo, J.; Elfadel, I.M. EEG-based emotion detection using unsupervised transfer learning. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 694–697.
14. Ding, Y.; Robinson, N.; Zeng, Q.; Chen, D.; Wai, A.A.P.; Lee, T.S.; Guan, C. TSception: A deep learning framework for emotion detection using EEG. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; pp. 1–7.
15. Yu, M.; Xiao, S.; Hua, M.; Wang, H.; Chen, X.; Tian, F.; Li, Y. EEG-based emotion recognition in an immersive virtual reality environment: From local activity to brain network features. *Biomed. Signal Process. Control* **2022**, *72*, 103349. [CrossRef]
16. Marín-Morales, J.; Higuera-Trujillo, J.L.; Greco, A.; Guixeres, J.; Llinares, C.; Scilingo, E.P.; Alcañiz, M.; Valenza, G. Affective computing in virtual reality: Emotion recognition from brain and heartbeat dynamics using wearable sensors. *Sci. Rep.* **2018**, *8*, 13657. [PubMed]
17. Ding, Y.; Wu, G.; Chen, D.; Zhang, N.; Gong, L.; Cao, M.; Qin, Z. DeepEDN: A deep-learning-based image encryption and decryption network for internet of medical things. *IEEE Internet Things J.* **2020**, *8*, 1504–1518.
18. Gupta, K.; Lazarevic, J.; Pai, Y.S.; Billinghurst, M. AffectivelyVR: Towards VR personalized emotion recognition. In Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology, Virtual Event, 1–4 November 2020; pp. 1–3.
19. Hofmann, S.M.; Klotzsche, F.; Mariola, A.; Nikulin, V.; Villringer, A.; Gaebler, M. Decoding subjective emotional arousal from EEG during an immersive virtual reality experience. *eLife* **2021**, *10*, e64812. [PubMed]
20. Marín-Morales, J.; Llinares, C.; Guixeres, J.; Alcañiz, M. Emotion recognition in immersive virtual reality: From statistics to affective computing. *Sensors* **2020**, *20*, 5163. [CrossRef]
21. Menezes, M.L.R.; Samara, A.; Galway, L.; Sant'Anna, A.; Verikas, A.; Alonso-Fernandez, F.; Wang, H.; Bond, R. Towards emotion recognition for virtual environments: An evaluation of eeg features on benchmark dataset. *Pers. Ubiquitous Comput.* **2017**, *21*, 1003–1013. [CrossRef]
22. Pei, G.; Shang, Q.; Hua, S.; Li, T.; Jin, J. EEG-based affective computing in virtual reality with a balancing of the computational efficiency and recognition accuracy. *Comput. Hum. Behav.* **2024**, *152*, 108085.
23. Suhaimi, N.S.; Mountstephens, J.; Teo, J. EEG-based emotion recognition: A state-of-the-art review of current trends and opportunities. *Comput. Intell. Neurosci.* **2020**, *2020*, 8875426.
24. Hosseini, S.F.; Ghabanchi, Z. What's in a Color? A neuropsycholinguistic study on the effect of colors on EEG brainwaves, immediate emotional responses, and English language vocabulary retention among Iranian young adults. *J. Neurolinguist.* **2022**, *63*, 101083. [CrossRef]
25. Hassib, M.; Braun, M.; Pfleging, B.; Alt, F. Detecting and influencing driver emotions using psycho-physiological sensors and ambient light. In Proceedings of the IFIP Conference on Human-Computer Interaction, Paphos, Cyprus, 2–6 September 2019; Springer: Cham, Switzerland, 2019; pp. 721–742.

26. Bower, I.S.; Clark, G.M.; Tucker, R.; Hill, A.T.; Lum, J.A.; Mortimer, M.A.; Enticott, P.G. Enlarged Interior Built Environment Scale Modulates High-Frequency EEG Oscillations. *Eneuro* **2022**, *9*, 1–16.
27. Schilling, T.; Sipatchin, A.; Chuang, L.; Wahl, S. Tinted lenses affect our physiological responses to affective pictures: An EEG/ERP study. In Proceedings of the 2nd International Neuroergonomics Conference: The Brain at Work and in Everyday Life, Philadelphia, PA, USA, 27–29 June 2018.
28. Peacock, E.J.; Wong, P.T.P. The stress appraisal measure (SAM): A multidimensional approach to cognitive appraisal. *Stress Med.* **1990**, *6*, 227–236.
29. Liu, X. *Methods and Applications of Longitudinal Data Analysis*; Elsevier: Amsterdam, The Netherlands, 2015.
30. Horvat, M.; Dobrinić, M.; Novosel, M.; Jerčić, P. Assessing emotional responses induced in virtual reality using a consumer EEG headset: A preliminary report. In Proceedings of the 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 21–25 May 2018; pp. 1006–1010.
31. Carrington, A.M.; Manuel, D.G.; Fieguth, P.W.; Ramsay, T.; Osmani, V.; Wernly, B.; Bennett, C.; Hawken, S.; Magwood, O.; Sheikh, Y.; et al. Deep ROC analysis and auc as balanced average accuracy, for improved classifier selection, audit and explanation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 329–341. [PubMed]
32. Huang, J.; Ling, C.X. Using AUC and accuracy in evaluating learning algorithms. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 299–310. [CrossRef]
33. Abraham, A.; Pedregosa, F.; Eickenberg, M.; Gervais, P.; Mueller, A.; Kossaifi, J.; Gramfort, A.; Thirion, B.; Varoquaux, G. Machine learning for neuroimaging with scikit-learn. *Front. Neuroinform.* **2014**, *8*, 14. [CrossRef] [PubMed]
34. Szczerbicki, E. Information-based integration for complex systems. In *Knowledge and Information Technology Management: Human and Social Perspectives*; IGI Global: Hershey, PA, USA, 2003; pp. 89–103.
35. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6.
36. Lawhern, V.J.; Solon, A.J.; Waytowich, N.R.; Gordon, S.M.; Hung, C.P.; Lance, B.J. EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces. *J. Neural Eng.* **2018**, *15*, 056013.
37. Yu, Y.; Abadi, M.; Barham, P.; Brevdo, E.; Burrows, M.; Davis, A.; Dean, J.; Ghemawat, S.; Harley, T.; Hawkins, P.; et al. Dynamic control flow in large-scale machine learning. In Proceedings of the Thirteenth EuroSys Conference, Porto, Portugal, 23–26 April 2018; pp. 1–15.
38. Ravanelli, M.; Bengio, Y. Interpretable convolutional filters with sincnet. *arXiv* **2018**, arXiv:1811.09725.
39. Huang, W.; Xue, Y.; Hu, L.; Liuli, H. S-EEGNet: Electroencephalogram signal classification based on a separable convolution neural network with bilinear interpolation. *IEEE Access* **2020**, *8*, 131636–131646.
40. Qiao, Y.; Alnemari, M.; Bagherzadeh, N. A two-stage efficient 3-D CNN framework for EEG based emotion recognition. In Proceedings of the 2022 IEEE International Conference on Industrial Technology (ICIT), Shanghai, China, 22–25 August 2022; pp. 1–8.
41. Zhu, Y.; Ozawa, K.; Kong, W. EEGNetT: EEG-based neural network for emotion recognition in real-world applications. In Proceedings of the 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), Nara, Japan, 9–11 March 2021; pp. 376–378.
42. Knez, I.; Kers, C. Effects of indoor lighting, gender, and age on mood and cognitive performance. *Environ. Behav.* **2000**, *32*, 817–831. [CrossRef]
43. Ma, C.; Wang, H.; Wu, J.; Xue, C. How the Quantity and Hue Contrast of Interface Color Coding Affect Human Perception: Evidence from Two EEG Studies. *Res. Sq.* **2022**. [CrossRef]
44. Li, Y.; Zhang, S.; Yin, Y.; Xiao, W.; Zhang, J. Parallel one-class extreme learning machine for imbalance learning based on Bayesian approach. *J. Ambient. Intell. Hum. Comput.* **2018**. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.