

Article

Decoding Brain Signals from Rapid-Event EEG for Visual Analysis Using Deep Learning

Madiha Rehman ^{1,*}, Humaira Anwer ¹, Helena Garay ^{2,3,4}, Josep Alemany-Iturriaga ^{4,5,6}, Isabel De la Torre Díez ⁷, Hafeez ur Rehman Siddiqui ¹ and Saleem Ullah ¹

- ¹ Institute of Computer Science, Khwaja Fareed University of Engineering & Information Technology, Rahim Yar Khan 64200, Pakistan; humaira.anwer@kfueit.edu.pk (H.A.); hafeez@kfueit.edu.pk (H.u.R.S.); saleem.ullah@kfueit.edu.pk (S.U.)
 - ² Universidad Europea del Atlantico, Isabel Torres 21, 39011 Santander, Spain; helena.garay@uneatlantico.es
 - ³ Universidade Internacional do Cuanza, Cuito EN 250, Angola
 - ⁴ Universidad de La Romana, Edificio G&G, C/Héctor René Gil, Esquina C/Francisco Castillo Marquez, La Romana 22000, Dominican Republic; josep.alemany@uneatlantico.es
 - ⁵ Facultad de Ciencias Sociales y Humanidades, Universidad Europea del Atlántico, Isabel Torres 21, 39011 Santander, Spain
 - ⁶ Departamento de Ciencias de Lenguaje, Educación y Comunicaciones, Universidad Internacional Iberoamericana Arecibo, Arecibo, PR 00613, USA
 - ⁷ Department of Signal Theory, Communications and Telematics Engineering, University of Valladolid, 47011 Valladolid, Spain; itordie@gmail.com
- * Correspondence: madiha.rehman@kfueit.edu.pk

Abstract: The perception and recognition of objects around us empower environmental interaction. Harnessing the brain's signals to achieve this objective has consistently posed difficulties. Researchers are exploring whether the poor accuracy in this field is a result of the design of the temporal stimulation (block versus rapid event) or the inherent complexity of electroencephalogram (EEG) signals. Decoding perceptive signal responses in subjects has become increasingly complex due to high noise levels and the complex nature of brain activities. EEG signals have high temporal resolution and are non-stationary signals, i.e., their mean and variance vary overtime. This study aims to develop a deep learning model for the decoding of subjects' responses to rapid-event visual stimuli and highlights the major factors that contribute to low accuracy in the EEG visual classification task. The proposed multi-class, multi-channel model integrates feature fusion to handle complex, non-stationary signals. This model is applied to the largest publicly available EEG dataset for visual classification consisting of 40 object classes, with 1000 images in each class. Contemporary state-of-the-art studies in this area investigating a large number of object classes have achieved a maximum accuracy of 17.6%. In contrast, our approach, which integrates Multi-Class, Multi-Channel Feature Fusion (MCCFF), achieves a classification accuracy of 33.17% for 40 classes. These results demonstrate the potential of EEG signals in advancing EEG visual classification and offering potential for future applications in visual machine models.

Keywords: BCI; EEG; visual classification; rapid-event design; block design



Citation: Rehman, M.; Anwer, H.; Garay, H.; Alemany-Iturriaga, J.; Díez, I.D.I.T.; Siddiqui, H.u.R.; Ullah, S. Decoding Brain Signals from Rapid-Event EEG for Visual Analysis Using Deep Learning. *Sensors* **2024**, *24*, 6965. <https://doi.org/10.3390/s24216965>

Academic Editor: Patricia A. Broderick

Received: 13 August 2024

Revised: 2 October 2024

Accepted: 28 October 2024

Published: 30 October 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

Electroencephalogram (EEG) imaging is a method used to assess the electrical activity of neurons in the brain. As the brain controls all bodily organs, brain signals change based on an individual's mental state, cognitive processes, visual inputs, and other influencing elements [1,2]. It is well established that brain activity recordings contain specific information about visual object categories [3,4]. However, recognizing object classes in textual or video data is simpler than in brain signals, which is still a challenge for researchers [5,6]. Studies on EEG signal processing have identified the occipital lobe as the region of the brain responsible for visual perception, including the recognition of objects, as well as their

shapes, colors, distances, and materials [7]. The occipital lobe can perform visuospatial processing and associated memory formation within a maximum of 200 ms [7,8]. Due to this biological connection, during EEG signal acquisition, visual stimuli are shown for 2 s, with 1 s breaks between consecutive stimuli. Research has shown that the rapid processing capabilities of the occipital lobe are crucial for accurate and timely visual perception [8]. This understanding has been pivotal in the development of EEG-based systems for object recognition [9].

Brain signals hold manifold information, reflecting a range of motor imagery tasks, emotional processes, sensory/auditory tasks, and cognitive behaviors [10,11]. This information can be utilized for a variety of endeavors, like for the recognition of emotions [12–14], sleep stages [15–17], prediction of critical thinking [18], speech activity detection in mute patients [19,20], etc. Various techniques are utilized for brain signals collection like fMRI (functional magnetic resonance imaging) [21,22], PET (positron emission tomography) [23], ECoG (electrocorticography) [24], MEG (magnetoencephalography) [25], and EEG [26]. FMRI and PET data provide great spatial resolution, but due to their lack of a temporal nature, cannot be used for visual object recognition [21,23]. The ECoG technique yields data with excellent temporal and spatial resolution but is highly invasive, as it requires the electrodes to be placed directly on the brain and not on the scalp [24]. MEG is used to measure the magnetic fields around the brain and is conducted inside a shielded room environment to avoid external electromagnetic noise. It provides high temporal and spatial resolutions and is a noninvasive technique [25]. However, due to the high cost and immobility of MEG, these devices are hard to use [27]. EEG also offers data with high temporal resolution that are well suited for object classification tasks [25].

EEG signals are classified into various frequency bands, like alpha (8–12 Hz), beta (13–25 Hz), theta (4–7 Hz), and gamma (30–80 Hz). Alpha and theta frequencies correspond to a person's relaxed state with their eyes closed. Beta and gamma bands are known for recognizing critical thinking, problem solving, and visual recognition in the brain [28] and are mostly used in EEG visual recognition tasks [29–32].

The Temporal Stimulation Design (TSD) for signal acquisition greatly affects the EEG signal. Studies have shown that TSD is performed using a block design or a rapid-event design [33,34]. In the former, a person is continuously shown a block of images from the same class without any rest between images for the signal waveform to return to its baseline, distorting the next waveform [35,36]. In the latter design, the person is shown random images from different classes with an interval between each image. This interval is provided so that the excited neurons can reach the baseline so as not to interfere with the waveform of the next signal [37,38].

The block design in temporal stimulation is helpful in the detection of brain activity, such as epilepsy, neural activity in a brain part, tumors, and motor imagery signals [39–41]. The rapid-event design is well suited for classification tasks to distinguish one brain signal from another [33].

A possible reason behind the low accuracy in the visual classification task is that the EEG signals are non-stationary [35,42,43]. Their statistical features, such as mean and variance, change over time. A study by Miladinović et al. [44] indicated that the non-stationarity of EEG signals can cause shifts in feature covariance with time. To effectively capture and understand these signal-shifting dynamics, sophisticated analytical techniques are needed.

The rest of this paper is structured as follows: Section 2 provides an overview of related work, highlighting the datasets used in this task to date, as well as their usages and design techniques. In Section 3, we provide a detailed description of the dataset, as well as the data processing and feature extraction processes and the proposed classifiers. Section 4 presents the experimental results in detail. In Section 5, a detailed discussion about the experiments, the achieved results, and comparisons with state-of-the-art approaches is presented. Finally, Section 6 concludes the paper, summarizing key findings and discussing future research.

2. Related Work

In 2017, Spampinato et al. [34] published an article on the classification of visual objects through EEG signals using their self-made dataset. The authors reported an accuracy of 93.91%. Subsequently, the code and data were made publicly available, leading to numerous publications [33,34,45–51]. All of these studies incorporated the same data and achieved improvements in accuracy up to 97.13%.

In 2020, Hammad et al. [52] and Renli et al. [33] used the same dataset and code initially released in [34], and claimed that the achievement of high classification accuracy on this task is not due to the model architecture or the EEG signal but the following:

1. Usage of a block design during signal acquisition.
2. No preprocessing employed, i.e., usage of unfiltered data, resulting in training on noisy data.
3. Test data are sourced from the same block as the training data.

Hammad et al. [45,52], Renli et al. [33], and Hari [32] substantiated their points by the following means:

1. EEG data collected on a set of object classes/images identical to that utilized by Spampinato et al. [34];
2. Application of the same block design technique;
3. Adoption of similar preprocessing methods;
4. Utilization of test data sourced from the same block as the training data.

Following these steps, they achieved the highest classification accuracy in KNN, i.e., 100%. Other models like SVM, MLP, 1D CNN, and LSTM also performed very well, proving that a block design results in an accuracy boost in the object classification task. The authors concluded that in a block design, the rise in electrical potential in the brain is not given enough time to return to its baseline, contaminating the waveform of the next signal [33,45,52].

Hammad et al. [45,52] and Renli et al. [33] gathered data using a rapid-event temporal stimulation design with the same object classes as used before and performed preprocessing; the achieved results achieved astonishing, i.e., the accuracy degraded to 5.6%. Renli et al. [33] claimed that the data collected, used, and released in [34] suffer from irreparable contamination, i.e., all the image signals are contaminated by the next signal.

With the use of EEG signals in an event-related design rather than a block design, the accuracy of the signal is severely compromised. Researchers [29,31,53–56] have collected data for the EEG classification task using the rapid-event approach, varying the number of classes and incorporating different object classes. All of these studies [5,32,47,50,57–59] suggest that higher accuracy in EEG object classification is possible for only a low number of classes using rapid-event temporal stimulation. The authors of [45] published comments on [46] in *Transactions on Pattern Analysis and Machine Learning*, substantiating the point that using test data from the same block as the train data resulted in high accuracy in block design. Using the same data and code, only applying cross validation by leaving one block out in each turn, resulted in very low accuracy.

Various studies have been conducted to collect data using rapid-event and block designs on different object classes. A summary of EEG dataset collection efforts for EEG visual classification tasks are provided in Table 1. Datasets collected over time have used a variety in stimuli, temporal stimulation designs, numbers of classes, numbers of images per class, numbers of subjects, devices, numbers of channels, and sampling rates. Classification accuracies achieved on varying datasets is presented in Table 2.

Table 1. Summary of datasets used for EEG visual classification.

Ref#	Name of Dataset	Journal and Year	Stimulus	TSD	No. of Classes	No. of Images/Clips per Class	No. of Subjects	Device/No. of Channels	Sampling Rate
[52]	ImageNet D1	Journal 2021	Image	Rapid Event	40	1000	01	BioSemi ActiveTwo recorder 104	4096 Hz
[34]	ImageNet D2	Journal 2017	Image	Block Design	40	50	06	ActiCap 128	1000 Hz
[33]	ImageNet D3	Journal 2021	Image	Rapid Event	40	50	6	BioSemi ActiveTwo recorder 104	4096 Hz
[33]	ImageNet V1	Journal 2021	Video	Rapid Event	12	32	06	BioSemi ActiveTwo recorder 104	4096 Hz
[53]	Stanford Dataset D4	Journal 2015	Image	Rapid Event	6	12	10	EGI HCGSN 128	1000 Hz
[29]	MPI DB D5	Journal 2010	Image	Rapid Event	3	4	04	ActiCap System 64	500 Hz
[54]	Things D6	Journal 2022	Image	Rapid Event	1854	10	10	Easy Cap 64	1000 Hz
[31]	ImageNet D7	Journal 2023	Image	Rapid Event	4	10	04	ActiCHamp 32	1000 Hz
[55]	MNIST D8	Journal 2024	Image	Rapid Event	11	116	01	Emotiv EPOC 14	128 Hz
[56]	Human dataset D9	Journal 2017	Image	Rapid Event	5	12	16	Easycap 74	1000 Hz
[60]	EEG-ImageNet D10	Journal 2024	Image (coarse-grained)	Block Design	40	50	16	-	1000 Hz
[60]	EEG-ImageNet D11	Journal 2024	Image (fine-grained)	Block Design	40	50	16	-	1000 Hz

Table 2. Summary of all the relevant literature.

Ref #	Year	Type	Dataset Utilized	Classes	TSD	Classifier	Accuracy
[60]	2024	Journal	EEG-ImageNet D11	40	Block Design	SVM MLP EEGNet RGNN	77.84% 81.63% 36.45% 70.57%
[60]	2024	Journal	EEG-ImageNet D10	40	Block Design	SVM MLP EEGNet RGNN	50.57% 53.39% 30.30% 47.03%
[61]	2024	Journal	ImageNet D1	40	Rapid Event	EEGVis_CMV (from EEG to Image)	17.9%
[57]	2024	Journal	MNIST D8	11	Rapid Event	RieManiSpectraNet	55%
[58]	2024	Journal	Human dataset D9	05	Rapid Event	LDA	68.75%

Table 2. Cont.

Ref #	Year	Type	Dataset Utilized	Classes	TSD	Classifier	Accuracy
[59]	2023	Journal	Stanford Dataset D4	06	Rapid Event	LSTM SVM	55.55% 66.67%
[31]	2023	Journal	ImageNet D7	04	Rapid Event	SVM CNN LSTM-CNN EEGNet	36.22% 64.49% 65.26% 79.29%
[54]	2022	Journal	Things D6	1854	Rapid Event	AlexNet ResNet-50 CORnet MoCo	15.4% 16.25% 21.05% 12.40%
[50]	2022	Journal	ImageNet D2	40	Block Design	SVM RNN-based Model Siamese Network Bi-LSTM HDRS-STF BiLSTM+AttGW	82.70% 84.00% 93.70% 92.59% 99.78% 99.50%
[5]	2023	Journal	Stanford Dataset D4	06	Rapid Event	LDA ShallowConvNet EENet LSTM EEG-Conv Transformer TSCNN	40.52% 46.51% 43.83% 38.06% 52.33% 54.28%
			Max Plank Institute Dataset [MPI DB]	03	Rapid Event	LDA ShallowConvNet EEGNet LSTM TSCNN	76.11% 77.42% 77.79% 60.61% 84.40%
[32]	2023	Journal	ImageNet D1	40	Rapid Event	LSTM k-NN SVM MLP 1D CNN EEGNet SyncNet	2.3% 2.1% 3.0% 2.8% 2.4% 17.6% 3.7%

Table 2. Cont.

Ref #	Year	Type	Dataset Utilized	Classes	TSD	Classifier	Accuracy
[45]	2022	Journal	ImageNet D2	40	Block Design	LSTM k-NN SVM MLP 1D CNN EEGNet SyncNet EEGChannelNet	2.7% 3.6% 3.0% 3.7% 3.3% 2.5% 3.8% 2.6%
[33]	2021	Journal	ImageNet D3	40	Rapid Event	LSTM k-NN SVM MLP 1D CNN	2.9% 3.2% 3.0% 3.7% 3.3%
[52]	2021	Conference	ImageNet D1	40	Rapid Event	1D CNN LSTM SVM k-NN	5.1% 2.2% 5.0% 2.1%
[62]	2019	Journal	ImageNet D2	40	Block Design	LSTM k-NN SVM MLP 1D CNN	63.1% 100% 100% 21.9% 85.9%
			ImageNet D3	40	Rapid Event	LSTM k-NN SVM MLP 1D CNN	0.7% 1.4% 2.7% 1.5% 2.1%
[46]	2020	Journal	ImageNet D2	40	Block Design	Inception v3 (from signals to images)	94.4%
[47]	2019	Journal	ImageNet D2	40	Block Design	Proposed LSTM-B	97.13%
[48]	2018	Conference	ImageNet D2	40	Block Design	Proposed Bidirectional LSTMs	97.3%
[49]	2018	Conference	ImageNet D2	40	Block Design	Proposed Region-level bi-directional LSTM	97.1%
[34]	2017	Conference	ImageNet D2	40	Block Design	GoogleNet VGG Proposed Method	92.6% 80.0% 89.7%

A systematic summary of literature review considering the used dataset, applied ML models, and achieved accuracy is provided in Table 2. Higher accuracy trends can be seen in the data when a block design is used and when a lower number of classes is applied in a rapid-event design. A block design applied to 40 classes achieved a maximum of 100% accuracy, while a rapid-event design applied to the same classes but with a differing class image achieved maximum accuracy of 5.6%. During this extensive literature survey, we have identified the following several factors that cause low accuracy in the EEG visual classification task:

1. Determination of the optimal number of object classes to increase accuracy, as a low number of classes results in higher accuracy and vice-versa;
2. Lack of exploration of the use of a rapid-event design versus a block design during EEG signal acquisition [32,45,52,62];
3. Selection of channels, which contributes to accuracy boosts using channel selection techniques such as the linear removal of channel drops in accuracy to chance, as reported in [27,32].
4. Ensuring accurate labeling of data as incorrect or arbitrary labeling of events in block and rapid-event designs, resulting in accuracy boosts, reported in an analysis by Ren Li et al. [33] (page 318, Section 2 point e).
5. Implementation of effective preprocessing techniques to enhance data quality, as raw data result in higher accuracy than filtered data, as reported in [32,34,45,52].

3. Materials and Methods

The proposed methodology, as presented in Figure 1, encompasses many essential steps in the processing and classification of EEG signals. First, the dataset is separated into EEG signals and annotations. Then, the signals are rereferenced to the mastoids to remove noise and artifacts. The signal is then bandpass-filtered to eliminate undesired frequencies. Subsequently, a notch-filtering technique is employed to eliminate any power-line interference. The EEG signals are divided into epochs, from which features are derived. Feature selection is a process that determines the most significant features for classification. This is accomplished using techniques like filters. Ultimately, classifiers are trained, then compared in order to evaluate their performance using metrics like accuracy and sensitivity. These metrics are crucial for the assessment and comparison of different classification methods.

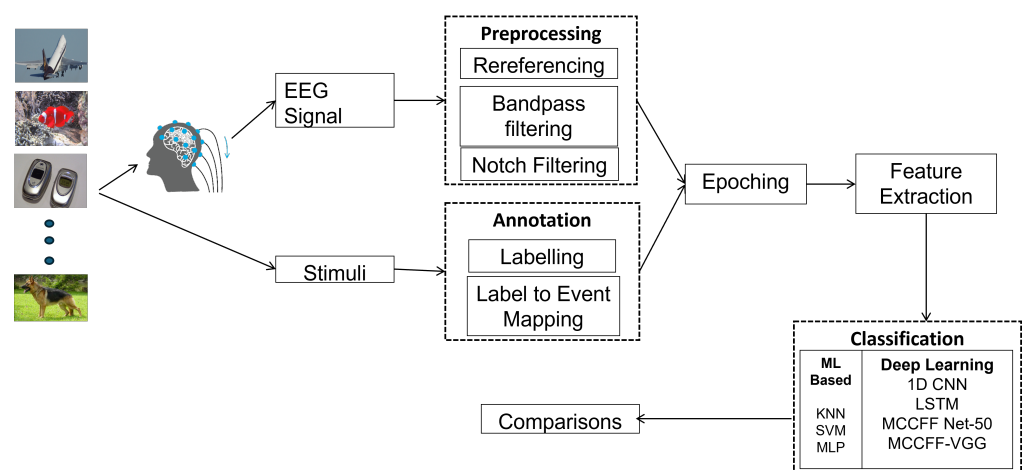


Figure 1. Diagram of the proposed methodology.

3.1. Dataset Description

The dataset used in this study comprises publicly available data originally collected by [52]. In this research, we refer to it as ImageNet comprises a subset of images taken from the ILSVRC (ImageNet Large-Scale Visual Recognition Challenge) dataset. This is one of the largest EEG Signal datasets. Comprehensive details about the dataset are provided

in Table 3. EEG signals were recorded while subjects viewed image stimuli from random object classes.

Table 3. Image-Net EEG data collection.

Device	BioSemi ActiveTwo recorder
Number of Subjects	1
Visual Stimuli	ILSVRC-2021
Total Classes	40
Images per Class	1000
Duration of Visual Stimuli	2 s with 1 s blanking
Sampling Frequency	4096 Hz
Data Resolution	24 bits
Temporal Stimulation Design	Rapid Event design

The dataset comprises of 40 classes with 1000 images each. The complete dataset is built on 40,000 images. Figure 2 shows the object classes used as visual stimuli. For each session, 10 images are randomly selected from each class, resulting in 100 sets of 400 images each. During each 1440 s session, the subject viewed 400 randomly ordered images. A total of 100 sessions were conducted, each approximately 1440 s in duration, employing a rapid-event temporal design. Figure 3 represents the baseline of these sessions. Each session begins and ends with 10 s of blank, followed by 2 s of visual stimulus and 1 s of a blank screen. The blank screen, displayed as a black screen, allows the subject's brain signal to return to baseline before presenting the next image. This method ensures that the signal from each new stimulus does not interfere with the preceding signal.

```
[ 'giant panda', 'parachute', 'mailbag', 'iron', 'jack-o-lantern', 'pajama', 'radio telescope', 'African elephant', 'German shepherd', 'lycaenid', 'golf ball', 'cellular telephone', 'revolver', 'grand piano', 'anemone fish', 'folding chair', 'running shoe', 'convertible', 'capuchin', 'missile', 'espresso maker', 'bolete', 'pizza', 'broom', 'coffee mug', 'mountain tent', 'canoe', 'sorrel', 'Egyptian cat', 'electric locomotive', 'mitten', 'pool table', 'electric guitar', 'banana', 'reflex camera', 'desktop computer', 'watch', 'daisy', 'airliner', 'mountain bike' ]
```

Figure 2. Classes used as visual stimulus.



Figure 3. Timeline of the visual stimuli shown to subjects.

3.2. Preprocessing and Feature Extraction

The raw EEG data from 99 brain data format (BDF) files were initially unprocessed. To manage computational resources effectively, the MNE library in Python was employed to read files in batches of two. Each file contained approximately 1440 s of EEG signals from 105 sensory positions, consisting of 104 channels and 1 stimulus channel providing event onset information. Separate event files were maintained using the same visual object sequence shown to the subjects, which was unique for each file/setting. The data were sampled at 4096 Hz, resulting in $((400 \times 2 \text{ s} + 1 \text{ s}) \times 4096) = 4,915,200$ time points, with an additional $(10 \text{ s} + 10 \text{ s}) \times 4096 = 81,920$ time points of start and end session blanks. The processing steps after reading the file include the following:

1. Raw EEG data are rereferenced to the mastoids to remove external noise and artifacts.
2. The data are bandpass-filtered by applying a zero-phase FIR filter from the MNE library. This filter eliminates phase shifts and gradually cuts off frequency components below 14 Hz and above 71 Hz, so no ringing artifacts remain in the signal.
3. A notch filter at 49–51 Hz is applied to remove any power-line noise.

4. The data are then epoched based on events, starting from -0.5 s and ending at 2.5 s for each event. The length of the epochs retrieved at this stage for a batch of files is $4,997,120 \times 104$. A total of $4,997,120$ of the time points are considered rows, while 104 represent the sensory positions.
5. Events corresponding to 400 visual stimuli are extracted from the stimulus channel and assigned unique class labels for all 40 classes.
6. The data are then annotated using the unique labels and epoch data.

After signal preprocessing, the subsequent crucial step is feature extraction. This step is particularly significant due to the inherent complexity of EEG signals, i.e., they are non-stationary, non-linear, and non-Gaussian [30,63]. Given the temporal nature of EEG signals, preserving their effectiveness requires a statistical feature analysis and extraction approach. EEG is recorded at a 4096 Hz sampling frequency. This constitutes a huge amount of data that cannot be fed directly into any model. Therefore, in order to group the data while preserving effectiveness, various statistical methods were evaluated to determine the most effective, such as the mean, standard deviation, and mean of absolute values of first and second differences [63,64] were applied. The best results were achieved by performing the standard deviation on each of the channels per epoch.

3.3. Proposed Classifiers

This study employed seven different classifier models, each with an architecture optimized for EEG data. The machine learning learning models include k nearest neighbor (KNN), support vector machine (SVM) and multilayer perceptron (MLP). Deep learning models used in this study include are 1D Convolutional Neural Networks (1D CNNs), Long Short-Term Memory (LSTM), and the proposed MCCFF with ResNet and VGG models. We only adopted the ResNet-50 and VGG architectures, i.e., the models were trained on the dataset from scratch. The input shape to the deep learning models is a 3D matrix consisting of events, channels, and features. Owing to their long-term dependencies the 1D CNN and LSTM architectures are excellent choices for time-series data. The parameters for all these models are presented in Table 4.

Table 4. Parameters of the classifiers used for EEG data after feature extraction.

Classifier	Parameters
KNN	$k = 5$
SVM	kernel = 'poly', $C = 20$, random_state = 1, gamma = 1, probability = True, class_weight = 'balanced'
MLP	Hidden_layers = 1500, Max_iterations = 2000, random_state = 42
1D CNN	Learning rate = 0.0005, batch size = 100, epochs = 200, optimizer = Adam, loss = sparse_categorical_crossentropy, metrics = Accuracy, no of Layers = 15, activation = Relu, Softmax
LSTM	Learning rate = 0.005, batch size = 100, epochs = 200, optimizer = Adam, loss = sparse_categorical_crossentropy, metrics = Accuracy, no of Layers = 14, activation = Relu, Softmax
MCCFF Net-50	Learning rate = 0.005, Batch size = 120, epochs = 150, optimizer = Adam, loss = sparse_categorical_crossentropy, metrics = Accuracy, no of layers = 51, activation = Relu, Sigmoid
MCCFF VGG	Learning rate = 0.001, Batch size = 100, epochs = 200, optimizer = Adam, loss = sparse_categorical_crossentropy, metrics = Accuracy, no of layers = 16, activation = Relu, Sigmoid

The best results with the KNN classifier were achieved with a k value of 5 . The SVM model used a polynomial kernel with a C value of 20 and gamma set to 1 . The MLP classifier was configured with 1500 hidden layers and a maximum of 2000 iterations. For the 1D CNN model, tuning the learning rate and epochs and incorporating a batch normalization layer with dropout layers significantly improved performance and lowered the chance of overfitting, respectively. The LSTM model adopted a similar approach to the 1D CNN, with a learning rate of 0.005 sparse categorical cross-entropy. It employed a batch size of 100 and 150 epochs, using batch normalization with an Adam optimizer. The number of epochs

and batch normalization were crucial factors in enhancing accuracy. The architectures of the 1D CNN and LSTM are presented in Table 5.

Table 5. 1D CNN and LSTM model architectures.

1D CNN Model			LSTM Model		
Layer	Neural Units/Kernel Size	Activation	Layer	Neural Units	Activation
Conv1D	8 (3, 3)	ReLU	LSTM	8	ReLU
Dropout	0.1	-	Batch Normalization	-	-
Batch Normalization	-	-	MaxPooling1D	-	-
MaxPooling1D	(4, 4)	-	LSTM	16	ReLU
Conv1D	16 (3, 3)	ReLU	Dropout	0.2	-
Dropout	0.2	-	Batch Normalization	-	-
Batch Normalization	-	-	MaxPooling1D	-	-
MaxPooling1D	(4, 4)	-	LSTM	32	ReLU
Conv1D	32 (3, 3)	ReLU	Dropout	0.4	-
Batch Normalization	-	-	Batch Normalization	-	-
Flatten	-	-	MaxPooling1D	-	-
Dense Layer	16	ReLU	Flatten	-	-
Dropout	0.4	-	Dense Layer	32	ReLU
Batch Normalization	-	-	Dense (Output Layer)	41	Softmax
Dense (Output Layer)	41	Softmax			

This study proposes a robust MCCFF model based on the ResNet-50 architecture. The MCCFF ResNet model is capable of handling 40 object classes, each with 1000 images. The model utilizes sets of EEG images as input and employs encoders to extract features in the time domain. The model architecture includes an initial layer of 1D convolution with a kernel size of 7×7 , a stride of 2, and 64 filters, using ReLU as the activation function, followed by batch normalization and a dropout. The fourth and fifth layers use 1D convolutions with a kernel size of 3×3 , a stride of 1, and 64 filters each. Batch normalization and dropout are applied between layers. The following layers use the same convolution with 128 filters, and subsequent layers increase to 256 and 512 filters each. The architecture concludes with a global pooling layer and a fully connected layer. The detailed architecture is diagrammatically explained in Figure 4.

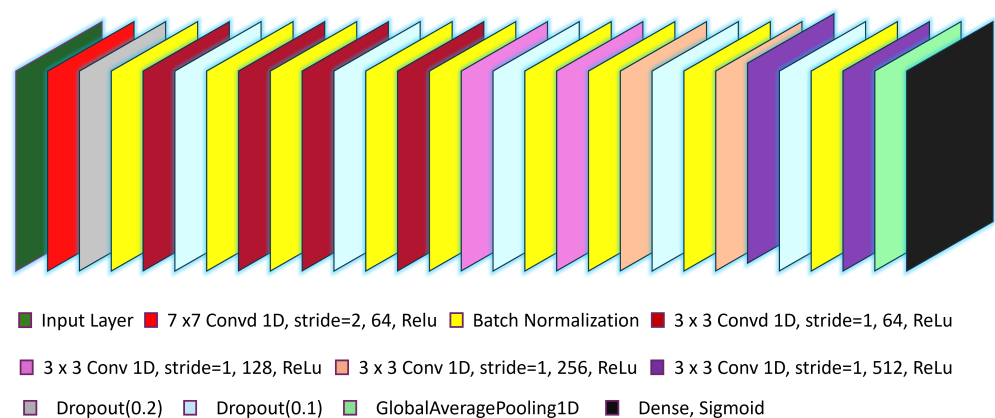


Figure 4. Proposed MCCFF model architecture based on ResNet-50.

The proposed MCCFF VGG model is structured into four blocks of layers. In the first block, there are two layers of 3×3 1D convolutions with 64 filters each, followed by dropout, batch normalization, and a max pooling layer with a pool size of 2 and a stride of 2.

The second block includes one layers of 3×3 1D convolution with 128 filters, followed by dropout, batch normalization, and a max pooling layer with the same pool size and stride. The third block consists of one layers of 3×3 1D convolution with 256 filters, followed by dropout and a max pooling layer with a pool size of 2 and a stride of 2. Each of these blocks uses the ReLU activation function. The final block contains two fully connected layers, each with 4096 dense units and dropout. The detailed architecture is diagrammatically explained in Figure 5.

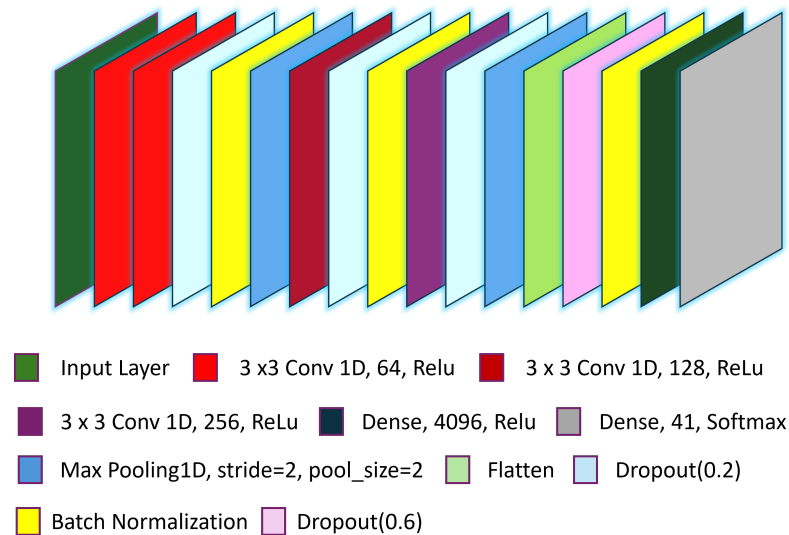


Figure 5. Proposed MCCFF model architecture based on VGG.

4. Experimental Results

This section presents the results of the classification performed in this study. A detailed examination of the dataset was conducted using two different approaches across all seven models, using EEG data without filtering and with filtering. In this study, the models were rigorously validated using a 5-fold cross-validation technique. This technique partitions the dataset into five subsets, where each subset serves as a validation set once, while the remaining subsets are used for training. By rotating through all subsets for validation, this approach provides a comprehensive evaluation of each model's generalizability and effectiveness. This methodological rigor enhances the study's confidence in the reported classification accuracies and ensures that the results are robust and statistically sound.

4.1. Results with a No-Filtering Approach

The complexity of EEG signals necessitates caution in preprocessing, as human-defined methods can potentially degrade signal performance [65,66]. Prior research has shown that achieving high performance on rapid-design EEG is only possible for a low number of classes [5]. In the no-filtering approach, we utilized the complete dataset only by excluding point 2, i.e., bandpass filtering from Section 3.2. All other preprocessing steps were incorporated. The results achieved using this approach are shown in Table 6. It is evident that the proposed MCCFF Net-50 and MCCFF VGG models outperformed the others in terms of the maximum number of channels and window size. However, accuracy degraded to chance with the traditional models, which validates the work reported in [52].

Table 6. Results on non-filtered data. All results were validated using 5-fold cross validation leaving one fold out.

Classifier	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
KNN	2.49	4.96	2.82	4.96
SVM	7.51	4.34	3.93	4.34
MLP	14.31	3.72	2.85	3.72
Proposed Models				
LSTM	21.86	10.97	9.32	10.97
1d-CNN	47.369	15.96	13.591	15.960
MCCFF Net-50	48.11	22.94	20.53	22.94
MCCFF VGG	62.05	33.16	34.59	33.17

4.2. Results with Filtered Data

The data were filtered using the steps outlined in Section 3.2. Table 7 summarizes the averaged results across various classifiers, revealing distinct performance patterns. Traditional models such as KNN, SVM, and MLP achieved relatively lower precision, recall, F1 scores, and accuracies, ranging from approximately 0.4% to 5.59%. In contrast, advanced neural network models like LSTM and 1D CNN exhibited improved performance, with LSTM achieving a precision of 47.96%, recall of 8.25%, F1 score of 5.86%, and accuracy of 8.25%, while 1D CNN demonstrated a precision of 52.42%, recall of 13.0%, F1 score of 11.07%, and accuracy of 12.99%. MCCFF Net-50 achieved a precision of 44.34%, recall of 13.5%, F1 score of 8.73%, and accuracy of 13.50%, while MCCFF VGG achieved a precision of 54.47%, recall of 14.57%, F1 score of 4.31%, and accuracy of 14.57%.

Table 7. Results achieved on filtered data. All the results were validated using 5-fold cross validation.

Classifier	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
KNN	40.0	4.9	3.4	4.96
SVM	5.1	4.34	4.1	4.34
MLP	13.5	5.59	6.51	5.59
Proposed Models				
LSTM	47.96	8.25	5.86	8.25
1d-CNN	52.42	13.0	11.07	12.99
MCCFF Net-50	44.34	13.5	8.73	13.50
MCCFF VGG	54.47	14.57	4.31	14.57

5. Discussion

5.1. Effect of the Sensor Selection Strategy

The bio-semi device used for signal acquisition uses four main sensory channels (A, B, C, and EXG), each equipped with 32 channels, except for EXG, which has 8. We used sequential feature selection (SFS) based on a backward elimination approach for channel selection [27]. Backward elimination eliminates the channels from the backward direction. The first elimination removes the sensory EXG group with channels. In the subsequent eliminations, eight channels were removed from each sensory group (A, B, and C), yielding 72, 48, and 24 channel configurations. In this way, SFS provided comprehensive coverage across all areas of the brain. This selection approach yielded significant results, with higher accuracies for 104 and 96 channels, whereas accuracy dropped notably with 72, 48, and 24 channels. The detailed results are provided in the Supplementary Materials. The accuracy trends of the 1D CNN, LSTM, and MCCFF models across a range of channels are shown in Figures 6 and 7 for unfiltered and filtered data, respectively. Higher numbers of channel configurations yielded superior accuracies across models, with MCCFF VGG achieving accuracies of 14.57% and 7.0% on 104 and 96 channels, respectively, compared to lower accuracies on fewer channels (e.g., 2.75% and 2.5% for MCCFF Net-50 and MCCFF VGG, respectively on 72, 48, and 24 channels). These visualizations provide insights into the various models' behaviors with varying numbers of channels, reinforcing the robustness of the MCCFF architectures in handling the complex temporal characteristics of EEG data.

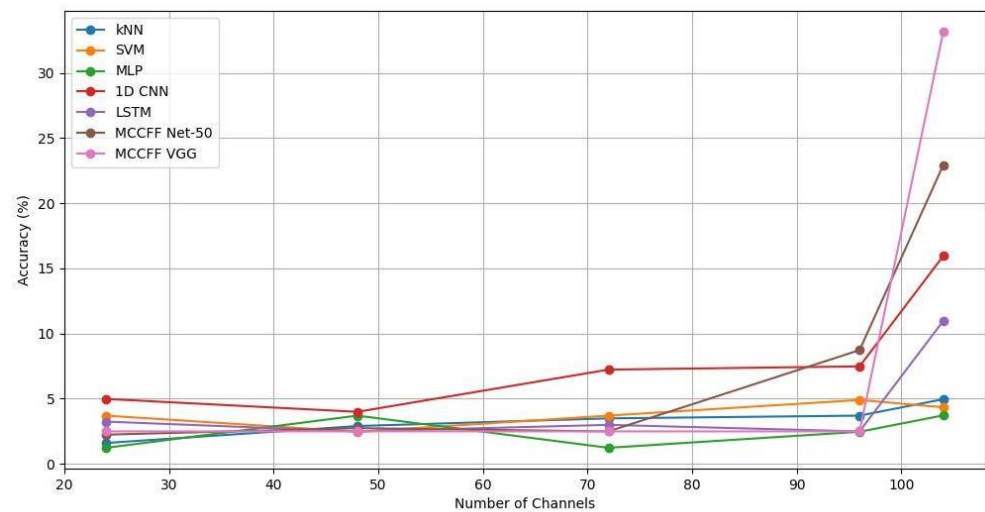


Figure 6. Accuracies of all the models for a 2500 ms time window and varying numbers of channels established on non-filtered data.

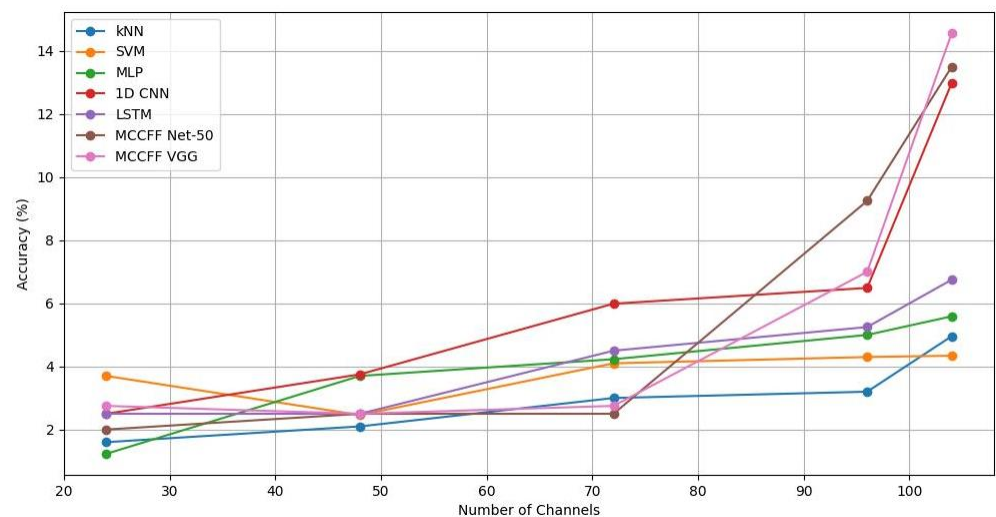


Figure 7. Accuracies of all the models for a 2500 ms time window and varying numbers of channels established on filtered data.

5.2. Effect of Window Sizes on Signal Accuracy

Figure 8 elucidates how varying window sizes influenced classifier performance. Longer window sizes resulted in higher accuracy and vice-versa. The analysis encompassed time-window slices of 2500 ms, 1500 ms, and 500 ms around events. Previous studies have used 1500 ms time windows, starting at 0s from the stimulus onset. In this study, we used time slices of -0.5 s before stimulus onset and 2.5 s, 1.5 s, and 0.5 s after stimulus onset to demonstrate the effects of longer and shorter window sizes on the data. The usage of -0.5 s before stimulus onset helps in capturing the baseline neural activity, thereby completing the signal waveform. This significantly improved the performance of the models. When shorter window sizes of 1.5 s and 0.5 s were used, crucial information needed for accurate recognition was missed, leading to the observed drop in performance for all models shown in Figure 8.

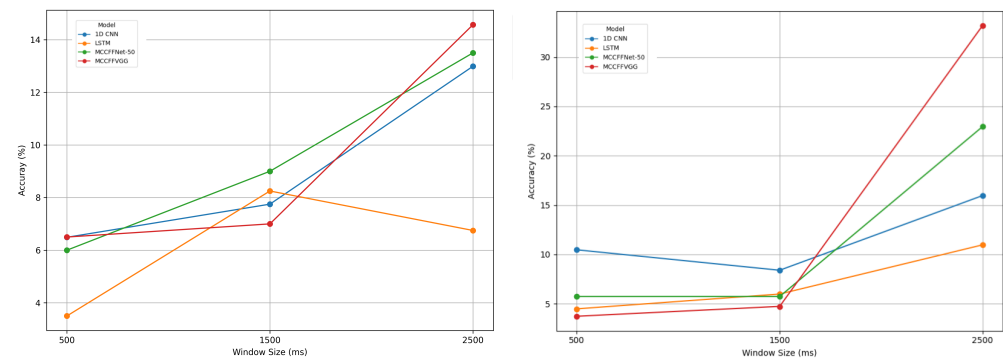


Figure 8. Effects of varying window sizes on filtered data (Left) and non-filtered data (right).

5.3. Comparison

A comparative study of the state of the art with the proposed method is provided in Table 8. Hari M Bharadwaj [32] achieved an accuracy of 17.6% on the same dataset using EEGNet. The experimental results suggest that the accuracies of the models (k-NN, SVM, and MLP) investigated in this study are in accordance with those reported in [32,52]. However, the proposed models following our multi-class, multi-channel feature fusion technique significantly improved the accuracies by up to 33.17%.

Table 8. Comparative analysis of the state of the art and the proposed methodology.

Study	Accuracy (%)
Hamad Ahmed et al. [52]	
1DCNN	5.1%
Hari M. Bharadwaj et al. [32]	
EEGNet	17.6%
Proposed Method	
MCCFF Net-50	22.94%
MCCFF VGG	33.17%

6. Conclusions

The field of EEG signal classification has seen significant advancements with the introduction of novel machine learning and deep learning approaches. However, due to the inherent complexity and non-stationary nature of EEG signals, achieving high classification accuracy remains challenging. Traditional models often struggle with the temporal and non-linear characteristics of these signals, necessitating the development of more sophisticated methods. The current study addresses these challenges by introducing the MCCFF-NET 50 and MCCFF-VGG models, which demonstrated substantial improvements in classification accuracy. The experimental results clearly show that these proposed approaches significantly outperform traditional models such as K-NN, SVM, and MLP, establishing a new benchmark in the field of object classification using visual EEG signals. While traditional models showed consistent but lower accuracies, the architectural enhancements in LSTM and 1D CNN led to notable performance improvements. Furthermore, this study revealed that the judicious selection of channels, covering the entire brain, significantly impacted classification accuracy. Future work will focus on exploring more sophisticated deep learning architectures and further optimizing preprocessing techniques. Additionally, the investigation of real-time applications and the expansion of the dataset to include more diverse stimuli could provide deeper insights and broader applicability in neuroscientific research and clinical diagnostics. The integration of multimodal data and the leveraging of transfer learning techniques also hold promise for future advancements in the field.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/s24216965/s1>.

Author Contributions: Writing—original draft: M.R.; supervision: S.U., H.u.R.S., H.G. and J.A.-I.; validation, M.R.; conceptualization, M.R.; funding acquisition, I.D.I.T.D.; writing—review and editing, M.R. and H.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset of EEG signals used in this study for visual classification of objects with a rapid-event temporal design is available at <https://iee-dataport.org/open-access/dataset-object-classification-randomized-eeg-trials> (accessed on 27 October 2024).

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

References

1. Sánchez-Reyes, L.M.; Rodríguez-Reséndiz, J.; AVECILLA-Ramírez, G.N.; García-Gomar, M.L. Novel algorithm for detection of cognitive dysfunction using neural networks. *Biomed. Signal Process. Control* **2024**, *90*, 105853. [CrossRef]
2. Sánchez-Reyes, L.M.; Rodríguez-Reséndiz, J.; AVECILLA-Ramírez, G.N.; García-Gomar, M.L.; Robles-Ocampo, J.B. Impact of eeg parameters detecting dementia diseases: A systematic review. *IEEE Access* **2021**, *9*, 60–74. [CrossRef]
3. Shen, G.; Horikawa, T.; Majima, K.; Kamitani, Y. Deep image reconstruction from human brain activity. *PLoS Comput. Biol.* **2019**, *15*, e1006633. [CrossRef] [PubMed]
4. Das, K.; Giesbrecht, B.; Eckstein, M.P. Predicting variations of perceptual performance across individuals from neural activity using pattern classifiers. *NeuroImage* **2010**, *51*, 1425–1437. [CrossRef]
5. Kalafatovich, J.; Lee, M.; Lee, S.-W. Learning Spatiotemporal Graph Representations for Visual Perception Using EEG Signals. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2023**, *31*, 97. [CrossRef]
6. Kersten, D.; Mamassian, P.; Yuille, A. Object Perception as Bayesian inference. *Annu. Rev. Psychol.* **2004**, *55*, 271–304. [CrossRef]
7. Katayama, O.; Stern, Y.; Habeck, C.; Coors, A.; Lee, S.; Harada, K.; Makino, K.; Tomida, K.; Morikawa, M.; Yamaguchi, R.; et al. Detection of neurophysiological markers of cognitive reserve: An EEG study. *Front. Aging Neurosci.* **2024**, *16*, 1401818. [CrossRef]
8. Rehman, A.; Khalili, Y.A. Neuroanatomy, Occipital Lobe. In *Medicine, Biology*; StatPearls Publishing: Treasure Island, FL, USA, 2019.
9. Holdaway, T. “Principals of Psychology PS200” Chapter 18: The Brain; PressBooks: Montreal, QC, Canada, 2024.
10. Cai, G.; Zhang, F.; Yang, B. Manifold Learning-Based Common Spatial Pattern for EEG Signal Classification. *IEEE J. Biomed. Health Inform.* **2024**, *28*, 1971–1981. [CrossRef]
11. Yang, K.; Hu, Y.; Zeng, Y.; Tong, L.; Gao, Y.; Pei, C.; Li, Z.; Yan, B. EEG Network Analysis of Depressive Emotion Interference Spatial Cognition Based on a Simulated Robotic Arm Docking Task. *Brain Sci.* **2024**, *14*, 44. [CrossRef]
12. Phukan, A.; Gupta, D. Deep Feature extraction from EEG Signals using xception model for Emotion Classification. *Multimed. Tools Appl.* **2024**, *83*, 33445–33463. [CrossRef]
13. Du, X.; Meng, Y.; Qiu, S.; Lv, Y.; Liu, Q. EEG Emotion Recognition by Fusion of Multi-Scale features. *Brain Sci.* **2023**, *13*, 1293. [CrossRef] [PubMed]
14. Wang, Y.; Zhang, B.; Di, L. Research Progress of EEG-Based Emotion Recognition: A Survey. *ACM Comput. Surv.* **2024**, *56*, 1–49. [CrossRef]
15. Krishnan, P.T.; Erramchetty, S.K.; Balusa, B.C. Advanced Framework for Epilepsy detection through image-based EEG Signal Analysis. *Front. Hum. Neurosci.* **2024**, *18*, 1336157. [CrossRef] [PubMed]
16. Su, K.-m.; Hairston, W.D.; Robbins, K. EEG-Annotate: Automated identification and labeling of events in continuous signals with applications to EEG. *J. Neurosci. Methods* **2018**, *293*, 359–374. [CrossRef]
17. Zhang, X.; Zhang, X.; Huang, Q.; Lv, Y.; Chen, F. A review of automated sleep stage based on EEG signals. *Biocybern. Biomed. Eng.* **2024**, *44*, 651–673. [CrossRef]
18. Jamil, N.; Belkacem, A.N. Advancing Real-Time Remote Learning: A Novel Paradigm for Cognitive Enhancement Using EEG and Eye-Tracking Analytics. *IEEE Access* **2024**, *12*, 93116–93132. [CrossRef]
19. Kocturova, M.; Jones, J. A Novel Approach to EEG Speech Activity Detection with Visual Stimuli and Mobile BCI. *Appl. Sci.* **2021**, *11*, 674. [CrossRef]
20. Craik, A.; He, Y.; Contreras-Vidal, J.L. Deep Learning for electroencephalogram (EEG) classification tasks: A review. *J. Neural Eng.* **2019**, *16*, 031001. [CrossRef]
21. Ruiz, S.; Lee, S.; Dalboni da Rocha, J.L.; Ramos-Murguialday, A.; Pasqualotto, E.; Soares, E.; García, E.; Fetz, E.; Birbaumer, N.; Sitaram, R. Motor Intentions Decoded from fMRI Signals. *Brain Sci.* **2024**, *14*, 643. [CrossRef]
22. Huettel, S.A. Event Related fMRI in cognition. *Neuroimage* **2012**, *63*, 1152–1156. [CrossRef]
23. Hahn, A.; Reed, M.B.; Vranka, C.; Godbersen, G.M.; Klug, S.; Komorowski, A.; Falb, P.; Nics, L.; Traub-Weidinger, T.; Hacker, M.; et al. High-temporal resolution functional PET/MRI reveals coupling between human metabolic and hemodynamic brain response. *Eur. J. Nucl. Med. Mol. Imaging* **2024**, *51*, 1310–1322. [CrossRef] [PubMed]

24. Chowdhury, E.; Mahadevappa, M.; Kumar, C.S. Identification of Finger Movement from ECoG Signal Using Machine Learning Model. In Proceedings of the IEEE 9th International Conference for Convergence in Technology (12CT), Pune, India, 5–7 April 2024; pp. 1–6.
25. Afnan, J.; Cai, Z.; Lina, J.M.; Abdallah, C.; Delaire, E.; Avigdor, T.; Ros, V.; Hedrich, T.; von Ellenrieder, N.; Kobayashi, E.; et al. EEG/MEG source imaging of deep brain activity within the maximum entropy on the mean framework: Simulations and validation in epilepsy. *Hum. Brain Mapp.* **2024**, *45*, e26720. [[CrossRef](#)] [[PubMed](#)]
26. Sharma, R.; Meena, H.K. Emerging Trends in EEG Signal Processing: A Systematic Review. *Springer Nat. Comput. Sci.* **2024**, *5*, 415. [[CrossRef](#)]
27. Dash, D.; Wisler, A.; Ferrari, P.; Davenport, E.M.; Maldjian, J.; Wang, J. MEG Sensor Selection for Neural Speech Decoding. *IEEE Access* **2020**, *8*, 182320–182337. [[CrossRef](#)]
28. Sari-Sarraf, V.; Vakili, J.; Tabatabaei, S.M.; Golizadeh, A. The brain function promotion by modulating the power of beta and gamma waves subsequent twelve weeks' time pressure training in chess players. *J. Appl. Health Stud. Sport Physiol.* **2024**. [[CrossRef](#)]
29. Simanova, I.; van Gerven, M.; Oostenveld, R.; Hagoort, P. Identifying Object Categories from Event-Related EEG: Toward Decoding of Conceptual Representations. *PLoS ONE* **2010**, *5*, e14465. [[CrossRef](#)]
30. Ashford, J.; Jones, J. Classification of EEG Signals Based on Image Representations of Statistical Features. In *Advances in Intelligent Systems and Computing*; Springer: Cham, Switzerland, 2020.
31. Deng, X.; Wang, Z.; Liu, K.; Xiang, X. A GAN Model Encoded by CapsEEGNet for Visual EEG Encoding and Image Reproduction. *J. Neurosci. Methods* **2023**, *384*, 109747. [[CrossRef](#)]
32. Bharadwaj, H.M.; Wilbur, R.B.; Siskind, J.M. Still an Ineffective Method with Supertrials/ERPs—Comments on 'Decoding Brain Representations by Multimodal Learning of Neural Activity and Visual Features'. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 14052. [[CrossRef](#)]
33. Li, R.; Johansen, J.S. The Perils and Pitfalls of Block Design for EEG Classification Experiments. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 316–333. [[CrossRef](#)]
34. Spampinato, C.; Palazzo, S. Deep Learning Human Mind for Automated Visual Classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
35. Raza, H.; Rathee, D.; Zhou, S.M.; Cecotti, H.; Prasad, G. Covariate shift estimation based adaptive ensemble learning for handling non stationarity in motor imagery related EEG-based brain computer interface. *Neurocomputing* **2019**, *343*, 154–166. [[CrossRef](#)]
36. Bashivan, P.; Rish, I.; Yeasin, M.; Codella, N. Learning Representations from EEG with Deep Recurrent Convolutional Neural Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
37. Bird, J.; Jones, L.; Milsom, D.; Malekmohammadi, A. A Study on mental state classification using eeg based brain machine interface. In Proceedings of the 9th International Conference on Intelligent Systems, Funchal, Portugal, 25–27 September 2018.
38. Nuthakki, S.; Kumar, S.; Kulkarni, C.S.; Nuthakki, Y. Role of AI Enabled Smart Meters to Enhance Customer Satisfaction. *Int. J. Comput. Sci. Mob. Comput.* **2022**, *11*, 99–107. [[CrossRef](#)]
39. Rehman, M.; Ahmed, T. Optimized k-Nearest Neighbor Search with Range Query. *Nuclues* **2015**, *52*, 45–49.
40. Wu, H.; Li, S.; Wu, D. Motor Imagery Classification for Asynchronous EEG-Based Brain-Computer Interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2024**, *32*, 527–536. [[CrossRef](#)]
41. Pasanta, D.; Puts, N.A. Functional Spectroscopy. In *Reference Module in Neuroscience and Biobehavioral Psychology*; Elsevier: Amsterdam, The Netherlands, 2024.
42. Padfield, N.; Zabalza, J.; Zhao, H.; Masero, V.; Ren, J. EEG based brain computer interfaces using motor imagery: Technique and challenges. *Sensors* **2019**, *19*, 1423. [[CrossRef](#)]
43. Nuthakki, S.; Kolluru, V.K.; Nuthakki, Y.; Koganti, S. Integrating Predictive Analytics and Computational Statistics for Cardiovascular Health Decision-Making. *Int. J. Innov. Res. Creat. Technol.* **2023**, *9*, 1–12. [[CrossRef](#)]
44. Miladinovic, A.; Ajsevic, M.; Jarmolowska, J.; Marusic, U.; Colussi, M.; Silveri, M.; Battaglini, G. A Effect of Power feature covariance shift on BCI spatial-filtering techniques: A comparative study. *Comput. Methods Program Biomed* **2019**, *198*, 105808. [[CrossRef](#)]
45. Ahmed, H.; Wilbur, R.B.; Bharadwaj, H.M.; Siskind, J.M. Confounds in the data—Comments on 'Decoding Brain Representations by Multimodal Learning of Neural Activity and Visual Features'. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 9217–9220. [[CrossRef](#)]
46. Palazzo, S.; Spampinato, C. Decoding Brain Representations by Multimodal Learning of Neural Activity and Visual Features. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 3833–3849. [[CrossRef](#)]
47. Zheng, X.; Chen, W. Ensemble Deep Learning for Automated Visual Classification Using EEG Signals. *Pattern Recognit.* **2019**, *102*, 107147. [[CrossRef](#)]
48. Fares, A.; Zahir, S.; Shedeed, H. Region Level Bi-directional Deep Learning Framework for EEG-based Image Classification. In Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine, Madrid, Spain, 3–6 December 2018.
49. Fares, A.; Zahir, S.; Shedeed, H. EEG-based image classification via a region level stacked bi-directional deep learning framework. In Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine, Madrid, Spain, 3–6 December 2018.

50. Guo, W.; Xu, G.; Wang, Y. Brain Visual Image signal classification via hybrid dilation residual shrinkage network with spatio temporal feature fusion. *Signal Image Video Process.* **2023**, *17*, 743–751. [[CrossRef](#)]
51. Abbasi, H.; Seyedarabi, H.; Razavi, S.N. A combinational deep learning approach for automated visual classification using EEG signals. *Signal Image Video Process.* **2024**, *18*, 2453–2464. [[CrossRef](#)]
52. Ahmed, H.; Wilbur, R.B.; Bharadwaj, H.M.; Siskind, J.M. Object Classification from Randomized EEG Trials. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; p. 3845.
53. Kaneshiro, B.; Guimaraes, M.P.; Kim, H.S.; Norcia, A.M. A Representational Similarity Analysis of the Dynamics of Object Processing Using Single-Trial EEG Classification. *PLoS ONE* **2015**, *10*, e0135697. [[CrossRef](#)]
54. Gifford, A.T.; Dwivedi, K.; Roig, G.; Cichy, R.M. A Large and Rich EEG Dataset for Modeling Human Visual Object Recognition. *NeuroImage* **2022**, *264*, 119754. [[CrossRef](#)]
55. Vivancos, D.; Cuesta, F. Mind Big Data 2022: A Large Dataset of Brain Signals. *arXiv* **2022**, arXiv:2212.14746.
56. Cichy, R.M.; Pantazis, D. Multivariate Pattern Analysis of MEG and EEG: A Comparison of Representational Structure in Time and Space. *NeuroImage* **2017**, *158*, 441–454. [[CrossRef](#)]
57. Falciglia, S.; Betello, F.; Russo, S.; Napoli, C. Learning Visual Stimulus-Evoked EEG Manifold for Neural Images Classification. *NeuroComputing* **2024**, *588*, 127654. [[CrossRef](#)]
58. Bhalerao, S.V.; Pachori, R.B. Automated Classification of Cognitive Visual Objects Using Multivariate Swarm Sparse Decomposition from Multichannel EEG-MEG Signals. *IEEE Trans. Hum.-Mach. Syst.* **2024**, *54*, 455–464. [[CrossRef](#)]
59. Ahmadi, H.; Gassemi, F.; Moradi, M.H. A Hybrid Deep Learning Framework for Automated Visual Classification Using EEG Signals. *Neural Comput. Appl.* **2023**, *35*, 20989–21005. [[CrossRef](#)]
60. Zhu, S.; Ye, Z.; Ai, Q. EEG-ImageNet: An Electroencephalogram Dataset and Benchmarks with Image Visual Stimuli of Multi-Granularity Labels. *arXiv* **2024**, arXiv:2406.07151.
61. Ye, Z.; Yao, L.; Zhang, Y.; Gustin, S. Self-Supervised Cross-Modal Visual Retrieval from Brain Activities. *Pattern Recognit.* **2024**, *145*, 109915. [[CrossRef](#)]
62. Li, R.; Johansen, J.S.; Ahmed, H.; Ilyevsky, T.V.; Wilbur, R.B.; Bharadwaj, H.M. Training on the Test Set? An Analysis of Spampinato et al. *Extraction* **2019** *31*, 6809–6810.
63. Singh, A.K.; Krishnan, S. Trends in EEG Signal feature extraction applications. *Front. Artif. Intell.* **2023**, *92*, 1072801. [[CrossRef](#)] [[PubMed](#)]
64. Ahmed, I.; Jahangir, M.; Iqbal, S.T.; Azhar, M.; Siddiqui, I. Classification of Brain Signals of Event Related Potentials using Different Methods of Feature Extraction. *Int. J. Sci. Eng. Res.* **2017**, *8*, 680–686. [[CrossRef](#)]
65. Badr, Y.; Tariq, U.; Al-Shargie, F.; Babiloni, F.; Mughairbi, F.A. A Review on Evaluating Mental Stress by Deep Learning Using EEG Signals. *Neural Comput. Appl.* **2024**, *36*, 12629–12654. [[CrossRef](#)]
66. Ari, E.; Tacgin, E. NF-EEG: A Generalized CNN Model for Multi-Class EEG Motor Imagery Classification Without Signal Preprocessing for Brain-Computer Interfaces. *Biomed. Signal Process. Control.* **2024**, *92*, 106081. [[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.