IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 22 June 2023, accepted 4 July 2023, date of publication 7 July 2023, date of current version 19 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3293421

RESEARCH ARTICLE

A Systematic Review of Electroencephalography Open Datasets and Their Usage With Deep Learning Models

ALBERTO NOGALES AND ÁLVARO J. GARCíA-TEJEDOR
CEIEC, Research Institute, Universidad Francisco de Vitoria, 28223 Madrid, Spain
Corresponding author: Alberto Nogales (alberto.nogales@ceiec.es)

ABSTRACT Data are the main headache for machine learning, both because of their varied nature and their limited availability. The medical field brings together both situations: tables, images, text, or signals that are difficult to acquire due to the number of patients, the complexity and time of acquisition, or ethical constraints. The existence of open datasets is the best option for researchers in this field. Electroencephalograms are a good example of this situation. This paper identifies the primary open datasets of electroencephalography tests and how they are used in deep learning models. The aim is to provide structured information that can be consulted by researchers in the field (both physicians and computer scientists) to know which datasets are available, which characteristics they have, or which deep learning models could be applied to them. The process followed the PRISMA methodology for systematic reviews applying different inclusion and exclusion criteria to obtain a set of high-quality papers on which the data sets used were analyzed. The databases included in the searches were Scopus, PubMed, Web of Science (WOS), Science Direct, IEEE Explorer, and SpringerLink. In total, 37 papers were selected which included 30 datasets that have been considered. Then, the DL models used in the papers and the different characteristics of the datasets have been statistically analyzed by obtaining different measures and graphs. The most relevant conclusions are the widespread use of convolutional neural networks (the less innovative among the different models) as the main tool for EEG data analysis. Against this position, we found the use of hybrid models and the family of RNNs as techniques to use in cases of brain stimuli, classification of levels of fatigue, and diagnosis of diseases. Related to the datasets' features, we demonstrate the difficulty in compiling this data due to the number of tests and that the minimum of channels or sampling frequency recommended to obtain good accuracies in the model should be studied.

INDEX TERMS Systematic review, deep learning, open datasets, electroencephalograms.

I. INTRODUCTION

Most people are connected every day through their mobile phones or computers. This entails the creation of vast amounts of data through organizations or private companies every day. According to [1], in 2020, 44 zettabytes were produced, and by 2025 is estimated to be between 163 and 175 zettabytes. The trend remains the same in the medical field due to new applications and the wide range of data from demographic information to images resulting in medical tests like radiographs or 3D scanners passing through those that collect the biomedical signal.

The associate editor coordinating the review of this manuscript and approving it for publication was Humaira Nisar.

Chang and Moura [2] define biosignal processing as extracting relevant information from biomedical signals. These are also described as physiological activities from organisms that can comprise neural, cardiac rhythms, or others. Among all the medical tests related to signals, electroencephalograms (EEGs) are considered the most beneficial for compiling brain signals.

EEGs are a type of data called time series, defined in [3] as sets of repeated observations of a single unit or individual at regular intervals over many instances. The case of EEGs corresponds to a test used to diagnose neurological diseases based on a set of electrodes placed around the scalp. EEGs compile a lot of data being very complex to analyze. They need professionals with high skills acquired through years of training. The problem with EEGs is that they are studied by eye, and due to their complexity, the professionals miss a lot of information.

There is a trend toward integrating and leveraging these enormous amounts of data to make medicine more personalized, efficient and focused on the patient. Nevertheless, classical methods like statistics are not powerful enough to manage that large number of variables and data. At this point, using more modern techniques, such as Artificial Intelligence (AI), is a great benefit. They allow us to find new patterns and predict how different variables behave or identify new ones not considered in complex medical problems, [4].

AI is a computer science field aiming to analyze and decipher human mechanisms related to intelligent behaviours that are, later, reproduced in machines, [5]. Among all the AI techniques, Machine Learning (ML) has stood out from the rest in recent years. ML is defined by [6] as a discipline that studies and develops algorithms that create systems that learn by finding patterns in datasets. ML comprises a wide range of techniques, with Artificial Neural Networks (ANN) obtaining the best results recently. Hecht-Nielsen [7] defines ANN as a computational model formed by several simple units that are strongly connected and can process information by responding to external stimuli. The benefits of ANN remained not very useful until deeper architectures, called Deep Learning (DL) models, arose. DL consists of ANN models with several layers that can learn data representation using more abstraction levels, [8]. Figure 1 depicts the hierarchy of the fields in AI described previously.

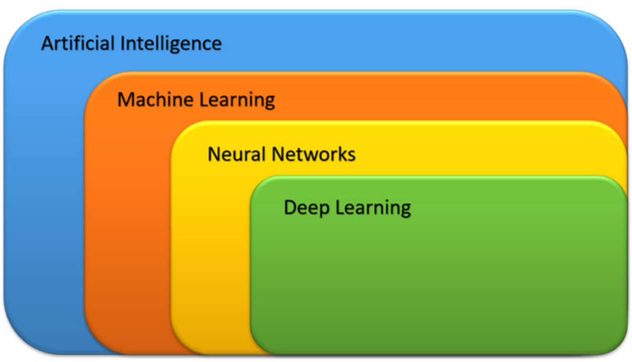

FIGURE 1. Hierarchy among artificial intelligence disciplines.

Roy et al. [9], a review of deep learning models with EEGs, states that there are a large number of works that cannot be reproduced as data is unavailable. It also points out that more than half of the studies use publicly available datasets. Considering also the difficulty of obtaining EEGs and the computational cost of developing deep neural models, it seems clear that there is a need to have a reference resource that can be consulted by researchers in this field (both physicians and computer scientists) to know which datasets are available or which models perform better.

This paper also presents an innovative character because, to the best of our knowledge, it is the only one that has studied what open EEG datasets can be found in the scientific community.

The main contribution of this paper is to present a systematic review of open EEG datasets used in works using DL techniques. The paper follows a methodology to obtain scientific papers utilizing this kind of dataset. Papers have been searched in the best-known scientific sources using a set of keywords to focus the searches. However, as EEGS datasets are scarce, not many open datasets are available, so there are not many papers that meet the selection criteria. After discarding some of them that either did not use an open EEG dataset with a DL model, did not provide model performance metrics, or did not include a description of the dataset and a link to download it, we remained with 37 works.

In the process, we provide a set of statistical metrics alongside some graphics that let to understand the information. This is useful in the following cases. The content will let researchers know which are the most used deep learning techniques and which accuracies they get depending on the dataset. It also could help scientists choose which models perform well with their specific data or use case. Another provided information is which datasets are available and how they perform, this is useful when researchers want to develop a new model/method and test it or know which models are not applied a lot. Another interesting usage is also to know which type of use cases does not have an open dataset that could be used by the scientific community, so people could create a new one. Finally, compiling the information of which are the most common values for the main features of the datasets (number of channels, sample rate, etc.) could help us to build a golden standard of the dataset.

This work has the following sections: Section II describes the DL techniques used in the papers. Section III details the method followed in compiling the documents. Section IV compiles all the studied features. Section V contains an in-depth discussion of the datasets and their use. Finally, Section VI presents some conclusions about the research.

II. STATE OF THE ART

By considering the creation of AlexNet as the main milestone in deep learning [10], the number of papers in medical bibliographic databases has been growing exponentially yearly. Figure 2 shows the number of publications from 2012 to 2023 containing the word deep learning. We can see that, in some cases, the number of publications doubled from one year to another. Then, considering the period from 2018 to 2023, we can see that most of the scientific production in this field is during those years. Even during this year, more than 4000 papers have been published in about 2 months and a half which means that at the end of the year, the number of papers will be greater than in 2022.

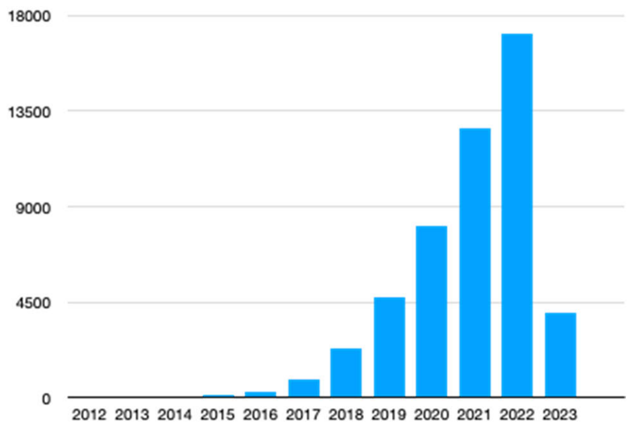

FIGURE 2. Distribution by the publication year of the deep learning papers indexed in PubMed from 2012 to 2023 (n=4524).

DL techniques are based on multiple models and architectures, although their effectiveness is directly related to the nature and quality of the data used in the training stage. This section describes the architectures and models that can use EEGs.

DL models can be classified depending on how they learn from the data. This case has three main classes: supervised, unsupervised, and semi-supervised.

Supervised models need labelled data to perform the training. In this case, the model knows the relation between input data and the expected output and uses the following classification.

Multilayer perceptron (MLP) is the simplest case of a DL model. The architecture comprises an input layer, several hidden layers, and an output layer. Lin et al. [11] use an MLP to classify EEG signals depending on the music some subjects are listening to.

Convolutional neural networks (CNNs) are the most used models with several applications in computer vision. Its primary ability is to detect patterns in a delocalized way. This characteristic lets to learn a particular pattern in an image that can later be seen in another part of another image. Recently, a specific type of CNN that manages graphs called Graph Convolutional Neural Networks (GCNN) has arisen. Kipf and Welling [12] presents this model as a method that encodes a graph's structure and its nodes' features using a special type of CNNs. CNNs are used by [13] to classify epileptic seizures. GCNN recognizes emotions by analyzing EEGs, [14].

Recurrent neural networks (RNNs) are defined by [15] as a model that uses an input vector of arbitrary length and applies a transition function recursively to its internal hidden state vector h_t. It uses data structures that are time series, for example, EEGs. Within RNNs, a particular type is called Short-Term Memory Networks (LSTM) or Gated Recurrent Unit (GRU). LSTMs were proposed to work with noisy or incomprehensible input data without information loss [16]. In the case of RNNs, [17] applies them to the prognosis in patients with neurodegeneration. Then, LSTMs have applications in emotion recognition (Alhagry et al. [18], 2017). Finally, GRU has been applied to emotion classification, [19].

The other leading group of models belongs to the category of unsupervised models. In this case, data is unlabeled, and there is no a priori knowledge about the final results [20].

Deep Autoencoders (DAE) use unsupervised learning. Defined in [21], its particular characteristic is that both the input and output layers have the same or similar size and two processing structures. The first one is the decoder which starts from the input data and reduces its size to a small piece that contains its main characteristics. The second part is the decoder which aims to upsample the previous small piece of data by upsampling it until reaches the input data size. In [22], autoencoders classify ictal EEG. We consider Restricted Boltzmann Machines (RBM) as a particular type of Autoencoder introduced by [23] that can learn a probability distribution. In DL, RBMs were used to implement Deep Boltzmann Machines (DBMs), [24]. The field of EEGs has applications like [25] that apply to motor imagery.

The previous learning types generate a new one by mixing them and are called semi-supervised. Generative adversarial networks (GAN) are under this class. GANs need neural models, the generator, and the discriminator. Both work in a training type called adversarial process, [26]. This architecture aims to learn and imitate a data distribution. The generative model is responsible for creating synthetic instances of the input data. Then, the discriminator evaluates these data and decides if it is similar enough to the input data or not. This task gives a probability of being authentic (input data) or synthetic (created by the generator). By repeating this process, the generator learns how to create data more like the input one. In this case, GANs are applied to perform data augmentation strategies with EEGs [27].

It is noteworthy that in recent years a trend in the creation of hybrid models has been detected. These types of models are seen as an important area of development within the DL soon, [28]. These architectures join two or more models generating a CNN-LSTM or an Autoencoder-LSTM.

To summarize all the information above, Figure 3 shows a taxonomy with all the deep learning models.

Metrics are an important aspect when evaluating a DL model. Four are the most important in this type of analysis: accuracy, specificity, sensitivity, and F-1 score. Accuracy is defined as the ratio between successful predictions and the total number of predictions. This metric is used as a way to measure the performance of a model in the first moment. Specificity measures the ratio between the number of true negatives (healthy people diagnosed correctly) and the total of those predicted as true negatives and false positives (healthy people diagnosed as sick). Sensitivity is similar to specificity but considers true positives instead true negatives and false negatives instead of false positives. F1-score considers true positives, false positives, and false negatives as described in

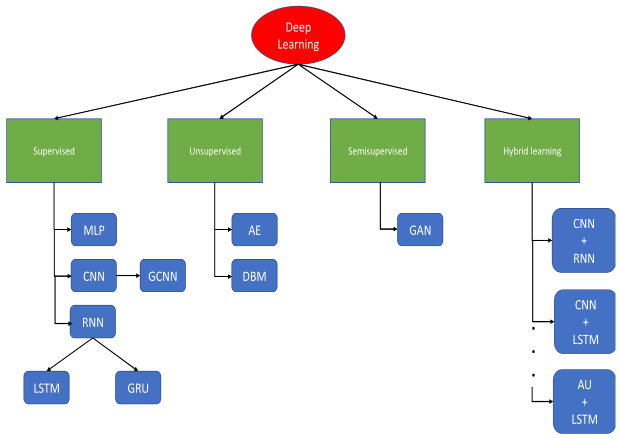

FIGURE 3. Taxonomy of deep learning models.

the following Equation.

$$\frac{2 * TP}{2 * TP + FP + FN} \tag{1}$$

III. METHODOLOGY

The method used to determine which research works are framed in a particular field or respond to the needs of certain research questions is called Systematic Literature Review (SLR) or just systematic review. There are different guidelines for conducting an SLR. For example [29] include the necessity of the review, research questions, development of the protocol, identifying the research works, establishing some inclusion and exclusion criteria, analyzing some features of the papers and creating the review as a paper. As the phases of the review process can differ, we have used [30], [31], [32], and [33] as references to design our process which is described in Figure 4. First, we formulate a set of research questions. Then, we start the process of finding and selecting the research works, from where we collect the datasets. Following, we analyze them, and the papers where they are used. Finally, we describe all this information in the presented paper.

A. FORMULATED RESEARCH QUESTIONS

As the main aim of this systematic review is compiling open datasets of EEGs that have been used with DL models, some information could be analyzed like the characteristics of the datasets or the deep learning models. In this way, the following research questions are proposed as a way to understand the purpose of the review and its utility.

Question 1: Which EEG datasets are freely available to researchers so they can perform their studies in deep learning?

Motivation 1: As EEGs are difficult to compile due to the time needed to do the test or the number of patients and controls, this data is scarce. This information source can be consulted by them to find data for their research.

Question 2: Which values have the main characteristics of the datasets? The number of channels, frequency, etc.

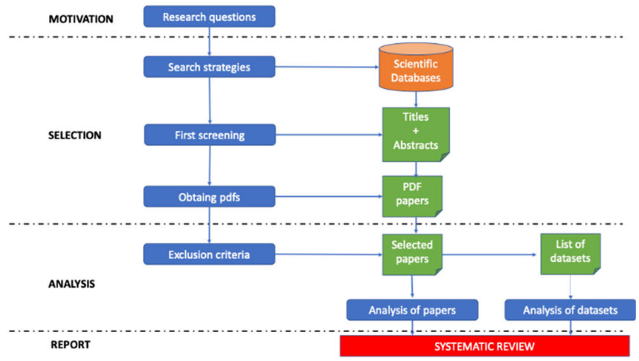

FIGURE 4. Review process conducted in the present research.

Motivation 2: This is key for researchers when establishing a protocol to compile their data. This decision must be taken by both profiles: physicians and computer scientists. This assures that the data accomplishes with a minimum quality so the deep learning models could be appropriate and cover a wide range of use cases.

Question 3: Which deep learning models perform better with electroencephalograms and their use cases?

Motivation 3: Given the metrics compiled in this work and the deep learning models that have been obtained, researchers can know which deep learning models best fit the different datasets depending on the characteristics and the use cases: diagnosis, motor imagery, etc. It also lets researchers know if the datasets are good enough to apply DL techniques.

B. SEARCH STRATEGY FOR IDENTIFYING THE STUDIES

To obtain the papers, we have set the following keywords to be used in every scientific source: ((''open dataset'') OR (''free dataset'') OR (''freely available dataset'') OR (''open data'') OR (''free data'') OR (''freely available data'')) AND ((''EEG'') OR (''electroencephalogram'')) AND ((''deep learning'') OR (''neural network'') OR (''neural networks'')). The search and collection of papers include everything published until March 15, 2023. The following sources were used to make the searches: Scopus,[1] PubMed,[2] Web of Science (WOS),[3] Science Direct,[4] IEEE Explorer,[5] and SpringerLink.[6] After discarding repeated items, conference papers surveys, or arxiv papers, the final selection of works has been made to apply more restrictive criteria.

C. CRITERIA FOR SELECTING PAPERS

A group of computer scientists has set out the following criteria to obtain the final set of papers.

The first selection of works consisted of a single screening where titles and abstracts are read to check if they meet the

[1]https://www.scopus.com/
[2]https://pubmed.ncbi.nlm.nih.gov/
[3]https://www.webofscience.com/
[4]https://www.sciencedirect.com/
[5]https://ieeexplore.ieee.org/
[6]https://link.springer.com/

minimum inclusion criterion of "a paper that uses an EEG open dataset to train a deep learning model". Searches in scientific resources were made. Then titles and abstracts were read, and those that did meet the criteria of including an open EEG dataset used with DL models to solve a particular use case were included for the following step.

As there is no way to automatize a more exhaustive process of selecting the papers, several quality requirements have been set out. This is a set of exclusion criteria that discard papers accomplishing the following:

1. The paper does not describe the DL model which is trained with a dataset of EEGs.
2. Metrics about the performance of the models are not included in the evaluation.
3. The paper does not include a detailed description of the dataset or a link to download it. Datasets available upon request are not considered. The EEGs are obtained from humans.

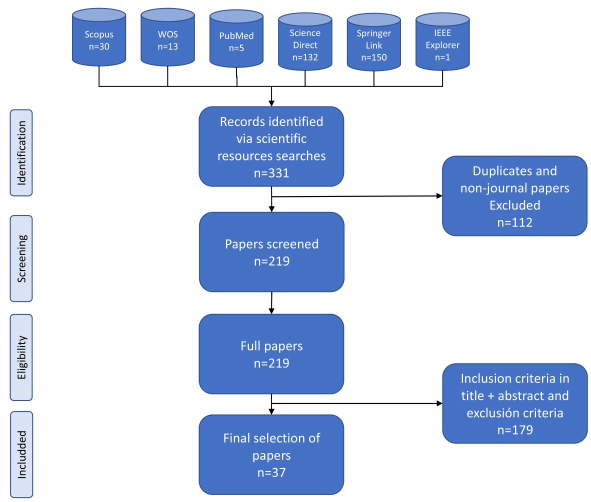

FIGURE 5. PRISMA diagram of the bibliographic review conducted.

D. PRISMA FLOW DIAGRAM

This systematic review compiles a set of papers by using the following methodology. Figure 5 contains a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram [34], which summarizes how we achieved the selection of papers used to compile the datasets reported in the paper.

From the first search, a total of 331 works has been obtained which are distributed as follows: Scopus (30), PubMed (5), WOS (13), Science Direct (132), IEEE Explorer (1), and SpringerLink (150). After eliminating duplicates and papers not published in journals (conferences, arxiv, etc.), it remained a total of 219 papers were. The next step was to analyze their titles and abstracts to check if the paper applies deep learning models in an open dataset of EEGs. If yes, we must check if they accomplish the three exclusion criteria. The previous decisions eliminated different papers, including one whose dataset was unavailable for download, MERTI-Apps [74]. After this process, the final set had 37 papers from which we are analyzing some features related to the used deep learning models and obtaining the report of the open EEG datasets.

E. THREATS TO VALIDITY

A systematic review can be put at risk due to potential biases and the imprecise application of the extraction method. To evaluate this, four dimensions are considered: internal validity, external validity, construct validity, and conclusion validity.

Internal validity: depending on the search process a validity threat can impact the representativeness of the selected scientific works. To avoid that, we have used [30], [31], [32], and [33] as guidelines to adjust our process. The research questions have been a guide to constructing the searches and thinking about the inclusion and exclusion criteria that best fit them. The selection of keywords and scientific resources could be a limitation, but we have tried to avoid it by using very clear terminology and using official resources like WOS, PubMed, etc.

External validity: We have limited the papers to works published in scientific journals discarding preprints, books, conferences, etc. This allows for obtaining strong conclusions that could be useful for scientists in related fields.

Construct validity: Research questions are one of the main pieces to guide the review and think about its utility. For this purpose, these research questions have been discussed among the authors and other researchers in the field.

Conclusion validity: To address the potential subjectivity of our study, the authors read the title and abstract for reviewing the first screening studies. This approach aimed to minimize bias in the extraction of data. In the event of disagreements between them, a consensus was reached through discussions. This approach ensured that the data collected was both reliable and objective.

F. DATA EXTRACTION AND CLASSIFICATION OF THE STUDIES

The paper comprises two types of studies on this set of papers. The first is on the deep learning models used with the datasets, the metrics to measure performance, the use case solved in the paper, and the paper's year of publication. The second compiles some characteristics related to the dataset: number of channels, number of individuals, distribution system, sampling frequency, or the format of the file.

IV. RESULTS

Some statistics have been obtained based on the previous characteristics used in the papers and datasets. This information was compiled by developing some Python scripts and using Matplolib, which provides a graphic set of charts, [35].

Apart from that, we have developed a VOSviewer[7] with the keywords in the selected paper to confirm their relation to the keywords used in the original searches. Figure 6 shows a VOSviewer with this information.

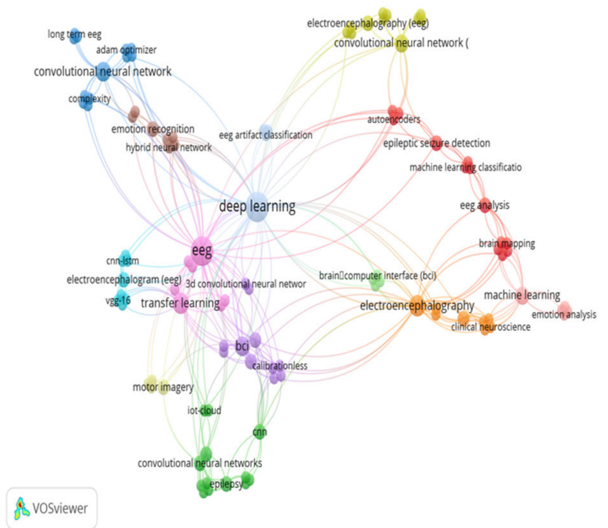

FIGURE 6. **VOSviewer network visualization.**

In the Figure above, we can see that the biggest nodes correspond to "deep learning" and "EEG" which makes sense as are key terms in our searches. Then, there are other keywords similar to the latter like "electroencephalogram" or "electroencephalography". Related to DL, some keywords describe the different models: "convolutional neural network", "autoencoders" or "cnn-lstm". Some keywords define the use cases: "emotion recognition", "motor imagery" or "epileptic seizure detection" which, as we will see later, are important in the study of datasets.

For the analysis of papers and datasets, the following graphs have been provided. Looking at the papers: bar diagrams with the publication year, percentages of the deep learning models used, DL metrics, the use case that has been solved, and the relation between the deep learning models and the use cases. By taking into account the datasets' characteristics: number of individuals and tests, length of the tests, number of channels, distribution of the electrodes systems that have been used, frequency in hertz during the test, and the file format provided to work with the data.

A. SUMMARY OF PAPERS

Table 1, attached at the end of the paper shows the selected set of papers with the following information: the paper's reference, the deep learning model used, the metrics applied during the experimentation, the tasks performed by the individuals while compiling the data, and the year of publication of the work. Apart from that, we have added a matrix of evaluation by using the Composite Quality Indicator (CQI) index of the i-th paper is determined by combining the normalized indicator values of the other indexes within the range of 0 to 1.

B. STATISTICS AND ANALYSIS OF THE INCLUDED STUDIES

This section provides graphs and statistics from analyzing the selected papers related to the use of open datasets. Figure 7 shows a bar chart distributing the 37 papers by year of publication.

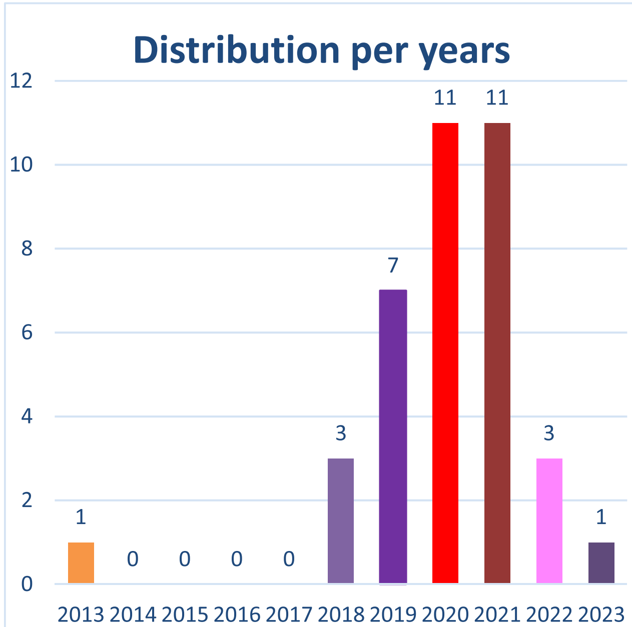

FIGURE 7. **Distribution by year of the selected papers (n=37).**

Another relevant piece of information that can be obtained from this preliminary analysis is related to the first research question; the type of DL model used. This knowledge is helpful for researchers to determine which are the most powerful models for processing EEGs. The usage percentages are collected in the following pie chart, Figure 8. As more than one model can be used in a paper, the number of instances is bigger than 37.

Another quality criterion for selecting a paper is the use of metrics to evaluate the performance of the models. Following, we compile a pair of aspects related to them. Again, it should be highlighted that a paper can use more than one metric. Figure 9 represents a diagram of bars that counts the times each DL metric appears in the set of selected papers which is interesting to know which of them researchers should apply in their works.

Another graphic related to metrics is the following boxplot, where we represent the distribution of the values obtained in the different papers after training the different DL models, Figure 10.

EEGs can solve several use cases. This information helps us know which application fields are less exploited, so there is scope for further research. Figure 11 uses a pie chart

[7]https://www.vosviewer.com/

TABLE 1. Summary of selected papers.

Paper	Deep learning model	Metrics	Aim	Year	CI	SJR (2021)	SNIP (2022)	CQI
[69]	CNN	74.2% Accuracy	Motor imagery (MI) electroencephalogram (EEG) signal classification	2019	98	0.927	1.33	0.891345
[70]	CNN	75.8 and 84.3 Accuracy	MI classification	2019	74	1.275	1.32	0.992178
[71]	CNN	91.15% Sensitivity	Classification of imagined or executed movements	2018	1157	1.719	1.55	2.169724
[72]	CNN	73.4% Accuracy	Classify emotion based on multimodal data	2018	110	0.803	1.42	0.889157
[73]	MLP	60.9%, 51.2%, and 68.4% Accuracy	Predict the levels of arousal, valance, and liking based on the learned features	2013	190	1.257	1.64	1.199843
[74]	RNN	91.3% and 94.8% Accuracy	Recognize emotions	2020	8	0.59	1.01	0.578499
[75]	CNN	86% and 77% Accuracy	Extract stimulus pattern features	2019	34	1.257	1.64	1.065012
[76]	AU	12.37% Sensitivity	Predict seizures	2020	5	0.243	0.00	0.091700
[77]	CNN	73.22 Accuracy	Detect drivers' drowsy states	2021	12	1.392	0.89	0.827637
[78]	CNN	71%, 72%, 70% and 72% Accuracy	MI classification	2020	48	0.859	1.25	0.795208
[79]	CNN	90% Accuracy	Classification of motor imagery EEG	2020	24	0.522	0.96	0.550083
[80]	AU+LSTM	98.79 Accuracy, 98.72 Sensitivity, 98.86 Specificity	Detecting seizures in pediatric patients	2021	27	0.797	1.11	0.704942
[81]	CNN	73.01% Accuracy and 68% Accuracy	Mental fatigue recognition	2020	31	1.601	2.20	1.385404
[82]	CNN	94.33%, 92.57 and 93% Accuracy	Detect drivers' fatigue	2021	2	0.195	0.39	0.210637
[83]	CNN+LSTM	83.9 and 83.7 Accuracy	Estimation of the sleep stages	2019	60	1.799	2.27	1.506577
[84]	CNN	65.62% and 85.66% Accuracy	Recognize EEG signals in imagined vowel tasks	2021	4	0.803	1.42	0.797540
[85]	CNN	82.2% Accuracy and 82.5% F1-Score	BCI Classification	2021	9	1.211	1.86	1.105155
[86]	AU	97% F1-Score	Sleep stage classification	2021	25	1.257	1.64	1.057233
[87]	CNN	66.5% Sensitivity, 97.9% Specificity and 67.9% Sensitivity, 97.0% Specificity	Classify sleep staging	2019	20	1.211	1.86	1.114663
[88]	CNN	72.3% and 79.4% Accuracy	EEG-based sleep staging and pathology detection	2020	35	1.504	1.66	1.161811
[89]	CNN+LSTM	87%, 86% and 86% Accuracy	Automatic Sleep Scoring	2020	9	0.754	0.88	0.592071
[90]	CNN	84.4%, 81.3% and 86.7% Accuracy	Sleep stage classification	2021	89	1.257	1.64	1.112549
[91]	CNN	90.89% Accuracy	Sleep Stage Classification	2020	11	1.497	2.24	1.344956
[92]	AU	89.49%, 92.86% and 96.77% Accuracy	EEG-Based Emotion Classification	2020	74	1.223	1.02	0.866718
[93]	CNN	99.5%, 96.5% and 95.7% Accuracy	Classification of epileptic EEG recordings	2019	90	1.309	1.94	1.238873
[94]	CNN	78.22% and 74.92% Accuracy	Emotion Recognition	2021	39	1.309	1.94	1.194793
[95]	CNN	70.15% Accuracy, 70.18 F1-Score and 77.07% Accuracy, and 75.48% F1-Score	Detection Attention Deficit and Hyperactivity Disorder (ADHD)	2021	11	2.781	2.81	2.009507
[96]	CNN	99.42%, 95.83% Accuracy. 99.55%, 96.29% Specificity. 97.55%, 89.57% Sensitivity 96.39%, 81.86% f1-score	Seizure Management	2022	1	2.781	2.81	2.000864
[97]	MLP	93.33% Accuracy and 95.0% Accuracy	Seizure Management	2019	11	2.781	2.81	2.009507
[98]	CNN, CNN+AU	99.2% Accuracy and 93.5% Sensitivity	Seizure Management	2018	95	0.871	1.37	0.882850
[97]	CNN+LSTM	93.3% Accuracy	MI EEG Classification	2022	0	0.761	1.05	0.647308
[99]	CNN	76.21% Accuracy	MIVEEG Classification	2023	0	1.275	1.32	0.928219
[100]	CNN	79.56% Accuracy	Brain Stimuli	2021	12	0.927	1.33	0.817015
[101]	MLP	99.44% Accuracy, 80.66% Sensitivity	Seizure Management	2021	0	0	0	0.000000
[102]	CNN+LSTM	94.66% Accuracy, 95.0% Precision, 95.0% Sensitivity, 95.0% F1-score. 84.53% Accuracy, 86.0% Precision, 85.0% Sensitivity, 83.0% F1-score.	Human activity recognition	2020	16	1.211	1.86	1.111205
[103]	CNN, CNN+LSTM	100% Accuracy and 100% Specificity	Seizure management	2020	6	0.833	1.39	0.799380
[60]	CNN	82.3% Accuracy, 75.0% Precision, 62.8% Recall, and 61.1% F1-score	Brain reaction after having sweet drinks	2022	44	1.024	1.58	0.968520

to describe this information which we have classified the datasets into 8 general categories:

- Motor imagery (MI) classification. This application aims to recognize a subject's intention, [25].
- Seizure management. EEGs of patients with epilepsy, a brain disorder that consists of abnormal cerebral activities.
- Estimation of sleep stages. Datasets collect the five possible stages a human can experiment with while sleeping.
- Recognize emotions. This task consists of classifying human emotional states as the domains of arousal and valence.
- Classify levels of fatigue. Mental fatigue happens when a subject has paid attention to a task for a long time. These datasets can measure different levels of fatigue, in some cases while driving.
- Disease diagnosis. In the medical field, we typically find datasets of epilepsy, but others can diagnose diseases such as Attention Deficit and Hyperactivity Disorder (ADHD).
- Brain stimuli. It measures how the brain responds to different perception tasks. For example, the response to images or the consumption of sweetened drinks.
- Human activity recognition. This is a way to detect artefacts while performing tasks such as reading, watching, and speaking.

Figure 12 combines the results of both previous analyses in a bubble diagram where the X-axis represents the deep learning model and the Y-axis the possible use cases. This information is interesting when a scientist needs to decide what DL models could be used depending on the use case they are working on. The bubble size and colour depend on the number of instances.

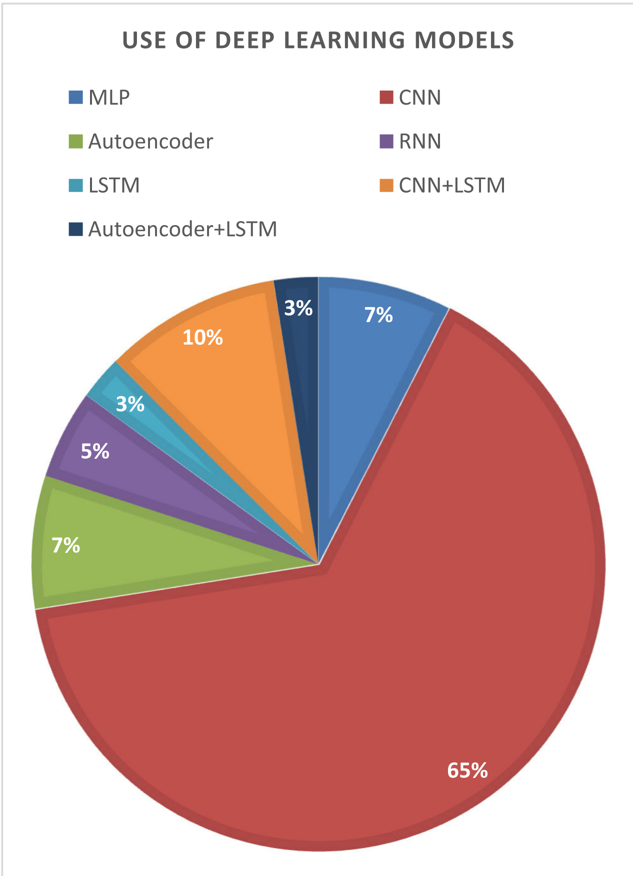

FIGURE 8. Distribution of the models by deep learning architecture type (n=40).

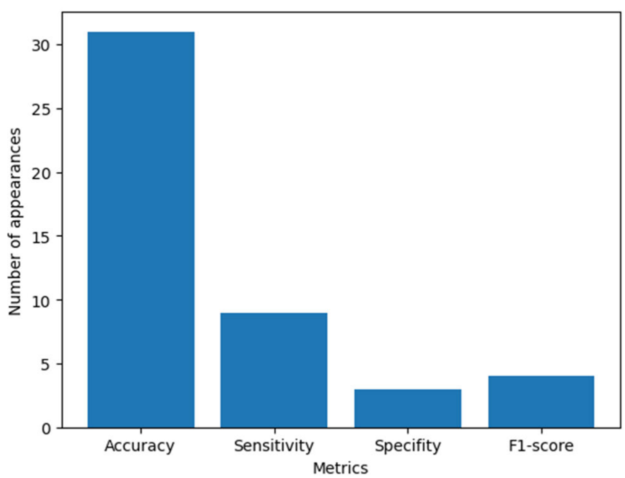

FIGURE 9. Distribution of the metrics (n=47).

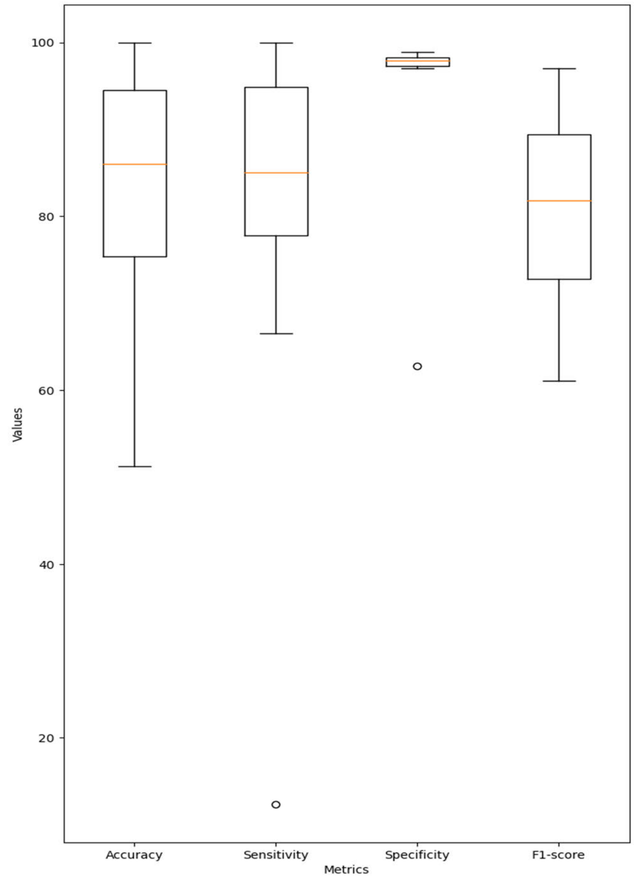

FIGURE 10. Distribution of the metrics (n=92).

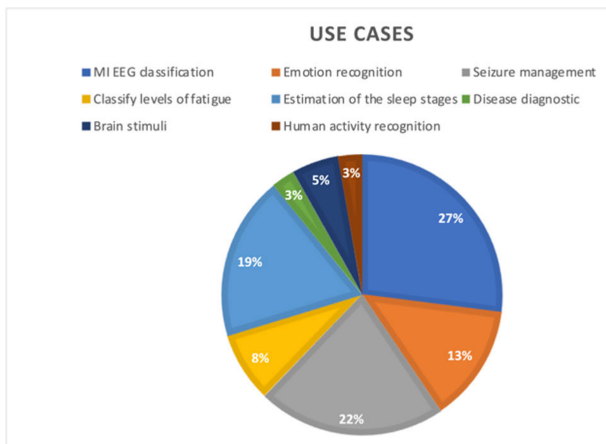

FIGURE 11. Distribution by use case (n=37).

C. SUMMARY OF DATASETS

As we have said before, we have applied the PRISMA method to obtain a set of papers from which we are analyzing the datasets used in them. The following is a brief description of them.

1. BCI competition IV 2a[8]: the imagination of movement of the left hand, right hand, both feet and tongue, [36]
2. BCI competition IV 2b[9]: motor imagery of left hand and right hand, [37]

[8] https://www.bbci.de/competition/iv/#dataset2a
[9] https://www.bbci.de/competition/iv/#dataset2b

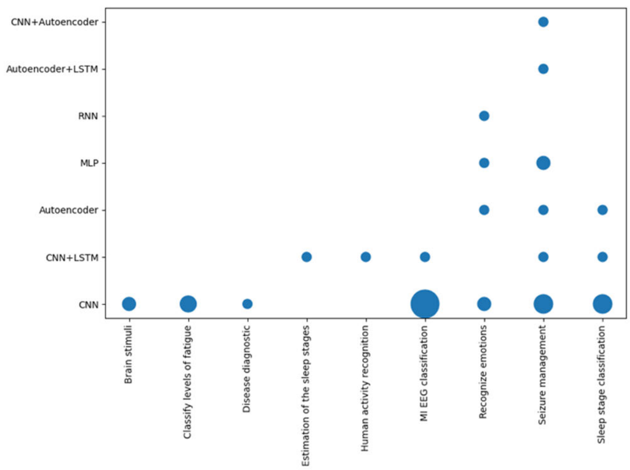

FIGURE 12. **Relationship between the deep learning models and use cases.**

3. DEAP and video signals[10]: emotion recognition of low arousal and low valence (LALV), high arousal and low valence (HALV), low arousal and high valence (LAHV,) and high arousal and high valence (HAHV), [38]
4. Multichannel EEG sustained attention driving task[11]: fatigue and non-fatigued during driving, [39]
5. Temple University EEG Corpus[12]: a compilation of different neural diseases, [40]
6. CHB-MIT Scalp EEG Database[13]: seizure and non-seizure states in epileptic patients, [41]
7. MAHNOB-HCI[14]: a scale of valence and arousal, [42]
8. Sleep EDF[15]: sleep stages after temazepam intake and after placebo intake, [43]
9. Motor Imagery dataset from Weibo et al. 2014[16]: simple MI (left hand, right hand, and feet) and compound MI (both hands, left hand combined with the right foot, right hand combined with the left foot), [44]
10. PhysioNet/CinC Challenge 2018[17]: wakefulness, stage 1, stage 2, stage 3, rapid eye movement (REM), and undefined, [45]
11. Open source SSVEP dataset[18]: healthy subjects focused on 40 characters flickering at different frequencies, [46]
12. BCI Competition III IVa[19]: MI of the left hand, right hand, and right foot, [47]
13. EEG data for driver fatigue detection[20]: drivers suffering fatigue or not, [48].
14. University of Bonn[21]: seizure and non-seizure states, [49].
15. Motor Imagery dataset from Zhou et al. 2016[22]: MI of the left hand, right hand, and feet, [50]
16. Sleep Heart Health Study[23]: sleep scores, [51]
17. EEG datasets for motor imagery brain-computer interface[24]: data for non-task-related and task-related states, [52]
18. DOD-O[25]: scored apnea patients, [53].
19. DOD-H[26]: scored sleep stages, [53].
20. CAP sleep database[27]: activity during NREM sleep, [54]
21. Bern-Barcelona EEG database[28]: patients have pharmacoresistant focal-onset epilepsy, [55]
22. MrOS Sleep[29]: sleep study, [56]
23. Database-Imaged-Vowels-1[30]: pronounce the five main vowels "a", "e", "i", "o", and "u" and six Spanish words, [57]
24. EEG+NIRS Single-Trial Classification[31]: it conducts two BCI experiments: left versus right-hand motor imagery; mental arithmetic versus resting state, [58].
25. MODMA[32]: this is a dataset for mental-disorder analysis which includes clinically depressed patients and controls, [59]
26. BehaveNET[33]: human task recognition of reading, speaking and watching TV.
27. EEG Sweeteners AI[34]: this study evaluated brain signals from 11 healthy subjects when they tasted passion fruit juice equivalently sweetened with sucrose, sucralose, and aspartame, [60]
28. MESA[35]: sleep study to understand how variations in sleep and sleep disorders vary across gender and ethnic groups and relate to measures of subclinical atherosclerosis, [56]

[10]https://www.eecs.qmul.ac.uk/mmv/datasets/deap/
[11]https://figshare.com/articles/dataset/Multi-channel_EEG_recordings_during_a_sustained-attention_driving_task/6427334/5
[12]https://isip.piconepress.com/projects/tuh_eeg/
[13]https://physionet.org/content/chbmit/1.0.0/
[14]https://mahnob-db.eu/hci-tagging/
[15]https://www.physionet.org/content/sleep-edfx/1.0.0/
[16]http://moabb.neurotechx.com/docs/generated/moabb.datasets.Weibo2014.html
[17]https://archive.physionet.org/physiobank/database/challenge/2018/
[18]http://bci.med.tsinghua.edu.cn/download.html
[19]https://www.bbci.de/competition/iii/desc_IVa.html
[20]https://figshare.com/articles/dataset/The_original_EEG_data_for_driver_fatigue_detection
[21]https://ebrary.net/59044/education/details_public_databases
[22]http://moabb.neurotechx.com/docs/generated/moabb.datasets.Zhou2016.html
[23]https://sleepdata.org/datasets/shhs
[24]https://academic.oup.com/gigascience/article/6/7/gix034/3796323
[25]https://dreem-dod-o.s3.eu-west-3.amazonaws.com/index.html
[26]https://dreem-dod-h.s3.eu-west-3.amazonaws.com/index.html
[27]https://archive.physionet.org/physiobank/database/capslpdb
[28]https://www.upf.edu/web/mdm-dtic/-/1st-test-dataset#.YfgOG1jMIUo
[29]https://sleepdata.org/datasets/mros
[30]http://www.ifp.illinois.edu/speech/speech_web_lg/data/mri/index.html
[31]http://doc.ml.tu-berlin.de/hBCI
[32]http://modma.lzu.edu.cn/data/index/
[33]https://zenodo.org/record/2552600#.ZBdONuzMJ
[34]https://github.com/Atzingen/EEG_Sweetners_AI
[35]https://sleepdata.org/datasets/mesa

29. CBICIC2019[36]: it comprises two subsets of MI with left and right-hand tasks.
30. Deep BCI[37]: classification of steady-state visual evoked potentials (SSVEPs) based BCI from ear EEG, [61]

All the information in the datasets has been collected in the following table attached at the end of the paper. The most used datasets are Sleep EDF and DEAP.

D. STATISTICS AND ANALYSIS OF THE OPEN EEGS'DATASETS

The first important feature in a dataset is the number of individuals which is directly related to the model behavior. Roy et al. [9] show that models increase their performance when the number of subjects exceeds 15. The number of tests is logically related to this feature. The values of both characteristics are compiled in Figure 13 which shows a double diagram bar with their distribution per dataset.

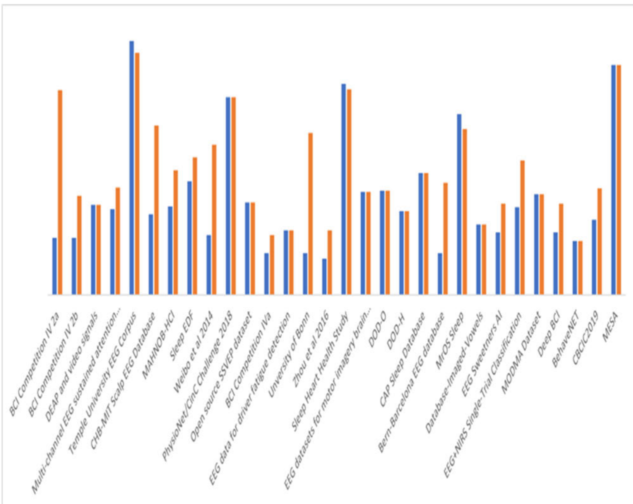

FIGURE 13. Distribution of the number of individuals and test (n=29).

The number of channels is also a critical decision depending on the use case. Jasper [62] tells that a minimum of 21 channels should be used to examine an adult brain. The usage of the number of channels can be seen in Figure 14.

Another particular feature of EEGs is that of the electrodes system which indicates how electrodes are placed around the scalp. Figure 15 shows a pie chart with the percentage of datasets according to the system.

Another interesting measure that will determine the performance of the model is sample frequency. The following bar diagram (Figure 16) represents the distribution of studies according to the frequency used to represent the data. This measure is directly related to the machine used to collect the data. In this case, different frequencies can be used in the same dataset.

[36]https://www.datafoundation.cn/competitions/342/datasets
[37]http://deepbci.korea.ac.kr/

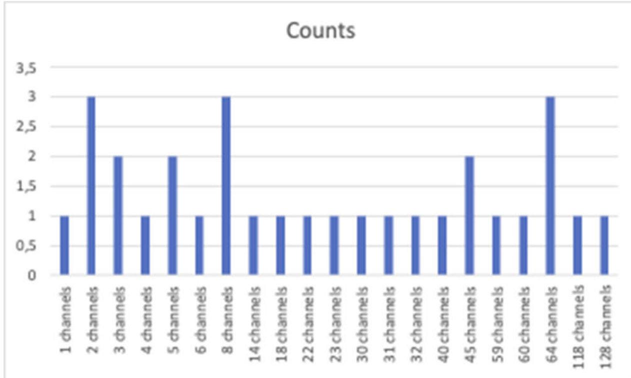

FIGURE 14. Distribution by the number of channels (n=29).

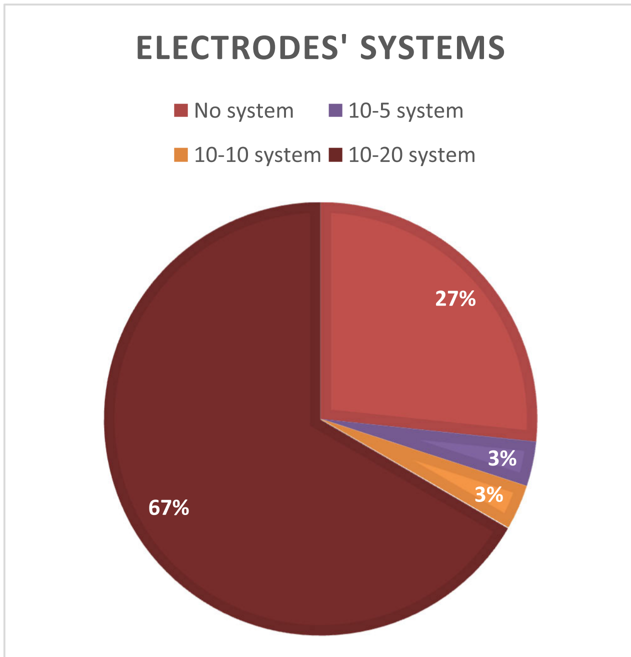

FIGURE 15. Electrodes' systems in percentages (n=30).

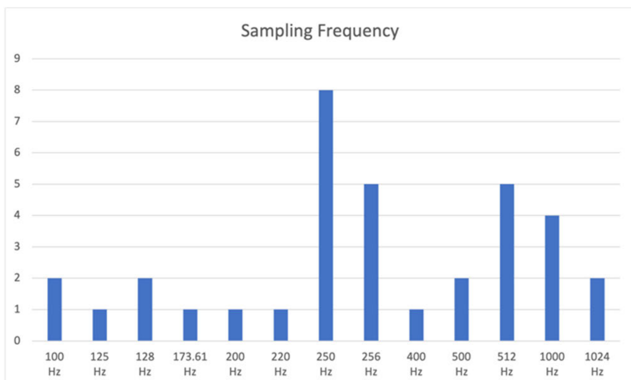

FIGURE 16. Distribution of datasets by sampling frequency (n=35).

Finally, we have a pie chart that compiles the file format used, Figure 17. This depends on the different software used when doing the test.

TABLE 2. Summary of selected datasets.

Dataset	Number of subjects	Total tests	Length per-test	Electrodes' system	Nº channels	Sampling frequency	Format	Papers
BCI Competition IV 2a	9	2,591	5 minutes	10-20 system	22 channels	250 Hz	GDF	3
BCI Competition IV 2b	9	45	5 minutes	No system	3 channels	250 Hz	GDF	3
DEAP and video signals	32	32	40 minutes	10-20 system	45 channels	512 Hz	BDF	5
Multi-channel EEG sustained attention driving task	27	62	90 minutes	10-20 system	32 channels	500 Hz	SET	2
Temple University EEG Corpus	16,986	10,874	20 minutes	10-20 system	31 channels	250 Hz (87%) 256 Hz (8.3%) 400 Hz (3.8%) 512 Hz (1%).	EDF	3
CHB-MIT Scalp EEG Database	22	664	1 to 4 hours	10-20 system	23 channels	256 Hz	EDF	3
MAHNOB-HCI	30	120	30 seconds	10-20 system	32 channels	256 Hz	BDF	1
Sleep EDF	78	197	9 hours	No system	2 channels	100 Hz	EDF	6
Weibo et al., 2014	10	320	8 seconds	10-20 system	60 channels	100 Hz	TXT	1
PhysioNet/CinC Challenge 2018	1,985	1,985	7.7 hours average	10-20 system	6 channels	200 Hz	MAT	2
Open source SSVEP dataset	35	35	4 minutes	10-20 system	64 channels	1000 Hz	MAT	2
BCI Competition **III** IVa	5	10	980 seconds	10-20 system	118 channels	1000 Hz	MAT	1
Driver fatigue detection	12	12	5 minutes	10-20 system	40 channels	1000 Hz	CNT	1
University of Bonn	5	500	23.6 seconds	10-20 system	Single-channel	173.61 Hz	TXT	2
Zhou et al., 2016	4	12	750 seconds	10-20 system	14 channels	250 Hz	CNT	1
Sleep Heart Health Study	3,295	2,651	About 8 hours	No system	2 channels	125 Hz	EDF	4
DOD-O	55	55	387 minutes	No system	8 channels	250 Hz	H5	1
DOD-H	25	25	427 minutes	No system	8 channels	250 Hz	H5	1
CAP Sleep Database	108	108	410 minutes	10-20 system	3 channels	From 128 to 512 Hz	EDF	1
Bern-Barcelona EEG database	5	3,740	20 seconds	10-20 system	64 channels	512 or 1024 Hz	TXT	1
MrOS Sleep	1,026	586	341 minutes	10-20 system	5 channels	256 Hz	EDF	1
Database-Imaged-Vowels	15	15	110 seconds	10-20 system	18 channels	1024 Hz	MAT	1
CBCI2019	18 and 6	60 and 40	4 seconds	No system	59 channels	250 and 1000 Hz	MAT	2
Deep BCI	11	33	1200 Seconds	10-20 system	8 channels	500 Hz	MAT	1
EEG Sweetners AI	11	33	16 seconds	10-20 system	2 channels	512 Hz	EDF	1
BehaveNET	8	8	300-380 seconds	No system	4 channels	220 Hz	CSV	1
EEG+NIRS Single-Trial	29	174	147 seconds	10-5 system	30 channels	128 Hz	MAT	2
MODMA Dataset	48	48	5 minutes	10-20 system	128 channels	250 Hz	MFF	1
MESA	6814	6814	6-8 hours	No system	5 channels	256 Hz	CSV	1
EEG datasets for motor imagery brain computer interface	52	52	51 minutes	10-10 system	64 channels	512 Hz	MAT	1

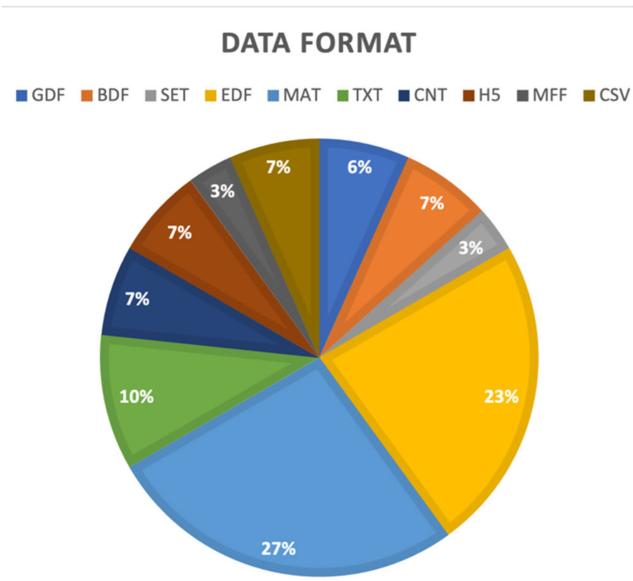

FIGURE 17. Distribution by file format (n=29).

V. DISCUSSION

This work provides a compilation of open EEG datasets analyzed using DL models in a set of papers selected by applying the PRISMA method in a systematic review. The results of the previous section are discussed below from a double perspective: on the one hand, the papers and the DL models used, and on the other, the datasets.

The first part of the statistical analysis starts with the year of publication of the papers. Figure 7 verifies, in part, the trend of papers in deep learning, mentioned in Figure 2. The number of articles published between 2018 and 2021 shows a significant increase. However, the total number is still small, and we can conclude that there is room for creating new open datasets available to the community. It is foreseeable that more articles on EEG and deep learning will be published in the coming years.

As can be seen in Figure 8, the most commonly used DL model, by a wide margin, is the CNN, which appears in 65% of the cases either as a 1-dimensional CNN (EEGs are processed channel by channel) or 2-dimensional CNNs, (EEGs are processed as a whole). Then there is a set of papers that use a hybrid model of CNN with LSTM, 10%. This is followed by hybrid models of Autoencoder and MLP (7%), RNNs (5%) and finally LSTMs or Autoencoder plus LSTM (3%). These numbers give us several ideas. First, using CNNs is successful but less innovative. This makes sense as EEGs can be managed as an image with convolutional filters. Second, using hybrid models seems an opportunity to make new contributions to the field. Finally, GAN and GCNN models are not used which is shocking. The first one has many applications in the creation of synthetic data (very useful considering the shortage of EEGs) or artefact removal (a typical task after collecting this data). The other can be used to model EEGs as graphs and study brain connectivity.

By analyzing the results in Figure 9, we can see that accuracy is the most used metric. This is meaningful as accuracy is the baseline metric to know if a deep learning model performs correctly. Otherwise, the fact of only working with accuracy leads to incomplete experiments as this metric only measures the number of hits. Accuracy has problems in models that use imbalanced datasets and does not give more interpretation of the performance of the model as it does not consider false positives and false negatives like sensitivity, specificity, and F1-score, [63]. Another conclusion obtained from the metrics is that none of the metrics measures the loss of the models which means that all the datasets are considered classification problems.

More information about the metrics is compiled in Figure 10. As we can see, all the metrics obtain values around 90% except specificity which performs near 100%, but with a small set of values. We can also see that the F1-score and sensitivity are more stable except for an outlier in the latter with a poor performance near 10%. We confirm that accuracy is the most used but with a wider range of values which indicates is not as precise as the others.

Regarding Figure 11, the most frequent use case is MI EEG classification, with more than 27% of the cases. This fact is related to BCI Competition IV,[38] a famous data resource in the field comprising a set of datasets for signal processing and BCI classification. Then, we can highlight three use cases among the rest: seizure management, sleep stage classification, and emotion recognition. The rest of the use cases only occur once, twice, or three times: disease diagnostic, human activity recognition, brain stimuli, and classification in levels of fatigue. We can conclude with this analysis that if we want to publish a dataset that brings value to the field, the last four use cases are not exploited a lot.

The features of DL models and use cases are represented in Figure 12. The biggest bubble representing MI EEG classification with CNN makes sense because the DL model is the most popular in its category and there are several MI datasets. For example, those that are part of the BCI Competition. Regarding this use case, we can see that only models with CNN are used, so there is room to experiment with other models. In the second position, we have papers using CNN in stages of fatigue, sleep classification, and seizure management. The latter has been studied with several DL models, so it seems there are not many opportunities to work with this data. The information on the chart can be used to identify what models can be used with our dataset. Also, to find combinations that have not been applied before to do new contributions to science. The rest of the combinations have few instances or none, so they can be considered niches to research. For example, using models that are not CNN in cases like levels of fatigue, brain stimuli, classification of sleep stages or diagnostic of diseases.

The second part of the statistical analysis comprises the datasets' features. We first find the number of subjects which ranges from 4 to 16,986. This is directly related to the number of tests going from 4 to 10,874, the mismatch with the previous values is because the dataset with the most individuals has not recorded a test for each of them, Figure 13. In fact, this a strange situation as most of the time the number of tests is greater than the number of individuals. As we can see most of the datasets are in the low range which confirms that compiling EEGs is not an easy task.

Test duration ranges from seconds to hours (usually, these are sleep studies or patients with epilepsy). The length of the tests in seconds occurred 6 times ranging from 4 to 30 with an average of 16.93 seconds. In the case of minutes, we found 14 experiments with lengths from 4 to 51 and an average value of about 14 minutes. Finally, there are 12 examples with tests lasting at least 1 hour ranging to 9 with an average of almost 6 hours. This is directly related to the use case as epileptic seizures only need seconds to be analyzed but sleep stages need hours.

The number of channels used in the datasets is shown in Figure 14. In our case, the different options are well distributed with works using only 1 channel and others using 128 channels. However, configurations of 64, 32, 16, and 8 channels, (Montoya-Martinez, Bertrand, and Francart 2019), which are recommended do not outstand. No analysis supports this recommendation for deep learning studies, so it could be a future work to be developed.

Other features that have not been studied under a minimum standard to be met are the electrodes system and the sampling frequency. As can be seen in Figure 15, 10-20 is the most used electrode system by far, which makes sense due to the following aspects. It is an international recommendation, [65]. Reference [66] highlights that it is also one of the most used. Other datasets do not provide this information or do not use one due to the number of channels. Figure 16 describes the use of sample frequency. In the first position, we can find 8 times a sampling frequency of 250 Hz. Then, datasets using 256, 512, and 1000 Hz are also noteworthy. Regarding the minimum Hz to obtain good performances in DL models, [67] demonstrate that a higher frequency does not provide better results. Nevertheless, there are no scientific papers that measure the minimum to obtain DL models that perform well, so it could be a future approach.

Finally, Figure 17 gives information about the file formats that have been used. In the first position, we find a format related to EEGLAB,[39] a well-known MATLAB tool for brain

[38] https://www.bbci.de/competition/iv/

[39] https://sccn.ucsd.edu/eeglab/index.php

signal processing. The second position is for European Data Format (EDF) a standard for storing multichannel biological and physiological signals, [68]. The rest of the formats are widely distributed.

VI. CONCLUSIONS AND FUTURE WORKS

This work provides a compilation of open EEG datasets from papers that apply deep learning models. It should be highlighted that the work doesn't consider datasets available upon request. We have used PRISMA to define a workflow for selecting a set of papers that uses these kinds of datasets. Our initial search returned 331 works which, after screening based on the inclusion/exclusion criteria, were reduced to 37. In these papers, 30 datasets were found. Some clear conclusions related to DL techniques are obtained: convolutional neural networks are widely used due to their link with the nature of the data, MI classification is the most common use case and accuracy is the most used metric, but others are more stable. By combining the first and second conclusions, we know that most of the papers apply CNNs to MI use cases. The conclusions related to the datasets comprise: EGGs are difficult to compile due to the low number of instances in general, the number of channels is not relevant so it should be studied, the most used electrode system is the 10-20 system, most relevant sample frequency should also be analyzed and EDF and MAT file formats stand out from the rest.

Further analysis concludes that the number of published papers per year is remarkable, but it is still worth working in the field. From 2018 to 2021, the amount has increased. But in the last 2 years have decreased a little. So, publishing open datasets is relevant for the scientific community. Related to the DL models, we can see that CNNs are a good solution which is why they have been widely applied. The graphics of the use cases are helpful to find application fields that have not been covered a lot or knowing which kind of datasets can obtain good results. The bubble diagram can be used by researchers to know which DL models should be involved in their datasets depending on the use case. In this way, there are several use cases not very exploited, but the use of CNN is not innovative in any case. The analysis of the dataset's characteristics leads us to conclude that the 10-20 system is the most widely used when collecting the data. No work supports the idea that this is the most efficient one. The sample rate of the datasets is very diverse; therefore, none is a priori better than the other. In the case of the number of channels and sample frequency, values are very distributed and again no works are supporting which values should be recommended.

The main limitation of the study is the number of selected works because there are not several papers accomplishing the criteria. As EEGs are medical data, people are reluctant to make them freely available, and researchers who compile the EEGs do not want to share them since they prefer to exploit them themselves. Another reason is the difficulty of collecting a good quality bank of EEGs as it is costly in terms of time. Another limitation of the work is that the authors of the papers using the datasets are not the same as those who have published them. This condition supposes a decoupling between the medical and computer science perspectives, not considering that both profiles are necessary.

Some niches to consider are the following. The use of Natural Language Processing (NLP) techniques such as Transformers and GCNN for not being so exploited. NLP models are one of the most advanced nowadays. If we make a parallelism between texts and EEGs, a sentence can be considered a channel, and a word in the sentence is a particular measure of the channel. This approach could be a starting point for applying these powerful models with EEGs. Another exciting application is studying the network connectivity that can be modelled by representing the EEGs as graphs. In this case, GCNNs are very useful and seem to be a niche.

As future work, the review shows that there is room for finding a gold standard of the characteristics of an EEG dataset to be used in multidisciplinary teams of physicians and computer scientists because sometimes the needs of some do not match those of others. Only one work has been found that studies a single characteristic of the datasets, the number of subjects, [9]. Thus, we propose to carry out different studies in the future to discover how the electrode system, the number of channels, or the sample rate influence obtaining good results when using DL models.

REFERENCES

[1] M. Völske, "Web archive analytics," 2021, *arXiv:2107.00893*.
[2] H.-H. Chang and J. M. F. Moura, "Biomedical signal processing," *Biomed. Eng. Des. Handb.*, vol. 2, pp. 559–579, Jan. 2010.
[3] W. F. Velicer and P. C. Molenaar, "Time series analysis for psychological research," in *Handbook of Psychology*, 2nd ed., I. Weiner, J. A. Schinka, and W. F. Velicer, Eds., 2012, doi: 10.1002/9781118133880.hop202022.
[4] K. Normandeau, "Beyond volume, variety and velocity is the issue of big data veracity," *Inside Big Data*, to be published. [Online]. Available: https://insidebigdata.com/2013/09/12/beyond-volume-variety-velocity-issue-big-data-veracity/
[5] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach. Malaysia*. London, U.K.: Pearson, 2016.
[6] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM J. Res. Develop.*, vol. 3, no. 3, pp. 210–229, Jul. 1959.
[7] R. Hecht-Nielsen, "Neurocomputing: Picking the human brain," *IEEE Spectr.*, vol. 25, no. 3, pp. 36–41, Mar. 1988.
[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
[9] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: A systematic review," *J. Neural Eng.*, vol. 16, no. 5, Oct. 2019, Art. no. 051001.
[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 25. Stateline, NV, USA, Dec. 2012, pp. 1097–1105.
[11] Y.-P. Lin, C.-H. Wang, T.-L. Wu, S.-K. Jeng, and J.-H. Chen, "Multilayer perceptron for EEG signal classification during listening to emotional music," in *Proc. TENCON IEEE Region Conf.*, Oct. 2007, pp. 1–3.
[12] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.
[13] M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo, and J. Xiang, "Epileptic seizure detection based on EEG signals and CNN," *Frontiers Neuroinform.*, vol. 12, p. 95, Dec. 2018.

[14] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG emotion recognition using dynamical graph convolutional neural networks," *IEEE Trans. Affect. Comput.*, vol. 11, no. 3, pp. 532–541, Jul. 2020.

[15] J. L. Elman, "Finding structure in time," *Cognit. Sci.*, vol. 14, no. 2, pp. 179–211, Mar. 1990.

[16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[17] G. Ruffini, D. Ibanez, M. Castellano, S. Dunne, and A. Soria-Frisch, "EEG-driven RNN classification for prognosis of neurodegeneration in at-risk patients," in *Proc. Int. Conf. Artif. Neural Netw.*, 2016, pp. 306–313.

[18] S. Alhagry, A. A. Fahmy, and R. A. El-Khoribi, "Emotion recognition based on EEG using LSTM recurrent neural network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 10, 2017, doi: 10.14569/IJACSA.2017.081046.

[19] J. X. Chen, D. M. Jiang, and Y. N. Zhang, "A hierarchical bidirectional GRU model with attention for EEG-based emotion classification," *IEEE Access*, vol. 7, pp. 118530–118540, 2019.

[20] R. Sathya and A. Abraham, "Comparison of supervised and unsupervised learning algorithms for pattern classification," *Int. J. Adv. Res. Artif. Intell.*, vol. 2, no. 2, pp. 34–38, 2013.

[21] D. H. Ballard, "Modular learning in neural networks," in *Proc. AAAI*, vol. 647, 1987, pp. 279–284.

[22] Y. Qiu, W. Zhou, N. Yu, and P. Du, "Denoising sparse autoencoder-based ictal EEG classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 9, pp. 1717–1726, Sep. 2018.

[23] D. E. Rumelhart and J. L. McClelland, "Information processing in dynamical systems: Foundations of harmony theory," in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*. MIT Press, 1987, pp. 194–281.

[24] R. Salakhutdinov and G. Hinton, "Deep Boltzmann Machines," in *Proc. 12th Int. Conf. Artif. Intell. Statist.*, vol. 5, 2009, pp. 448–455. [Online]. Available: https://proceedings.mlr.press/v5/salakhutdinov09a.html

[25] N. Lu, T. Li, X. Ren, and H. Miao, "A deep learning scheme for motor imagery classification based on restricted Boltzmann machines," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 6, pp. 566–576, Jun. 2017.

[26] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, vol. 2. Cambridge, MA, USA: MIT Press, 2014, pp. 2672–2680.

[27] Y. Luo and B. Lu, "EEG data augmentation for emotion recognition using a conditional Wasserstein GAN," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 2535–2538.

[28] S. Kurz, H. D. Gersem, A. Galetzka, A. Klaedtke, M. Liebsch, D. Loukrezis, S. Russenschuck, and M. Schmidt, "Hybrid modeling: Towards the next level of scientific computing in engineering," *J. Math. Ind.*, vol. 12, no. 1, pp. 1–12, Dec. 2022.

[29] B. A. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Keele Univ., Durham Univ. Joint Rep. EBSE 2007-001, 2007.

[30] A. Barua, M. U. Ahmed, and S. Begum, "A systematic literature review on multimodal machine learning: Applications, challenges, gaps and future directions," *IEEE Access*, vol. 11, pp. 14804–14831, 2023.

[31] V. Thongchotchat, Y. Kudo, Y. Okada, and K. Sato, "Educational recommendation system utilizing learning styles: A systematic literature review," *IEEE Access*, vol. 11, pp. 8988–8999, 2023.

[32] M. N. Y. Marican, S. A. Razak, A. Selamat, and S. H. Othman, "Cyber security maturity assessment framework for technology startups: A systematic literature review," *IEEE Access*, vol. 11, pp. 5442–5452, 2023.

[33] S. D. A. Bujang, A. Selamat, O. Krejcar, F. Mohamed, L. K. Cheng, P. C. Chiu, and H. Fujita, "Imbalanced classification methods for student grade prediction: A systematic literature review," *IEEE Access*, vol. 11, pp. 1970–1989, 2023.

[34] D. Moher, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *Ann. Internal Med.*, vol. 151, no. 4, p. 264, Aug. 2009.

[35] P. Barrett, J. Hunter, J. T. Miller, J.-C. Hsu, and P. Greenfield, "matplotlib—A portable Python plotting package," in *Astronomical Data Analysis Software and Systems XIV* (ASP Conference Series), Pasadena, CA, USA, vol. 347, P. Shopbell, M. Britton, and R. Ebert, Eds. San Francisco, CA, USA: Astronomical Society of the Pacific, 2005, p. 91.

[36] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b," *Frontiers Neurosci.*, vol. 6, p. 39, Mar. 2012.

[37] C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, G. Pfurtscheller, and B. C. I. Competition, *Graz Data Set A, Provided by the Institute for Knowledge Discovery (Laboratory of Brain–Computer Interfaces)*. Graz, Austria: Graz Univ. Technology, 2008.

[38] S. Koelstra, C. Muhl, M. Soleymani, J. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis;Using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jan. 2012.

[39] Y. Zhao, X. Cao, J. Lin, D. Yu, and X. Cao, "Multimodal emotion recognition model using physiological signals," 2019, *arXiv:1911.12918*.

[40] I. Obeid and J. Picone, "The temple university hospital EEG data corpus," *Frontiers Neurosci.*, vol. 10, p. 196, May 2016.

[41] A. H. Shoeb, *Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment*. Cambridge, MA, USA: Massachusetts Institute of Technology, 2009.

[42] J. Lichtenauer, J. Shen, M. Valstar, and M. Pantic, "Cost-effective solution to synchronised audio-visual data capture using multiple sensors," *Image Vis. Comput.*, vol. 29, no. 10, pp. 666–680, Sep. 2011.

[43] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000.

[44] W. Yi, S. Qiu, K. Wang, H. Qi, L. Zhang, P. Zhou, F. He, and D. Ming, "Evaluation of EEG oscillatory patterns and cognitive process during simple and compound limb motor imagery," *PLoS ONE*, vol. 9, no. 12, Dec. 2014, Art. no. e114853.

[45] M. M. Ghassemi, B. E. Moody, L. H. Lehman, C. Song, Q. Li, H. Sun, R. G. Mark, M. B. Westover, and G. D. Clifford, "You snooze, you win: The PhysioNet/Computing in cardiology challenge 2018," in *Proc. Comput. Cardiol. Conf. (CinC)*, vol. 45, Sep. 2018, pp. 1–4.

[46] Y. Wang, X. Chen, X. Gao, and S. Gao, "A benchmark dataset for SSVEP-based brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1746–1752, Oct. 2017.

[47] B. Blankertz, K.-R. Müller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schlogl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schroder, and N. Birbaumer, "The BCI competition 2003: Progress and perspectives in detection and discrimination of EEG single trials," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1044–1051, Jun. 2004.

[48] J. Min, P. Wang, and J. Hu, "Driver fatigue detection through multiple entropy fusion analysis in an EEG-based system," *PLoS ONE*, vol. 12, no. 12, Dec. 2017, Art. no. e0188756.

[49] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 64, no. 6, Nov. 2001, Art. no. 61907.

[50] B. Zhou, X. Wu, Z. Lv, L. Zhang, and X. Guo, "A fully automated trial selection method for optimization of motor imagery based brain-computer interface," *PLoS ONE*, vol. 11, no. 9, Sep. 2016, Art. no. e0162657.

[51] S. Quan, "The sleep heart health study: Design, rationale, and methods," *Sleep*, vol. 20, pp. 1077–1085, Jan. 1998.

[52] H. Cho, M. Ahn, S. Ahn, M. Kwon, and S. C. Jun, "EEG datasets for motor imagery brain–computer interface," *GigaScience*, vol. 6, no. 7, Jul. 2017, Art. no. gix034.

[53] A. Guillot, F. Sauvet, E. H. During, and V. Thorey, "Dreem open datasets: Multi-scored sleep datasets to compare human and automated sleep staging," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 9, pp. 1955–1965, Sep. 2020.

[54] M. G. Terzano, L. Parrino, A. Smerieri, R. Chervin, S. Chokroverty, C. Guilleminault, M. Hirshkowitz, M. Mahowald, H. Moldofsky, A. Rosa, R. Thomas, and A. Walters, "Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep," *Sleep Med.*, vol. 3, no. 2, pp. 187–199, Mar. 2002.

[55] R. G. Andrzejak, K. Schindler, and C. Rummel, "Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 86, no. 4, Oct. 2012, Art. no. 46206.

[56] G.-Q. Zhang, L. Cui, R. Mueller, S. Tao, M. Kim, M. Rueschman, S. Mariani, D. Mobley, and S. Redline, "The national sleep research resource: Towards a sleep data commons," *J. Amer. Med. Inform. Assoc.*, vol. 25, no. 10, pp. 1351–1358, Oct. 2018.

[57] G. A. P. Coretto, I. E. Gareis, and H. L. Rufiner, "Open access database of EEG signals recorded during imagined speech," in *Proc. 12th Int. Symp. Med. Inf. Process. Anal.*, vol. 10160, 2017, Art. no. 1016002.
[58] J. Shin, A. von Lühmann, B. Blankertz, D. Kim, J. Jeong, H. Hwang, and K. Müller, "Open access dataset for EEG+NIRS single-trial classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1735–1745, Oct. 2017.
[59] H. Cai, "MODMA dataset: A multi-modal open dataset for mental-disorder analysis," 2020, *arXiv:2002.09283*.
[60] G. V. V. Atzingen, H. Arteaga, A. R. D. Silva, N. F. Ortega, E. J. X. Costa, and A. C. D. S. Silva, "The convolutional neural network as a tool to classify electroencephalography data resulting from the consumption of juice sweetened with caloric or non-caloric sweeteners," *Frontiers Nutrition*, vol. 9, Jul. 2022, Art. no. 901333.
[61] N. Kwak and S. Lee, "Error correction regression framework for enhancing the decoding accuracies of ear-EEG brain–computer interfaces," *IEEE Trans. Cybern.*, vol. 50, no. 8, pp. 3654–3667, Aug. 2020.
[62] H. H. Jasper, "The ten-twenty electrode system of the international federation," *Electroencephalogr. Clin. Neurophys.*, vol. 10, pp. 370–375, Feb. 1958.
[63] E. Mortaz, "Imbalance accuracy metric for model selection in multi-class imbalance classification problems," *Knowl.-Based Syst.*, vol. 210, Dec. 2020, Art. no. 106490.
[64] J. Montoya-Martínez, J. Vanthornhout, A. Bertrand, and T. Francart, "Effect of number and placement of EEG electrodes on measurement of neural tracking of speech," *PLoS ONE*, vol. 16, no. 2, 2021, Art. no. e0246769, doi: 10.1371/journal.pone.0246769.
[65] S. Yang and F. Deravi, "On the usability of electroencephalographic signals for biometric recognition: A survey," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 6, pp. 958–969, Dec. 2017.
[66] Management Association, *Assistive Technologies: Concepts, Methodologies, Tools, and Applications: Concepts, Methodologies, Tools, and Applications*. Hershey, PA, USA: IGI Global, 2013.
[67] T. Wen, Y. Du, T. Pan, C. Huang, and Z. Zhang, "A deep learning-based classification method for different frequency EEG data," *Comput. Math. Methods Med.*, vol. 2021, pp. 1–13, Oct. 2021.
[68] B. Kemp and J. Olivan, "European data format 'plus' (EDF+), an EDF alike standard format for the exchange of physiological data," *Clin. Neurophysiol.*, vol. 114, no. 9, pp. 1755–1761, Sep. 2003.
[69] G. Xu, X. Shen, S. Chen, Y. Zong, C. Zhang, H. Yue, M. Liu, F. Chen, and W. Che, "A deep transfer convolutional neural network framework for EEG signal classification," *IEEE Access*, vol. 7, pp. 112767–112776, 2019.
[70] H. Wu, Y. Niu, F. Li, Y. Li, B. Fu, G. Shi, and M. Dong, "A parallel multiscale filter bank convolutional neural networks for motor imagery EEG classification," *Frontiers Neurosci.*, vol. 13, p. 1275, Nov. 2019.
[71] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Hum. Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, Nov. 2017.
[72] Y.-H. Kwon, S.-B. Shin, and S.-D. Kim, "Electroencephalography based fusion two-dimensional (2D)-convolution neural networks (CNN) model for emotion recognition system," *Sensors*, vol. 18, no. 5, p. 1383, Apr. 2018.
[73] D. Wang and Y. Shang, "Modeling physiological data with deep belief networks," *Int. J. Inf. Educ. Technol.*, vol. 3, no. 5, p. 505, 2013.
[74] J.-H. Maeng, D.-H. Kang, and D.-H. Kim, "Deep learning method for selecting effective models and feature groups in emotion recognition using an Asian multimodal database," *Electronics*, vol. 9, no. 12, p. 1988, Nov. 2020.
[75] J. J. Podmore, T. P. Breckon, N. K. N. Aznan, and J. D. Connolly, "On the relative contribution of deep convolutional neural networks for SSVEP-based bio-signal decoding in BCI speller applications," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 611–618, Apr. 2019.
[76] J. Pedoeem, S. Abittan, G. B. Yosef, and S. Keene, "TABS: Transformer based seizure detection," in *Proc. IEEE Signal Process. Med. Biol. Symp. (SPMB)*, Dec. 2020, pp. 1–6.
[77] J. Cui, Z. Lan, Y. Liu, R. Li, F. Li, O. Sourina, and W. Müller-Wittig, "A compact and interpretable convolutional neural network for cross-subject driver drowsiness detection from single-channel EEG," *Methods*, vol. 202, pp. 173–184, Jun. 2022.
[78] L. Xu, M. Xu, Y. Ke, X. An, S. Liu, and D. Ming, "Cross-dataset variability problem in EEG decoding with deep learning," *Frontiers Hum. Neurosci.*, vol. 14, p. 103, Apr. 2020.
[79] M. Miao, W. Hu, H. Yin, and K. Zhang, "Spatial-frequency feature learning and classification of motor imagery EEG based on deep convolution neural network," *Comput. Math. Methods Med.*, vol. 2020, pp. 1–13, Jul. 2020.
[80] A. Abdelhameed and M. Bayoumi, "A deep learning approach for automatic seizure detection in children with epilepsy," *Frontiers Comput. Neurosci.*, vol. 15, p. 29, Apr. 2021.
[81] Y. Liu, Z. Lan, J. Cui, O. Sourina, and W. Müller-Wittig, "Inter-subject transfer learning for EEG-based mental fatigue recognition," *Adv. Eng. Informat.*, vol. 46, Oct. 2020, Art. no. 101157.
[82] W. M. Shalash, "A deep learning CNN model for driver fatigue detection using single EEG channel," *J. Theor. Appl. Inf. Technol.*, vol. 99, no. 2, pp. 462–477, 2021.
[83] H. Korkalainen, J. Aakko, S. Nikkonen, S. Kainulainen, A. Leino, B. Duce, I. O. Afara, S. Myllymaa, J. Töyräs, and T. Leppänen, "Accurate deep learning-based sleep staging in a clinical population with suspected obstructive sleep apnea," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 7, pp. 2073–2081, Jul. 2020.
[84] L. C. Sarmiento, S. Villamizar, O. López, A. C. Collazos, J. Sarmiento, and J. B. Rodríguez, "Recognition of EEG signals from imagined vowels using deep learning methods," *Sensors*, vol. 21, no. 19, p. 6503, Sep. 2021.
[85] P. R. A. S. Bassi, W. Rampazzo, and R. Attux, "Transfer learning and SpecAugment applied to SSVEP based BCI classification," *Biomed. Signal Process. Control*, vol. 67, May 2021, Art. no. 102542.
[86] A. Guillot and V. Thorey, "RobustSleepNet: Transfer learning for automated sleep staging at scale," 2021, *arXiv:2101.02452*.
[87] F. Li, R. Yan, R. Mahini, L. Wei, Z. Wang, K. Mathiak, R. Liu, and F. Cong, "End-to-end sleep staging using convolutional neural network in raw single-channel EEG," *Biomed. Signal Process. Control*, vol. 63, Jan. 2021, Art. no. 102203.
[88] H. Banville, O. Chehab, A. Hyvärinen, D.-A. Engemann, and A. Gramfort, "Uncovering the structure of clinical EEG signals with self-supervised learning," *J. Neural Eng.*, vol. 18, no. 4, Aug. 2021, Art. no. 046020.
[89] R. Yan, F. Li, D. D. Zhou, T. Ristaniemi, and F. Cong, "Automatic sleep scoring: A deep learning architecture for multi-modality time series," *J. Neurosci. Methods*, vol. 348, Jan. 2021, Art. no. 108971.
[90] E. Eldele, Z. Chen, C. Liu, M. Wu, C. Kwoh, X. Li, and C. Guan, "An attention-based deep learning approach for sleep stage classification with single-channel EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 809–818, 2021.
[91] X. Huang, K. Shirahama, F. Li, and M. Grzegorzek, "Sleep stage classification for child patients using DeConvolutional neural network," *Artif. Intell. Med.*, vol. 110, Nov. 2020, Art. no. 101981.
[92] J. Liu, G. Wu, Y. Luo, S. Qiu, S. Yang, W. Li, and Y. Bi, "EEG-based emotion classification using a deep neural network and sparse autoencoder," *Frontiers Syst. Neurosci.*, vol. 14, p. 43, Sep. 2020.
[93] R. San-Segundo, M. Gil-Martín, L. F. D'Haro-Enríquez, and J. M. Pardo, "Classification of epileptic EEG recordings using signal transforms and convolutional neural networks," *Comput. Biol. Med.*, vol. 109, pp. 148–158, Jun. 2019.
[94] M. R. Islam, M. M. Islam, M. M. Rahman, C. Mondal, S. K. Singha, M. Ahmad, A. Awal, M. S. Islam, and M. A. Moni, "EEG channel correlation based model for emotion recognition," *Comput. Biol. Med.*, vol. 136, Sep. 2021, Art. no. 104757.
[95] Y. Zhang, H. Cai, L. Nie, P. Xu, S. Zhao, and C. Guan, "An end-to-end 3D convolutional neural network for decoding attentive mental state," *Neural Netw.*, vol. 144, pp. 129–137, Dec. 2021.
[96] O. Baser, M. Yavuz, K. Ugurlu, F. Onat, and B. U. Demirel, "Automatic detection of the spike-and-wave discharges in absence epilepsy for humans and rats using deep learning," *Biomed. Signal Process. Control*, vol. 76, Jul. 2022, Art. no. 103726.
[97] J. Zhang, B. Xu, and H. Yin, "Depression screening using hybrid neural network," *Multimedia Tools Appl.*, vol. 82, pp. 26955–26970, Mar. 2023.
[98] M. Alhussein, G. Muhammad, M. S. Hossain, and S. U. Amin, "Cognitive IoT-cloud integration for smart healthcare: Case study for epileptic seizure detection and monitoring," *Mobile Netw. Appl.*, vol. 23, no. 6, pp. 1624–1635, Dec. 2018.
[99] Y. Li, X. Zhang, and D. Ming, "Early-stage fusion of EEG and fNIRS improves classification of motor imagery," *Frontiers Neurosci.*, vol. 16, p. 149, Jan. 2023.

[100] Y. Zhu, Y. Li, J. Lu, and P. Li, "EEGNet with ensemble learning to improve the cross-session classification of SSVEP based BCI from ear-EEG," *IEEE Access*, vol. 9, pp. 15295–15303, 2021.
[101] R. N. Bairagi, M. Maniruzzaman, S. Pervin, and A. Sarker, "Epileptic seizure identification in EEG signals using DWT, ANN and sequential window algorithm," *Soft Comput. Lett.*, vol. 3, Dec. 2021, Art. no. 100026.
[102] A. Salehzadeh, A. P. Calitz, and J. Greyling, "Human activity recognition using deep electroencephalography learning," *Biomed. Signal Process. Control*, vol. 62, Sep. 2020, Art. no. 102094.
[103] M. Ma, Y. Cheng, X. Wei, Z. Chen, and Y. Zhou, "Research on epileptic EEG recognition based on improved residual networks of 1-D CNN and indRNN," *BMC Med. Informat. Decis. Making*, vol. 21, no. S2, pp. 1–13, Jul. 2021.

ALBERTO NOGALES received the master's degree in software engineering and artificial intelligence and the Ph.D. degree from the University of Alcalá, Spain, with a focus on semantic web and social network analysis. He has experience in several European research projects and publications in conferences and journals from JCR. His education has been completed working at Technische Universität Wien, Austria; Tallinna Tehnikaülikool, Estonia; and the University of Málaga, Spain. Since November 2017, he has been a Postdoctoral Researcher and a Lecturer with the CEIEC, Research Institute, Universidad Francisco de Victoria. He is a Computer Scientist. He is working on deep learning and its use in signal processing, predictive models, and computer vision in medical applications.

ÁLVARO J. GARCÍA-TEJEDOR received the Ph.D. degree in biochemistry. He has been a Lecturer with UCM, UC3M, and Universidad Francisco de Vitoria (UFV), where he is currently an Associate Professor in artificial intelligence. He has been involved in more than 20 research and development projects mainly in AI. As a researcher, he started doing mathematical modeling of biological systems. Since 2007, he heads CEIEC, Research Institute, UFV, for technological innovation with a social accent. He has produced several serious games for the transmission of educational and cultural content and integration of people with intellectual disabilities. His current work is focused on neural models, deep learning techniques, and their applications in several fields.