IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 1 December 2023, accepted 23 January 2024, date of publication 30 January 2024, date of current version 16 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3360328

SURVEY

A Comprehensive Survey of EEG Preprocessing Methods for Cognitive Load Assessment

KONSTANTINA KYRIAKI[1], DIMITRIOS KOUKOPOULOS[2], AND CHRISTOS A. FIDAS[1]
[1]Department of Electrical and Computer Engineering, University of Patras, 265 04 Patras, Greece
[2]Department of History and Archaeology, University of Patras, 265 04 Patras, Greece

Corresponding author: Konstantina Kyriaki (up1070027@upatras.gr)

This work was supported by the Hellenic Foundation for Research and Innovation (HFRI) through the 2nd Call for proposals for H.F.R.I. Research Projects to Support Faculty Members and Researchers, under the project "Electroencephalography and Eye Gaze driven Framework for Intelligent and Real-Time Human Cognitive Modelling (CogniX)," under Grant 3849. The publication of the article in open access mode was financially supported by HEAL-Link Greece.

ABSTRACT Preprocessing electroencephalographic (EEG) signals during computer-mediated Cognitive Load tasks is crucial in Human-Computer Interaction (HCI). This process significantly influences subsequent EEG analysis and the efficacy of Artificial Intelligence (AI) models employed in Cognitive Load Assessment. Consequently, it stands as an indispensable procedure for developing dependable systems capable of adapting to users' cognitive capacities and constraints. We systematically analyzed fifty-seven (57) research papers on computer-mediated Cognitive Load EEG experiments published between 2018 and 2023. The preprocessing methods identified were multiple, controversial, and strongly dependent on the particularities of each experiment and the derived experimental dataset. Our investigation involved the meticulous classification of preprocessing methods based on distinct parameters, namely the degree of user intervention, the noise level, and the subject pool size. Particular attention was paid to semi-automated denoising technology since conventional methods, advanced approaches, and standardized pipelines overwhelm research, but no optimum solution is available yet. This survey is anticipated to provide a valuable contribution to the rising demand for an efficient and fully automated preprocessing approach in EEG-based computerized Cognitive Load experiments.

INDEX TERMS Cognitive load, denoising, electroencephalography, preprocessing, real-time human cognitive modelling, user studies, working memory.

I. INTRODUCTION

Preprocessing is a vital step in analyzing EEG signals, enhancing the quality and reliability of the experimental data before further analysis and interpretation. It involves a series of carefully designed procedures to prepare raw data for further processing and remove noise, artifacts, and other elements that may obscure the underlying neural activity of interest. By preparing the signal in this manner, researchers can improve the accuracy of subsequent feature extraction techniques and the overall performance of the AI models used in Cognitive Load Assessment.

Consequently, preprocessing involves data organization and denoising, that is, cleaning data from noise without distorting the actual signal triggered by neuronal activity [1]. Preprocessing is a flexible process combining several techniques with advantages and disadvantages. Many standardized semi-automated and fully automated pipelines have been suggested in the literature [2], [3], [4], [5], [6], [7], [8], [9], [10]; however, there is no consensus on the optimal approach, method, or protocol. In light of the strong dependency of preprocessing on the nature of the problem and its variations across different research areas, we undertake a comprehensive analysis of Cognitive Load assessment within the HCI field.

Similarly, the definition of artifacts and noise also depends on the nature of the problem [11]. In general, artifacts are non-neuronal factors that degrade the quality of EEG signals, obscure analysis, and bias interpretations. In particular, they can eliminate the classification accuracy, distort BCI devices,

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

and make the diagnosis of brain disorders difficult [12]. They can be classified into internal and external artifacts [13].

Internal artifacts originate from unwanted physiological signals. In this study, ocular and myogenic artifacts prevailed. Ocular artifacts, such as eye blinks, eye saccades, and eye movements, are captured mostly by frontal electrodes. Muscle artifacts originate from any type of body muscle. Head and body movements such as clenching, chewing, swallowing, talking [12], scratches on the head, sneezes [14], and cardiac pulse (1,2 Hz) comprise muscle movements. Myogenic artifacts contaminate Cognitive Load experiments, and their removal is challenging [15]. Other internal artifacts include respiration and sweating.

External artifacts are triggered by environmental factors. They incorporate electromagnetic interference such as line noise 50/60 Hz, radio frequency from nearby devices [12], [16], equipment malfunction artifacts such as improper contact of the EEG leads, positional drift of the leads with time, slow drifts at the skin/electrolyte/electrode interface [17], impedance variations [18], and high-impedance electrodes that result in strong fluctuations and bad channels [19].

To describe the level of mental effort while processing information, the terms ''Cognitive Load'', ''Mental Workload'', and ''Cognitive Workload'' are often used interchangeably. They are all interrelated but are placed in different contexts and stem from different theories. ''Cognitive Load'' stems from the Cognitive Load Theory (CLT), a learning theory whose objective is to elucidate how the load of learning tasks can impact students' capacity to comprehend new information and effectively construct knowledge. When processing new information, students handle it differently from information already stored. Intrinsic and extraneous are the two basic types of Cognitive Load. Intrinsic cognitive load is the essential load required for learning and cannot be altered. Conversely, extraneous cognitive load is unproductive for learning and can be modified, for instance,by environmental distractions [20].

The terms ''Mental workload'' and ''Cognitive workload'' are employed by the Multiple Resource Theory (MRT), which centers around workplace performance and engagement in multitasking activities within a working environment. This theory presents a cube model that illustrates the mental resources needed, their intersections, and their distribution across different stages, codes, modalities, and visual channels [21].

EEG preprocessing approaches have been discussed across various contexts exploring the functioning of the brain, such as neuroscience and cognitive science [22], [23], [24], [25], [26], [27]. The performance of a range of preprocessing methods has been evaluated in bench-marking studies [10], [28], [29], [30]. Some reviews have been particularly focused on removing movement [31], muscle [15], or physiological (ocular, muscle, cardiac) artifacts [12]. Moreover, noise decontamination research in the Brain-Computer Interface (BCI) domain is broad [32]. Comprehensive reviews on the entire EEG signal analysis, specifically on the assessment of Cognitive Load [33], described in detail the preprocessing stage [34], [35]. However, to the best of our knowledge, none of the documented works is specified in reviewing the EEG preprocessing approaches in the context of the cognitive effort required by a user to accomplish a specific computerized task.

This paper begins by presenting the background theory concerning EEG signal preprocessing methods. A two-tier categorization of signal preprocessing methods for Cognitive Load assessment is adopted, specifically a) semi-automated preprocessing and b) fully-automated preprocessing. Given the significance of the semi-automated preprocessing, we comprehensively examined it across three dimensions: a) basic preprocessing methods, b) advanced denoising methods encompassing the time-frequency domain, spatial domain, Blind Source Separation algorithms, deep learning algorithms, and hybrid denoising techniques, and c) standardized preprocessing approaches. Next, we discuss the motivation for conducting this survey and outline research questions. A thorough analysis of the findings for each research question is then introduced. The paper concludes with a comprehensive discussion of the main findings, future directions, and limitations.

II. THEORETICAL BACKGROUND

A. SEMI-AUTOMATED PREPROCESSING

The semi-automated preprocessing protocol is shaped by the researcher's input. The researcher possesses a significant degree of autonomy to visually inspect the recorded EEG signals and define the necessary steps and parameters for their unique preprocessing pipeline. Semi-automated preprocessing extends from a fundamental cleaning and organization of raw EEG data to sophisticated denoising solutions that operate beyond basic cleaning and remove complex patterns. Filtering, segmentation, manual noise rejection, and averaging are some of the methods employed to perform basic preprocessing. More advanced approaches, such as Blind Source Separation (BSS), Empirical Mode Decomposition (EMD), Deep learning, and hybrid solutions, have enriched the quiver of denoising options, resulting in a higher signal-to-noise ratio. Yet, standardized solutions greatly contribute to the overall semi-automated preprocessing of the EEG signal. In essence, the researcher's individual choices determine the signal quality and affect interpretations in subsequent analysis. Although semi-automated preprocessing is subjective, time-intensive, and requires experience and experimentation, it has played a prominent role in Cognitive Load research. Next, we provide a concise overview of commonly employed semi-automated preprocessing techniques.

1) SEMI-AUTOMATED BASIC PREPROCESSING

Several conventional methods exist that organize the raw EEG data, such as resampling, re-referencing, and segmentation. Researchers use *resampling* to either increase (upsampling) or decrease (downsampling) the number of data

points in EEG data. This is done to facilitate frequency reconstruction for further analysis or to reduce computer resource demands related to storage and processing speed [36].*Re-referencing* is employed to improve the interpretation of recorded EEG signals. After data acquisition, the reference site can be altered to give the researcher a different option. Any other channel or scheme can be selected for re-reference. The new reference signal is calculated and subsequently subtracted from each channel [17]. The most popular re-referencing approach is the Common Average Reference (CAR), which is the average of all the channels. According to an experiment cited by Lahane et al., CAR is the optimal reference and achieves a high classification accuracy [1]. Moreover, a study assessing functional Brain-Heart Interplay (BHI) compared CAR with mastoid average, Laplacian reference, Cz reference, and the Reference Electrode Standardization Technique (REST) and concluded that CAR was more consistent. It is independent of specific scalp regions and experimental alterations [37]. Furthermore, it is reliable for identifying artifacts that equally influence all channels, such as line noise [3]. In addition to its dependence on the number and position of the active electrodes, CAR may affect the spectral analysis. Nevertheless, it is proposed as the most appropriate choice among the existing methods [37].*Segmentation* divides the original EEG data into segments. These segmented portions, referred to as epochs, have different durations and alignments on specific events, depending on the experimental design. They may overlap and encompass events within their time frame, such as stimulus presentations or participants' responses. They can be created by combining channels from various brain regions [38]. Segmentation is commonly used as it makes the signal more amenable to analysis.

In addition to organizing EEG data, traditional denoising approaches are employed to clean the EEG data. These approaches include manual denoising, filtering, baseline correction, detrending, interpolation, regression, and averaging. *Manual denoising* is cumbersome and subjective, and it may lead to the loss of valuable data segments and reduction of the statistical power of the data [39]. However, it contributes significantly to preprocessing [40, Chapter 7]. Researchers visually examine signal waveforms, spectrograms, and topographic maps to identify artifacts. Eye blinks and saccades induce large and transient artifacts with high amplitude in the signal. EEG activity in the gamma band (30-100Hz) can cause similar effects, leading to misinterpretations [19]. Ocular artifacts, such as eye blinks, eye saccades, and eye movements, are primarily captured by frontal electrodes. Cardiac activity results in small and constant artifacts [33]. Muscle artifacts affect EEG signals differently depending on the muscle, its contraction level, and the participant's sex. They exhibit a wide frequency spectrum [28]. The firing of muscle fibers may induce frequencies from 7 to 20Hz [19], striated muscles can extend from 20Hz to 300Hz, and the temporal, masseter, and frontalis muscles may range from [40-80Hz], [50-60Hz], [30-40Hz] respectively [15] (for a very informative presentation, see [41]). For noisy epochs, the thresholds are determined based on the amplitude, time, or standard deviation from the mean. Other types of thresholds are used to determine bad channels or noisy participants' data.

Filtering has been thoroughly examined for denoising [29]. Using the percentage of significant channels as a data quality metric, filtering is evaluated as the most important preprocessing step [10]. This can improve the Signal-to-Noise ratio and facilitate analysis [42]. Filtering is necessary for Event-Related Potentials (ERPs). Disregarding filtering may decrease the statistical power [40, Chapter 4]. Various filtering approaches have been proposed in the literature. Common approaches, such as the Median filter, Savinsky-Golay Filter, and adaptive filtering, have been frequently applied to smooth noisy EEG signals. A Median Filter is a non-linear filter that is effective for denoising. It is easy to implement and has low area and power requirements [43]. The Savinsky-Golay Filter uses convolution to smooth the signal by determining the window size and polynomial-order parameters [44]. Moreover, Wiener filters eliminate the mean square error between the desired signal and its estimate [22]. It requires calibration, and it is not adequate for online applications [22]. Yet, no reference signal is required [24]. Contrariwise, adaptive filtering uses a reference signal and subtracts it from the raw EEG signal. The least mean square (LMS) method and recursive least square (RMS) method are popular adaptive filtering algorithms. [24]. However, it performs worse than the Wiener filter. In addition, multi-step filtering is an interesting approach. It includes nine band-pass filters followed by spatial filtering. Recently, a motor imagery experiment used a refined two-step filtering method based on gradient-weighted class activation mapping (Grad-CAM). In the first step, raw EEG data were fed into a Convolutional Neural Network (CNN). The resulting Grad-CAM data were used to determine filtered frequencies. In the second stage, the filtered data were fed into the CNN. More stable training data were extracted, and the classification performance was high. Moreover, it was proved robust among different subjects. This promising denoising approach was proposed for denoising EEG signals across contexts because the filtering range is determined by the performance of the network [45]. In general, filtering strongly impacts the signal and requires further consideration. It may distort the signal [42] and cause misinterpretations [40]. EEG data vary across participants [10] and experiments requiring different filtering implementations. Filter design and customization are time-consuming and require expertise. An effective denoising filter requires the specification of various parameters, such as the domain of the filter response, type of impulse response (FIR/IIR), type of frequency response, transition bandwidth, roll-off, filter order, and cut-off frequency [42]. An optimal and automated filtering procedure across studies remains an ongoing debate [29].

Baseline correction often focuses on correcting DC offset. Usually, a time interval, often before stimulus presentation, is specified as a baseline and used to correct the waveform upward or downwards [46], [40, Chapter 2.7]. The effectiveness of baseline correction in data quality is disputed. It may not result in a higher data quality and may be substituted by a high-pass filter [10]. Another simple denoising technique is *Detrending*, where a smoothing function is fitted and subtracted from the signal. By determining the function parameters, researchers can select the type of trends they want to reject [17]. The rejection of contaminated data portions may lead to discontinuities and loss of essential information [11]. A solution is *Interpolation*, which recreates a rejected portion of the data from the remaining healthy data [40, Chapter 7.5]. A threshold is necessary to differentiate the data segments for either rejection or interpolation. The interpolation of channels with a threshold of four standard deviations more line noise than the average line noise of all channels improved the overall quality of the EEG signal [10]. Another classical denoising technique is *Regression*. It involves the calculation of artifact propagation coefficients, which are subsequently subtracted from the raw signal to mitigate their impact. Regression algorithms are simple but require one or more reference channels and suffer from bidirectional contamination. One limitation is that it is effective for denoising the EEG signal only from ocular artifacts [12], [23], [47]. *Averaging* is another commonly employed denoising technique. It diminishes random fluctuations caused by noise and enhances the signal-to-noise ratio when noise overlaps with the signal of interest. Several studies have integrated it into their preprocessing pipelines, primarily to extract features such as ERPs by recording repeated trials of the same task and then averaging the extracted epochs. It does not depend on thresholding like signal decomposition methods. However, time misalignment and individual differences resulting in different epochs can minimize the accuracy of the averaging method [48], [49].

2) SEMI-AUTOMATED ADVANCED DENOISING TECHNIQUES

a: TIME-FREQUENCY DOMAIN

Wavelet Transform (WT) is an advanced denoising method that decomposes the EEG signal into wavelet coefficients in the time-frequency domain by shifting and dilating the mother wavelet function. The artifactual signal is rejected using a threshold, and the clean signal is reconstructed. The user must specify the mother wavelet function, the threshold algorithm, and the level of decomposition [26]. For a cognitive task, the Discrete Meyer (dmey) function proved its denoising efficiency among forty-six (46) mother wavelet functions [50]. Regarding thresholding, there is no consensus. Hard thresholding causes discontinuity, while soft thresholding suffers from deviation issues. Hence, novel approaches, such as grand-based adaptive algorithms and functions related to the decomposition level, have been introduced [51], [52]. Particularly for cognitive tasks, Rigsure hard thresholding demonstrated optimal performance [50]. Various decomposition levels, between 1 and 9, have been employed [53], [54], [55]. In EEG denoising, the Discrete Wavelet Transform (DWT) has been frequently applied [24]. In this method, the signal is decomposed at different levels, and the mother wavelet function is dilated by powers of two [26].

Empirical Mode Decomposition (EMD) is a data-driven denoising method. It decomposes the signal into intrinsic mode functions (IMFs), applying a sifting process. The extraction of an IMF is an iterative procedure where upper and lower envelopes are constructed through interpolation of minimum and maximum points. The mean of the envelopes is subtracted from the input signal. When the minimum and maximum points are equal or differ by one, and the mean of the envelopes at all points is zero, the iteration stops, and an IMF is extracted and subtracted from the input signal. The algorithm ends when the signal becomes a monotonic function [56]. EMD is flexible and efficient and performs better than WT. The Cubic Spline Interpolation and the Akima Spline Interpolation have been performed highly for eye blink removal. However, EMD is not recommended for signals that exhibit abrupt changes or sudden shifts [26]. In addition, it is not recommended for real-time scenarios, especially for large datasets, due to the computer demands during interpolating data points and constructing the envelopes [56].

b: SPACIAL DOMAIN

Signal Space Projection (SSP) decomposes signal into separate spacial components. The amplitude of these components may change in the time domain, but their spatial characteristics remain constant. The method is based on the hypothesis that the subspace of the true EEG signal differs from the noisy subspace. PCA has been widely used to calculate the noisy subspace. SSP causes spacial distortions that may obstruct further analysis. Source Informed Reconstruction (SIR) has been effectively used to minimize such distortions [12], [57].

Beamforming is a robust, spatial filtering denoising technique offered by various EEG analysis software toolboxes. Beamforming algorithms calculate the contribution of a specific location to the space where measurements are done by strengthening the signal of this specific location and weakening the signal of the remaining areas. A vector beamformer calculates output along the three directions,x, y, and z, whereas a scalar beamformer calculates output only for the orientation identified by the forward field. For EEG data, the Common Average Reference is recommended. The beamforming technique is a data-driven method that does not require a priori specification of the location or configuration of artifacts. It performs highly when artifacts have different spatial characteristics from the signal of interest. In addition, it is efficient in denoising signal from external artifacts. Yet, beamforming is ineffective for correlated sources, and choosing the most appropriate algorithm remains elusive [12], [58], [59].

c: BLIND SOURCE SEPARATION (BSS)

A substantial section in the field of denoising is Blind Source Separation (BSS) algorithms. They can denoise signals without prior knowledge of the source and traits of each artifact [60] but require data from many channels. Yet, some single-channel versions, such as single-channel ICA, have been explored [61]. The dominant BSS algorithms are Independent Component Analysis (ICA), Principal Component Analysis (PCA), Canonical Correlation Analysis (CCA), Morphological Component Analysis (MCA) [12], [24], [47], Artifact Subspace Reconstruction (ASR) [11], and Spatio-Spectral Decomposition (SSD). The optimal choice is defined by the type of captured data [15].

ICA decomposes EEG signal into temporally independent sources (Independent Components - ICs). Noisy ICs are manually or automatically detected and rejected or corrected. ICA performs well, and it has been widely adopted across contexts. However, it requires a large amount of data and may result in valuable data loss owing to the removal of whole artifactual ICs [26]. Moreover, the manual detection and removal of ICs is time-consuming, subjective, and requires expert knowledge [11]. Even when automated algorithms are employed for the detection and rejection/correction of ICs, ICA has certain limitations. If there are more time-independent sources than the active electrodes, signal decomposition into ICs is not applicable [39]. In addition, they may have under- or over-clean artifacts [62]. Manual detection and rejection or correction of noisy ICs is cumbersome. To address this, a range of algorithms have been developed for automatic detection (e.g., Infomax-Runica, FastICA, SOBI, and AMICA) and rejection or correction of ICs (e.g., ICLabel, Multiple Artifact Rejection Algorithm (MARA)) [12]. Runica, namely the automated versions of Infomax, FastICA, SOBI, AMUSE, and JADE, have similar denoising performances for myogenic artifacts [15]. Overall, AMICA is the most effective compared to Infomax (runica), FastICA, and SOBI [10]. MARA is an automated and efficient classifier that utilizes a binary linear classifier to determine whether an IC is an artifact or a neuronal signal, thereby enabling researchers to retain or reject it. MARA has demonstrated a strong performance online and in various experimental contexts. It handles effectively different types of artifacts [63], particularly myogenic artifacts [3].

PCA decomposes the EEG signal into uncorrelated components called Principal Components (PCs). A set of PCs indicates artifacts, except for the case of similar amplitudes in the same band between artifacts and signals. PCA is rarely used for denoising [23], [39]. *CCA* calculates the correlation between the true EEG signal and its time-delayed version. These components are derived from uncorrelated sources. It is less computationally demanding than ICA because it uses second-order statistics [64]. In addition, it performs well for muscle artifacts due to its low auto-correlation characteristics [61].*MCA* extracts linearly combined morphological components with different waveforms into dictionaries and calculates the sparsest component. It may perform better than the Wavelet transform, but it is recommended for datasets with a variety of artifact morphology [12]. *ASR* [65] can be performed offline or online. Initially, clean epochs are defined as a reference. Then, it decomposes the signal into PCs according to that reference. It aims to minimize transient artifacts with a large amplitude. The burst cut-off parameter is specified. A small value leads to highly aggressive artifact removal, whereas a larger value results in less aggressive artifact removal [29]. It is powerful, with optimal cut-off parameters in the range of [20,30], and removes more internal artifacts than neuronal activity [66]. Moreover, it requires less computational resources, is faster than ICA, and is preferred for online denoising [11]. Finally, *SSD* generates components by increasing the power at the frequency of interest and minimizing the power and noise at the neighboring frequencies. The SSD performs better than the ICA-driven SOBI algorithm, requiring only a few milliseconds to run. It is applicable even for 1:10 SNR [67].

d: DEEP LEARNING

Finally, the innovative application of deep learning on denoising is promising. Deep Learning, computationally, imitates the behavior of the brain to learn from a large amount of data. It has gained a lot of attention in a variety of contexts, such as image and language processing, motor imagery methods, EEG reconstruction, and creation [68], automatic detection of alcoholism [69], and epileptic seizures [70], because it seems to outperform state-of-the-art methods [71]. Recently, it has been effectively employed for denoising [72]. It is less time-consuming and can handle much data [73]. Moreover, it is flexible once trained and can cope with non-linear and non-stationary artifacts [68]. A major limitation is the large amount of data required for training [28].

Convolutional Neural Networks (CNNs), a subset of deep learning models, are promising in the field of denoising [74]. A CNN is a feed-forward network. It incorporates convolutional layers that perform convolutions to identify patterns, pooling layers for faster parameterization, and fully-connected layers with full connections to the previous layer for high-level reasoning [71]. Given the high performance of CNN architectures for image data, using images extracted by sequential raw EEG data as an input to a CNN constitutes a trend in EEG preprocessing [34], [75]. These images are called topo-maps and preserve the spatial and spectral traits of the signal, namely, the electrode location and adequate frequency information [76]. A CNN can effectively learn the most important features from a particular task: it is fast, accurate, and effectively investigates the spatiotemporal characteristics of the EEG signal [74], [77]. However, without a robust and generalized framework and appropriate performance metrics, CNN's performance depends on parameters defined by the user [77]. Moreover, the CNN approach suffers from over-fitting [28].

Different CNN denoising approaches have been evaluated and compared. A deep one-dimensional residual convolutional neural network (1D-ResCNN), which has been applied

to raw EEG signals, outperformed ICA, Fast Independent Component Analysis (FICA), Recursive least squares(RLS) filter, Wavelet Transform (WT), and Deep neural network (DNN) denoising methods and maintained the non-linear attributes of the signal [74]. CNN denoising approaches have been suggested for the removal of ocular [28], [78] and myogenic artifacts [72]. MultiResUNet3+, a one-dimensional convolutional neural network (1D-CNN), proved superior to other CNN approaches, even for high-noise levels [68].

Recurrent Neural Networks (RNN) have also been applied to denoising. They incorporate layers for pattern recognition. These layers are recursive, meaning that their output influences their input in the following steps. An efficient type of recurrent layer is the Long Short-term Memory (LSTM) layer [79], [80]. RNN performs well for ocular artifacts but not for myogenic artifacts [28].

e: HYBRID DENOISING TECHNIQUES

Researchers have developed hybrid denoising approaches that blend two or more individual methods to achieve more effective denoising outcomes. By doing so, they harness the strengths of each while mitigating their respective limitations [26]. The subsequent text outlines some of the most widely adopted hybrid methods.

WT is often coupled with a BSS technique or vice versa [26]. Usually, researchers employ the WT as the preprocessing agent, followed by applying a BSS method for further denoising [26]. For instance, WT has been coupled with ICA to remove large movement artifacts [37]. In Multiscale Principal Component Analysis (MSPCA), which showed remarkable results in experiments with healthy participants [81], [82], [83], the signal was decomposed initially with WT, and the wavelet coefficients were further decomposed using PCA. PCs were retained or rejected by applying the Kaiser rule. The Invert Wavelet Transform (IWT) was then employed to reconstruct the clean signal, which was submitted once more to PCA [83]. In other studies, the signal was first decomposed with a BSS method, the resulting components were further decomposed using Wavelet Transform, the noisy components were removed by thresholding, and the clean signal was reconstructed [12].

BSS techniques are also mixed with EMD. Specifically, EMD decomposes the signal into IMFs, while the artifactual components are removed using a BSS algorithm, or vice versa [12], [26]. This approach was adopted in a few-channel settings with promising results. A Multivariate Empirical Mode Decomposition (MEMD) was employed to decompose the signal into multivariate IMFs, which were further decomposed with CCA [61].

Blending adaptive filtering with BSS, WT, or EMD has also been suggested. In general, the signal is initially decomposed with WT, BSS, or EMD, and the noisy components are further processed and removed using adaptive filtering [12]. Moreover, a two-step BSS for EEG denoising has been documented. For instance, PCA was initially used to eliminate the dimension problem and assist in further denoising using ICA [26]. Finally, ICA mixed with a CNN or a Support Vector Machine (SVM) is a promising hybrid denoising technique [11], [24].

3) STANDARDIZED SEMI-AUTOMATED PREPROCESSING PIPELINES

A plethora of *standardized* semi-automated preprocessing pipelines are available. EEGLab's preprocessing chain stands out as a widely recognized solution, exhibiting superior performance in comparison to Fieldtrip, Brainstorm, MNE, and Harvard Automated Processing Pipeline for Electroencephalography (HAPPE) [10].

Makoto's preprocessing pipeline [6] is a constantly updated approach. It generally suggests the following steps: downsampling, high-pass filtering, rejection of bad channels and interpolation, rereferencing to the CAR, cleanline for line noise removal, epoch extraction, epoch rejection, ICA, and bilateral dipole estimation. Inspired by Makoto's preprocessing pipeline, two pipelines are Danielle Gruber's pipeline with ASR and one without ASR [7].

Fieldtrip [84] recommends the following preprocessing routine: trial definition, high-pass filter, low-pass filter, notch filter, bad trials, and channel rejection.

Another chain, Brainstorm, suggests the following steps: manual/ automatic detection/ rejection of bad trials and channels, definition of events, baseline correction, frequency filtering, resampling, averaging, and estimation of noise statistics [9].

Furthermore, the MNE preprocessing routine includes assessing data quality, identifying problematic channels, applying low- and high-pass filters, extracting data segments (epochs), employing ICA to correct artifacts, utilizing SSP for artifact correction, generating artifact-free epochs from ICA and SSP, and eliminating epochs based on peak-to-peak (PTP) amplitudes [8].

Autoreject is another semi-automated solution. It automatically detects noisy segments that the researcher decides to accept or discard. It provides complete transparency in the definition of parameters, and it has been suggested for heterogeneous EEG setups [64].

B. FULLY AUTOMATED PREPROCESSING

A growing number of standardized, fully automated preprocessing pipelines require zero user interaction. They were developed to alleviate the cumbersome task of selecting and combining various denoising methods and specifying their parameters [64]. The preprocessing algorithms' type, order, and parameters are adjusted automatically. The following automated preprocessing chains have been reported in the literature for healthy young adults.

Automatic Artifact removal(AAR) is a fully automated toolbox for artifact removal offered by EEGLab [85]. The EPOS is a holistic pipeline based on ICLabel. It is inspired by Delorme and Makeig, but it is fully

automated with machine-learning artifact detection [86]. Another ICLabel-based chain is the Reduction of Electroencephalographic Artifacts (RELAX). The alpha frequency band oscillations related to working memory have high performance [62]. The PREP pipeline proposes regression for line noise removal and re-reference of the signal to estimate the actual average reference, detection, and interpolation of bad channels [2]. It is based on robust statistics but leaves contaminated channels that require manual inspection afterward [5]. A novel pipeline based on PREP principles is Automagic [87]. HAPPE is a highly cited solution that can handle short recording lengths and data contaminated with significant noise. Initially, it offers a semi-automated setting to the user, but the final setting is fully automated. HAPPE achieves high performance but is not recommended for ERP preprocessing [3]. Over- or under-rejection of data [62] and poor performance owing to automatic re-referencing [10] were reported. For ERPs from large datasets, the Automatic preprocessing pipeline (APP) has been suggested [5]. Other automated preprocessing chains for healthy adults include Computational Testing for Automated Preprocessing (CTAP), Batch Electroencephalography Automated Processing Platform [86], and TAPEEG for the resting state [5].

Fully automated preprocessing protocols appear to outperform many semi-automated methods. Specifically, HAPPE was compared to four semi-automated preprocessing methods: ICA, manual segment rejection with ICA, ASR with ICA, and Semiautomatic Selection of ICA (SASICA). The HAPPE rejected 42% of the components and retained 67.8% of the data variance after rejection. Moreover, the retained components' mean and median artifact probability were 0.13 and 0.1, respectively. The ICA rejected far more components (86.3%) and retained significantly less data variance (24.8 %). Regarding data quality, ICA had almost double the mean (0.23) and median artifact probability (0.15). Manual Segment Rejection with ICA rejected significantly more data segments (73.6%) and retained less data variance (35.6%) than HAPPE did. The mean (0.26) and median artifact probability (0.27) doubled those of the HAPPE. The ASR with ICA rejected significantly more data segments (76.25%) and retained less data variance (43.9%). The mean (0.24) and median artifact probability (0.23) were once more double compared to HAPPE. On the contrary, SASICA did not differ significantly from HAPPE regarding the rejection of data segments(29.5%) and the maintenance of data variance (80.4%). However, Mean (0.33) and median artifact probability (0.27) were almost triple compared to HAPPE. Overall, ICA, manual segmentation with ICA, and ASR with ICA rejected more EEG data and maintained higher artifact levels, resulting in poor performance compared with HAPPE. HAPPE outperformed SASICA because it retained more EEG data but with higher levels of artifacts [3].

The fully-automated APP solution was compared to a semi-automated preprocessing approach consisting of DC correction, band-pass filtering [1, 40Hz], a 50 Hz Notch filter, visual detection of bad epochs and bad channels, interpolation of bad channels with 3D splines, and CAR. For ERP data captured from healthy and schizophrenia patients, the semi-automated pipeline rejected more trials (5.16±1.53%) than the APP (3.62±2 %). In addition, it interpolated a similar number of channels (1.00±1%) as APP (1.05±0.66%). For resting data, the semi-automated pipeline rejected fewer trials (4.99±2.73%) than APP (8.32±2.24%) and interpolated a similar number of channels (1.15±2.51%) as APP (1.2±1.87%). In general, it is claimed that the two preprocessing solutions perform similarly [5].

III. MOTIVATION AND METHODOLOGY

A. RESEARCH MOTIVATION AND CONTRIBUTION

Due to the significant advancements and reduced costs of EEG devices, there has been widespread interest in monitoring human cognitive states, particularly the Cognitive Load. However, challenges persist in preprocessing EEG signals from cognitive load experiments, with existing literature offering diverse suggestions lacking a consensus [65].

The widely used customized semi-automated preprocessing protocols are based on experience, can be time-consuming, and reduce statistical power. Recently, there has been a surge in interest in automated protocols [29]. Although they are known for their reliability, they are not tailored for Cognitive Load EEG Experiments. Only the RELAX method has shown high performance concerning Working Memory [62], a broader research area encompassing Cognitive Load.

Achieving high accuracy in determining cognitive load requires incorporating into the preprocessing protocols the interaction modalities specific to the experiment medium, such as Virtual Reality (VR), Extended Reality (XR), Augmented Reality (AR), or a standard computer interface. The different noise artifacts that are generated need to be handled accordingly.

Our objective is to identify relevant preprocessing attributes and provide recommendations to enable researchers to make more informed decisions when selecting a preprocessing approach, contributing to the standardization of preprocessing methods in the evaluation of Cognitive Load in EEG experiments.

B. RESEARCH QUESTIONS

We formulated the following research questions:

1) RESEARCH QUESTION (A)

Preprocessing varies significantly based on the specific task and research questions. How can a diverse array of Cognitive Load tasks be consolidated to offer valuable and applicable guidelines to researchers? Moreover, in what context does each task category find relevance and usefulness? Does a universally accepted semi-automated denoising approach exist when classifying tasks of low-noise and high-noise? Is there a widely agreed-upon automated method for denoising

when tasks are differentiated into low- and high-noise categories?

2) RESEARCH QUESTION (B)

How do researchers decide whether a fully automated or semi-automated preprocessing approach should be utilized? Which data organization approaches are usually employed from resampling, referencing, and segmentation? When do researchers prefer basic denoising, combining manual detection and rejection of noisy data, filtering, baseline correction, or averaging? How do they manually detect and reject artifactual data? When do they apply filtering, in which range and type of impulse response? When are more sophisticated algorithms, such as BSS and Deep Learning, employed for preprocessing EEG signals? Do they exist independently, or are they amalgamated, creating hybrid approaches? Which BSS algorithm is suggested for different types of artifacts? What are the types of deep learning that have particular interest? How do researchers parameterize deep learning algorithms to achieve the best results? Which customized semi-automated preprocessing pipeline is documented? Are standardized semi-automated preprocessing pipelines applied in Cognitive Load EEG experiments?

3) RESEARCH QUESTION (C)

To evaluate the effectiveness of the preprocessing methods, researchers tested their approaches on datasets with many subjects [88]. In the Cognitive Load domain, which are the semi-automated and automated denoising methods that have been proven effective in many subjects? Which are the semi-automated and automated denoising methods that have been proven effective in many subjects?

C. RESEARCH METHODOLOGY

We conducted a systematic search in the ACM and IEEE databases using specific keywords, namely [''EEG'' and ''Cognitive load''], spanning from January 2018 to November 2023. We restricted our focus to this six-year window to capture the most significant research developments and advancements in EEG signal processing.

In particular, we applied the PRISMA method [89]. Our search began in the ACM Digital Library, initially yielding 503 records within the specified time frame. We then removed articles based on their title and abstract and screened 140 records. Among these 117 records were deemed ineligible for our survey. Ineligible were the studies if they did not involve EEG experiments if the cognitive load was measured using alternative techniques (e.g., eye-tracking), if they explored different cognitive processes (e.g., attention), if the participants did not fit the category of healthy young adults, or if the primary task was not relevant to the field of Human-Computer Interaction (HCI). This left us with a final selection of 14 full-length articles. Additionally, we searched IEEE Xplore. This search retrieved 87 papers. After removing duplicates and applying exclusion criteria,43 papers were identified as eligible for inclusion. In this comprehensive survey, we compiled a total dataset of 57 papers covering January 2018 to November 2023.

IV. ANALYSIS OF RESULTS

A. RQ (A): CLASSIFICATION OF COGNITIVE LOAD EXPERIMENTS BASED ON THE LEVEL OF NOISE GENERATED FROM VARIOUS INPUT MODALITIES

The Cognitive Load research landscape is filled with a plethora of tasks, ranging from single- or dual-standardized tasks to workload and learning tasks, games, and presentation tasks. Each task requires a distinct preprocessing approach, which adds to the workload of researchers. We believe categorizing those tasks based on the potential level of noise they introduce is a valuable endeavor.

Single tasks, executed on a computer by just pressing a mouse, were classified under the *low-noise* category. Computerized tasks from cognitive psychology [90], [91], [92], [93], such as N-back tasks [94], [95], [96], [97] and Sternberg tasks [76], [98], [99], [100], [101], [102] were considered low-noise. Computer games [85], [103], [104] were included in the low-noise category. Data presentation tasks, visual [105], [106], [107], [108] or auditory [109], [110], [111], and multiple learning tasks [13], [18], [112], [113], including code comprehension tasks [114], [115], [116] and mathematical tasks [117], [118], [119], [120], were categorized as low-noise ones. Additionally, dual tasks with low noise were a Tetris game coupled with an oddball task [121], a Stroop task coupled with an oddball task [122], a SIMKAP task [123], and a threefold task performed on a computer and a customized AR system [14].

On the other hand, tasks that engaged more muscles than those typically employed to click and move the mouse were classified as high-noise. We further divide this category into two subcategories: *reduced-high-noise* and *elevated-high-noise*.

Tasks classified as *reduced-high-noise* require participants to activate multiple muscles beyond those required for simple mouse movements and clicking while standing or sitting. For instance, participants may need to speak, turn their heads, or continuously move their arms while performing such tasks. A learning task that involved creating concept maps and lists was categorized as reduced-high-noise because of the requirement of writing with a keyboard [124]. Similarly, code writing is considered as a reduced-high-noise task [38]. A simultaneous Interpretation task from Japanese to English was classified as reduced-high-noise because it required more muscle movements for speaking [125]. Sketching in Auto-CAD using a mouse and keyboard or speech and gestures introduced significant contamination to the EEG signal, and they were considered reduced-high-noise tasks [126]. Driving on a simulator [127], [128], [129], [130], [131], was also classified as a reduced-high-noise task. A Robotic-Assisted Surgery (RAS) task likely induced reduced-high noise due

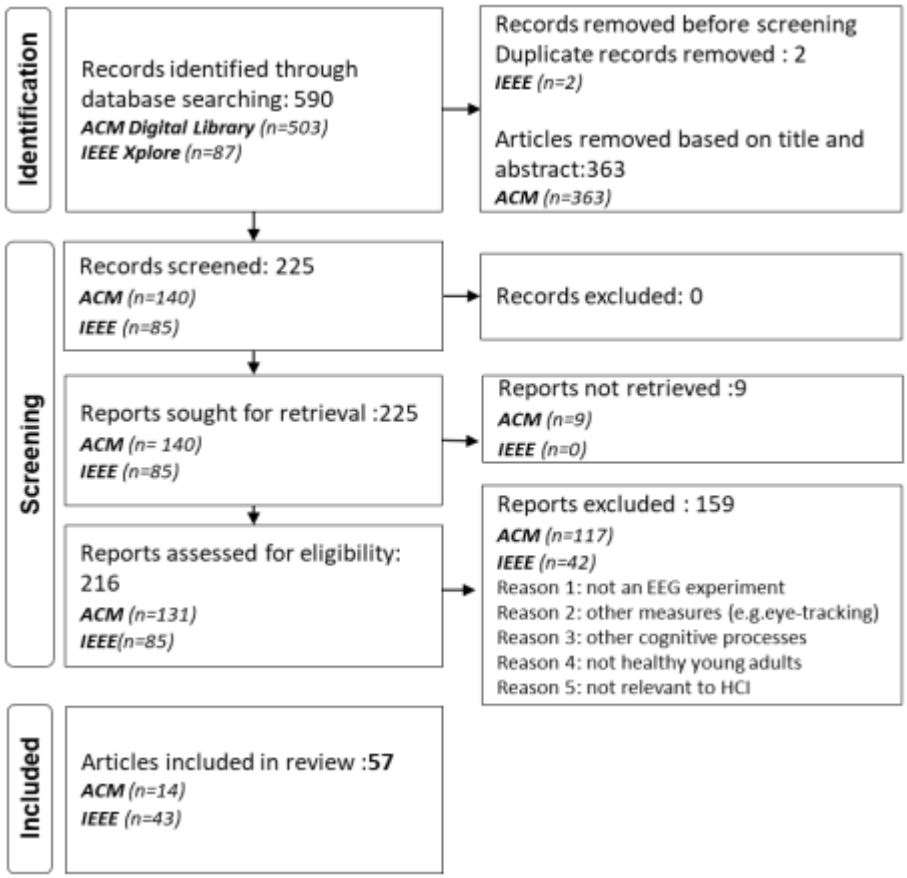

FIGURE 1. Using the PRISMA method, we conducted a systematic literature review that resulted in 57 final articles from an original count of 590.

to the robotic console's manipulations [132]. VR tasks were also considered reduced-high-noise tasks, given the accompanying head movements [133], [134], [135], [136], [137].

Elevated-high-noise tasks were characterized by the participants being in motion. An oddball task performed while sitting, standing, or walking on a treadmill was considered elevated-high-noise owing to motion. The treadmill and microcontroller likely caused additional noise [138]. Another task, playing a puzzle while walking, was also included in this category [139].

Apart from three *low-noise* tasks that were preprocessed automatically [18], [85], [140], the majority employed semi-automated pipelines customized by the researcher. Particularly, low-noise tasks were preprocessed with simple denoising methods in the time domain (baseline correction, downsampling, re-referencing of CAR, and averaging), followed by a band-pass Finite Impulse Response (FIR) filter between 0.1-1 Hz and 30-45 Hz. Next, among BSS techniques, methods guided by ICA prevailed, and PCA was used to discard ocular artifacts. Afterward, one-second, non-overlapping epochs around the stimulus presentation were extracted. Epoch rejection with the amplitude range as a threshold was common, compared to rejection of channels and all participants' data.

Researchers have built their semi-automated pipelines to preprocess signals from *reduced-high-noise* tasks. After downsampling, FIR or IIR-Butterworth filters with cut-off frequencies in the range of 0.5-1 Hz and 45-100 Hz were documented. Next, ICA was exclusively adopted. Finally, epoching and manual rejection of noisy epochs, such as large amplitudes, missing signals, and streams with a fixed constant value completed the EEG signal preprocessing.

Given that only two *elevated-high-noise* tasks were documented, no safe preprocessing suggestions can be made. Yet, it could be stated that Infinite Impulse Response (IIR) Butterworth band-pass filters, the BSS techniques ICA or ASR, epoching, and manual epoch rejection were present in both studies.

B. RQ (B): STATE OF THE ART PREPROCESSING FOR COGNITIVE LOAD ASSESSMENT

Studies have overwhelmingly favored using semi-automated protocols despite the remarkable diversity of the available fully automated preprocessing pipelines. Notably, only two automated solutions, AAR by EEGLab [85] and BESA software [18], [140], were mentioned. However, it is important to highlight that some highly effective and influential fully automated solutions within the research community, such as EPOS, RELAX, HAPPE, and PREP, are notably absent from

TABLE 1. Semi-automated basic preprocessing of low noise (L), reduced-high noise (R), and elevated-high noise (E) tasks in descending order by the number of subjects.

Study	Year	Subj.	Noise Level	Resampling	Re-referencing	Baseline Correction	Averaging
Spender et al. [111]	2021	75	L	sampling:1000Hz			
Kwak et al. [98]	2020	65	L	downsampling:100Hz			
Zhu et al. [123]	2021	50	L		CAR		
Pollmann et al. [93]	2019	30	L		avg mastoids		
Jung et al. [120]	2023	29	L		CAR		
Kuanar et al. [101]	2018	25	L				
Volmer et al. [14]	2022	23	L				
Yoghourdjian et al. [106]	2021	22	L		CAR		
Hijazi et al. [116]	2023	21	L		CAR		
Parekh et al. [94]	2018	20	L				
Bilalpur et al. [110]	2018	20	L			0.5s during fixation	
Kosch et al. [105]	2020	18	L				
Meng et al. [107]	2022	16	L	downsampling	yes	to remove the influence of data drifts	
Beiramvand et al. [97]	2023	15	L				
Tamura et al. [112]	2019	15	L			mean amplitude, 1s before stimulus	avg epochs per page and participant
Jimenez-Guarneros and Gomez-Gil [100]	2020	15	L	downsampling:250Hz	CAR	remove avg amplitude of each channel	
Zhang et al. [99]	2018	15	L	downsampling:250Hz	CAR	remove avg amplitude of each channel	
Havugimana et al. [76]	2021	15	L				ERP: eigenspace-based bootstrap sampling
Havugimana et al. [76]	2023	15	L				ERP: eigenspace-based bootstrap sampling
Kujur et al. [119]	2022	13	L				avg per segment
Ortiz et al. [121]	2020	11	L			200ms before stimulus	ERP per auditory tone
Diwakar et al. [117]	2020	10	L	downsampling:1000Hz			
Kwon et al. [96]	2018	10	L		CAR	avg amplitude of each channel	ERP per participant
Song et al. [109]	2023	10	L	downsampling:250Hz(3 channels); 40Hz(28 channels)			
Vargas et al. [95]	2021	5	L	upsampling: 256Hz			
Swerdloff and Hargrove [122]	2021	4	L				grand avg ERP:per epoch type and participant
Alyan et al. [128]	2023	32	R	downsampling:256Hz	CAR		
Medeiros et al. [38]	2019	30	R				
Pulver et al. [129]	2023	23	R				
Yagura et al. [125]	2020	22	R	sampling: 500 Hz			
Ahmadi et al. [135]	2023	19	R	downsampling:256Hz	avg mastoids		
Gupta et al. [134]	2019	13	R				avg alpha (8-13 Hz)
Baig and Kavakli [126]	2018	12	R			yes	
Ahmed et al. [127]	2022	2	R	downsampling: 250 Hz			
Swerdloff and Hargrove [138]	2020	10	E			yes	grand avg ERP:avg 37 trials per task
Namakura et al. [139]	2021	3	E	downsampling:256Hz	CAR		

the literature. Consequently, we delved into the analysis of semi-automated approaches.

1) HOW WAS STANDARD PREPROCESSING EXECUTED IN SEMI-AUTOMATED APPROACHES BY RESEARCHERS?

Exclusive standard preprocessing was frequently documented [13], [85], [90], [95], [98], [108], [109], [113], [118], [119], [120], [125], [129], [132], [134], [136]. Despite the advances in denoising techniques, standard methods still have their place in most preprocessing pipelines.

Specifically, *resampling* was reported in thirteen studies (Table 1). Downsampling at about 250 Hz has been frequently documented [99], [100], [113], [128], [135], [139]. Song et al. downsampled data from 3 channels to 250 Hz and data from 28 channels to 40 Hz [109]. Upsampling to 256 Hz after a 250 Hz recording sampling frequency was employed by Vargas et al. [95].

Re-reference to the average mastoids was not common [93], [135] compared to Common Average Reference (CAR) [96], [99], [100], [106], [107], [116], [120], [123], [128], [139]. If the contact of a channel becomes poor at any single point during recording, the re-reference may generate further noise [23]. To deal with this problem, Namakura et al. detected and removed bad channels before re-reference to CAR [139].

Averaging event trials [96], [112], [119], [121], [122], [138] or frequency bands [134] were performed to eliminate noise and obtain grand average ERPs. Havugimana et al. selected a shallow CNN architecture for predicting Cognitive Load and denoising. The available data were not sufficient enough to train and test the model. Therefore, eigenspace-based bootstrap sampling was employed, and ERPs were calculated by averaging a participant's single-trial EEG signals, resulting in 44000 topo-maps fed into a shallow CNN architecture [76], [102].

Regarding *baseline correction*, a short time interval during fixation [110], the mean amplitude, a short time window before stimulus presentation [112], [121], or after stimulus presentation [109] were used as a baseline to remove the influence of data drifts [107]. Moreover, the average amplitude of each channel [99], [100], [141] before stimulus presentation [96] has also been suggested. *Detrending* was observed in three studies [93], [104], [117]. A few papers reported *interpolation* [38], [116], [139]. Symmetrical electrodes within a 5 cm radius were reported as an interpolation scheme for channel reconstruction [106].

Segmentation and the manual rejection of epochs, channels, and whole participant data are summarized in Tables 2 and 3.

Segmentation was widely used but differed significantly across contexts. Mostly, segments were constructed around the task [93] or the task level [97], [135]. Another segmentation was based on brain region dissociation, [38] or on the peak of eye blinks [128]. Epoching in particular was documented as an early step in preprocessing [14], [107], [109], [110], [117], [129], yet most studies applied it as a later step after basic filtering [101], [122], [138] or after filtering and BSS algorithms [93], [94], [95], [96], [99], [100], [106], [112], [121], [125], [126], [127], [135], [139], [141]. Overlapping was rare, with an overlapping percentage of 80% [38] or 50% [105]. Regarding the epoch duration, short epochs of 1-second were preferred for rapid changes in EEG data [96], [105], [112], [121], [122], [124], [138] with a time interval approximately 200ms before stimulus presentation until 800ms after stimulus presentation [14], [96], [109], [121], [122], [138]. Epochs around 2 or 3 seconds were the case in quite a few studies [93], [94], [95], [110], [125], [126], [128], [135], and [139]. Epochs of 5 seconds duration [38], [106], [119], 9-second epochs [99], [100], and 10-second epochs were also reported [97], [117], [120], [129].

The specification of a threshold was a trend during *manual epoch rejection*. Several researchers have employed an amplitude threshold [93], [96], [98], [103], [107], [112], [128], [135]. Epochs with amplitudes exceeding the range of $[-70\mu V, 70\mu V]$ [98], $[-75\mu V, 75\mu V]$ [96], or $[-100\mu V, 100\mu V]$ [112] were discarded. Epochs with amplitudes outside the range $[-200\mu V, 200\mu V]$ in any of the frontal channels Fp1 and Fp2 were rejected due to their correlation with heavy eye movement artifacts [93]. Another common threshold type was the standard deviation (SD) of the mean channel value. For a Tetris computer game, if any channel during an epoch was out of the range of $[-4, 4]$ SD (standard deviation) from the mean channel value, this epoch was discarded [121]. To play Puzzle 2048 on an iPod while walking, Namakura et al. rejected epochs from the $[-5, 5]$ SD range from the mean channel value [139]. The EEGLab function pop_autorej was applied to eliminate eye blinks by rejecting epochs with amplitudes outside the range $[-500\mu V, 500\mu V]$ and SD out of the range of $[-5, 5]$ [128]. Other manual epoch rejections have been reported, such as short windows with high peaks [134], trials with large interruptions such as missing signals, large continuous spikes, and streams of a fixed constant value [136], eyeblinks detected using the EyeTribe [13], and occasional noise spikes, namely sudden, short-duration fluctuations in the voltage [132].

For the rejection of the entire channel, the threshold was the correlation of a particular channel with its neighboring channels. Namakura et al. removed whole channels according to a correlation threshold of 0.08 between one channel and its neighboring channels. Specifically, they rejected channels with less than 0.08 correlation [139]. Corrupted EEG recordings [94], bad channels, and flat data [107], [116], [134] were also manually discarded.

Rejection of all participant data was documented when the participant could not cope with the task [98], [106]. Specifically, data from one participant with fewer than 30 trials left per task after denoising with baseline correction, band-pass filtering, and ICA were excluded [138]. In a code review task, the entire data set from one participant was

TABLE 2. Segmentation and manual denoising of low-noise(L) tasks in descending order by the number of subjects.

Study	Year	Subj.	Noise level	Segmentation	Epoch rejection	Channel rejection	Participant rejection
Mercier et al. [103]	2020	82	L		large artifacts		
Kwak et al. [98]	2020	65	L		amplitudes exceeding $[-70, 70]\mu V$, incorrect trials		3 due to their difficulty to adapt to high cognitive load
Peitek et al. [116]	2022	39	L				1 participant: many slow/high-response-time data points
Pollmann et al. [93]	2019	30	L	task data:6 min per participant,resting data:3 min per participant,2s non-overlapping epochs	amplitudes exceeding $[-200, 200]\mu V$ in Fp1 or Fp2 (heavy eye-movement artifacts)		
Jung et al. [120]	2023	29	L	10-sec epochs			
Sinha et al. [13]	2019	26	L		corresponding to eyeblinks detected by Eyetribe		
Kuanar et al. [101]	2018	25	L	0.5s epochs			3 participants with noisy signal
Volmer et al. [14]	2022	23	L	$[-200, 700]ms$ in 3 time windows per epoch			
Yoghourdjian et al. [106]	2021	22	L	5s epochs prior the participant signaled they had an answer		bad channels	1 dropped out during online recording, 1 with reference problems
Hijazi et al. [116]	2023	21	L			flat and noisy channels	
Parekh et al. [94]	2018	20	L	2s epochs preceding users response from each trial		corrupted recordings	
Bilalpur et al. [110]	2018	20	L	2.5s epochs from each trial(0.5s:fixation+ 2s:stimulus)	noisy epochs		
Kosch et al. [105]	2020	18	L	1s epochs with 0.5s overlap	4 first seconds, 4 last seconds		
Meng et al. [107]	2022	16	L	non-overlapping segments before and after stimulation	epochs with excessive amplitude	yes	
Beiramvand et al. [97]	2023	15	L	3 segments for each level (30s); 3 subsegments (10s each)			
Tamura et al. [112]	2019	15	L	1s epochs	amplitudes exceeding $[-100, 100]\mu V$		
Jimenez-Guarneros and Gomez-Gil [100]	2020	15	L	9s epochs: from 2000 ms before presentation of SET to 3500 ms after presentation of TEST			
Zhang et al. [99]	2018	15	L	9s epochs: from 2000 ms before presentation of SET to 3500 ms after presentation of TEST			
Kujur et al. [119]	2022	13	L	5s segments	avg per segment		
Ortiz et al. [121]	2020	11	L	1s epochs: $[-200, 800]ms$ around the stimulus	with channel outliers 4 SD from the mean channel value		
Diwakar et al. [117]	2020	10	L	10s epochs			
Kwon et al. [96]	2018	10	L	1s epochs: $[-200, 800]ms$ around the stimulus	amplitudes exceeding $[-75, 75]\mu V$		
Song et al. [109]	2023	10	L	700ms epochs:[0,700]ms			
Shakti et al. [104]	2019	7	L		bad blocks after filtering		
Vargas et al. [95]	2021	5	L	2s non-overlapping epochs			
Swerdloff and Hargrove [122]	2021	4	L	1s:$[-200, 800]ms$ around auditory tone			

TABLE 3. Segmentation and manual denoising of reduced-high noise (R) and elevated-high noise (E) tasks in descending order by the number of subjects.

Study	Year	Subj.	Noise level	Segmentation	Epoch rejection	Channel rejection	Participant rejection
Alyan et al. [128]	2023	32	R	epochs: [-500- 1200]ms relative to the peak of the eye blink	epoch rejection with amplitudes exceeding $[-500, 500]\mu V$ and 5 SD ("pop autorej" EEGLab function)	identification of bad channels ("clean artifacts" EEGLab function)	
Medeiros et al. [38]	2019	30	R	grouping 60 channels into 13 brain regions, 5s epochs with 80% overlap			
Pulver et al. [129]	2023	23	R	10s epochs			
Yagura et al. [125]	2020	22	R	3s epochs, shifted by 1s.			
Gupta et al. [136]	2020	24	R		trials with large interruptions in the signal (missing signals, large continuous spikes, streams of a fixed constant value): 5,56% rejection of all recorded data		
Ahmadi et al. [135]	2023	19	R	bins per difficulty level;2-3 s epochs		voltage threshold, rejection rate=15 (saccades, muscle movements)%	
Gupta et al. [134]	2019	13	R		noise spikes,last 5 game levels per condition	flat data	
Zhou et al. [132]	2020	12	R		occasional noise spikes		
Baig and Kavakli [126]	2018	12	R	back-to-back epoching for keyboard and gesture tasks with a 0.5s difference between epochs	bad blocks after filtering+ICA		
Ahmed et al. [127]	2022	2	R	epoching			
Swerdloff and Hargrove [138]	2020	30	E	1s epochs:$[-200, 800]ms$ around the stimulus			1 participant with < 30 trials left per task after filtering+ICA
Namakura et al. [139]	2021	3	E	2.7 sec epochs: starting 0.5 sec before the heel contact	with channel outliers 5 SD from the mean channel value	channel correlation with neighboring channels < 0.8	

rejected because it contained many data points with very slow or very high response times, which was translated into the absence of code comprehension [115].

Below, we attempt to clarify the optimal filters and their specifications for Cognitive Load applications (see Tables 4 and 5).

The Median Filter was employed once by Gupta et al. in a VR shape selection task to further denoise the signal after manual denoising and averaging [134]. In their next study, a Wiener filter was added to the semi-automated pipeline to remove muscle artifacts caused by head turnings during the VR experience [136]. The 3rd order parameter in a Savinsky-Golay Filter was characterized as being optimal for denoising [113].

Low-pass filters [38], [104], [116], [126], [128], [133], high-pass filters [13], [38], [104], [112], [116], [123], [126], [127], [128], [132], [142], and band-pass filters have been constantly used [13], [14], [90], [93], [94], [96], [97], [98], [99], [100], [101], [106], [110], [112], [115], [119], [120], [121], [122], [125], [129], [130], [136], [137], [138], [139], [141]. High-pass cut-off frequencies were clustered in the range [0.1,1] Hz [13], [14], [38], [90], [93], [94], [96], [97], [98], [99], [100], [106], [107], [110], [116], [119], [120], [121], [122], [125], [128], [129], [130], [135], [136], [137], [138], [139]. Low-pass cut-off frequencies ranged mainly between 30Hz and 50Hz [13], [14], [93], [94], [96], [97], [98], [99], [106], [110], [112], [119], [120], [121], [122], [125], [126], [130], [136], [138], [141]. Yet, several studies preferred higher values, between 70 Hz and 100 Hz [38], [90], [107], [110], [116], [129], [137].

FIR filtering was documented in many experimental situations [14], [93], [95], [99], [100], [112], [115], [116], [125], [126], [128], [141]. IIR filtering [132], particularly a 2nd-6th order Butterworth type, was also cited [38], [97], [109], [120], [129], [130], [138], [139].

TABLE 4. Filtering of low-noise (L) tasks in descending order by the number of subjects.

Study	Year	Subj.	Noise level	Filter type	Low	High	FIR/IIR	Goal
Kwak et al. [98]	2020	65	L	band-pass	0.5	40		
Zhu et al. [123]	2021	50	L	high-pass	1			
Zhu et al. [123]	2021	50	L					line noise
Peitek et al. [115]	2022	39	L				FIR	
Peitek et al. [115]	2022	39	L	Notch	49	51		
Peitek et al. [115]	2022	39	L	band-pass	4	200		
Pollmann et al. [93]	2019	30	L	band-pass	0.5	45	FIR	
Jung et al. [120]	2023	29	L	band-pass	1	40	IIR:5th order Butterworth	
Sinha et al. [13]	2019	26	L	band-pass	0.5	30		DC-offset, muscle artifacts,line noise
Sinha et al. [13]	2019	26	L	high-pass	7			remove eyeblinks artifacts
Volmer et al. [14]	2022	23	L	band-pass	0.1	30		
Yoghourdjian et al. [106]	2021	22	L	band-pass	1	30	FIR	reduction of slow wave potentials
Hijazi et al. [116]	2023	21	L	high-pass	1		FIR	DC component, slow-wave drifts
Hijazi et al. [116]	2023	21	L	low-pass		90	FIR	
Hijazi et al. [116]	2023	21	L	Notch				50 Hz
Parekh et al. [94]	2018	20	L	band-pass	0.1	45		
Bilalpur et al. [110]	2018	20	L	band-pass	0.1	45		muscle artifacts at 40-100Hz
Meng et al. [107]	2022	16	L	band-pass	0.1	100		
Beiramvand et al. [97]	2023	15	L	band-pass	0.5	40	IIR: 0-phase Butterworth	
Tamura et al. [112]	2019	15	L	high-pass	1			
Tamura et al. [112]	2019	15	L	band-pass	3	50	FIR	
Jimenez-Guarneros and Gomez-Gil [100]	2020	15	L	band-pass	1	45	FIR	
Zhang et al. [99]	2018	15	L	band-pass	1	45	FIR	
Kujur et al. [97]	2022	13	L	band-pass	0.5	40		
Ortiz et al. [121]	2020	11	L	band-pass	0.1	30	FIR	
Ahmed et al. [113]	2020	10	L	Savitzsky-Golay				
Diwakar et al. [117]	2020	10	L	Savitzsky-Golay				for smoothing
Diwakar et al. [117]	2020	10	L	Notch				50 Hz
Swerdloff and Hargrove [122]	2020	10	L	band-pass	1	30	IIR: 2nd order Butterworth	
Liang et al. [108]	2018	10	L	filter (no details)				
Kwon et al. [96]	2018	10	L	band-pass	0.5	30		
Song et al. [109]	2023	10	L	band-pass	0.5	10	IIR: 3rd order Butterworth	
Fard et al. [114]	2021	8	L	filter(no details)				
Shakti et al. [104]	2019	7	L	low-pass				
Shakti et al. [104]	2019	7	L	high-pass				
Vargas et al. [95]	2021	5	L				FIR	anti-aliasing filter for upsampling
Vargas et al. [95]	2021	5	L	high-pass	0.5			for DC offset removal
Vargas et al. [95]	2021	5	L	Notch				50 Hz
Grubov et al. [90]	2020	-	L	band-pass	0.016	70		
Grubov et al. [90]	2020	-	L	Notch				50 Hz

Power line interference was frequently removed from the signal [123] using Notch filter [126], [133] surrounding the frequencies of 50 Hz [38], [90], [95], [115], [116], [117], [135] or 60 Hz [129]. The Cleanline procedure at 50 Hz and 100 Hz was reported once [137].

2) HOW WAS THE UTILIZATION OF ADVANCED DENOISING APPROACHES EXECUTED?

Despite the variety of sophisticated denoising algorithms, only a few of those approaches were used in Cognitive Load Experiments, namely BSS and a few deep learning techniques. Discrete Wavelet Transform (DWT) was employed only by Beiramvand et al. employed a Discrete Wavelet Transform (DWT), who defined db4 as the mother wavelet due to its resemblance with blinking [97]. Other advanced algorithms such as EMD, SSP, and Beamforming that have proved their efficiency were completely absent from the Cognitive Load domain.

a: BLIND SOURCE SEPARATION METHODS (BSS)

ICA (see Table 6) was a popular BSS denoising approach. Infomax(rubica) [93], [99], [100], [112], [116], [138], [141],

TABLE 5. Filtering of reduced-high-noise (R) and elevated high-noise (E) tasks in descending order by the number of subjects.

Study	Year	Subj.	Noise level	Filter type	Low	High	FIR/IIR	Goal
Aygun et al. [130]	2022	80	R	band-pass	0.1	32	IIR: 6th order Butterworth	
Baceviciute et al. [137]	2020	78	R	band-pass	0.5	100		
Baceviciute et al. [137]	2020	78	R	CleanLine				50 Hz, 100 Hz
Alyan et al. [128]	2023	32	R	high-pass	0.1		FIR	environmental, muscular artifacts
Alyan et al. [128]	2023	32	R	low-pass	16		FIR	environmental, muscular artifacts
Alyan et al. [128]	2023	32	R	high-pass	1.5		FIR	
Pulver et al. [128]	2023	23	R	pass-band	0.4	75	IIR: 2nd order Butterworth	
Medeiros et al. [38]	2019	30	R	high-pass	0.5		IIR: 4th order Butterworth	trend removal
Medeiros et al. [38]	2019	30	R	low-pass		90	IIR: 7th order Butterworth	
Medeiros et al. [38]	2019	30	R	Notch				50 Hz
Gupta et al. [136]	2020	24	R	band-pass	1	40		
Gupta et al. [136]	2020	24	R	Wiener				for motion artifacts (VR head turnings)
Yagura et al. [125]	2020	22	R	band-pass	1	50	FIR	
Ahmadi et al. [135]	2023	19	R	high-pass	1			
Ahmadi et al. [135]	2023	19	R	low-pass	50			
Gupta et al. [134]	2019	13	R	Median				
Zhou et al. [132]	2020	12	R	high-pass			IIR	
Zhou et al. [132]	2020	12	R	2s Hanning				fast Fourier transform wrap-around step artifacts
Baig and Kavakli [126]	2018	12	R	low-pass		45	FIR	
Baig and Kavakli [126]	2018	12	R	high-pass	0.1			
Ahmed et al. [127]	2022	2	R	high-pass				
Gerry et al. [133]	2018	2	R	low-pass				
Gerry et al. [133]	2018	2	R	Notch				
Swerdloff and Hargrove [138]	2020	10	E	band-pass	0.5	30	IIR (Butterworth)	baseline drifts, walking noise and electric power frequency
Namakura et al. [139]	2021	3	E	band-pass	1	200	IIR (Butterworth)	

FastICA [38], and AMICA were used to extract ICs. The detection and rejection of noisy ICs has been performed manually [93], [94], [104], [110], [112], [126], [130] based on the ICs' scalp topography [93], [99], [100], [116], [141], time course [93], [116], spectral properties [93], [99], [100], [116], [122], [138], [141], location of source foci [99], [100], [141], dipole estimations [99], [100], [139], [141], and other statistical criteria such as range, linearity, probability, and kurtosis [122], [138]. In one study, artifact detection and rejection/correction were completely automated by employing the MARA approach [137]. Noisy ICs were also rejected automatically [115].ICLabel and a modified version of the BLINKER algorithm were employed [128].

ICA was selected mostly to identify ocular artifacts (eye blinks and eye movements) [14], [93], [94], [103], [110], [112], [122], [126], [130], [135], [136], [138], muscle artifacts [93], [94], [107], [107], [110], [116], [116], [126], [128], [128], [138], cardiac artifacts [93], [128], and artifacts caused by sweat and movements [138]. Ocular artifacts were associated with concentrated signal distribution in the occipital regions. In contrast, muscle movements were associated with major signal distribution in specific brain areas. Finally, movement artifacts were related to increased power distribution in high-frequency regions [104].

Apart from ICA, three more BSS techniques were documented, namely PCA, ASR, and SSD. Specifically, despite that PCA is nowadays rarely used directly in denoising [23], some EEG experiments on Cognitive Load reported its use to remove ocular artifacts [96], [106], [121], such as saccades and blinks [99], [100], [141]. Ocular artifacts were associated with the first PC [96]. In two studies, PCA was initially applied to the continuous signal to correct ocular artifacts, and ICA was then applied to the epoched data for further denoising [99], [100]. To remove non-stationary or large amplitude noise, ASR was employed [123]. Additionally, ASR was applied to the continuous data to eliminate non-stationary and prominent artifacts, while ICA was used in a subsequent preprocessing stage to denoise the epoched data [139]. A notable approach was the SSD algorithm. It was applied in the range of $[8Hz, 12Hz]$ for the extraction of alpha oscillations and in the range of $[3Hz, 7Hz]$ for the extraction of theta oscillations. Thus, the noise outside the alpha and theta band powers was decreased, and the SNR was optimized [105].

TABLE 6. Blind Source Separation (BSS) of low-noise (L),reduced-high-noise (R), and elevated-high-noise (E) tasks in descending order by the number of subjects.

Study	Year	Subj.	Noise Level	BSS	Components removal	Target
Elkerdawy et al. [92]	2020	127	L	ICA		
Mercier et al. [103]	2020	82	L	ICA		eye blinks, lateral eye movements
Zhu et al. [123]	2021	50	L	ASR		large amplitude artifacts
Peitek et al. [115]	2022	39	L	ICA	automatic rejection	eye, muscle movements
Pollmann et al. [93]	2019	30	L	ICA (infomax-runica)	manual based on scalp topography, time course and power spectral intensity	ocular, muscular, cardiac artifacts
Volmer et al. [14]	2022	23	L	ICA		occular
Yoghourdjian et al. [106]	2021	22	L	PCA		blinks and eye movements
Hijazi et al. [116]	2023	21	L	ICA(Infomax)	manual based on topographic maps, time course, and activity power spectrum	ocular, myogenic artifacts
Parekh et al. [94]	2018	20	L	ICA	manual	muscle/head/eye movements
Bilalpur et al. [110]	2018	20	L	ICA	manual	eye-blinks, eye-movements, and few muscle artifacts
Kosch et al. [105]	2020	18	L	SSD		
Meng et al. [107]	2022	16	L	ICA		blinking, muscle tension
Tamura et al. [112]	2019	15	L	ICA (infomax-runica)	manual	eye-movements
Jimenez-Guarneros and Gomez-Gil [100]	2020	15	L	PCA		saccades and blink artifacts
Jimenez-Guarneros and Gomez-Gil [100]	2020	15	L	ICA (infomax)	manual based on scalp topography and spectral density, source foci located inside the head boundary and cerebral cortex, source dipoles are bilateral	
Zhang et al. [99]	2018	15	L	PCA		saccades and blink artifacts
Zhang et al. [99]	2018	15	L	ICA (infomax)	manual based on scalp topography and spectral density, source foci located inside the head boundary and cerebral cortex, source dipoles are bilateral	
Ortiz et al. [121]	2020	11	L	PCA	manual	occular artifacts
Kwon et al. [96]	2018	10	L	PCA		occular artifacts
Fard et al. [114]	2021	8	L	ICA		
Shakti et al. [104]	2019	7	L	ICA	manual	muscle,movement,eye artifacts
Swerdloff and Hargrove [122]	2021	4	L	ICA	manual based on statistical criteria (i.e., range, linearity, probability, kurtosis, and spectral properties)	
Aygun et al. [130]	2022	80	R	ICA	manual rejection	blink artifacts
Baceviciute et al. [137]	2020	78	R	MARA (ICA-guided)		
Alyan et al. [128]	2023	32	R	ICA (AMICA)	automatic: ICLabel	muscle, eye, heart artifacts
Alyan et al. [128]	2023	32	R	ICA (AMICA)	automatic: modified BLINKER algorithm	eye blinks
Medeiros et al. [38]	2019	30	R	ICA (FastICA)		
Gupta et al. [136]	2020	24	R	ICA		
Ahmadi et al. [135]	2023	19	R	ICA		blinking
Baig and Kavakli [126]	2018	12	R	ICA	manual	eye blinking and muscles artifacts
Ahmed et al. [127]	2022	2	R	ICA		
Swerdloff and Hargrove [138]	2020	10	E	ICA (infomax)	manual based on statistical thresholds (i.e., range, linearity, probability, kurtosis, and spectral properties)	eye blinks, eye movements, sweat, muscle contraction, movement
Namakura et al. [139]	2021	3	E	ASR		non-stationary and large noises
Namakura et al. [139]	2021	3	E	ICA		

b: DEEP LEARNING

Kuanar et al., after band-passing the signal and removing data from three participants, combined a CNN with three different types of RNN to predict Cognitive Load (see Table 7). 2D images were fed to a CNN with the following specifications: four convolutional layers with 3×3 kernel size and 32 filters, a max-pooling layer with 2×2 kernel size, four convolutional layers with 3×3 kernel size and 64 filters, a max-pooling

TABLE 7. Deep learning architectures, robust to signal noise and distortion, for Cognitive Load classification. Conv: Convolutional Layer; pool: Pooling layer; FC: Fully Connected Layer; BiLSTM:Bidirectional.

Study	Year	Subj.	Noise level	Input	CNN	RNN	Activation function	Accuracy
Kuanar et al. [101]	2018	25	L	spectral topomaps-3 channel EEG images	4 Conv-32 (3x3)- max pool (2x2)-4 Conv-64(3x3)- max pool.(2x2)-Conv-128 (3x3)- max pool.(2x2)-FC-512	LSTM	ReLU(conv) Softmax (FC)	84.48%
Kuanar et al. [101]	2018	25	L	spectral topomaps-3 channel EEG images	4 Conv-32 (3x3)- max pool (2x2)-4 Conv-64(3x3)- max pool.(2x2)-Conv-128 (3x3)- max pool.(2x2)-FC-512	LSTM 1D Conv.	ReLU(conv) Softmax (FC)	87.68%
Kuanar et al. [101]	2018	25	L	spectral topomaps-3 channel EEG images	4 Conv-32 (3x3)- max pool (2x2)-4 Conv-64(3x3)- max pool.(2x2)-Conv-128 (3x3)- max pool.(2x2)-FC-512	BiLSTM	ReLU(conv) Softmax (FC)	92.5%

layer with 2×2 kernel size, one convolutional layer with 3×3 kernel size and 128 filters, a max-pooling layer with 2×2 kernel size and a fully-connected layer. ReLU accompanied all the hidden layers. The output of the CNN was fed into three different RNN architecture models: LSTM, LSTM with 1D convolution, and bidirectional LSTM. Their output was carried into a fully connected layer, and a 4-node Softmax layer made the final Cognitive Load prediction. The classification accuracies of the three models were 84.48%, 87.68%, and 92.5%, respectively; however, this is not an appropriate metric for denoising performance [101].

3) STANDARDIZED SEMI-AUTOMATED PREPROCESSING PIPELINES

EEGLab and Makoto's preprocessing pipeline was employed by Lim et al. [142], whose dataset was used by Zhu et al. [123]. The EEGLab preprocessing pipeline was also used in a Simultaneous Interpretation task [125]. MNE's autoreject package, coupled with a custom preprocessing script, was used by Volmer et al. to remove noisy artifacts and reconstruct them with interpolation [14].

Regarding software toolboxes, the EEGLab package was preferred [90], [93], [99], [100], [104], [107], [108], [112], [114], [115], [117], [121], [122], [123], [125], [126], [128], [133], [135], [137], [138]. Additionally,ERPLab [122], [138],a customized Graphical User Interface (GUI) in MATLAB [113], MNE [14], [76], [102], BESA software [18], [140], ABM B-Alert Live software [124], and Fieldtrip were reported [106].

C. RQ (C): ON THE EFFICACY OF DENOISING APPROACHES BASED ON THE NUMBER OF SUBJECTS

It is important to highlight that the efficiency of preprocessing protocols is closely linked to the number of subjects participating in the EEG experiment. A novel classification framework that denoised signals using MSPCA and combined Variational Mode decomposition (VMD) with Linear Regression (LR) and Cascade feed-forward Neural Network (CFNN) showed the highest classification accuracy for five subjects. Its effectiveness was confirmed using a large database of 52 subjects [88]. Another robust framework that combines MSPCA denoising methods along with Empirical Fourier Decomposition (EFD) and a Feed-Forward Neural Network (FFNN) proved its efficiency in four datasets consisting of 5, 1,1, and 3 subjects [143].

Five low-noise tasks exceeded the threshold of 50 participants. Specifically, ICA proved effective in denoising EEG signal captured during a cognitive assessment task performed by 127 university students [92]. A serious computer game played by 82 university students proved that manual epoch rejection of large artifacts and ICA are efficient denoising techniques [103]. A sonification low-noise task with 75 participants validated resampling to 1000 Hz [111]. A Sternberg task with 65 participants validated downsampling to 100 Hz, manual rejection of incorrect trials, difficulty managing high cognitive load and amplitudes out of the range of $[-70\mu V, 70\mu V]$, and band-pass filtering with a high-pass cut-off frequency of 0.5Hz and a low-pass cut-off frequency of 40 Hz [98]. Another low-noise task of 50 participants proved the significance of CAR re-referencing, high-pass filtering of 1 Hz, line noise removal, and ASR algorithm [123].

A reduced-high-noise task was performed with 78 participants, proving the efficiency of a band-pass filter with 0.5 Hz high-pass frequency and 100 Hz low-pass frequency of CleanLine to reduce power line noise of 50 and 100 Hz and MARA, an ICA-guided denoising approach [137]. Another reduced-high-noise task on a driving simulator proved the effectiveness of IIR band-pass filtering in the range of [0.1-32] Hz, and ICA with the manual rejection of noisy ICs related to eye blinks [130].

Elevated-high-noise tasks were performed with fewer than 50 subjects, failing to validate the efficiency of the preprocessing methods.

V. MAIN FINDINGS

A notable contribution of this study is the proposal of a noise-oriented preprocessing approach for cognitive load assessment (which is grounded on the artifacts stemming from the user's computer-mediated activity), along with a comprehensive analysis of existing works that are founded on this innovative approach. Accordingly, we classified

tasks into two distinct categories: low-noise and high-noise, primarily based on the potential for internal artifacts. Low-noise tasks encompass activities where users interact with a computer and involve relatively simple movements executed by a mouse. Conversely, high-noise tasks entail substantial muscle engagement, including walking, head movements, mouse-based drawing, keyboard typing, and speech production. We further subdivided this high-noise category into two distinct subcategories: ''reduced-high-noise'' and ''elevated-high-noise''.

This classification also considers the potential influence of external artifacts. Based on the aforementioned noise-centered classification approach, we synthesized and categorized tasks according to their likely noise levels. The research on denoising low-noise tasks has reached a satisfactory level. Similarly, denoising reduced-high-noise tasks with stationary participants, such as VR, AR, and MR tasks, has also seen sufficient progress.

The first finding is that semi-automated methods remain in the foreground despite time and experience requirements. Simple and advanced preprocessing methods were sequentially combined and parameterized according to the particularities of each dataset. The prevailing denoising approach is filtering, followed by a BSS approach, mostly using ICA models. Automated ICA versions are a trend. Moreover, novel and promising denoising approaches, such as deep neural networks and fully automated solutions, have not yet been investigated in computerized Cognitive Load experiments. Their performance and reliability have been demonstrated in the literature. We think that it is time to prove their power and ease of use.

With regards to *semi-automated preprocessing protocols*, we note that *standard preprocessing* mainly included resampling methods, particularly downsampling to 250Hz for recording frequencies greater than 500 Hz, re-referencing to CAR, averaging event trials, baseline correction with a short pre-stimulus portion of the signal (200ms to 1s), segmentation, manual detection and correction/rejection of artifacts, and filtering. The effects of baseline correction on the signal are open to question. However, several studies have included baseline correction in their semi-automated processes.

The prevailing *segmentation* was epoching and was performed mostly after the basic filtering and BSS methods. Non-overlapping epochs related to the events were the norm. The duration varied according to the task. One or two seconds was usually sufficient. Longer tasks or trials resulted in longer epochs. For tasks with a stimulus presentation, the extracted epochs were extended to approximately 800ms after the stimulus presentation. Before the stimulus presentation, 200ms was sufficient to identify the baseline. For any level of noise, performing epoching at a later stage during preprocessing is recommended. For the duration of each epoch, it is preferable to use short epochs centered around the presentation of the stimulus. In addition, it is advisable to consider the rejection of epochs using an amplitude threshold.

Specifically, *manual interventions* and visual inspections still have value for researchers. Manual detection and rejection of epochs, channels, or whole participant data were apparent in almost every study. Nevertheless, their effectiveness is open to question and prone to data loss and expert bias. Regarding the manual rejection of data, the threshold for epoch rejection was mainly an amplitude range, a standard deviation from the mean, or a time threshold. However, diverse threshold values were employed depending on the task and visually detected noise. Two tasks with computer puzzle games similarly rejected epochs. Almost equal threshold values for the standard deviation from the mean were selected. However, the first was a sitting task, and the second was a walking task. In general, specific manual rejection techniques can differ considerably, contingent on the characteristics of the task at hand.

Regarding *filtering,*the Median filter, Savitzsky- Golay filter, and Wiener filter were employed only once. Although the Wiener filter is not recommended for motion situations with a large amount of noise, it was selected for a VR task. A 50Hz Notch filter was sometimes employed to remove line noise. Concerning the design of the filters, FIR was the dominant type of impulse response, and band-pass was the dominant type of frequency response. The most cited high-pass cut-off frequencies were 0.1 Hz, 0.5 Hz, or 1 Hz. The low-pass cut-off frequencies varied significantly, yet 30 Hz or 45 Hz were the most popular values. Information regarding transition band, roll-off, and filter order was extremely poor. Because researchers are not always aware of the filtering effects on the signal and choose filter designs that may distort the signal, more detailed specifications are required.

The most prevalent *advanced denoising* option in semi-automated preprocessing protocols was the BSS algorithms and particularly ICA. To separate the signal into ICs, Infomax and its automated version, Runica, stand out. Many tools have been presented to detect noisy ICs manually. Topographic maps, spectral properties, and dipole estimation were common. Some general guidelines extracted were that ocular artifacts were usually concentrated signals in the occipital regions, whereas muscle artifacts were concentrated in specific brain areas. Movement artifacts increased in distribution in high-frequency regions. An ICA-guided MARA was selected in one study to automatically detect and remove/correct noisy ICs. PCA was typically used to remove ocular contamination and may be coupled with ICA to remove the signal from other artifacts further. ASR was used alone or followed by ICA to remove non-stationary and large amplitude artifacts. SSD, an eigenvalue decomposition method, was an interesting and reliable approach to denoise signals.

It should be noted that given the trend of using deep learning to denoise EEG, no such method was employed directly in the works we studied. However, one work combined CNN with RNN, particularly with different LSTM models, to indirectly address noise issues. Despite the efficacy

of deep learning on mental workload EEG classification tasks [34], no clear assumptions can be safely deduced from the denoising performance of this deep learning approach in our review. Since this deep learning approach gave high accuracy in predicting Cognitive Load, it could be inferred that noise consisted of no obstacles and was effectively eliminated during deep learning calculations.

Regarding the task noise level, the prevailing preprocessing pipeline for low-noise tasks included epoching, manual rejection of noisy epochs, a band-pass FIR filter ranging from 0.5 Hz to 30 Hz, and ICA. Similarly, for reduced-high-noise tasks and elevated-high-noise tasks, the dominant preprocessing pipeline encompassed epoching, manual rejection of epochs, IIR Butterworth filter with a frequency range of 0.5 Hz to 50 Hz, and ICA. These suggestions address the lack of recommendations for denoising EEG signal captured during different intensities of movement [31].

Regarding the number of subjects, there is a need to conduct Cognitive Load experiments with more subjects. Only seven out of 57 experiments were performed with more than 50 subjects, which is considered a valid threshold for proving the efficacy of the preprocessing methods.

In a global sense, despite the advances in denoising field research, researchers of Cognitive Load in the realm of HCI avoid novel techniques. Several experiments were found to exclusively employ standard preprocessing, omitting more advanced and efficient techniques such as BSS or WT. Most studies used traditional filtering methods with BSS methods followed by segmentation. Modern approaches such as deep learning and semi-automated or fully-automated pipelines remain in research and are rarely applied in real experiments.

Effective EEG denoising for healthy subjects holds significant implications across multiple research fields. Within the realms of Human-Computer Interaction (HCI) and Cognitive Neuroscience, robust EEG preprocessing enhances technologies connected to the brain, enabling the monitoring of various health and cognitive states, such as emotions, attention, drowsiness, and workload, both offline and in real-time scenarios. Furthermore, in the rapidly expanding domain of Brain-Computer Interface (BCI) research, a cleaner EEG signal is pivotal for the development of more autonomous BCI applications and devices. These applications span a wide range, including affective computing, artistic applications, gaming BCIs, industrial robotics, and auxiliary traffic monitoring systems designed to assist drivers and pilots [15], [32]. Regarding passive BCI, a cleaner EEG signal can facilitate the online estimation of an operator's workload and enhance higher working performance and human safety [144]. Moreover, in the dynamically developing scientific field of Biomedical Engineering, a more effective denoising approach can lead to better monitoring and analysis of biomedical data [25]. Finally, an optimal EEG preprocessing approach can enhance interdisciplinary collaboration in the understanding of brain activity.

VI. CONCLUSION AND FUTURE DIRECTION

The current study aimed to elucidate the trends in preprocessing EEG data acquired during computerized Cognitive Load experiments. We envisage an optimal fully automated pipeline for all types of computerized Cognitive Load experiments, severing the ties between the task and the selected preprocessing methods.

We gathered studies spanning six years, from 2018 to 2023, and compiled data on the semi-automated preprocessing routines utilized.

We argue that in the future, well-structured EEG datasets such as EEGdenoiseNet [28] are required for evaluating denoising approaches. Regarding Cognitive Load, there are publicly available EEG datasets for healthy adults, such as EEGLearn, where 15 subjects performed a Sternberg task; EEGMAT, where 66 subjects performed serial subtractions; Hybrid EEG-NIRS, where 26 subjects performed standardized cognitive tasks; and STEW, where 48 subjects performed the SIMKAP task [33].

The EEGLearn dataset was used for a Custom Domain Adaptation (CDA) approach for Cognitive Load prediction [100], parameter-optimized CNN method for Cognitive Load assessment [76], and center loss function deep-learning model for Cognitive Load recognition [99]. Moreover, one multitasking EEG experiment used the STEW dataset to assess the Cognitive Load using graph methods [123]. Another EGG experiment employed Hybrid EEG-NIRS to assess Cognitive Load through an audified EEG [111]. These open-source datasets were not designed for studying EEG denoising, and only EEGMAT enhanced their credibility by recording EEG signal from more than 50 subjects. Thus, it is a future challenge to provide the research community with standardized EEG datasets, particularly for Cognitive Load, which has been tested on many participants.

Typically, tasks are designed for laboratories and not for real-world applications. In our review, an experiment targeted Cognitive Load prediction during ambulation and different postural tasks [138]. Another EEG experiment on gait cycle variability was also related to activity [139]. Comfortable and reliable EEG devices [13], [112], [121], [132] and prototype caps [95], [111] are gaining ground. In conclusion, our recommendation is to design ecologically valid tasks, test the performance of the preprocessing methods in real-time scenarios, give more emphasis on single-channel approaches, and prefer novel and easy-to-apply EEG devices for conducting experiments.

This study proposes a denoising framework tailored to the EEG data and user interaction context, classifying tasks into three noise levels. Our analysis yields valuable observations and recommendations, particularly relevant for diverse digital interactions such as VR, AR, mixed reality (MR), and Metaverse.

Many denoising techniques achieve remarkable accuracy; however, none of them completely solves the denoising problems for diverse experiments. Approaches with robust results in other fields and classification problems may be

appropriate for denoising EEG signals. One solution is to focus on geometrical features, as was done in research on the diagnosis of depression [145]. Specifically, the second-order differential plot (SODP) was employed to extract 26 geometrical features, such as standard descriptors, angles between consecutive vectors, and distances to coordinate. After feature selection, classification was performed using the K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) algorithms. A remarkable classification accuracy of 98.79% was cited, significantly higher than the classification based on features from the time, frequency, or time-frequency domains. Furthermore, in a study to identify epileptic seizures, four geometrical features were extracted from the Poincaré plot of DWT coefficients [146]. Binary particle swarm optimization was employed to select features, and KNN and SVM were used to classify subjects as seizure or seizure-free to achieve an accuracy of 99.3 %. Moreover, geometrical features have been used to detect alcoholism. After applying the phase space dynamic technique, 34 geometrical features were extracted. After selecting the most significant of these, 11 common classifiers, such as neural network, multilayer neural network, recurrent neural network, generalized regression neural network, and feedforward neural network were evaluated, resulting in a 99.16% accuracy [69].

Conclusively, given the increased interest in HCI for monitoring Cognitive Load with EEG across diverse computer interaction domains, the preprocessing of EEG data is rather challenging. Conventional approaches have probably reached their performance limits, and novel approaches are taking center stage. However, there is no consensus on a standardized solution. Within this perspective, the suggested three-tier noise-oriented preprocessing approach offers a structured alternative, providing information for developing an optimized protocol for computerized Cognitive Load EEG experiments.

Limitations: Previous reviews on EEG classification included denoising but focused on ''Cognitive Workload'', extracting works more related to the workplace [33], [34]. This study is a portion of bigger research regarding working memory; therefore, the keywords for extracting papers targeted the broad concept of Working Memory EEG experiments. Hereby, the portion related to one basic concept of Working Memory, Cognitive Load, was used. Since the term ''Cognitive Load'' stems from a learning theory [20], the keywords ''EEG'' AND ''Cognitive load'' adopted by our study may elicit works more related to learning. In the future, to overcome this issue, more precise keywords would enrich our study. Specifically, the terms ''Mental Workload'' and ''Cognitive Workload'' that stem from a workplace theory [21] could extract more preprocessing approaches targeting the working environment and performance.

EEG preprocessing and Cognitive Load are multidisciplinary topics. Related reviews extracted articles from multidisciplinary databases, such as Web of Science [15], [34], Scopus [22], PubMed [34], and Google Scholar [22], [24]. Our research was conducted in two libraries, IEEE Xplore and ACM Digital Library, addressing more computer scientists and electrical engineers. Thus, neuroscientists and clinicians might be skeptical about adopting suggested techniques. An additional search in other digital databases would enrich our study with impactful preprocessing techniques from other disciplines, such as biomedical engineering.

Another limitation is that we focused on gathering preprocessing strategies from the last six years to encompass the recent developments. Thus, there may be older noteworthy approaches in the field of HCI that might have been overlooked. Numerous novel EEG preprocessing methodologies have originated as variations or amalgamations of pre-existing approaches. Therefore, a comprehensive understanding of the current landscape could elicit promising future solutions. So we conclude that in future a wider time window could prove beneficial.

REFERENCES

[1] P. Lahane, J. Jagtap, A. Inamdar, N. Karne, and R. Dev, ''A review of recent trends in EEG based brain-computer interface,'' in *Proc. Int. Conf. Comput. Intell. Data Sci. (ICCIDS)*, Feb. 2019, pp. 1–6.

[2] N. Bigdely-Shamlo, T. Mullen, C. Kothe, K.-M. Su, and K. A. Robbins, ''The PREP pipeline: Standardized preprocessing for large-scale EEG analysis,'' *Frontiers Neuroinform.*, vol. 9, p. 16, Jun. 2015.

[3] L. J. Gabard-Durnam, A. S. Mendez Leal, C. L. Wilkinson, and A. R. Levin, ''The Harvard automated processing pipeline for electroencephalography (HAPPE): Standardized processing software for developmental and high-artifact data,'' *Frontiers Neurosci.*, vol. 12, p. 97, Feb. 2018.

[4] A. R. Levin, A. S. Méndez Leal, L. J. Gabard-Durnam, and H. M. O'Leary, ''BEAPP: The batch electroencephalography automated processing platform,'' *Frontiers Neurosci.*, vol. 12, p. 513, Aug. 2018.

[5] J. R. da Cruz, V. Chicherov, M. H. Herzog, and P. Figueiredo, ''An automatic pre-processing pipeline for EEG analysis (APP) based on robust statistics,'' *Clin. Neurophysiol.*, vol. 129, no. 7, pp. 1427–1437, 2018.

[6] M. Miyakoshi, *Makoto's Preprocessing Pipeline*.

[7] D. Gruber, *Streamline Your EEGLAB Experience*.

[8] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, L. Parkkonen, and M. S. Hämäläinen, ''MNE software for processing MEG and EEG data,'' *NeuroImage*, vol. 86, pp. 446–460, Feb. 2014.

[9] F. Tadel, S. Baillet, J. C. Mosher, D. Pantazis, and R. M. Leahy, ''Brainstorm: A user-friendly application for MEG/EEG analysis,'' *Comput. Intell. Neurosci.*, vol. 2011, pp. 1–13, Jan. 2011.

[10] A. Delorme, ''EEG is better left alone,'' *Sci. Rep.*, vol. 13, p. 2372, Feb. 2023.

[11] S. Sadiya, T. Alhanai, and M. M. Ghassemi, ''Artifact detection and correction in EEG data: A review,'' in *Proc. 10th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, May 2021, pp. 495–498.

[12] M. M. N. Mannan, M. A. Kamran, and M. Y. Jeong, ''Identification and removal of physiological artifacts from electroencephalogram signals: A review,'' *IEEE Access*, vol. 6, pp. 30630–30652, 2018.

[13] A. Sinha, D. Roy, R. Chaki, B. K. De, and S. K. Saha, ''Readability analysis based on cognitive assessment using physiological sensing,'' *IEEE Sensors J.*, vol. 19, no. 18, pp. 8127–8135, Sep. 2019.

[14] B. Volmer, J. Baumeister, S. Von Itzstein, M. Schlesewsky, I. Bornkessel-Schlesewsky, and B. H. Thomas, ''Event related brain responses reveal the impact of spatial augmented reality predictive cues on mental effort,'' *IEEE Trans. Vis. Comput. Graphics*, vol. 29, no. 12, pp. 4990–5007, Dec. 2023.

[15] X. Chen, X. Xu, A. Liu, S. Lee, X. Chen, X. Zhang, M. J. McKeown, and Z. J. Wang, ''Removal of muscle artifacts from the EEG: A review and recommendations,'' *IEEE Sensors J.*, vol. 19, no. 14, pp. 5353–5368, Jul. 2019.

[16] J. Mateo, E. Sánchez-Morla, and J. Santos, "A new method for removal of powerline interference in ECG and EEG recordings," *Comput. Elect. Eng.*, vol. 45, pp. 235–248, Jul. 2015.
[17] A. de Cheveigné and D. Arzounian, "Robust detrending, rereferencing, outlier detection, and inpainting for multichannel data," *NeuroImage*, vol. 172, pp. 903–912, May 2018.
[18] A. Qayyum, M. K. A. A. Khan, M. Mazher, and M. Suresh, "Classification of EEG learning and resting states using 1D-convolutional neural network for cognitive load assesment," in *Proc. IEEE Student Conf. Res. Develop. (SCOReD)*, Nov. 2018, pp. 1–5.
[19] R. Abreu, A. Leal, and P. Figueiredo, "EEG-informed fMRI: A review of data analysis methods," *Frontiers Hum. Neurosci.*, vol. 12, p. 29, Feb. 2018.
[20] J. Sweller, J. J. G. van Merriënboer, and F. Paas, "Cognitive architecture and instructional design: 20 years later," *Educ. Psychol. Rev.*, vol. 31, no. 2, pp. 261–292, Jun. 2019.
[21] C. D. Wickens, "Multiple resources and mental workload," *Hum. Factors*, vol. 50, no. 3, pp. 449–455, Jun. 2008.
[22] J. A. Urigüen and B. Garcia-Zapirain, "EEG artifact removal—State-of-the-art and guidelines," *J. Neural Eng.*, vol. 12, Apr. 2015, Art. no. 031001.
[23] S.-P. Kim, *Preprocessing of EEG*. Singapore: Springer, 2018, pp. 15–33.
[24] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of artifacts from EEG signals: A review," *Sensors*, vol. 19, no. 5, p. 987, Feb. 2019.
[25] R. Martinek, M. Ladrova, M. Sidikova, R. Jaros, K. Behbehani, R. Kahankova, and A. Kawala-Sterniuk, "Advanced bioelectrical signal processing methods: Past, present and future approach—Part II: Brain signals," *Sensors*, vol. 21, no. 19, p. 6343, 2021.
[26] M. Sheoran, S. Kumar, and S. Chawla, "Methods of denoising of electroencephalogram signal: A review," *Int. J. Biomed. Eng. Technol.*, vol. 18, p. 385, Jan. 2015.
[27] M. K. Islam, A. Rastegarnia, and Z. Yang, "Methods for artifact detection and removal from scalp EEG: A review," *Clin. Neurophysiol.*, vol. 46, nos. 4–5, pp. 287–305, Nov. 2016.
[28] H. Zhang, M. Zhao, C. Wei, D. Mantini, Z. Li, and Q. Liu, "EEGdenoiseNet: A benchmark dataset for deep learning solutions of EEG denoising," *J. Neural Eng.*, vol. 18, no. 5, Oct. 2021, Art. no. 056057.
[29] K. A. Robbins, J. Touryan, T. Mullen, C. Kothe, and N. Bigdely-Shamlo, "How sensitive are EEG results to preprocessing methods: A benchmarking study," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1081–1090, May 2020.
[30] M. Grobbelaar, S. Phadikar, E. Ghaderpour, A. F. Struck, N. Sinha, R. Ghosh, and M. Z. I. Ahmed, "A survey on denoising techniques of electroencephalogram signals using wavelet transform," *Signals*, vol. 3, no. 3, pp. 577–586, Aug. 2022.
[31] D. Gorjan, K. Gramann, K. De Pauw, and U. Marusic, "Removal of movement-induced EEG artifacts: Current state of the art and guidelines," *J. Neural Eng.*, vol. 19, no. 1, Feb. 2022, Art. no. 011004.
[32] M. F. Mridha, S. C. Das, M. M. Kabir, A. A. Lima, M. R. Islam, and Y. Watanobe, "Brain-computer interface: Advancement and challenges," *Sensors*, vol. 21, no. 17, p. 5746, Aug. 2021.
[33] Y. Zhou, S. Huang, Z. Xu, P. Wang, X. Wu, and D. Zhang, "Cognitive workload recognition using EEG signals and machine learning: A review," *IEEE Trans. Cognit. Develop. Syst.*, vol. 14, no. 3, pp. 799–818, Sep. 2022.
[34] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *J. Neural Eng.*, vol. 16, no. 3, Apr. 2019, Art. no. 031001.
[35] 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, Aug. 2019, Art. no. 051001.
[36] M. Hassan and S. Md. Rabiul Islam, "Design and implementation of preprocessing chip for brain computer interface machine," in *Proc. Int. Conf. Robot., Elect. Signal Process. Techn. (ICREST)*, Jan. 2019, pp. 428–433.
[37] D. Candia-Rivera, V. Catrambone, and G. Valenza, "The role of electroencephalography electrical reference in the assessment of functional brain–heart interplay: From methodology to user guidelines," *J. Neurosci. Methods*, vol. 360, Aug. 2021, Art. no. 109269.
[38] J. Medeiros, R. Couceiro, J. Castelhano, M. C. Branco, G. Duarte, C. Duarte, J. Durães, H. Madeira, P. Carvalho, and C. Teixeira, "Software code complexity assessment using EEG features," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 1413–1416.
[39] A. Borowicz, "Using a multichannel Wiener filter to remove eye-blink artifacts from EEG data," *Biomed. Signal Process. Control*, vol. 45, pp. 246–255, Aug. 2018.
[40] S. J. Luck, *Applied Event-Related Potential Data Analysis*. LibreTexts, 2022.
[41] ERP Boot Camp, *Example of Scrolling Through a Continuous Dataset to Assess Bad Channels and Artifacts*, Ch. 7.
For this Youtube video no publisher name and location are provided. The channel name is ERP Boot Camp.
[42] A. Widmann, E. Schröger, and B. Maess, "Digital filter design for electrophysiological data—A practical approach," *J. Neurosci. Methods*, vol. 250, pp. 34–46, Jul. 2015.
[43] K. Sundaram, "FPGA based filters for EEG pre-processing," in *Proc. 2nd Int. Conf. Sci. Technol. Eng. Manage. (ICONSTEM)*, Mar. 2016, pp. 572–576.
[44] U. Shukla, G. J. Saxena, M. Kumar, A. S. Bafila, A. Pundir, and S. Singh, "An improved decision support system for identification of abnormal EEG signals using a 1D convolutional neural network and Savitzky-Golay filtering," *IEEE Access*, vol. 9, pp. 163492–163503, 2021.
[45] Y. Yan, H. Zhou, L. Huang, X. Cheng, and S. Kuang, "A novel two-stage refine filtering method for EEG-based motor imagery classification," *Frontiers Neurosci.*, vol. 15, Sep. 2021, Art. no. 657540.
[46] M. Z. I. Ahmed, N. Sinha, E. Ghaderpour, S. Phadikar, and R. Ghosh, "A novel baseline removal paradigm for subject-independent features in emotion classification using EEG," *Bioengineering*, vol. 10, no. 1, p. 54, 2023.
[47] A. Chaddad, Y. Wu, R. Kateb, and A. Bouridane, "Electroencephalography signal processing: A comprehensive review and analysis of methods and techniques," *Sensors*, vol. 23, no. 14, p. 6434, Jul. 2023.
[48] R. Martinez Orellana, B. Erem, and D. H. Brooks, "Time invariant multi electrode averaging for biomedical signals," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 1242–1246.
[49] H. Mir, I. Prasad, K. Yu, N. Thakor, and H. Al-Nashash, "ERP signal estimation from single trial EEG," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 2989–2992.
[50] T. A. Suhail, K. P. Indiradevi, E. M. Suhara, S. A. Poovathinal, and A. Anitha, "Performance analysis of mother wavelet functions and thresholding methods for denoising EEG signals during cognitive tasks," in *Proc. Int. Conf. Power, Instrum., Control Comput. (PICC)*, Dec. 2020, pp. 1–6.
[51] P. Divya and B. A. Devi, "Epileptic EEG signal denoising enhancement using improved threshold based wavelet method," in *Proc. Int. Conf. Syst., Comput., Autom. Netw. (ICSCAN)*, Jul. 2021, pp. 1–4.
[52] M. N. Tibdewal, M. Mahadevappa, A. K. Ray, M. Malokar, and H. R. Dey, "Power line and ocular artifact denoising from EEG using notch filter and wavelet transform," in *Proc. 3rd Int. Conf. Comput. for Sustain. Global Develop. (INDIACom)*, Mar. 2016, pp. 1654–1659.
[53] S. N. S. S. Daud, R. Sudirman, N. H. Mahmood, and C. Omar, "Denoising semi-simulated EEG signal contaminated ocular noise using various wavelet filters," in *Proc. 13th Int. Conf. Inf. Commun. Syst. (ICICS)*, Jun. 2022, pp. 452–457.
[54] C. I. Salis, A. E. Malissovas, P. A. Bizopoulos, A. T. Tzallas, P. A. Angelidis, and D. G. Tsalikakis, "Denoising simulated EEG signals: A comparative study of EMD, wavelet transform and Kalman filter," in *Proc. 13th IEEE Int. Conf. Bioinf. BioEngineering*, Nov. 2013, pp. 1–4.
[55] X.-B. Qin, Y.-Z. Zhang, M.-T. Huang, and M. Wang, "EEG signal recognition based on wavelet transform and neural network," in *Proc. Int. Symp. Comput., Consum. Control (IS3C)*, Jul. 2016, pp. 523–526.
[56] A. Egambaram, N. Badruddin, V. S. Asirvadam, and T. Begum, "Comparison of envelope interpolation techniques in empirical mode decomposition (EMD) for eyeblink artifact removal from EEG," in *Proc. IEEE EMBS Conf. Biomed. Eng. Sci. (IECBES)*, Dec. 2016, pp. 590–595.
[57] J. Vosskuhl, T. P. Mutanen, T. Neuling, R. J. Ilmoniemi, and C. S. Herrmann, "Signal-space projection suppresses the tACS artifact in EEG recordings," *Frontiers Hum. Neurosci.*, vol. 14, Dec. 2020, Art. no. 536070.
[58] B. U. Westner, S. S. Dalal, A. Gramfort, V. Litvak, J. C. Mosher, R. Oostenveld, and J.-M. Schoffelen, "A unified view on beamformers for M/EEG source reconstruction," *NeuroImage*, vol. 246, Feb. 2022, Art. no. 118789.

[59] M. Uji, N. Cross, F. B. Pomares, A. A. Perrault, A. Jegou, A. Nguyen, U. Aydin, J.-M. Lina, T. T. Dang-Vu, and C. Grova, "Data-driven beamforming technique to attenuate ballistocardiogram artefacts in electroencephalography–functional magnetic resonance imaging without detecting cardiac pulses in electrocardiography recordings," *Hum. Brain Mapping*, vol. 42, no. 12, pp. 3993–4021, 2021.

[60] A. Bera, N. Das, and M. Chakraborty, "Optimal filtering of single channel EEG data using linear filters," in *Proc. 6th Int. Conf. Bio Signals, Images, Instrum. (ICBSII)*, Feb. 2020, pp. 1–6.

[61] X. Chen, X. Xu, A. Liu, M. J. McKeown, and Z. J. Wang, "The use of multivariate EMD and CCA for denoising muscle artifacts from few-channel EEG recordings," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 2, pp. 359–370, Feb. 2018.

[62] N. Bailey, M. Biabani, A. Hill, A. Miljevic, N. Rogasch, B. McQueen, O. Murphy, and P. Fitzgerald, "Introducing RELAX: An automated pre-processing pipeline for cleaning EEG data—Part 1: Algorithm and application to oscillations," *Clin. Neurophysiol.*, vol. 149, pp. 178–201, May 2023.

[63] I. Winkler, S. Haufe, and M. Tangermann, "Automatic classification of artifactual ICA-components for artifact removal in EEG signals," *Behav. Brain Functions*, vol. 7, p. 30, Dec. 2011.

[64] M. Jas, D. A. Engemann, Y. Bekhti, F. Raimondo, and A. Gramfort, "Autoreject: Automated artifact rejection for MEG and EEG data," *NeuroImage*, vol. 159, pp. 417–429, Oct. 2017.

[65] P. Bloniasz, "Artifact subspace reconstruction (ASR) for electroencephalography artifact removal must be optimized for each unique dataset," *Qeios*, 2022.

[66] C. Chang, S. Hsu, L. Pion-Tonachini, and T. Jung, "Evaluation of artifact subspace reconstruction for automatic artifact components removal in multi-channel EEG recordings," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 4, pp. 1114–1121, Apr. 2020.

[67] V. V. Nikulin, G. Nolte, and G. Curio, "A novel method for reliable and fast extraction of neuronal EEG/MEG oscillations on the basis of spatio-spectral decomposition," *NeuroImage*, vol. 55, no. 4, pp. 1528–1535, Apr. 2011.

[68] M. S. Hossain, S. Mahmud, A. Khandakar, N. Al-Emadi, F. A. Chowdhury, Z. B. Mahbub, M. B. I. Reaz, and M. E. H. Chowdhury, "MultiResUNet3+: A full-scale connected multi-residual unet model to denoise electrooculogram and electromyogram artifacts from corrupted electroencephalogram signals," *Bioengineering*, vol. 10, no. 5, p. 579, 2023.

[69] M. T. Sadiq, H. Akbari, S. Siuly, Y. Li, and P. Wen, "Alcoholic EEG signals recognition based on phase space dynamic and geometrical features," *Chaos, Solitons Fractals*, vol. 158, May 2022, Art. no. 112036.

[70] X. Xu, Y. Zhang, R. Zhang, and T. Xu, "Patient-specific method for predicting epileptic seizures based on DRSN-GRU," *Biomed. Signal Process. Control*, vol. 81, Mar. 2023, Art. no. 104449.

[71] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Comput. Intell. Neurosci.*, vol. 2018, pp. 1–13, Feb. 2018.

[72] H. Zhang, C. Wei, M. Zhao, Q. Liu, and H. Wu, "A novel convolutional neural network model to remove muscle artifacts from EEG," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 1265–1269.

[73] Y. An, H. K. Lam, and S. H. Ling, "Auto-denoising for EEG signals using generative adversarial network," *Sensors*, vol. 22, no. 5, p. 1750, Feb. 2022.

[74] W. Sun, Y. Su, X. Wu, and X. Wu, "A novel end-to-end 1D-ResCNN model to remove artifact from EEG signals," *Neurocomputing*, vol. 404, pp. 108–121, Sep. 2020.

[75] C. Yildirim, "A review of deep learning approaches to EEG-based classification of cybersickness in virtual reality," in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Dec. 2020, pp. 351–357.

[76] F. Havugimana, M. B. Muhammad, K. A. Moinudin, and M. Yeasin, "Predicting cognitive load using parameter-optimized CNN from spatial–spectral representation of EEG recordings," in *Proc. 20th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2021, pp. 710–715.

[77] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update," *J. Neural Eng.*, vol. 15, Apr. 2018, Art. no. 031005.

[78] N. Mashhadi, A. Z. Khuzani, M. Heidari, and D. Khaledyan, "Deep learning denoising for EOG artifacts removal from EEG signals," in *Proc. IEEE Global Humanitarian Technol. Conf. (GHTC)*, Oct. 2020, pp. 1–6.

[79] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Comput.*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019.

[80] R. C. Staudemeyer and E. R. Morris, "Understanding LSTM—A tutorial into long short-term memory recurrent neural networks," 2019, *arXiv:1909.09586*.

[81] M. T. Sadiq, X. Yu, Z. Yuan, and M. Z. Aziz, "Motor imagery BCI classification based on novel two-dimensional modelling in empirical wavelet transform," *Electron. Lett.*, vol. 56, no. 25, pp. 1367–1369, 2020.

[82] M. T. Sadiq, X. Yu, Z. Yuan, M. Z. Aziz, N. U. Rehman, W. Ding, and G. Xiao, "Motor imagery BCI classification based on multivariate variational mode decomposition," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 6, no. 5, pp. 1177–1189, Oct. 2022.

[83] M. T. Sadiq, X. Yu, Z. Yuan, F. Zeming, A. U. Rehman, I. Ullah, G. Li, and G. Xiao, "Motor imagery EEG signals decoding by multivariate empirical wavelet transform-based framework for robust brain–computer interfaces," *IEEE Access*, vol. 7, pp. 171431–171451, 2019.

[84] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Comput. Intell. Neurosci.*, vol. 2011, pp. 1–9, Jan. 2011.

[85] Y. Liu, Y. Yu, Z. Ye, M. Li, Y. Zhang, Z. Zhou, D. Hu, and L.-L. Zeng, "Fusion of spatial, temporal, and spectral EEG signatures improves multilevel cognitive load prediction," *IEEE Trans. Human-Mach. Syst.*, vol. 53, no. 2, pp. 357–366, Apr. 2023.

[86] J. Rodrigues, M. Weiß, J. Hewig, and J. J. B. Allen, "EPOS: EEG processing open-source scripts," *Frontiers Neurosci.*, vol. 15, Jun. 2021, Art. no. 660449.

[87] A. Pedroni, A. Bahreini, and N. Langer, "Automagic: Standardized preprocessing of big EEG data," *NeuroImage*, vol. 200, pp. 460–473, Oct. 2019.

[88] M. T. Sadiq, X. Yu, Z. Yuan, M. Z. Aziz, S. Siuly, and W. Ding, "Toward the development of versatile brain–computer interfaces," *IEEE Trans. Artif. Intell.*, vol. 2, no. 4, pp. 314–328, Aug. 2021.

[89] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *Int. J. Surg.*, vol. 8, no. 5, pp. 336–341, 2010.

[90] V. Grubov, A. Badarin, and V. Maksimenko, "Analysis of information perception and processing during long-term and intense cognitive load using combined EEG and NIRS," in *Proc. Int. Conf. Nonlinearity, Inf. Robot. (NIR)*, Dec. 2020, pp. 1–2.

[91] S. M. Salaken, I. Hettiarachchi, L. Crameri, S. Hanoun, T. Nguyen, and S. Nahavandi, "Evaluation of classification techniques for identifying cognitive load levels using EEG signals," in *Proc. IEEE Int. Syst. Conf. (SysCon)*, Aug. 2020, pp. 1–8.

[92] M. Elkerdawy, M. Elhalaby, A. Hassan, M. Maher, D. Shawky, and A. Badawi, "Building cognitive profiles of learners using EEG," in *Proc. 11th Int. Conf. Inf. Commun. Syst. (ICICS)*, Apr. 2020, pp. 027–032.

[93] K. Pollmann, O. Stefani, A. Bengsch, M. Peissner, and M. Vukelić, "How to work in the car of the future?: A neuroergonomical study assessing concentration, performance and workload based on subjective, behavioral and neurophysiological insights," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, 2019, pp. 1–14.

[94] V. Parekh, M. Bilalpur, C. V. Jawahar, S. Kumar, S. Winkler, and R. Subramanian, "Investigating the generalizability of EEG-based cognitive load estimation across visualizations," in *Proc. 20th Int. Conf. Multimodal Interact., Adjunct*, New York, NY, USA, 2018, pp. 1–5.

[95] J. F. Vargas, B. Zhou, H. Bello, and P. Lukowicz, "Brainwear: Towards multi-modal garment integrated EEG," in *Proc. ISWC*, New York, NY, USA, 2021, pp. 113–118.

[96] J.-H. Kwon, E. Kim, C.-H. Im, and D.-W. Kim, "Classification of different cognitive load using electroencephalogram(EEG): Preliminary study," in *Proc. Joint 10th Int. Conf. Soft Comput. Intell. Syst. (SCIS) 19th Int. Symp. Adv. Intell. Syst. (ISIS)*, Dec. 2018, pp. 205–208.

[97] M. Beiramvand, T. Lipping, N. Karttunen, and R. Koivula, "Mental workload assessment using low-channel prefrontal EEG signals," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2023, pp. 1–5.

[98] Y. Kwak, K. Kong, W. Song, B. Min, and S. Kim, "Multilevel feature fusion with 3D convolutional neural network for EEG-based workload estimation," *IEEE Access*, vol. 8, pp. 16009–16021, 2020.

[99] W. Zhang and Q. Liu, "Using the center loss function to improve deep learning performance for EEG signal classification," in *Proc. 10th Int. Conf. Adv. Comput. Intell. (ICACI)*, Mar. 2018, pp. 578–582.

[100] M. Jiménez-Guarneros and P. Gómez-Gil, "Custom domain adaptation: A new method for cross-subject, EEG-based cognitive load recognition," *IEEE Signal Process. Lett.*, vol. 27, pp. 750–754, 2020.

[101] S. Kuanar, V. Athitsos, N. Pradhan, A. Mishra, and K. R. Rao, "Cognitive analysis of working memory load from EEG, by a deep recurrent neural network," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 2576–2580.

[102] F. Havugimana, K. A. Moinudin, and M. Yeasin, "Deep learning framework for modeling cognitive load from small and noisy EEG data," *IEEE Trans. Cogn. Developmental Syst.*, early access, 2023.

[103] J. Mercier, K. Whissell-Turner, A. Paradis, and I. Avaca, "Good vibrations: Tuning a systems dynamics model of affect and cognition in learning to the appropriate frequency bands of fine-grained temporal sequences of data: Frequency bands of affect and cognition," in *Proc. 9th Int. Conf. Softw. Develop. Technol. Enhancing Accessibility Fighting Info-Exclusion*, New York, NY, USA, 2021, pp. 194–202.

[104] D. Shakti, S. Saha, V. Sardana, P. Akula, C. M. Ananda, and S. Tewary, "EEG as a tool to measure cognitive load while playing sudoku: A preliminary study," in *Proc. 3rd Int. Conf. Electron., Mater. Eng. Nano-Technology (IEMENTech)*, Aug. 2019, pp. 1–5.

[105] T. Kosch, A. Schmidt, S. Thanheiser, and L. L. Chuang, "One does not simply RSVP: Mental workload to select speed reading parameters using electroencephalography," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, 2020, pp. 1–13.

[106] V. Yoghourdjian, Y. Yang, T. Dwyer, L. Lawrence, M. Wybrow, and K. Marriott, "Scalability of network visualisation from a cognitive load perspective," *IEEE Trans. Vis. Comput. Graphics*, vol. 27, no. 2, pp. 1677–1687, Feb. 2021.

[107] X. Meng, W. Zheng, and K. Huang, "Cognitive load evaluation of human-computer interface based on EEG multi-dimensional feature," in *Proc. IEEE 25th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2022, pp. 1536–1541.

[108] Y. Liang, W. Liang, J. Qu, and J. Yang, "Experimental study on EEG with different cognitive load," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2018, pp. 4351–4356.

[109] Z. Song, M. Li, L. Wu, Y. Wu, G. Xu, F. Lin, and W. Liao, "An adaptive-load classifier based on cognitive load similarity," *IEEE Sensors J.*, vol. 23, no. 18, pp. 21978–21988, Sep. 2023.

[110] M. Bilalpur, M. Kankanhalli, S. Winkler, and R. Subramanian, "EEG-based evaluation of cognitive workload induced by acoustic parameters for data sonification," in *Proc. 20th ACM Int. Conf. Multimodal Interact.*, New York, NY, USA, 2018, pp. 315–323.

[111] R. Spender, T. C. Davies, and S. D. Pinder, "Detecting changes in cognitive load through audified EEG," in *Proc. IEEE Region 10 Conf. (TENCON)*, Dec. 2021, pp. 289–292.

[112] K. Tamura, T. Okamoto, M. Oi, A. Shimada, K. Hatano, M. Yamada, M. Lu, and S. Konomi, "Pilot study to estimate 'difficult' area in e-learning material by physiological measurements," in *Proc. 6th ACM Conf. Learn. @ Scale*, New York, NY, USA, 2019, pp. 1–4.

[113] S. Ahmed, Md. A. A. Walid, and M. Islam, "EEG-based cognitive load assessment in MATLAB GUI and impact on learning system," in *Proc. 2nd Int. Conf. Adv. Inf. Commun. Technol. (ICAICT)*, Nov. 2020, pp. 484–487.

[114] M. H. Fard, K. Petrova, N. Kasabov, and G. Y. Wang, "Studying transfer of learning using a brain-inspired spiking neural network in the context of learning a new programming language," in *Proc. IEEE Asia–Pacific Conf. Comput. Sci. Data Eng. (CSDE)*, Dec. 2021, pp. 1–6.

[115] N. Peitek, A. Bergum, M. Rekrut, J. Mucke, M. Nadig, C. Parnin, J. Siegmund, and S. Apel, "Correlates of programmer efficacy and their link to experience: A combined EEG and eye-tracking study," in *Proc. 30th ACM Joint Eur. Softw. Eng. Conf. Symp. Found. Softw. Eng.*, New York, NY, USA, Nov. 2022, pp. 120–131.

[116] H. Hijazi, J. Duraes, R. Couceiro, J. Castelhano, R. Barbosa, J. Medeiros, M. Castelo-Branco, P. de Carvalho, and H. Madeira, "Quality evaluation of modern code reviews through intelligent biometric program comprehension," *IEEE Trans. Softw. Eng.*, vol. 49, no. 2, pp. 626–645, Feb. 2023.

[117] A. Diwakar, T. Kaur, C. Ralekar, and T. K. Gandhi, "Deep learning identifies brain cognitive load via EEG signals," in *Proc. IEEE 17th India Council Int. Conf. (INDICON)*, Dec. 2020, pp. 1–5.

[118] M. O. Zhuravlev, R. V. Ukolov, A. E. Runnova, and A. A. Titova, "The method of spatial–temporal analysis of various oscillational modes in the EEG brain activity," in *Proc. Int. Conf. Quality Manage., Transp. Inf. Secur., Inf. Technol. (IT&QM&IS)*, Sep. 2020, pp. 562–565.

[119] A. Kujur, A. Bhattacharya, G. Sharma, and J. Kumar, "Prediction of workload under distraction using supervised learning algorithms," in *Proc. 3rd Int. Conf. Issues Challenges Intell. Comput. Techn. (ICICT)*, Nov. 2022, pp. 1–5.

[120] M.-K. Jung, H. Kim, S. Lee, J. B. Kim, and D.-J. Kim, "Interpretability of hybrid feature using graph neural networks from mental arithmetic based EEG," in *Proc. 11th Int. Winter Conf. Brain-Comput. Interface (BCI)*, Feb. 2023, pp. 1–5.

[121] O. Ortiz, D. Blustein, and U. Kuruganti, "Test-retest reliability of time-domain EEG features to assess cognitive load using a wireless dry-electrode system," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 2885–2888.

[122] M. M. Swerdloff and L. J. Hargrove, "Identifying the onset of increased cognitive load using event-related potentials in electroencephalography," in *Proc. 10th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, May 2021, pp. 909–912.

[123] G. Zhu, F. Zong, H. Zhang, B. Wei, and F. Liu, "Cognitive load during multitasking can be accurately assessed based on single channel electroencephalography using graph methods," *IEEE Access*, vol. 9, pp. 33102–33109, 2021.

[124] E. M. Barrella, C. Cowan, J. Girdner, M. K. Watson, and R. Anderson, "Measuring connections: Engineering students' cognitive activities and performance on complex tasks," in *Proc. IEEE Frontiers Educ. Conf. (FIE)*, Oct. 2019, pp. 1–8.

[125] H. Yagura, H. Tanaka, T. Kinoshita, H. Watanabe, S. Motomura, K. Sudoh, and S. Nakamura, "Analysis of selective attention processing on experienced simultaneous interpreters using EEG phase synchronization," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 66–69.

[126] M. Z. Baig and M. Kavakli, "Analyzing novice and expert user's cognitive load in using a multi-modal interface system," in *Proc. 26th Int. Conf. Syst. Eng. (ICSEng)*, Dec. 2018, pp. 1–7.

[127] Y. Ahmed, J. Gaber, H. A. Gabbar, and J. Ren, "EEG-based integrated solution for drivers and safe transportation," in *Proc. IEEE 4th Int. Conf. Cybern., Cognition Mach. Learn. Appl. (ICCCMLA)*, Oct. 2022, pp. 227–232.

[128] E. Alyan, S. Arnau, J. E. Reiser, S. Getzmann, M. Karthaus, and E. Wascher, "Decoding eye blink and related EEG activity in realistic working environments," *IEEE J. Biomed. Health Inform.*, vol. 27, no. 12, pp. 5745–5754, Dec. 2023.

[129] D. Pulver, P. Angkan, P. Hungler, and A. Etemad, "EEG-based cognitive load classification using feature masked autoencoding and emotion transfer learning," in *Proc. 25th Int. Conf. Multimodal Interact.*, New York, NY, USA, 2023, pp. 190–197.

[130] A. Aygun, B. Lyu, T. Nguyen, Z. Haga, S. Aeron, and M. Scheutz, "Cognitive workload assessment via eye gaze and EEG in an interactive multi-modal driving task," in *Proc. Int. Conf. Multimodal Interact.*, New York, NY, USA, 2022, pp. 337–348.

[131] D. He, B. Donmez, C. C. Liu, and K. N. Plataniotis, "High cognitive load assessment in drivers through wireless electroencephalography and the validation of a modified N-back task," *IEEE Trans. Human-Mach. Syst.*, vol. 49, no. 4, pp. 362–371, Aug. 2019.

[132] T. Zhou, J. S. Cha, G. Gonzalez, J. P. Wachs, C. P. Sundaram, and D. Yu, "Multimodal physiological signals for workload prediction in robot-assisted surgery," *ACM Trans. Hum.-Robot Interact.*, vol. 9, no. 2, pp. 1–26, Jan. 2020.

[133] L. Gerry, B. Ens, A. Drogemuller, B. Thomas, and M. Billinghurst, "Levity: A virtual reality system that responds to cognitive load," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, 2018, pp. 1–6.

[134] K. Gupta, R. Hajika, Y. S. Pai, A. Duenser, M. Lochner, and M. Billinghurst, "In AI we trust: Investigating the relationship between biosignals, trust and cognitive load in VR," in *Proc. 25th ACM Symp. Virtual Reality Softw. Technol.*, New York, NY, USA, 2019, pp. 1–10.

[135] M. Ahmadi, H. Bai, A. Chatburn, M. A. Najatabadi, B. C. Wünsche, and M. Billinghurst, "Comparison of physiological cues for cognitive load measures in VR," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Abstr. Workshops (VRW)*, Mar. 2023, pp. 837–838.

[136] K. Gupta, R. Hajika, Y. S. Pai, A. Duenser, M. Lochner, and M. Billinghurst, "Measuring human trust in a virtual assistant using physiological sensing in virtual reality," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces (VR)*, Mar. 2020, pp. 756–765.
[137] S. Baceviciute, A. Mottelson, T. Terkildsen, and G. Makransky, "Investigating representation of text and audio in educational VR using learning outcomes and EEG," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, 2020, pp. 1–13.
[138] M. M. Swerdloff and L. J. Hargrove, "Quantifying cognitive load using EEG during ambulation and postural tasks," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 2849–2852.
[139] A. Nakamura, Y. Suzuki, S. Yano, and T. Nomura, "EEG activity related to decrease in persistency of gait cycle variability during distracted walking," in *Proc. IEEE 3rd Global Conf. Life Sci. Technol. (LifeTech)*, Mar. 2021, pp. 183–184.
[140] A. Qayyum, I. Faye, A. S. Malik, and M. Mazher, "Assessment of cognitive load using multimedia learning and resting states with deep learning perspective," in *Proc. IEEE-EMBS Conf. Biomed. Eng. Sci. (IECBES)*, Mali, Dec. 2018, pp. 600–605.
[141] P. Bashivan, G. M. Bidelman, and M. Yeasin, "Spectrotemporal dynamics of the EEG during working memory encoding and maintenance predicts individual behavioral capacity," *Eur. J. Neurosci.*, vol. 40, no. 12, pp. 3774–3784, Dec. 2014.
[142] W. L. Lim, O. Sourina, and L. P. Wang, "STEW: Simultaneous task EEG workload data set," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 11, pp. 2106–2114, Nov. 2018.
[143] X. Yu, M. Z. Aziz, M. T. Sadiq, Z. Fan, and G. Xiao, "A new framework for automatic detection of motor and mental imagery EEG signals for robust BCI systems," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.
[144] N. Salimi, M. Barlow, and E. Lakshika, "Mental workload classification using short duration EEG data: An ensemble approach based on individual channels," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2019, pp. 393–398.
[145] H. Akbari, M. T. Sadiq, M. Payan, S. S. Esmaili, H. Baghri, and H. Bagheri, "Depression detection based on geometrical features extracted from SODP shape of EEG signals and binary PSO," *Traitement du Signal*, vol. 38, no. 1, pp. 13–26, Feb. 2021.
[146] H. Akbari, M. T. Sadiq, N. Jafari, J. Too, N. Mikaeilvand, A. Cicone, and S. Serra-Capizzano, "Recognizing seizure using Poincaré plot of EEG signals and graphical features in DWT domain," *Bratislava Med. J.*, vol. 124, pp. 12–24, Jan. 2023.

KONSTANTINA KYRIAKI received the B.S. degree in applied informatics from the University of Macedonia, Thessaloniki, Greece, in 2002, and the M.S. degree in information systems from The University of Sheffield, Sheffield, U.K., in 2004.

She is currently a Research Assistant with the Interactive Technologies Laboratory, Department of Electrical and Computer Engineering, University of Patras, Patras, Greece. Her research interests include working memory and EEG analysis.

DIMITRIOS KOUKOPOULOS received the Diploma degree in computer engineering and informatics from the University of Patras, Greece, the M.Sc. degree in communications and signal processing from the Department of Electrical Engineering and Electronics, Imperial College of Science, Technology and Medicine, and the University of London, U.K., and the Ph.D. degree in computer engineering and informatics from the University of Patras. He was a Faculty Member with the Department of Cultural Heritage Management and New Technologies, University of Patras, from 2013 to 2019; the University of West Greece, from 2008 to 2013; and in other tertiary education institutions. He is currently an Associate Professor with the Department of History and Archaeology, University of Patras. He participated as a researcher in many EU and national-funded projects. He is the author of more than 75 papers in well-known international journals, such as *Theoretical Computer Science*, *ACM Transactions on Computer Systems*, *Journal of Parallel and Distributed Computing*, *Journal of Educational Computing Research*, *Multimedia Tools and Applications*, and *Heritage*, and conferences, such as Euromed, IMET, SPAA, HiPC, and DISC. His research interests include algorithms, trustworthy multimedia networks, and cultural technology. He has served as a member of the program committee in various conferences and as a reviewer in many journals. He is an Editor of *Handbook for Applications in Cultural Heritage Environments*.

CHRISTOS A. FIDAS received the Diploma degree in electrical and computer engineering and the Ph.D. degree in information systems from the University of Patras, Greece. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of Patras, and an Expert Member of the International Federation for Information Processing (IFIP) on the TC-13 Human-Computer Interaction Group. He has an extensive publication record in reputable scientific journals and conferences, and has been granted a patent. His research interests include information systems, emphasizing human–computer interaction, usable-security, and cultural heritage. He has received several scientific awards and research grants for his contributions to the field. For more information, visit the link (http://cfidas.info).