IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 19, 2020, accepted November 9, 2020, date of publication November 16, 2020, date of current version November 25, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3037995

A Dynamic Filtering DF-RNN Deep-Learning-Based Approach for EEG-Based Neurological Disorders Diagnosis

GHAITH BOUALLEGUE[1], RIDHA DJEMAL[2], SALEH A. ALSHEBEILI[3], AND HESHAM ALDHALAAN[4]

[1]University of Sousse, ENISo of Sousse, Sousse 4002, Tunisia
[2]University of Sousse, ISSAT of Sousse, Sousse 4002, Tunisia
[3]Department of Electrical Engineering, King Saud University, Riyadh 11421, Saudi Arabia
[4]Center For Autism Research: CFAR Riyadh, King Faisal Specialist Hospital and Research Center, Riyadh 12713, Saudi Arabia

Corresponding author: Saleh A. Alshebeili (dsaleh@ksu.edu.sa)

This work was supported by the Researchers Supporting Project number (RSP-2020/46), King Saud University, Riyadh, Saudi Arabia.

ABSTRACT Filtering of unwanted signals has a great impact on the performance of EEG signal processing applied to neurological disorders diagnosis. It is so difficult to remove undesirable noises using static filtering approaches as the performance of such techniques is strongly relying on specific EEG signal sub-bands, whose locations differ from one subject to another. In this paper, we present a novel dynamic filtering approach, which makes use of Finite and Infinite Impulse Response (FIR and IIR) filters along with a Recurrent Neural Networks using a Gated-Recurrent Unit (RNN-GRU), to identify and preprocess the most informative sub-bands pertaining to a particular neurological disorder. This combination of RNN with GRU requires more hidden layers than for conventional NN structures, and therefore offers much higher capacity to learn fitting and extract features from highly complex EEG data recording to afford better harmonization of the diagnosis process. Followed by an Independent Component Analysis (ICA) algorithm, all extracted features become independent to facilitate classification of clinical disorders using Convolutional Neural Network (CNN). The proposed diagnosis system achieves an average of 100% classification accuracy for epilepsy according to an offline diagnosis process using Bonn and MIT datasets, and when the same system is applied to autism provides an average accuracy of 99.5% using KAU dataset. The presented dynamic deep-learning approach applied to EEG classification pipeline, which includes artifact removal, feature extraction and classification, leads to significant improvements in the accuracy of the diagnosis classification regarding the targeted neurological pathologies.

INDEX TERMS Electroencephalography (EEG), deep learning, recurrent neural networks (RNN), convolutional neural network (CNN), dynamic filtering.

I. INTRODUCTION

Brain-computer interface is a new technology that could provide a wide variety of useful functions related to control, communication, and medical applications. Indeed, Electroencephalography (EEG) signal can measure electrical activities generated by large numbers of neurons, and is widely used for the diagnosis of neural disorders such as epilepsy, autism spectrum disorder (ASD), Alzheimer, and other complex disorders. Epilepsy and autism affect, respectively, nearly 2% [1] and 0.7% [2] percent of all people; both disorders share a common characteristic in that they can simply be diagnosed using EEG signals. Even if these neurological pathologies are quite different, it seems there are some particular links between epilepsy and autism for two main reasons: First, both of them mostly affect the same rhythms; namely, alpha and beta. Second, the epilepsy affects nearly 30 percent of all people with autism, according to the American Epilepsy Society. Note that the proposed diagnosis process developed for autism and epilepsy can deal with other relevant brain activity disorders like Alzheimer and sleep disorders based on the capabilities of deep learning analysis. Currently, the diagnosis of such disorders is mainly carried out manually by a limited number of neurologists and medical experts.

The associate editor coordinating the review of this manuscript and approving it for publication was Venkata Rajesh Pamula.

In some cases, the neurologists need several hours to conduct a complete diagnosis process and get a final decision for a single patient especially when the medical data are heterogeneous in certain cases. All these facts highlight the need to build a smart and a reliable diagnosis system that would provide important assistance in diagnosing and analyzing such disorders in order to reduce the risk of errors and to improve medical data analysis. The diagnosis system has a great potential in helping neurologists and medical experts during the diagnosis process, notably in terms of time-saving and increasing the diagnosis accuracy.

Despite its low spatial resolution, EEG imaging has several advantages over other imaging techniques in that it is simple to implement, of high temporal resolution, and of wide availability for use without the need for a great neurological expertise. These advantages make EEG a powerful diagnostic tool. However, the EEG recording is rather complex, nonlinear, non-stationary, imbalanced and buried with artifacts and interferences which bias results and related studies. Furthermore, some artifacts might also imitate cognitive impairments or other neurological pathologies and therefore biases the visual interpretation conducted during clinical diagnosis as in sleep or autism disorders [3], etc. So, this signal requires sophisticated signal processing to clean it up from all possible unpredictable contaminations or interferences before any further analysis.

The typical pipeline EEG-based diagnosis system includes EEG signal pre-processing, features extraction, and decoding. The preprocessing block is built to remove unwanted EEG signals. The most significant features are extracted from the filtered EEG signal, and passed to the decoding or classification stage. Once the EEG signal is acquired, the existing automated diagnosis approaches make use of machine learning to implement a couple of signal processing techniques, whose performance is highly application and patient dependent.

Many automated diagnosis approaches, relying on studies pertaining to EEG abnormalities due to neurological disorders, have been reported in literature in order to provide an efficient and early diagnosis of such disorders with a better harmonization and less contradictory diagnosis. Related approaches include techniques based on Shannon entropy [4], spectral entropy [5], multi-scale entropy [6], empirical mode decomposition (EMD) [7], and second-order difference plot (SODP) modeling [8], [9]. Band power (BP) [10], fast Fourier transform (FTT) [11], and discrete wavelet transform (DWT) techniques [12], [13] as well as common spatial patterns (CSP) have been all tried for EEG features extraction [14].

Independent component analysis (ICA) has been also widely used for brain disorder features extraction. Thus, ICA allows to maximize the variance of the projected signal from one class while minimizing it for the other one by applying dedicated spatial filters. Therefore, the selection of an appropriate preprocessing band-pass filter, which captures most of the power (variance) of changes resulting from neurological disorders, is essential to achieve great improvements in the performance of features extraction. However, the selection of such a preprocessing band-pass filter is not a trivial task, because the band of appropriate filter is subject and channel specific, and hardly be identified by visual inspection manner. In addition, a poor selection of the filter band may result in a decrease in the accuracy of brain disorders diagnosis. Although a filter with wide band (i.e., 8-30 Hz) [15] is usually adopted for ICA in brain disorder diagnosis classification, an increasing number of studies suggests that optimizing the filter band could significantly improve the classification accuracy [16].

In this paper, the deep learning diagnosis process is applied to both autism and epilepsy. In fact, according to the American epilepsy Society, epilepsy affects nearly 30 percent of all people with autism spectrum disorder. However, there is no certainty on the existence of similarity between subjects with autism and those with epilepsy. In this context, we have addressed these two different neurological disorders diagnosis problems to measure the capability of our proposed deep learning approach to self-adapt to different neurological pathologies while keeping a high classification accuracy. Three alternatives have been explored to fix the appropriate filter band selection accordingly The first approach optimizes the filter band concurrently while the features extraction stage is being executed using; e.g., the ICA, CSP, or BP techniques. The second approach is based on the concept of extracting features from multiple frequency bands, while the third approach considers the implementation of adaptive or dynamic filtering techniques, thus making it possible to choose the appropriate filter band for each user and/or each channel [17]. This, however, requires the use of filter design techniques to customize the filter type and order during the training phase.

On the other hand, multi-neurological disorders are based on different biomedical EEG data usually imbalanced and nonstationary integrating certain atypical diseases, which requires high complexity characterization. In this context, the deep learning model represents a tremendous opportunity to support medical authorities to exploit and significantly improve medical data analysis and to reduce the risk of medical errors. The proposed model integrates a significant number of hidden layers, which is trained using large sets of labelled EEG data to allow learning features directly without the need for manual extraction. Furthermore, applied deep learning approach provides a better harmonization of the diagnosis and prognosis protocols [18], [19]. In this context, the recurrent neural networks (RNN) play an important role. These networks can be realized using long short-term memory (LSTM) cells. However, these cells suffer from the vanishing gradient problem. Note that this problem can be prevented using the recently developed gated recurrent units (GRU) with easier and time-saving implementation; thanks to the less complex GRU structure compared with the LSTM approach [20], which has encouraged us to harness its tools to create an enhanced signal processing chain.

A significant contribution of the present work is in the preprocessing step, where basic FIR/IIR and state-of-the-art techniques are attempted to present a comprehensive filter design. A set of six filters are customized and applied at the epoch level of the EEG data belonging to each channel to generate the filtered signal, which is used to train an RNN with GRUs. Once trained, our proposed intelligent network (RNN) is able to apply the appropriate filter with a dedicated order leading to a minimization of its MSE at the epoch level. Thus, GRU-RNN combined with FIR and IIR filter design represent our dynamic filtering approach to help maximizing the classification accuracy. Furthermore, we have combined this filtering approach with the ICA and convolutional neural network (CNN) to maximize the accuracy of the decoding operation.

This paper is organized as follows. Section 2 provides a brief description for the work related to our research topic. Section 3 introduces the EEG datasets used in this study, and details the computerized diagnosis for both epilepsy and autism. Section 4 presents and discusses the findings of the proposed algorithm.

II. RELATED WORK

Feature extraction and classification techniques for EEG signal analysis are well established and provide very good results, while the preprocessing techniques still remain an open issue requiring more attention because the final accuracy of any classification analysis depends deeply on the applied filtering techniques. Many signal processing techniques have been widely used for pre-processing purposes based on different approaches, as described below:

- Applying regression methods which represent traditional techniques for removing physiological artifacts, like ocular and muscle artifacts. These approaches are based on subtracting estimated artifacts from the acquired signal. In fact, the average artifacts subtraction based techniques require high sampling frequencies and are just capable of eliminating only repetitive artifact patterns. Such subtraction can be made in both time or frequency domain [3].
- Using time and frequency analysis methods. Usually, the FFT is applicable to the frequency analysis of stationary signals where wavelets like DWT provide a flexible time-frequency grid to analyze signals whose spectral contents changes over time [21]. Indeed, the FFT is the fast way to compute the magnitude and phase decomposition applied to each channel of EEG recording. Thus, for neurological disorder diagnosis, the frequency components belonging to alpha and beta rhythms are located during the pre-processing phase, and relevant components are then reconstructed using the inverse FFT [22]. Unlike the Fourier transform components which are exclusively localized in frequency, wavelet transforms provide a trade-off in time-frequency localization due to the better tunable time-frequency features especially for non-stationary EEG signals.
- Applying static filtering where the same filter belonging to FIR or IIR family is applied for all subjects. Different EEG signal preprocessing techniques using FIR filters such as Equiripple and Kaiser-window can be applied to find an optimal fit between the desired and the available frequency responses. On the other hand, elliptic, Butterworth, and Chebyshev filters constitute the well known classical IIR filters, each of which can be optimal for a particular use [17].
- Using unsupervised learning techniques where no prior information about EEG recording and extra reference channels are required. For example, the ICA represents a typical processing algorithm applied to multichannel EEG signal to decompose the original recording into multiple independent source components. As explained in [23], this technique extracts information from electrodes with respect to eyes by estimating the interference of the Electrooculography (EoG) using the recursive least squares (RLS) algorithm. Unfortunately, some significant non-ocular frequencies can also be removed and the global classification accuracy is strongly affected. Similar adaptive filtering techniques have been used in [24], where the best sub-band is fixed via an appropriate objective function belonging to CSP technique. Some other works reported in [25], [26] used multi-scale entropy function or multi scale ranked organizing maps to eliminate noise from each feature vector, but this approach can be optimal for some users but not for the remaining ones.

The design of efficient EEG-based filtering technique still remains an open issue. Many other filtering techniques have been recently published. For example in [27], the EEG filtering is applied with feature extraction by combining a discriminative feature extractor with new strategy to filter out unwanted features from EEG signals. This approach is based on filtering out signals related to one property of the EEG signal while retaining the remaining features. Another filtering alternative, as presented in [28], is developed at the classification level, where a combination of a discriminative feature extractor with new strategy to filter out unwanted features from EEG signals have been proposed. Thus, a channel selection algorithms is applied that provides a possibility to work with fewer channels to increase the system performance by removing the noisy channels. This selection is done at the classification level which is called embedded technique. Other heuristic techniques based on genetic algorithms can also be applied for selecting relevant channels and eliminate the other ones and their performances depend on the applied classification techniques. In [29], authors proposed the wavelet decomposition to get sub-bands (SBs) of EEG signals. Subsequently, fuzzy entropy, logarithmic of the squared norm, and fractal dimension are computed for each SB for the purpose of classification. Since the validation is done using a single-channel EEG dataset, the accuracy is relatively limited which is a round of 79%.

TABLE 1. EEG-based diagnosis architectures for Epilepsy (ANN: Artificial Neural Network, PNN: Probabilistic neural network, KNN: k-nearest neighbors algorithm, MLPNN: Multilayer perceptron neural network).

Author	Feature Extraction	Classifier	Accuracy in %
Ocak [30]	Apen on DWT	ANN	96
Bao et al [31]	FFT, FD and others	PNN	94.1
Guo et al [32]	Relative wavelet energy	ANN	95.2
Guo et al [33]	Genetic programming based	KNN	99
Wang et al [34]	Wavelet packet entropy	KNN	100
Orham et al [35]	DWT	ANN	100
Djemili et al [36]	EMD	MLPNN	100
Khalil et al [37]	DWT, Shannon entropy	ANN	100
Ibrahim et al [38]	DWT, Shannon entropy	KNN	100

TABLE 2. EEG-based diagnosis architectures for autism (FD: Fractal Dimension, mMSE: modified Multi-Scale Entropy, VG: Visibility graph).

Author	Feature Extraction	Classifier	Dataset	Acc (%)
Sheikhani et al [39]	STFT	NN	Own dataset	82.4
Ahmadlou et al [40]	wavelet, FD	RBNN	Iranian	90
Bols et al [41]	mMSE	SVM	USA	70-100
Sheikhani et al [42]	STFT	NN	Iranian	96.4
Ahmadlou et al [43]	wavelet, VG	EPNN	Iranian	95.5
Ahmadlou et al [44]	wavelet, fuzzy SL	EPNN	Iranian	95.5
Alhaddad et al [12]	FFT	FLDA	Saudi	90
Alsaggaf et al [45]	FFT	FLDA	Saudi	80.3
Djemal et al [13]	DWT-entropy	ANN	Saudi	99.7

Note that even if the aforementioned techniques are suitable for particular subjects and for specific datasets, they may not provide the same performance when applied to subjects belonging to other datasets. To address this issue, an exploration of filter design is proposed to find the appropriate filter candidate based on the Signal-to-noise ratio (SNR) minima between the original and the filtered EEG signal for neurological diagnosis applications. We have considered dynamic filtering, capable of adjusting their parameters in order to minimize the mean squared error (MSE) of the filtered EEG epochs and their raw EEG signals, along the line of a recurrent neural network to auto-customize the filtering technique for each epoch of data and for the purpose of improving the classification accuracy.

After applying the training of the RNN to select the suitable filter with the optimized order for the current epoch, we evaluate the effectiveness of our proposed dynamic filtering approach on two neurological disorders, epilepsy and autism. The automated EEG signal processing allowing to detect epileptic form discharges would be of great value in epilepsy diagnosis. The system should be accurate for a wide range of epileptic patients having different types of seizures (focal, generalized, etc.) or having different seizure severity levels (mild, moderate, and high). Although several promising EEG-based automatic epilepsy diagnosis methods have been reported in literature; see Table 1, further research efforts are still needed so that automated diagnosis methods become well-suited for clinical use.

For the autism diagnosis, different EEG-based processing techniques have been investigated and evaluated using different specific datasets. Figure 1 presents a typical autistic EEG recording with patterns quite similar to a healthy subject, which requires more sophisticated pre-processing techniques to achieve high classification rate.

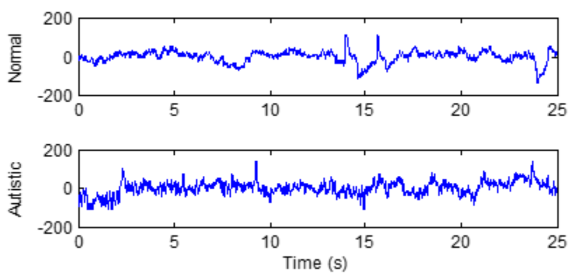


FIGURE 1. EEG recording samples for normal and autistic subject.

Studies relevant to autism diagnosis are presented in Table 2. Authors in [40] investigated wavelets combined with fractal dimension (FD) to estimate the complexity and any dynamical changes in autism recording. This approach is evaluated using KAU dataset of two groups of patients: 9 ASD and 8 non-ASD children acquired within eyes-closed condition. A classification accuracy of 90% is obtained using a radial basis neural network (RBNN) classifier. An optimized version of ASD diagnosis has been presented using visibility graph (VG) and fuzzy synchronization likelihood (Fuzzy SL) method [44]. Both of the aforementioned techniques provided a classification accuracy close to 95.5%.

The implementation of MSE as a basic feature followed by different classifiers such as KNN, SVM and Naive Bayesian (NB) to discriminate between a normal and a high risk autistic signal has been presented in [41]. This signal processing combination intends to optimize the classification accuracy based on specific dataset extracted by Net Station, and reached 80% of accuracy conducted on a group of patients at the age of 9 months. Similar works presented in [12] used a local dataset recorded by 16 channels, where the most significant artifacts are identified and removed by a simple visual inspection. Time and frequency feature extraction analysis have been proposed (raw data and FFT) before feeding these feature vectors to the Fisher linear discriminant analysis (FLDA) for discrimination purpose between normal and autistic states. Thus, the evaluation of the autism diagnosis measurement provided an accuracy rate of 90%. However, another application of these same techniques, reported in [45], provided lower performances although they used the same dataset. This performance degradation could be attributed to bad artifacts removal during the preprocessing phase. Recent deep learning diagnosis techniques using DWT-entropy-ANN based method provided much higher accuracy around 99% [13].

According to Lotte *et al.* [23], deep learning networks are less effective for EEG signals classification in BCI, given the limited training data available. However, shallow convolutional neural networks are more promising, which present our case. Moreover, current works have lack of designing a dynamic multitask algorithm that is adaptable to multiple diseases (ASD/ Epilepsy) [46].

As for the preprocessing and feature extraction tools, the recent findings in BCI technologies have developed massive amounts of brain data featuring high dimensionality, multiple modalities (e.g. physical modes such as frequency or time, multiple brain imaging techniques or conditions), and multiple couplings as functional connectivity data. On the other hand, tensors offer promising BCI tools such as for the analysis tasks and for the fusion of massive data using a mathematical back-end for hidden complex pattern recognition, thanks to its multi-way nature [46].

However, these methods present a complexity to vector machine learning and standard matrix methods [46]. Moreover, since band power and time point features methods are the most commonly used tools for feature extraction [47], they can be effective for a signal with a higher signal-to-noise ratio. Nonetheless, they do not require measuring the correlation or synchronization between signals from different sensors and/or frequency bands which can result in higher classification accuracies.

III. METHODS

As mentioned above, our diagnosis system deals with epilepsy and autism disorders. For epilepsy disorder, we have used the two available datasets, Bonn and MIT, to evaluate the effectiveness of our proposed system. On the other hand, for autism, we have evaluated our signal processing chain using the dataset provided by King Abdulaziz University (KAU) lab belonging to King Abdulaziz Hospital in Jeddah, Kingdom of Saudi Arabia (KSA) the dataset is publicly available and can be found in [48].

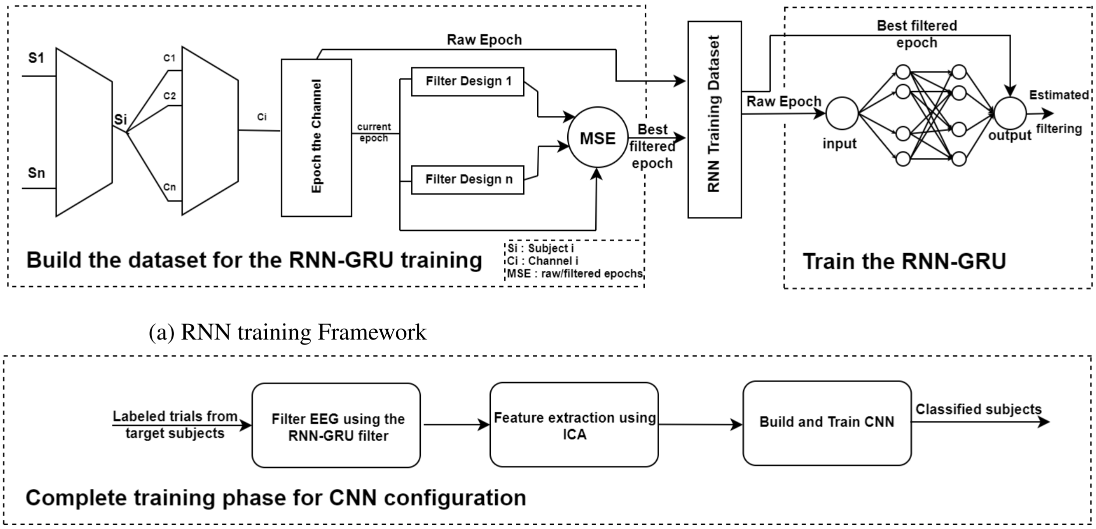
A. DATA DESCRIPTION

The autism dataset used in our work was recorded by the KAU lab with the assistance of King Abdulaziz Hospital neurological service. To preserve the patient's medical confidentiality, all personnel information are omitted from the EEG recording which is conducted during relaxing state. There are two groups of subjects which consist of ten healthy volunteers and nine autistic subjects (six males and three females) who are aged from 6 to 16 years. The g.tec EEG cap integrating high resolution Ag/AgCl electrodes and USB amplifiers operating using BCI2000 software is used to record EEG signals by means of all 16 available channels. According to dataset description presented in [12], the relaxing state corresponds to a weak state and all channels are referenced according to the 10-20 international acquisition system.

The dataset is cleaned by the acquisition system using band pass filter with a range of frequencies from 0.1 to 60Hz. Furthermore, a dedicated notch filter with stop band frequency of 60Hz is integrated in the acquisition system where all EEG signals were digitized with a sampling frequency of 256Hz. The registration time for normal subjects is between 12 and 30 minutes with a total duration of about 173 minutes, while for autistic subjects, the acquisition time is between 5 and 27 minutes with a total duration of 148 minutes. During the acquisition process, it is of an utmost importance that the patient be relaxed, therefore, volunteers were asked to sit comfortably, to stop talking and to refrain from any muscular activities.

As of the epilepsy, we have considered Bonn and MIT datasets, as described below:

- CHB-MIT EEG database: This EEG recording includes 906 hours of EEG data for 23 epileptic subjects having a total of 163 seizure events [49]. All patients are aged under 18 years and undergoing medication withdrawal for epilepsy evaluation. As in the autism case, the EEG recording was sampled at 256Hz in the Boston Hospital using 18-channel according to 10-20 standard bipolar montage.
- University of Bonn EEG dataset: This dataset contains EEG for normal and epileptic subjects with a sampling rate of 173.61Hz. These signals are obtained at the Epilepsy Center at the University of Bonn, Germany [50]. Five subsets of EEG signals (A, B, C, D and E), which include healthy and epileptic EEG records have been included in this dataset. Subsets A and B of



(a) RNN training Framework

(b) CNN training framework

FIGURE 2. Proposed framework for EEG deep learning analysis.

EEG data have been acquired from five healthy volunteers, with open and closed eyes, respectively.

B. DIAGNOSIS SYSTEM FRAMEWORK

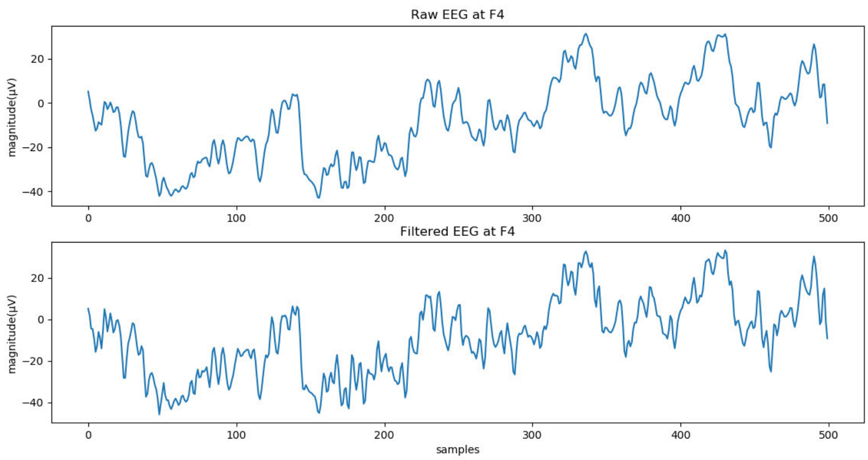
The proposed diagnosis system techniques stand on two main complementary steps, as described below:

- The first step consists of segmenting the EEG data for each channel belonging to each subject, with a set of epochs each of which is of 50 samples duration, where the choice of this size will be discussed in Section 4. Then, we apply to each epoch six types of predefined FIR and IIR filters to select the best filter with the appropriate order. The best filter is the one which provides the minimum MSE computed between the original epoch and filtered signal. Next, we design different filters; 6 types (Equiripple and Kaiser window FIR filters and Butterworth, Chebyshev 1 and 2 and Elliptic IIR filters), where the order of each can be tailored with different number of ripples and cut-off frequency bands according to minimisation of the MSE for each epoch. Once we complete the filtering process for all channels and for all subjects, we use these results to build a new dataset consisting of pairs of the original and best filtered epochs. These data are then fed to an RNN for training purposes, where the networks inputs are the raw epochs and the networks targets are the best filtered epochs. The trained RNN includes 3 hidden layers, which is capable to provide efficient EEG filtering without any prior information about the signal itself. Figure 2.a depicts the schematic diagram of the training step, where Si (i=1,2,.., n) denotes the ith subject, and Ci (i=1,2, …, m) denotes the ith channel.
- The second step is dedicated for the training of the CNN classifier. Specifically, the outputs of trained RNN are fed to the ICA to get independent features. The CNN is trained to classify a given epoch whether it belongs to a normal state or a seizure state in MIT dataset, and whether it belongs to a healthy subject or a non-healthy subject in Bonn and KAU datasets. The inputs to the CNN are the ICA extracted features and its targets are the labels, 0s for normal states (healthy subjects) and 1s for seizure states (non-healthy subjects). Figure 2.b depicts the flow for obtaining a complete trained EEG-based diagnosis system.

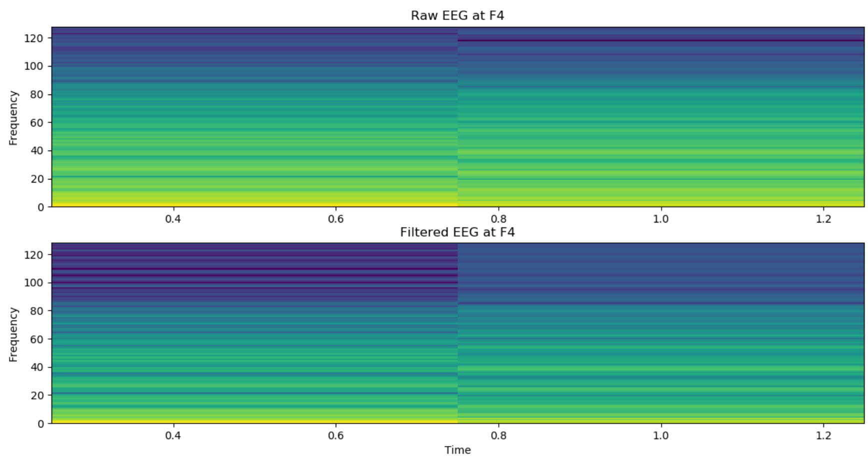
C. FILTER DESIGN

In this study, we have considered two sets of FIR and IIR filters represented by: Equiripple, Kaiser-window for FIR category and Butterworth, Chebyshev I and II, and Elliptic belonging to IIR filters. Although all these techniques serve the same purpose, there are many differences in functionalities and performances. Indeed, the IIR filters are quite efficient since they can provide similar amplitude responses with fewer coefficients, or lower side lobes for the same number of coefficients, while providing a lower delay than that obtained by the FIR filters. Butterworth, Chebyshev type I, type II, and elliptic filters are optimal for use in a specific application context. For example, Butterworth provides the best representation of an ideal band-pass filter response, while elliptic filters allow to get equal ripples in both the pass-band and stop-band filter regions. For Chebyshev filters, the frequency response has pass-band ripples in type I and stop-band ripples in type I.

On the other hand, FIR filters are the immediate choice when linear phase is a requirement, and they are inherently stable. FIR filters, like the Equiripple (FFe) and Kaiserwindow (FFk) filters, are also helpful to achieve



(a) EEG channel (F4) temporal behaviour



(b) EEG channel (F4) frequency behaviour

FIGURE 3. **Raw & filtered EEG using our proposed approach.**

fractional constant delays. Designing of such filters can be accomplished, as described below:

- Equiripple Filter: The design is based on Parks-McClellan algorithm, which utilizes the Remez exchange algorithm. This technique minimizes the maximum weighted error between the desired response and the actual filter response for a given order. Because the resulting filter will exhibit an equiripple behavior in its pass-band and stop-band frequency responses, it is sometimes called equiripple filter. The ripple is controlled in the design process by the weighting function and the filter order.
- Kaiserwindow Filter: In this filter, the window length mainly affects the transition band, while the coefficients roll off (the shape factor)controls the pass-band and stop-band ripple sizes. Therefore, by keeping the window length constant, we can adjust the shape factor to have the appropriate level of pass-band and stop-band ripples.

The aforementioned filters have been designed so that the pass band region lies between (8.0-30.0 Hz) to cover exclusively α-band and β-band of the EEG frequency spectrum, which has proven to contain ASD patterns [13], [45]. The filter order is ranging from 3 to 33, a varying number of ripples and several cut-off frequencies giving, hundreds of proposed filtering approaches for each 50 samples EEG epoch. Note that these filters will only be used during the training phase. During the dynamic filtering, we preserve the most significant sub-bands, alpha and beta by specifying the pass-band frequencies accordingly. Figure 3 shows that the filtered signal is almost close to the non filtered one where the MSE is about $1.13\mu V^2$. Moreover, the visual inspection of the frequency spectrum of both raw and filtered signals using our RNN filtering technique proves that it preserves alpha and beta bands frequencies.

D. RECURRENT NEURAL NETWORK DESIGN

It is the intention here to design a robust (machine-learning) network that performs filtering, which focuses exclusively on the informative frequencies of EEG signals. Indeed, this can be achieved by utilizing an RNN. Once perfectly trained, this RNN will be capable of filtering any EEG signal without prior knowledge of its artifacts. The training process is accomplished by considering the raw epochs as the network's inputs and the best filtered signals of FIR and IIR filters as the network's targets.

RNNs exhibit a highly non-linear dynamic mapping, which can be promoted in a number of interesting applications, including spatiotemporal pattern classification, control, optimization, forecasting. The RNNs architectures integrate learning algorithms that can deal with time-varying input and/or output in non-trivial ways. Initially, all weight values are chosen randomly and are optimized during the stage of training.

To build the RNN, a training database has been created that includes raw and best filtered EEG signal epochs, each is with a duration of 50 samples. All these original and best filtered epochs are fed to the RNN for training purposes.

RNNs have an internal state (memory), which allows to exhibit dynamic temporal behavior for time sequence series. Considering an input x_t, an output y_t and a hidden state h_t, RNN's basic system is defined by:

$$\begin{aligned} h_t &= F(h_{t-1}, x_t) \\ y_t &= G(h_t) \end{aligned} \tag{1}$$

where F represents the state transition function, and G is the output function. RNN offers different variants, such as the Gated-Recurrent Unit (GRU) that solves the "vanishing" or "exploding" gradient problems, which commonly occur during the training of RNNs. Note that RNN-GRU has been successfully applied to EEG-based seizure detection, Talathi [51]; therefore, we have used it here to design a robust (machine-learning) filter. The discrete dynamical equations governing the operation of RNN-GRU:

$$\begin{aligned} \tilde{c_t} &= tanh(W^T(r_t \odot c_{t-1}) + U^T x_t) \\ z_t &= \sigma(W_z^T c_{t-1} + U_z^T x_t) \\ c_t &= (1 - i_t) \odot c_{t-1} + i_t \odot \tilde{c_t} \end{aligned} \tag{2}$$

With W, U and V are the transition, input and output matrices respectively. $z = \{i, f, r\}$ is representing the gating functions: input gate, the forget gate and the internal gate, $\odot$ is the Hadamard product and σ is the Sigmoid function. The trainable model parameters are $\{W, W_z, U, U_z\}$.

From the above mentioned training database, we have created the RNN-GRU, which will be fed by a set of inputs (raw EEG slice) and targets (Best filtered sample). This block will reconstruct an optimally filtered EEG that will be post-processed to extract features.

The RNN architecture stands as shown in Table 3, where the hyperparameters were set as follows: a fixed learning rate of 10^{-3}, MSE as the loss function and, a considerably low batch size of 10. As for the optimizer, "Adam Optimizer" has been chosen as being one of the best used optimizers in the machine learning field [52].

TABLE 3. Recurrent Neural Network Model Architecture (GRU: Gated Recurrent Unit hidden layer; FC: Fully Connected hidden layer).

Layer type	output shape	activation function
input layer	(10,50,1)	——
GRU1	(10,100,1)	Relu
GRU2	(10,100,1)	Relu
FC	(10,50)	linear

Figure 3 shows the raw EEG signal of F4 channel and its corresponding filtered signal at the output of the trained RNN. The results are obtained by using the KAU dataset. Among the 19 subjects in this dataset, we have generated a new dataset which consists of 269,500 pairs of signals of raw epochs and their corresponding optimum filtered versions obtained at the outputs of FIR and IIR filters. The 10-fold cross-validation strategy has been used, where the dataset is randomly divided into 10 equal parts (10 subsets), nine of them were used for training and one for testing. Specifically, we have extracted 111 100 of 50-sample segments from MIT dataset, 53 900 of 50-sample segments from Bonn dataset, and 104 500 of 50-sample segments from KAU dataset.

E. FEATURE EXTRACTION

This is the second step in our proposed signal processing chain, which implements ICA for features extraction. ICA is presented as an extension of the principal component analysis (PCA) technique [53]. Yet, PCA enhances the covariance matrix of the data which portray second-order statistics, whilst ICA optimizes higher-order statistics such as kurtosis. Thus, PCA finds uncorrelated components while ICA finds independent components [53], [54]. Therefore, PCA can obtain independent sources when the higher-order correlations of mixture data are small or insignificant [54]. ICA has multiple algorithms such as FastICA [55], projection pursuit [54], and Infomax [54]. The main purpose of these methods is to extract independent components using

1) the maximization of the non-Gaussianity.
2) the minimization of the mutual information.
3) the implementation of the maximum likelihood (ML) estimation method [56].

For our case, we have considered FastICA as the algorithm to extract the independent components. FastICA algorithm extracts independent components by maximizing the non-Gaussianity by maximizing the negentropy for the obtained signals using a fixed-point iteration algorithm [55]. FastICA has a cubic or at least quadratic convergence speed, thus, it is much faster than gradient-based algorithms that have linear convergence. Moreover, FastICA has no learning

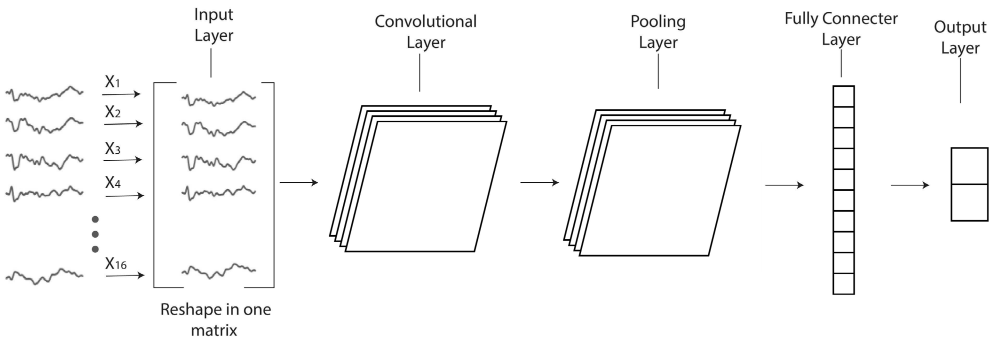


FIGURE 4. The proposed CNN architecture.

rate or other flexible parameters which makes it simple to use [57].

The ICA accepts at its inputs the already filtered signals of RNN. Specifically, it processes 16 filtered channels to generate 16 features sequences. The signal of each input channel is of 50 samples duration and the corresponding output of ICA is also of 50 samples duration. The ICA will not only help to extract features but also will force this set of features to be totally independent, giving more information to the classifier. Table 4 compares the performance of our RNN-GRU to other similar neural networks such as simple RNN, and LSTM models. We have conducted the following experiment through which we have tested other recurrent neurons in the RNN architecture along with different CNN models by modifying the number of layers in the CNN architecture. The table below shows that only LSTM and GRU cells can achieve accuracies higher than 90%, which eliminates the consideration of simple RNN for the problem at hand. This can be explained since GRU and LSTM use different ways of gating information to prevent vanishing gradient problem. As for the classification models, on the other hand, the CNN showed decaying performance by increasing the number of layers except for the case of simple RNNs.

TABLE 4. Performance evaluation of different RNN and CNN models.

	Simple RNN	LSTM	GRU
1 CNN layer	87.51%	95.47%	99.50%
2 CNN layers	89.75%	94.28%	99.03%
3 CNN layers	90.22%	91.0%	96.81%

F. CLASSIFICATION TECHNIQUE

This final step in our proposed tool-chain system is implemented using the CNN. This network has been widely used in many fields, including image and video processing, biomedical engineering, wireless and optical communication systems. Motivated by the work of Schirrmeister *et al.* [58], where CNN was used to decode and visualize EEG patterns, this network is also used here for classification. The classification is of two classes. For autism, the CNN output layer is of 2 classes: Autistic and Healthy. For epilepsy, the CNN output layer is also of 2 classes: seizure and normal.

Figure 4 shows the main architecture of the CNN block. The input takes the form of a matrix of size 16×400, where 16 represents the number of independent signals created by the ICA, and 400 defines the number of samples of each signal.

The CNN contains 5 layers, the first one is the input layer that feeds the features matrix followed by a convolutional layer and a max-pooling. The remaining are 2 fully connected layers, where the first has a dropout of 0.8. The CNN architecture is detailed in Table 5. Hyperparameters were set as follows: a fixed learning rate of 10^{-3}, a batch size of 50, a stride of 1 for the convolutional and the max pooling layers, MSE as the loss function, and ''Adam optimizer'' as the optimizer component. Finally, from the 19 subjects, we have extracted 12,000 labeled matrices to train the CNN, and applied the 10-fold cross-validation technique as previously discussed.

TABLE 5. CNN Model Architecture (Conv2d: 2d convolutional layer; Fc: Fully Connected hidden layer).

Layer type	output shape	activation function
Input layer	(50,16,400,1)	—
Conv2d	(50,16,400,23)	Relu
Max-Pooling	(50,8,200,23)	—
FC	(50,120)	Relu
FC	(50,2)	Softmax

IV. RESULTS AND DISCUSSION

In this section, we will discuss several findings based on the obtained results after the implementation of the deep learning EEG signal diagnosis. The proposed EEG filter design (RNN-GRU) and the classifier (CNN) were created using Keras built on top of Tensorflow installed on python 3.6 running on a GPU Nvidia Geforce GTX 950M.

A. FILTER DESIGN EXPLORATION

To evaluate the performance of our proposed RNN-GRU dynamic filtering approach, we have built a first draft of

RNN classification architecture according to the scheme presented in Table 6 with four layer different activation functions. Then, we applied different well known preprocessing techniques providing adapative EEG filetring. The following five filtering techniques have been tested using the three available datasets dedicated for epilepsy and autism diagnosis.

TABLE 6. Recurrent Neural Network Model Architecture for the Comparison Task.

Layer type	output shape	activation function	Dropout
input layer	(5,500,1)	—	—
GRU1	(5,100,1)	Relu	0.2
GRU2	(5,100,1)	Relu	0.2
FC	(5,2)	Softmax	—

(a) A classification using a static filter: a 4$^{\text{th}}$ order bandpass Butterworth filter.
(b) An LDA-Mahalanobis distance classification based on adaptive filtering techniques reported by Belwafi *et al* [16] (2014).
(c) A classification based on adaptive filtering technique of Correa *et al* [59] (2019).
(d) A classification filtered with the cascaded filter based on OWA operator (COWA) of Pander.T [60] (2019).
(e) A classification based on our proposed RNN filtering.

For the CHB-MIT dataset, we created 2 labeled groups of EEG signals (Normal/Seizure) states, and for the Bonn and KAU datasets, the labels were (Normal/Epileptic)and (Normal/Autistic) subjects, respectively. Using the RNN-based classifier detailed in Table 6, the CAD network is trained to build a model for each dataset independently. Table 7 shows a significant improvement in classification accuracy of proposed diagnosis system architecture across all three datasets compared with all other implemented classification techniques. Since the neurological classification differs for each case, we have to validate the performance of our proposed filtering technique on different neurological disorders without using the ICA algorithm. Thus this proposed platform represents a first evaluation of the contribution of the proposed dynamic filtering techniques on the average accuracy improvement. We tracked every filter used in our work to show how it contributes to the overall dataset filtering. We concluded that our proposed method outperforms all the presented competing techniques in terms of accuracy for all available datasets and for both epilepsy and autism where the improvement rate is more significant autism. Also, we measured the percentage of the selection of each of six FIR and IIR filters. The contribution rate is the number of times a particular filter is selected for the preprocessing stage, during the training phase, divided by the total number of utilized epochs. As depicted in Table 8, no single filter type can present an optimal filtering for all data cases. Kaiser window filter, which has shown outstanding results in Belawfi's work [16], has also shown here great contribution to the filtering process but did not manage to present optimality in all filtering cases. Therefore, it is important that all filters are implemented particularly for large scale databases. Table 8 also shows specifications of employed filters.

In all experiments, we have segmented the data of each channel to epochs, each of which has a length of 50 samples (corresponding to 0.2 seconds in KAU and MIT datasets). The choice of this fixed length was a critical task, since it greatly affects the convergence of the RNN filter. Figure 5 shows the performance of both the RNN-filter loss and the CNN classifier accuracy with regard to the pre-selected epoch length. In order to evaluate the performance of the diagnosis process, some statistical parameters were calculated: true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). In addition, the metrics of sensitivity, specificity, and accuracy were also calculated according to the following equations:

$$\begin{aligned} Sensitivity &= \frac{TN}{FP+TN}100 \\ Specificity &= \frac{TP}{TP+FN}100 \\ Accuracy &= \frac{TP+TN}{TP+FP+TN+FN}100 \end{aligned} \tag{3}$$

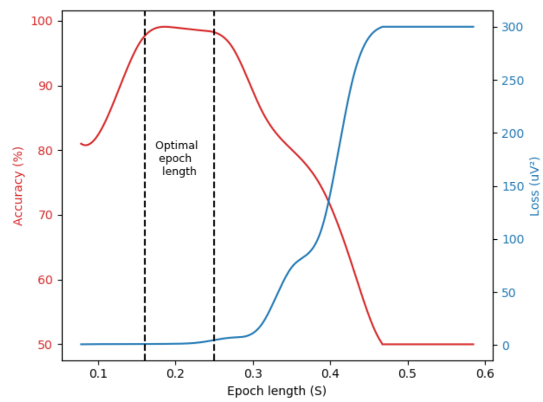


FIGURE 5. RNN Loss/CNN Accuracy performance along with the variation of the filtering window size.

These values were computed for all subjects belonging to the same dataset and for all our three used datasets which are listed in Table 9. According to this table, the results obtained show that the metrics values are quite similar for all data sets during the validation process of our diagnosis tool. Furthermore, the values of sensitivity and specificity are quite close. It is obvious from Figure 5 that as the epoch length increases, the classification accuracy increases but upto a certain limit, and then starts to decrease. The RNN-filter loss, on the other hand, maintains low value as long as the epoch length is less than 0.3 second, and then increases dramatically to a high value where it becomes constant. This is intuitively not surprising due to the (low/non)-convergence

TABLE 7. Filtering techniques and classification performances.

Filtering Technique	CHB-MIT dataset	Bonn dataset	KAU dataset
Static filter: Butterworth-4	96.86%	96.41%	79.56%
Belwafi's	98.20%	97.16%	82.22%
Adaptive Filtering	98.93%	98.31%	85.16%
COWA	99.50%	96.89%	85.74%
Porposed RNN-GRU filter	100.0%	97.26%	86.67%

TABLE 8. Filters contribution in EEG pre-processing and their specifications.

Filter type	Contribution rate (%)	Order	Passband Cutoff Freq. (Hz)	Stopband Cuttof Freq. (Hz)	Passband Atten-uation (dB)	Stopband Atten-uation (dB)
Kaiser window	92.0	3-33	[8-29.5]	[7.5-30]	0.1-0.2	30
Chebyshev 2	5.0	3-14	[8-29.5]	[7.5-30]	0.25	25
Chebyshev 1	1.0	3-14	[8-29.5]	[7.5-30]	0.25	25
elliptic	1.0	3-14	[8-29.5]	[7.5-30]	0.25	25
butterworth	0.5	3-14	[8-29.5]	[7.5-30]	0.25	25
equiripple	0.5	3-33	[8-29.5]	[7.5-30]	0.1-0.25	30

TABLE 9. Performance evaluation of the proposed design.

Dataset	Sensitivity	Specificity	Accuracy
Bonn	55.02	94.44	98.923
MIT	44.92	95.0	99.73
KAU	52.13	96.21	99.98

of the RNN internal weights as, in this case, the RNN is calling for more layers to go deeper. Fortunately, there is a region for the epoch length, where the classification accuracy is the maximum and the RNN-filter loss is the minimum. The choice of this fixed length was a critical task since it greatly affects the convergence of the RNN filter; therefore, an optimization is required to determine its value. The proposed size of 50 samples (0.2 seconds for KAU and MIT datasets) is the result of trade-off optimization between the filter design (which is kept of order less than 34) applied before the construction of our GRU-RNN and the convergence of RNN filter, which requires large number of data segments for training and testing. This optimization process is demonstrated in Figure 5.

Based on Figure 5, we notice that the optimal size of an epoch varies between 0.16 and 0.25. Therefore, we use 50 samples corresponding to 0.1953125 seconds for 256 Hz sampling rate, and is approximated by its rounded value of 0.2 seconds whenever needed for ease of presentation.

In this study, the RNN-GRU network is trained using a huge number of iterations(2200K). Figure 6 shows the convergence of the loss function toward the optimal weights with a final loss of 1.13 for the function categorical cross-entropy

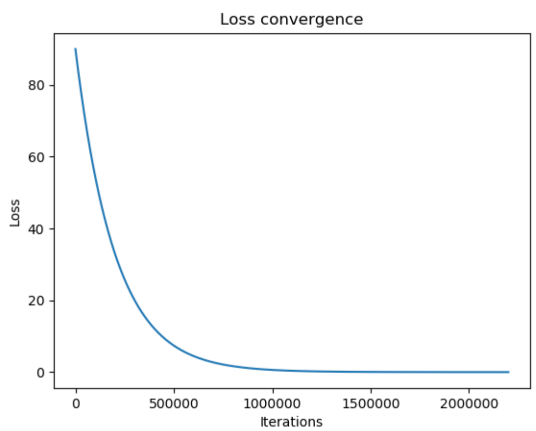


FIGURE 6. RNN-GRU loss/iteration performance.

which confirms the efficiency of our proposed machine learning filter.

B. DESIGN OPTIMIZATION

After conducting the dynamic filter design in the previous section, we turn to optimize the overall diagnosis system architecture. Our proposed filter has shown remarkable results with competing filtering techniques in the previous subsection. Therefore, the remaining task is to set our filter in the diagnosing architecture and analyze its performance in order to achieve high classification accuracy. Therefore, in this subsection, we fix our filtering technique based on RNU-GRU filter and study and the CNN network, preceded by the ICA module. Despite the fact that the size of the features has not been reduced, the application of the

TABLE 10. k-fold cross validation evaluation.

k	Training Accuracy (%)	Validation Accuracy (%)	Training loss	Validation loss
5	99.58	99.56	0.043	0.020
6	99.65	99.71	0.037	0.013
7	99.47	99.44	0.058	0.036
8	99.51	99.04	0.054	0.027
9	99.74	99.64	0.020	0.011
10	99.93	99.82	0.009	0.009

ICA has made it possible to achieve better classification using our designed CNN. Figure 7 shows the variation of accuracy against the number of epochs during the training and the validation phases when the proposed system architecture employed the KAU dataset. As noticed, the accuracy of both phases maintains the same behavior, and reaches to a high steady-state value in very small number of epochs. Thanks to the ICA, we reached 99.5% of accuracy in only 10 epochs because independence of features ensures our CNN not to fall in the ''Overfitting'' problem.

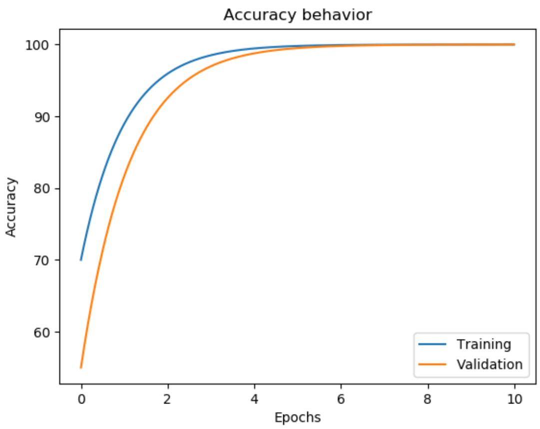


FIGURE 7. Training and validation accuracy behavior.

To investigate more the performance of our feature extraction technique, we have considered a comparison with different feature extraction tools such as Principal Component Analysis (PCA) and Common spatial pattern (CSP) along with the ICA using 10 subjects from the KAU dataset. Figure 8 presents the variation of accuracy for the mentioned selected feature extraction techniques. Thanks to its independent components, the ICA generally offers the highest accuracies among the other features extraction methods, which justifies its choice in the feature extraction block.

For further evaluation of robustness of our classifier, the k-fold cross validation is applied where k is varying from five to ten. Thus, we have used the KAU dataset and we fixed the number of epochs to 10. Table 10 shows the obtained results, where more than 99% accuracy has been achieved for both the training and validation sets. The loss is also barely reached 0.06 for the training sets and 0.04 for the validation

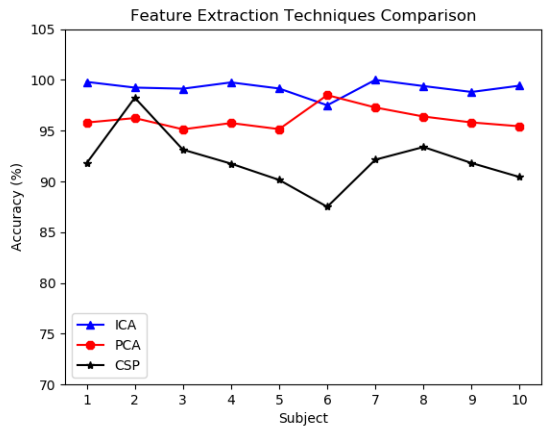


FIGURE 8. Performance of different feature extraction techniques.

sets. These results confirm the high efficiency and ensure the robustness of our classifier.

To prove the necessity of using such complex deep neural network classifier, we have considered simpler supervised and unsupervised classification algorithms such as Support Vector Machine(SVM), k-nearest neighbors algorithm (KNN), k-means clustering and CNN along with different feature extraction techniques. Table 11 shows the performance of these classifiers trained using the fixed dataset. By referring to Table 11, we observe that KNN can give competing results when CSP features are used. However, CNN provides better classification with ICA, which makes CNN/ICA the most adequate tool for our design.

TABLE 11. Classification performances of several classifiers using different feature extraction techniques.

	ICA	PCA	CSP
SVM	60.0%	54.11%	56.0%
KNN	60.20%	71.0%	87.44%
K-means	47.82%	51.70%	49.39%
CNN	99.50%	96.12%	92.40%

To show the efficiency of applying our dynamic filtering to the EEG recording for both epilepsy and autism, we propose to use the topographic maps representation of the power spectral density with their related electrodes. Furthermore, the electrode map is shown for the all different frequencies:

TABLE 12. Computational results of the proposed algorithm.

Operation	Number of parameters	Computational results in FLOPS
RNN Filter Design	96 550	1 512 395
ICA	6418	57 627
CNN Classification	52 874	792 841
Total Diagnosis Computation	155 842	2 362 863

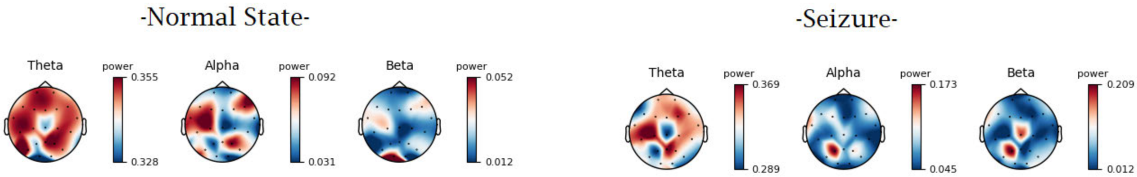


FIGURE 9. Power spectral density distribution on electrode map for normal and epileptic seizure states.

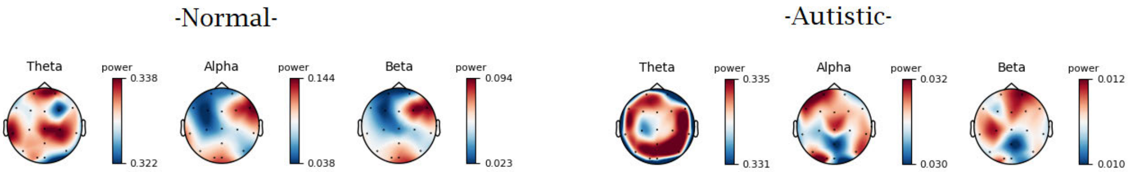


FIGURE 10. Power spectral density distribution on electrode map for normal and autistic states.

from delta to beta. Only the three main frequencies: theta (6Hz), alpha (10 Hz) and beta (22 Hz) are considered for this deep analysis.

For the epilepsy diagnosis process, we have applied the same power spectral density measurements related to the abovementioned rhythms. This topographic representation shows the distribution of each of the three sub bands along the electrode locations. As a first analysis of Figure 9, the maps of filtered EEG signals to the epileptic subject shows that the power spectral density is not localized in a specific location in the scalp but it is distributed in all regions.

It is relevant here to consider the complexity of our filters design combined with RNN-GRU, where we evaluate the computational cost for each phase through the diagnosis process. According to the obtained results presented in Table 12, the proposed dynamic filtering combined with the RNN-GRU handled an important number of network parameters during the training phase close to 96 550 and required a total computation cost of almost 1.5 mega flops per second (MFLOPS). The classification needs less parameters and also less computation amount with only 0.8 MFLOPS. According to these reasonable computational complexities, real time diagnosis process of the proposed algorithm could be implemented in advanced electronic systems.

As depicted in Figure 10, the log power spectral density shows a clear spike for normal subject around the alpha rhythm. However, this spike is passing away for autistic subjects. On the other hand, and according to the same figure, the power spectral density with respect to alpha and beta rhythms for normal subjects are quite similar and are localized in frontal and parietal regions of the head. However, for autistic subject this similarity is lost and the power spectral density for the same rhythms are localized in the occipital region. This assumption is also neatly verified for normal and autistic registration after applying dynamic filtering.

Table 13 presents a comparison of our proposed autism CAD system with other existing methods, whose details have been discussed in Section 2. It should be noted that most of proposed methods were validated by different datasets, which makes fair comparison for all methods slightly difficult. By virtue of Table 11, we observe that our method achieves an average training accuracy of 99.6 and an average classification accuracy of 99.5, which is greater than any other existing method that has used the same database. Also, our method stands competitive to other methods which were tested using different datasets. Moreover, the RNN-GRU dynamic filtering approach proposed in this paper is more universal than the filtering approaches proposed in other works in the sense that its design has been based on the utilization of combined data from the three datasets (MIT, Bonn, KAU). Therefore, our proposed filtering approach can preprocess the most informative sub-bands of EEG signals whether they belong to healthy, epileptic, or autism subjects.

TABLE 13. Several EEG-based CAD of autism spectrum disorder.

Author	Feature extraction	Classifier	Dataset	Acc (%)
Alhaddad et al. [12]	FFT	FLDA	KSA	90
Djemal et al. [13]	DWT, Shannon entropy	ANN	KAU	99.7
Grossi et al. [26]	squashing phase, noise elimination phase	MS-ROM	Villa Santa Maria Institute	100
Thapaliya et al. [61]	Mean, Standard Deviation, DFFT, Shannon entropy	SVM, DNN, Logistic and Naive Bayes	University of Miami	100
Bosl et al. [62]	recursive feature elimination algorithm	SVM	Boston Children's Hospital	100
Haputhanthri et al. [63]	DWT, Mean, Standard Deviation	logistic regression, SVM, Naïve Bayes, random forest	University of Miami	93.33
Our work	ICA	CNN	KAU	99.5

V. CONCLUSION

In this paper, we have proposed an EEG-based diagnosis system well suited for two neurological disorders, epilepsy and autism. The design of such a system requires a preprocessing stage, which is very critical for achieving reliable performance. It is a common practice in similar system design to use static (predefined) filters to remove artifacts and reduce the effect of noise. Unfortunately, these types of static filters may not perform well in all cases due to the intrinsic features of the EEG recording and its associated neurological disorder. In this paper, we have proposed a machine learning filter based on the RNN-GRU to be implemented in the preprocessing stage. It has been demonstrated in this work that this type of machine learning networks is so effective in filtering EEG signals, with a loss rate (MSE) as low as 1.13. In particular, when this filtering approach is combined with ICA for features extraction and CNN for classification, a high average accuracy rate can be achieved, which makes our proposed architecture competitive with other existing approaches. Thus, the proposed diagnosis system achieves an average of 100 % for epilepsy according to an offline diagnosis using Bonn and MIT datasets, where the same system applied to autism provides an average accuracy of 99.5% using KAU dataset.

REFERENCES

[1] I. B. Slimen, "EEG epileptic seizure detection and classification based on dual-tree complex wavelet transform and machine learning algorithms," *J. Biomed. Res.*, vol. 34, no. 3, p. 151, 2020, doi: 10.7555/JBR.34.20190026.

[2] C. Lord, T. S. Brugha, T. Charman, J. Cusack, G. Dumas, T. Frazier, E. J. H. Jones, R. M. Jones, A. Pickles, M. W. State, J. L. Taylor, and J. Veenstra-VanderWeele, "Autism spectrum disorder," *Nature Rev. Disease Primers*, vol. 6, no. 1, p. 5, Jan. 2020, doi: 10.1038/s41572-019-0138-4.

[3] 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.

[4] M. Tanveer, R. B. Pachori, and N. V. Angami, "Entropy based features in FAWT framework for automated detection of epileptic seizure EEG signals," in *Proc. IEEE Symp. Comput. Intell. (SSCI)*, Nov. 2018, pp. 1946–1952, doi: 10.1109/SSCI.2018.8628733.

[5] H. Helakari, J. Kananen, N. Huotari, L. Raitamaa, T. Tuovinen, V. Borchardt, A. Rasila, V. Raatikainen, T. Starck, T. Hautaniemi, T. Myllylä, O. Tervonen, S. Rytky, T. Keinänen, V. Korhonen, V. Kiviniemi, and H. Ansakorpi, "Spectral entropy indicates electrophysiological and hemodynamic changes in drug-resistant epilepsy–A multimodal MREG study," *NeuroImage: Clin.*, vol. 22, 2019, Art. no. 101763, doi: 10.1016/j.nicl.2019.101763.

[6] J. Kapica, M. Jolanta, A. Eugene, and K. Ziniuk, "Multiscale entropy analysis of EEG signals," *Current Problems Psychiatry*, vol. 13, no. 1, pp. 31–33, 2012.

[7] V. Bajaj and R. B. Pachori, "EEG signal classification using empirical mode decomposition and support vector machine," in *Proc. Int. Conf. Soft Comput. Problem Solving (SocProS)* (Advances in Intelligent and Soft Computing), vol. 131, K. Deep, A. Nagar, M. Pant, J. Bansal, Eds. New Delhi, India: Springer, 2012.

[8] E. Carlino, M. Sigaudo, A. Pollo, F. Benedetti, T. Mongini, F. Castagna, S. Vighetti, and P. Rocca, "Nonlinear analysis of electroencephalogram at rest and during cognitive tasks in patients with schizophrenia," *J. Psychiatry Neurosci.*, vol. 37, no. 4, pp. 259–266, Jul. 2012.

[9] M. Duda, J. A. Kosmicki, and D. P. Wall, "Testing the accuracy of an observation-based classifier for rapid detection of autism risk," *Transl. Psychiatry*, vol. 4, no. 8, p. e424, Aug. 2014, doi: 10.1038/tp.2014.65.

[10] A. Ameera, A. Saidatul, and Z. Ibrahim, "Analysis of EEG spectrum bands using power spectral density for pleasure and displeasure state," in *Proc. IOP Conf., Mater. Sci. Eng.*, vol. 557, Jun. 2019, Art. no. 012030, doi: 10.1088/1757-899X/557/1/012030.

[11] M. K. Delimayanti, B. Purnama, N. G. Nguyen, M. R. Faisal, K. R. Mahmudah, F. Indriani, M. Kubo, and K. Satou, "Classification of brainwaves for sleep stages by high-dimensional FFT features from EEG signals," *Appl. Sci.*, vol. 10, no. 5, p. 1797, Mar. 2020, doi: 10.3390/app10051797.

[12] M. J. Alhaddad, M. I. Kamel, and H. M. Malibary, "Diagnosis autism by Fisher linear discriminant analysis FLDA via EEG," *Int. J. Bio-Sci. Bio-Technol.*, vol. 4, no. 2, pp. 45–54, 2012.

[13] R. Djemal, K. AlSharabi, and S. Ibrahim, "EEG-based computer aided diagnosis of autism spectrum disorder using wavelet, entropy, and ANN," *BioMed Res. Int.*, vol. 2017, Apr. 2017, Art. no. 9816591.

[14] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Optimizing spatial patterns with sparse filter bands for motor-imagery based brain–computer interface," *J. Neurosci. Methods*, vol. 255, pp. 85–91, Nov. 2015.

[15] M. Congedo, R. E. John, D. De Ridder, and L. Prichep, "Group independent component analysis of resting state EEG in large normative samples," *Int. J. Psychophysiol.*, vol. 78, no. 2, pp. 89–99, Nov. 2010.

[16] K. Belwafi, R. Djemal, F. Ghaffari, and O. Romain, "An adaptive EEG filtering approach to maximize the classification accuracy in motor imagery," in *Proc. IEEE Symp. Comput. Intell., Cognit. Algorithms, Mind, Brain (CCMB)*, Orlando, FL, USA, Dec. 2014, pp. 121–126.

[17] K. Belwafi, F. Ghaffari, R. Djemal, and O. Romain, "A Hardware/Software prototype of EEG-based BCI system for home device control," *J. Signal Process. Syst.*, vol. 89, no. 2, pp. 263–279, Nov. 2017.

[18] M. Mahmud, M. S. Kaiser, A. Hussain, and S. Vassanelli, "Applications of deep learning and reinforcement learning to biological data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2063–2079, Jun. 2018.

[19] L. A. W. Gemein, R. T. Schirrmeister, P. Chrabąszcz, D. Wilson, J. Boedecker, A. Schulze-Bonhage, F. Hutter, and T. Ball, "Machine-learning-based diagnostics of EEG pathology," *NeuroImage*, vol. 220, Oct. 2020, Art. no. 117021.

[20] J. Chung, C. Gulcehre, K. H. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," in *Proc. Deep Learn. Represent. Learn. Workshop (NIPS)*, 2014. [Online]. Available: https://arxiv.org/abs/1412.3555

[21] O. Faust, U. R. Acharya, H. Adeli, and A. Adeli, "Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis," *Seizure*, vol. 26, pp. 56–64, Mar. 2015.

[22] A. Jacquin, E. Causevic, R. John, and J. Kovacevic, "Adaptive complex wavelet-based filtering of EEG for extraction of evoked potential responses," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 2005, pp. 393–396.

[23] F. Lotte, L. Bougrain, A. Cichocki2, 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, no. 3, 2018, Art. no. 031005.

[24] A. Mobaien and R. Boostani, "ACSP: Adaptive CSP filter for BCI applications," in *Proc. 24th Iranian Conf. Electr. Eng. (ICEE)*, May 2016, pp. 466–471.

[25] P. D. Velu and V. R. de Sa, "Single-trial classification of gait and point movement preparation from human EEG," *Frontiers Neurosci.*, vol. 7, p. 84, Jan. 2013.

[26] E. Grossi, C. Olivieri, and M. Buscema, "Diagnosis of autism through EEG processed by advanced computational algorithms: A pilot study," *Comput. Methods Programs Biomed.*, vol. 142, pp. 73–79, Apr. 2017.

[27] Y. Yao, J. Plested, and T. Gedeon, "Information-preserving feature filter for short-term EEG signals," *Neurocomputing*, vol. 408, pp. 91–99, Sep. 2020.

[28] M. Z. Baig, N. Aslam, and H. P. H. Shum, "Filtering techniques for channel selection in motor imagery EEG applications: A survey," *Artif. Intell. Rev.*, vol. 53, no. 2, pp. 1207–1232, Feb. 2020.

[29] M. Sharma, S. Patel, and U. R. Acharya, "Automated detection of abnormal EEG signals using localized wavelet filter banks," *Pattern Recognit. Lett.*, vol. 133, pp. 188–194, May 2020.

[30] H. Ocak, "Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 2027–2036, Mar. 2009.

[31] F. Sheng Bao, J.-M. Gao, J. Hu, D. Lie, Y. Zhang, and K. J. Oommen, "Automated epilepsy diagnosis using interictal scalp EEG," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2009, pp. 6603–6607.

[32] L. Guo, D. Rivero, and A. Pazos, "Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks," *J. Neurosci. Methods*, vol. 193, no. 1, pp. 156–163, Oct. 2010.

[33] L. Guo, D. Rivero, J. Dorado, C. R. Munteanu, and A. Pazos, "Automatic feature extraction using genetic programming: An application to epileptic EEG classification," *Expert Syst. Appl.*, vol. 38, no. 8, pp. 10425–10436, Aug. 2011.

[34] D. Wang, D. Miao, and C. Xie, "Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection," *Expert Syst. Appl.*, vol. 38, no. 11, pp. 14314–14320, 2011.

[35] U. Orhan, M. Hekim, and M. Ozer, "EEG signals classification using the K-means clustering and a multilayer perceptron neural network model," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 13475–13481, 2011.

[36] R. Djemili, H. Bourouba, and M. C. Amara Korba, "Application of empirical mode decomposition and artificial neural network for the classification of normal and epileptic EEG signals," *Biocybern. Biomed. Eng.*, vol. 36, no. 1, pp. 285–291, 2016.

[37] K. AlSharabi, S. Ibrahim, R. Djemal, and A. Alsuwailem, "A DWT-entropy-ANN based architecture for epilepsy diagnosis using EEG signals," in *Proc. 2nd Int. Conf. Adv. Technol. Signal Image Process. (ATSIP)*, Mar. 2016, pp. 288–291.

[38] S. Ibrahim, K. AlSharabi, R. Djemal, and A. Alsuwailem, "An adaptive learning approach for EEG-based computer aided diagnosis of epilepsy," in *Proc. Int. Seminar Intell. Technol. Appl. (ISITIA)*, Jul. 2016, pp. 55–60.

[39] A. Sheikhani, H. Behnam, M. R. Mohammadi, M. Noroozian, and P. Golabi, "Connectivity analysis of quantitative electroencephalogram background activity in autism disorders with short time Fourier transform and coherence values," in *Proc. Congr. Image Signal Process.*, May 2008, pp. 207–212.

[40] M. Ahmadlou, H. Adeli, and A. Adeli, "Fractality and a Wavelet-Chaos-Neural network methodology for EEG-based diagnosis of autistic spectrum disorder," *J. Clin. Neurophysiol.*, vol. 27, no. 5, pp. 328–333, Oct. 2010.

[41] W. Bosl, A. Tierney, H. Tager-Flusberg, and C. Nelson, "EEG complexity as a biomarker for autism spectrum disorder risk," *BMC Med.*, vol. 9, no. 1, p. 18, Dec. 2011.

[42] A. Sheikhani, H. Behnam, M. R. Mohammadi, M. Noroozian, and M. Mohammadi, "Detection of abnormalities for diagnosing of children with autism disorders using of quantitative electroencephalography analysis," *J. Med. Syst.*, vol. 36, no. 2, pp. 957–963, Apr. 2012.

[43] M. Ahmadlou, H. Adeli, and A. Adeli, "Improved visibility graph fractality with application for the diagnosis of autism spectrum disorder," *Phys. A, Stat. Mech. Appl.*, vol. 391, no. 20, pp. 4720–4726, Oct. 2012.

[44] M. Ahmadlou, H. Adeli, and A. Adeli, "Fuzzy synchronization likelihood-wavelet methodology for diagnosis of autism spectrum disorder," *J. Neurosci. Methods*, vol. 211, no. 2, pp. 203–209, Nov. 2012.

[45] E. A. Alsaggaf and M. I. Kamel, "Using EEGs to diagnose autism disorder by classification algorithm," *Life Sci. J.*, vol. 11, no. 6, pp. 305–308, 2014.

[46] 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, Jun. 2019, Art. no. 031001, doi: 10.1088/1741-2552/ab0ab5.

[47] H. U. Amin, W. Mumtaz, A. R. Subhani, M. N. M. Saad, and A. S. Malik, "Classification of EEG signals based on pattern recognition approach," *Frontiers Comput. Neurosci.*, vol. 11, p. 103, Nov. 2017, doi: 10.3389/fncom.2017.00103.

[48] *EEG KAU Dataset*. Accessed: 2012. [Online]. Available: https://malhaddad.kau.edu.sa/Pages-BCI-Datasets.aspx

[49] A. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Harvard Univ.–MIT Division Health Sci. Technol., Cambridge, MA, USA, Sep. 2009.

[50] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E*, vol. 64, Nov. 2001, Art. no. 061907.

[51] S. S. Talathi, "Deep recurrent neural networks for seizure detection and early seizure detection systems," 2017, *arXiv:1706.03283*. [Online]. Available: http://arxiv.org/abs/1706.03283

[52] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent.*, San Diego, CA, USA, 2015.

[53] A. Tharwat, "Principal component analysis-a tutorial," *Int. J. Appl. Pattern Recognit.*, vol. 3, no. 3, pp. 197–240, 2016.

[54] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, vol. 46. Hoboken, NJ, USA: Wiley, 2004.

[55] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Netw.*, vol. 10, no. 3, pp. 626–634, May 1999.

[56] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Netw.*, vol. 13, nos. 4–5, pp. 411–430, Jun. 2000.

[57] A. Tharwat, "Independent component analysis: An introduction," *New England J. Entrepreneurship*, Aug. 2018, doi: 10.1016/j.aci.2018.08.006.

[58] R. Schirrmeister, J. Springenberg, L. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgardand T. Ball, "Deep learning with convolutional neural networks for brain mapping and decoding of movement-related information from the human EEG," *Hum. Brain Mapping*, vol. 38, no. 11, pp. 1–30, 2017.

[59] A. G. Correa, L. L. Orosco, P. Diez, and E. L. Leber, "Adaptive filtering for epileptic event detection in the EEG," *J. Med. Biol. Eng.*, vol. 39, no. 6, pp. 912–918, 2019.

[60] T. Pander, "EEG signal improvement with cascaded filter based on OWA operator," *Signal, Image Video Process.*, vol. 13, no. 6, pp. 1165–1171, Sep. 2019.

[61] S. Thapaliya, S. Jayarathna, and M. Jaime, "Evaluating the EEG and eye movements for autism spectrum disorder," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Seattle, WA, USA, Dec. 2018, pp. 2328–2336.

[62] W. J. Bosl, H. Tager-Flusberg, and C. A. Nelson, "EEG analytics for early detection of autism spectrum disorder: a data-driven approach," *Sci. Rep.*, vol. 8, no. 1, p. 6828, Dec. 2018. [Online]. Available: https://www.nature.com/scientificreports

[63] D. Haputhanthri, G. Brihadiswaran, S. Gunathilaka, D. Meedeniya, Y. Jayawardena, S. Jayarathna, and M. Jaime, "An EEG based channel optimized classification approach for autism spectrum disorder," in *Proc. Moratuwa Eng. Res. Conf. (MERCon)*, Jul. 2019, pp. 123–128.

GHAITH BOUALLEGUE received the degree in electrical and control engineering from the University of Gabes, ENIG of Gabes, Tunisia. He is currently pursuing the Ph.D. degree with the University of Sousse, ENISo of Sousse, Tunisia. His current research interests include machine learning, signal processing in biomedical applications, and brain-computer interface design.

RIDHA DJEMAL received the B.S. and M.S. degrees in electrical engineering from the Joseph Fourier University of Grenoble, France, in 1991 and 1992, respectively, and the Ph.D. degree in microelectronics from the National Institute of Grenoble, France, in 1996. From 1997 to 2005, he was an Assistant Professor with Monastir University, Tunisia. He was an Associate Professor with the Electrical Engineering Department, King Saud University, Riyadh, Saudi Arabia, from 2007 to 2017. He is currently a Professor with the Electronics Department, National Institute for Applied Sciences and Technology (ISSAT), Sousse University, Tunisia. His current area of interest includes the design of FPGA-based hardware/software architectures related to high-speed networks, image and video and signal processing. His current research is devoted to the development of many EEG signal processing applications related to Motor imagery and neurological disorder diagnosis. He is also involved in the review process of a number of ISI research journals, and national and international conferences.

SALEH A. ALSHEBEILI was the Chairman of the Electrical Engineering Department, King Saud University, Riyadh, Saudi Arabia, from 2001 to 2005. He has more than 27 years of teaching and research experience in the area of communications and signal processing. He was a member of the Board of Directors with the King Abdullah Institute for Research and Consulting Studies from 2007 to 2009; a member of the Board of Directors with Prince Sultan Advanced Technologies Research Institute from 2008 to 2017, where he was the Managing Director from 2008 to 2011; and the Director of the Saudi-Telecom Research Chair from 2008 to 2012. He has been the Director of the Technology Innovation Center, RF and Photonics in e-Society, funded by the King Abdulaziz City for Science and Technology (KACST) since 2011. He is currently a Professor with the Electrical Engineering Department, King Saud University. He has been in the Editorial Board of the *Journal of Engineering Sciences* of King Saud University from 2009 to 2012. He also has an active involvement in the review process of a number of research journals, KACST general directorate grants programs, and national and international symposiums and conferences.

HESHAM ALDHALAAN is currently working at the King Faisal Specialist Hospital Research Centre, Riyadh, Saudi Arabia, holding the following positions: Consultant Pediatric Neurologist; Epileptologist since 2002; the Program Director of Pediatric Neurology Residency from 2006 to 2014; the Director of the Center for Autism Research since November 2013; the Chairman of the Pediatric Neurology Training program for Saudi Commission for Health since 2019; and Program Director of Neuro Technology from 2003 to 2010. He is also holding a position outside the King Faisal Specialist Hospital Research Centre, i.e., a member of the Executive Board for Autism Center of Excellence since 2019; a member of the Scientific Committee with the King Salman Centre for Disability Research (KSCDR) since 2019; and a member of the Scientific Committee with Prince Mohammad Bin Salman Autism Centre since 2017. He also has multiple publications ongoing research activities in Autism Disability.