Seizure Types Classification by Generating Input Images With in-Depth Features From Decomposed EEG Signals for Deep Learning Pipeline

Anand Shankar, Samarendra Dandapat, *Member, IEEE*, and Shovan Barma, *Member, IEEE*

***Abstract*—Electroencephalogram (EEG) based seizure types classification has not been addressed well, compared to seizure detection, which is very important for the diagnosis and prognosis of epileptic patients. The minuscule changes reflected in EEG signals among different seizure types make such tasks more challenging. Therefore, in this work, underlying features in EEG have been explored by decomposing signals into multiple subcomponents which have been further used to generate 2D input images for deep learning (DL) pipeline. The Hilbert vibration decomposition (HVD) has been employed for decomposing the EEG signals by preserving phase information. Next, 2D images have been generated considering the first three subcomponents having high energy by involving continuous wavelet transform and converting them into 2D images for DL inputs. For classification, a hybrid DL pipeline has been constructed by combining the convolution neural network (CNN) followed by long short-term memory (LSTM) for efficient extraction of spatial and time sequence information. Experimental validation has been conducted by classifying five types of seizures and seizure-free, collected from the Temple University EEG dataset (TUH v1.5.2). The proposed method has achieved the highest classification accuracy up to 99% along with an *F*1-score of 99%. Further analysis shows that the HVD-based decomposition and hybrid DL model can efficiently extract in-depth features while classifying different types of seizures. In a comparative study, the proposed idea demonstrates its superiority by displaying the uppermost performance.**

***Index Terms*—Convolution neural network, continuous wavelet transform, electroencephalogram, hilbert vibration decomposition, long short-term memory, seizure types.**

Manuscript received 31 July 2021; revised 13 February 2022 and 1 March 2022; accepted 10 March 2022. Date of publication 16 March 2022; date of current version 5 October 2022. This work was supported by the Sunrise Career project with Ref: NECBH/2019-20/118 under the North East Centre for Biological Sciences and Healthcare Engineering (NECBH) Twinning Outreach Programme hosted by the Indian Institute of Technology Guwahati (IITG), Guwahati, Assam funded by the Department of Biotechnology (DBT), Ministry of Science and Technology, Government of India under Grant BT/COE/34/SP28408/2018 for providing necessary financial support. *(Corresponding author: Anand Shankar.)*

Anand Shankar and Shovan Barma are with the Department of Electronics and Communication Engineering, Indian Institute of Information Technology Guwahati, Assam 781015, India (e-mail: anand.shankar@iiitg.ac.in; shovan@iiitg.ac.in).

Samarendra Dandapat is with the Electronics and Electrical Department, Indian Institute of Technology Guwahati, Assam 781039, India (e-mail: samaren@iitg.ac.in).

Digital Object Identifier 10.1109/JBHI.2022.3159531

I. Introduction

ELECTROENCEPHALOGRAM (EEG) based epileptic seizure detection and analysis have been considered as a reliable tool in clinical diagnosis due to its simple, low cost, portable, and easy to use system, compared to other modalities [1], [2]. Generally, the seizure diagnosis is performed by the medical professionals based on visual interpretation of long EEG records along with behavioral features, which makes the task time-consuming and error-prone [1]–[4]. For this purpose, several machine learning based methods have been proposed and found very suitable [3], [4]. However, in the machine learning framework, the detection of different seizure types has not been given enough attention, but it could play a significant role in epileptic seizure diagnosis followed by appropriate drug selection and treatment processes [5], [6]. Indeed, to discriminate different types of seizures accurately, defining distinctive features among different seizure types need to be well-defined, which is very challenging [5], [7]. In this context, the use of traditional machine learning might not be suitable as it fully relies on pre-defined hand-crafted features [1], [2]. In this direction, an advanced machine learning technique, called deep learning (DL), a data-driven technique that has evidenced remarkable performance in image-based classification may be useful [1], [2], [8]–[10]. It automatically extracts high-level features from input data, which is not possible manually [3], [4], [9]. However, its performance highly relies on a large number of diverse input data along with a suitable DL pipeline [1], [2], [8], [10]. Certainly, the successful application of DL in image classification can be involved in seizure types discrimination by framing 2D images from 1D EEG signals, while preserving the relevant characteristics [5], [6], [11]–[14]. Therefore, in this work, a new 2D image generation technique from a 1D EEG signal has been proposed to obtain in-depth diverse features among different seizures types. Besides, a hybrid DL pipeline capable of extracting temporal and spatial features has also been developed for effective classification.

In the literature, most of the works have employed numerous traditional machine learning based methods aiming at epileptic seizure detection; on the contrary, very few works address the important concern of different types of seizures discrimination [5], [6], [11], [17]. For instance, Saputro, *et al.* [15], have discriminated three types of seizures along with seizure-free by considering different features including Mel frequency cepstral coefficient, Hjorth descriptor, and independent component analysis for support vector machine (SVM) classifier and achieved accuracy up to 91.4%; Wijayanto, *et al.* [16], have proposed an empirical mode decomposition technique for suitable statistical features followed by SVM based classification to discriminate four types of seizures and achieved the classification accuracy up to 95%; Kassahun, *et al.* [17], have done similar kinds of work based on ontology and genetics to classify two types of seizure. Nevertheless, the performance of all these methods fully relies on the selection of appropriate hand-crafted features manually, which makes the task very challenging, especially for distinguishing minuscule differences among different seizure types [5], [6], [11], [14]. In this regard, the DL-based technique that bypasses the manual feature selection process could be useful [5], [6]. Besides, it has already evidenced remarkable performance in image-based classification including biomedical signals [1], [2], [8], [10]. Nevertheless, the efficacy of the DL-based method for particular applications mainly depends on two crucial issues — the requirement of a large number of diverse input samples for discovering apposite features automatically and extracting those features through different layers of the DL [1], [2], [8]–[14]. It becomes quite challenging when analyzing 1D EEG signals by converting them into 2D images [9], [11]–[14]. In this direction, few researchers have proposed different DL-based pipelines for the discrimination of different types of seizures; for example, Roy, *et al.* [13], have discriminated eight different types of seizures by employing a basic convolution neural network (CNN) and achieved an $F1$-score of 72.2% without addressing the vital issue of input image generation technique; Sriraam, *et al.* [11], have employed four different CNN models including basic CNN, AlexNet, VGG16, and VGG19 to discriminate eight types of seizures types where the input images are generated by vertically concatenating the 2D spectrogram of raw EEG and achieved the classification accuracy of 81.14%, 84.06%, 79.71%, and 76.81% respectively; in very similar types of work, Asif, *et al.* [6], have adopted saliency-encoded spectrogram technique for 2D images generation and fed them into SeizureNet which is a combination of several CNN blocks. Nevertheless, their image generation mechanism is favorable, but the DL pipeline having only CNN blocks might not be suitable. Indeed, the CNN efficiently extracts spatial features, whereas for seizure detection temporal information needs to be taken into account especially while discriminating different seizures types. Ahmedt, *et al.* [12], have used stacked auto-encoder, CNN, recurrent neural network (RNN), and recurrent CNN to classify eight types of seizures, achieving a maximum weighted $F1$-score of 94.0%. Liu, *et al.* [5], have employed a hybrid bilinear DL architecture to classify eight different types of seizures by using the 2D representation of raw EEG as input, and for the generation of 2D representation, a short-time Fourier transform (STFT) approach has been used. In the above studies, the proposed methods have either focused on the generation of images directly from raw EEG signals or chosen different DL architectures, which can be further improved by comprising a more efficient 2D image generation technique along with an appropriate DL pipeline. In this context, the generation of 2D images from 1D EEG while preserving the relevant characteristics and dynamic patterns of the original signal, and an efficient DL framework that can extract spatial as well as temporal features, could be more appropriate.

In this view, the benefit of decomposition of a signal into several subcomponents to mine underlying characteristics has already been proven by numerous applications including epileptic seizures analysis [1]–[4], [18]–[21]. Such decomposition can be accomplished by predefined basis functions that are suitable for stationary signals [20], [21]. Meanwhile, like all physiological signals, EEG signals are strictly non-linear and non-stationary [7], [22]. Hence, non-linear decomposition techniques, such as empirical mode decomposition (EMD) [18], ensemble EMD, vibrational mode decomposition (VMD) [19], Hilbert vibration decomposition (HVD) [20]–[21] could be more efficient, which have been successfully applied in EEG signal processing [3], [4], [18]. Nonetheless, the EMD decomposes a signal into several band-limited intrinsic mode functions (IMFs) recursively by removing local oscillations; where the VMD extracts IMFs by dynamically calculating center frequencies concurrently [18], [19]. Eventually, these methods fail to preserve the phase information, which plays a very significant role in the analysis of different types of seizures [20], [21]. Meanwhile, the HVD, an adaptive signal decomposition technique, disintegrates a signal into a certain number of subcomponents by preserving phase information and relevant characteristics of the parent signal [23], [24]. Consequently, potential discriminative features can be generated from the decomposed subcomponents, which can further improve the classification performance of the model. Therefore, in this work, the HVD method has been adopted to decompose the EEG signals, which might discriminate distinctive features among different types of seizures.

Generally, in seizure detection, the long recorded EEG signals (sometimes more than hours) are considered, where the appearances of epileptic seizures are infrequent [7]. Indeed, developing any signal processing technique that fully relies on such long EEG signals, might not be efficient [25]. Instead, it is more convenient to consider EEG segments, which could not only provide insightful information more efficiently but also reduce the computational burden [1]–[4], [25]–[27]. In addition, it could solve the demand for large input datasets for DL-based classification [25]. In this regard, several researchers have found convincing results by developing seizure detection techniques by considering the different lengths of EEG [8], [25]–[27]. For instance, Emami, *et al.*, have considered the various duration of EEG segments (0.5s, 1s, 2s, 5s, and 10s) to classify epileptic seizure [25]; Gao *et al.*, have divided EEG signal into frames of 4s long to discriminate epileptic seizures [8]; Shankar, *et al.*, have sectioned EEG signal into 5.9s and 10s each for analysis of epileptic seizure and types of seizure [14], [26], [27] and so on. Therefore, in this study, the HVD technique has been employed on the EEG segments that have been sectioned from long EEG recordings. Subsequently, 2D images have

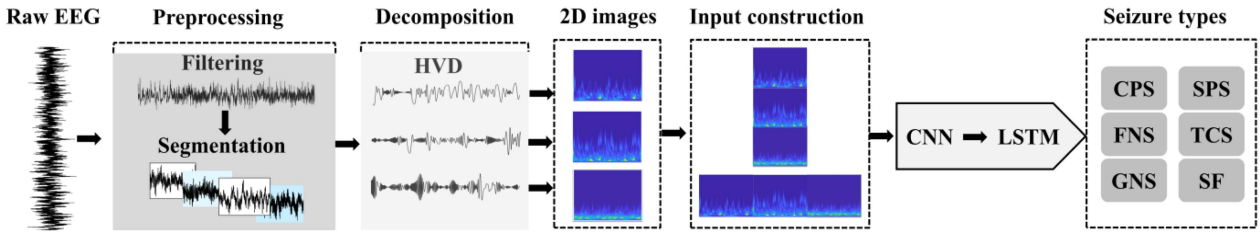

Fig. 1. The outline of the proposed idea to classify different types of seizures (complex partial seizure (CPS), Focal non-specific seizure (FNS), Generalized non-specific seizure (GNS), Simple partial seizure (SPS), Tonic-clonic seizure (TCS)) along with seizure-free using EEG recordings.

been encoded from the decomposed subcomponents of the 1D signals for the DL-based classification.

In the literature, several researchers have proposed different 2D image generation techniques from 1D EEG signals for various DL pipelines [8]–[11], [25]–[27]. For instance, Emami, *et al.* [25], have constructed 2D images by taking direct snapshots of EEG segments of certain durations; Gao, *et al.* [8], have generated 2D images based on power spectrum density energy diagram, which may not efficient due to the low-frequency resolution of EEG; Shankar, *et al.* [14], [26], [27], have proposed time series based techniques including recurrence plot and gramian angular field for 2D image construction for their DL-based seizure classification, which ignores the spatial features. The other methods of 2D image generation are the time-frequency representation (TFR) of 1D EEG signals, which is well-suited as it preserves the inherent time and frequency information [11], [12], [29], [30]. Among different TFRs, the continuous wavelet transform (CWT) which can provide better resolution in quick frequency fluctuations has been found very much suitable for this purpose and has displayed noticeable results in seizure analysis [31]. Therefore, in this work, 2D images have been constructed from 1D EEG signals by TFR involving CWT.

In this view, the proposed idea has been developed to distinguish five different types of seizure along with seizure-free. First, the large EEG signals have been segmented into several sections which are further decomposed into multiple subcomponents by the HVD technique. Then, 2D images have been constructed for specific decomposed subcomponents of 1D EEG signal by the TFR involving CWT. Next, the generated images have been restructured for the hybrid DL pipeline. For experimental validation, the Temple University EEG dataset (TUH v1.5.2) for analysis of different types of seizures has been considered [32]. Several experiments including appropriate subcomponents selection, image quality assessment, classification performances, and comparative study have been performed. The main contributions of this work are as follows:

- The 2D input images generation from decomposed subcomponents of 1D EEG signals for in-depth features extraction.
- Restructuring of the generated 2D images for the DL pipeline for effective performance.
- A hybrid DL pipeline (CNN+LSTM) for the classification of different types of seizures.
- Analysis of how signal decomposition improves the classification performance in the DL framework.

The rest of the paper has been organized as: Section II describes the proposed method in detail. The experimental methodology has been presented in Section III, followed by results and discussion in Sections IV and V respectively. Finally, the conclusions have been drawn in Section VI.

II. Proposed Method

A system overview of the proposed method has been displayed in Fig. 1. Firstly, the EEG signals have been preprocessed and segmented for decomposition by HVD. Next, the 2D images have been constructed by CWT for certain decomposed subcomponents. After that, for particular subcomponents, the generated images have been restructured into a single image by stacking in two ways — horizontally and vertically, and directly fed into a hybrid DL (CNN followed by LSTM) pipeline separately for classification of different types of seizures. The following subsections elaborate the proposed idea in detail by describing each sub-block.

A. Preprocessing

In preprocessing step, the recorded signals need to be processed for noise and artifact removal [1]–[4]. However, the EEG dataset used in this study are already free from artefacts and noise [32], [33], hence, the collected EEG signals have been considered for the next step of processing. In the literature, it has been reported that most epileptic seizure activities occur between 0.5Hz to 30Hz [14], [28]. Therefore, a band pass filter having cut-off frequencies 0.5Hz and 30Hz has been employed to remove the unnecessary higher frequencies [5], [6], [11]–[14]. Now, the filtered signals have been segmented into multiple pieces with a pre-defined length of 10s. The segmentation has been performed by 50% overlapping to minimize information loss [25]. It provides a large number of diverse samples, which is highly essential for successful deep learning based classification [1], [2], [25]–[27]. Next, each of the segmented EEG signals has been used for decomposition by the HVD technique.

B. Hilbert Vibration Decomposition (HVD)

The HVD decomposes a non-linear signal into a certain number of subcomponents having slowly varying instantaneous amplitude (IA) and frequency (IF) [23]. Indeed, it is very much suitable for analyzing non-linear and non-stationary signals such as EEG signals [20], [21], [24]. The decomposed subcomponents certainly preserve the phase information of the actual signal,

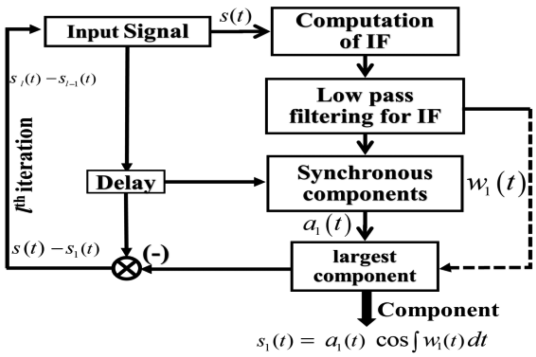

Fig. 2. A systematic diagram of the HVD method.

which enables it to track seizure events reliably [24]. However, the HVD method can be employed, if two fundamental signal conditions hold — a) the fundamental signal is obtained by the superposition of at least one quasi-harmonic with several full cycles and b) each vibrating component's frequency and envelope are different from each other [23], [24]. Indeed, the EEG signal satisfies these two primary conditions by displaying several quasi-harmonic full-cycle components, and their IFs differ [24]. Certainly, it has already been successfully employed for the analysis of EEG signals [18], [21].

In HVD, a multi-component signal is decomposed into a certain number of subcomponents with slow varying IA and IF in an iterative way [23]. In each iteration, the slowly varying synchronous components are separated by filtering the IFs [23], [24]. Certainly, each of the components keeps the phase information intact with the parent signal. During the execution of the HVD, the IF estimation and synchronous detection operation are performed. Eventually, the synchronous detection enables it to retain the components having small varying amplitudes which are usually dominated by other large components [24]. It can be estimated by analyzing the IFs. Therefore, first, the IA and IF of an EEG signal are computed. Then, the IF is filtered by a low-pass filter (LPF) having a very low cut-off frequency [23], [24]. The synchronous component is created by the IA and filtered IF [24]. The component having the largest amplitude is separated from its parent signal and the residue is used for the next iteration [23], [24]. Different stages of the HVD process have been depicted in Fig. 2.

Now, a multicomponent signal, $s(t)$ which consists of l number of subcomponents of slow varying frequency, $\omega(t)$ and amplitude, $a(t)$ can be defined by (1);

$$s(t) = \sum_{l} a_l(t)\ \cos\left(\int \omega_l(t)dt\right) \tag{1}$$

Next, the IF can be estimated by analytical representation, $A(t)$ of $s(t)$ (2);

$$A(t) = s(t) + j\,\tilde{s}(t) \tag{2}$$

where $\tilde{s}(t)$is the Hilbert transform (HT) of $s(t)$, and evaluated by (3), where integral is the Cauchy principal value (*P.V.*).

$$\tilde{s}(t) = HT\left[s(t)\right] = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{s(\tau)}{t-\tau}\, d\tau \tag{3}$$

Finally, the IF has been estimated by (4);

$$\omega(t) = \frac{d}{dt}\varphi(t) = \frac{d}{dt} tan^{-1}\left(\frac{\tilde{s}(t)}{s(t)}\right) \tag{4}$$

where, $\omega(t)$ and $\varphi(t)$ denote the IF and instantaneous phase (IP) respectively. Next, synchronous demodulation has been performed to detect the subcomponent having the largest envelope [24]. Indeed, synchronous detection enables in detecting the small changes in $s(t)$. Now, in-phase $s_{l=r}(t)$ and HT phase $(\tilde{s}_{l=r}(t))$ in r–reference of the l^{th} component can be calculated by (5) shown at the bottom of this page. where, $a_l(t)$, $\omega_l(t)$, and $\phi_l(t)$ denote the IA, IF, and IP of the l^{th} components respectively; whereas, $\omega_r(t)$ refers to the IF of the largest component in r-reference. Next, the subcomponent, $s_1(t)$ having the largest energy, has been separated from the initial compositions $s(t)$. This signal, $s_{l\text{-}1}(t) = s_l(t)$–$s_1(t)$ contains the lower energy and can be disintegrated in the subsequent iterations. Eventually, the decomposed subcomponents are with energy in decreasing order, *i.e.*, the first subcomponent has the largest energy than its very next, and so on [24]. Thus, the EEG signals have been decomposed into multiple subcomponents, which have been further considered for 2D image generation.

C. 2D Image Construction and Stacking

1) Image Construction by CWT: For discriminating different types of seizures, tracing minor variations in instantaneous time and frequency in the EEG must appear in the 2D image [5], [6], [11]–[14]. Therefore, the 2D images have been generated for individual decomposed subcomponents by employing CWT which preserves the instantaneous time and frequency information very efficiently [29]–[31]. The CWT decomposes a signal into wavelets (mother wavelets) which are nothing but small oscillations at the highly localized time [29], [30]. The mother wavelet is scaled and shifted along the time axis of the target signals, which generates a time-frequency texture, and offers a very good time and frequency localization [31]. The CWT has been found very suitable and efficient for representing nonlinear and non-stationary signals including EEG [1], [2], [29]–[31]. Mathematically, the CWT for signal $x(t)$ can be defined by (6);

$$CWT_s(d,\,\Delta) = \frac{1}{\sqrt{d}} \int x(t)\ \psi\left(\frac{t-\Delta}{d}\right) dt \tag{6}$$

where, $CWT(d, \Delta)$, and ψ denote wavelet coefficient and basis wavelet function respectively. The $\psi(t)$ is shifted by Δ and contracted or dilated by d. In case of contraction, *i.e.*, $d < 1$, $\psi(t)$ provides high temporal resolution which is effective

$$\left.\begin{aligned} s_{l=r}(t) &= \tfrac{1}{2}\, a_l(t)\ \left[\cos\left(\phi_l(t)\right) + \cos\left(\int\left(\omega_l(t) + \omega_r(t)\, dt + \phi_l(t)\right)\right)\right] \\ \tilde{s}_{l=r}(t) &= \tfrac{1}{2}\, a_l(t)\ \left[\sin\left(\phi_l(t)\right) - \sin\left(\int\left(\omega_l(t) + \omega_r(t)\, dt + \phi_l(t)\right)\right)\right] \end{aligned}\right\} \tag{5}$$

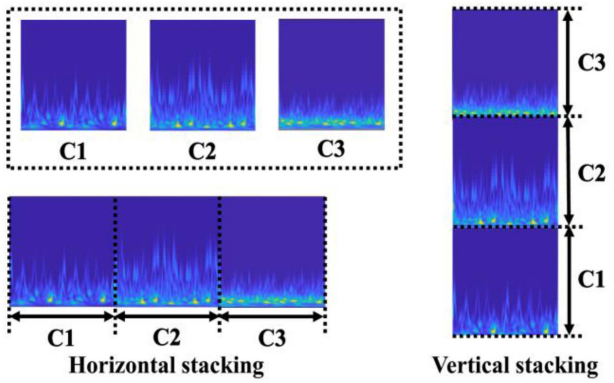

Fig. 3. The construction of input images for the proposed DL pipeline by stacking of 2D images generated from decomposed subcomponents of EEG segments vertically and horizontally.

for measuring the short-time activities, whereas, for dilation, *i.e.*, $d > 1$, the $\psi(t)$ gives high spectral resolution [29]–[31]. In this work, the Morlet wavelet has been chosen, which works efficiently for analyzing the complex signal, like EEG [30], and displayed reasonable performance in several EEG based studies [11], [12], [29]–[31]. It is capable of recording short spines of repeating and alternating spikes with start and end time neatly [29]. Mathematically, the Morlet wavelet function is defined by (7), where t denotes the sampling instants.

$$\psi(t) = e^{-t^2/2} \cos(5t) \tag{7}$$

Eventually, in the HVD method, the first few decomposed subcomponents conserve most of the energy of the parent signal [23], [24]. Hence, for 2D image generation by CWT for each EEG segment, selection of the first few subcomponents might be appropriate, which preserves the underlying variations competently. Indeed, the generated images are with different energy levels, which cannot be directly fed into the deep learning model. Therefore, a single image needs to be constructed by restructuring the images of the subcomponents for each of the EEG segments. In this regard, the single image generation has been performed by an image stacking mechanism.

2) Image Stacking: Input data preparation is one of the vital stages for the successful application of DL-based classifiers [1], [2], [9], [10]. In this work, the first three decomposed subcomponents (C1, C2, and C3) have been taken into account and stacked in two ways — vertically and horizontally, to construct a single image as displayed in Fig. 3. The 2D images of C1, C2, and C3 subcomponents are shown in the top left corner of Fig. 3, whereas, the horizontally and vertically stacked images have been displayed on the bottom left and right side respectively. Finally, stacked images have been fed into the proposed hybrid DL pipeline separately to investigate the best image stacking mechanism.

D. Hybrid Deep Learning Pipeline

The proposed hybrid DL pipeline consists of two well-known models CNN and LSTM. The CNN is very efficient in extracting the spatial features, while the LSTM discriminates time related events very competently, which can play a very significant role in epileptic seizure analysis [8], [9], [20], [34]–[38]. In earlier works, the combination of both models has been found very suitable in seizure analysis [20], [21], [34], [36], [38]. They have been combined in two ways — parallel and serial, where, the serial combination has been found more efficient [20], [34]–[36]. Therefore, in this work, the serial combination, *i.e.*, CNN followed by LSTM has been considered.

In CNN, multiple hidden layers consist of convolution, pooling, batch normalization, which automatically extracts the relevant features from raw input data [8]–[10], [14]. A pre-defined kernel performs the convolution with the inputs and extracts structural information, which passes through a non-linear activation function. Mathematically, the convolution outcome is estimated by (8),

$$C_p^m = f\left(\sum_{q=1}^{N_{m-1}} conv2D\left(w_{q,p}^m,\ F_q^{m-1}\right) + b_p^m\right) \tag{8}$$

where, C_p^m refers to convolution outcome, *i.e.*, feature map of p^{th} node in m^{th} layer, while 2D convolution operation is represented by *conv2D*. The F_q^{m-1} and $N_{m\text{-}1}$ refer to the p^{th} node and number of nodes in the $(m\text{–}1)^{\text{th}}$ layer respectively. The weight of the trainable kernel is represented by $w_{q,p}^m$. The convolution outcome followed by a non-linear activation ($f(y) = max(0, y)$). After that, the outcome of the convolution layer has been passed through the batch normalization layer, followed by the max-pooling layer. The max-pooling layer reduces the spatial size, feature variance, and complexity of the model and provides a smooth and sharp set of features [8], [9], [14]. Besides, the internal covariance shift issue can be minimized by batch normalization [14], [25]–[27]. Next, the outcome has been flattened and fed into the LSTM for temporal feature extraction.

The LSTM network is a special class of recurrent neural network that is capable to characterize events disjointed by the time [20], [34]–[36], [38]. It estimates long-term dependencies in input data and is well suited for time sequence event analysis due to its short-term memory [20]. The key component in LSTM is known as cell and its activities have been characterized by three important gates — input, forget, and output [34], [35]. The gates can detect long-term dependencies in the input stream [36]. Further, the gates in the LSTM network collaborate with each other to preserve the initial information and can improve the learning ability [38]. Certainly, appropriate estimation of the learning rate for cells is very important as it could reduce the vanishing gradients issue that is preserved or elapsed over a long time [34]. The mathematical expressions inside the LSTM are (9);

$$\left.\begin{aligned} z_f &= f\left(W_f\left[I_t, H_{t-1}\right]\right) \\ z_i &= f\left(W_i\ \left[I_t, H_{t-1}\right]\right) \\ z &= \tanh\left(W\left[I_t, H_{t-1}\right]\right) \\ z_o &= f\left(W_o\ \left[I_t, H_{t-1}\right]\right) \\ x_t &= z_f \times x_{t-1} + z_i \times z \\ H_t &= z_o \times \tanh\left(x_t\right) \\ O_t &= f\left(W\ H_t\right) \end{aligned}\right\} \tag{9}$$

In LSTM, the cell gate (z) is responsible for remembering the information across time; z_f denotes the forget gate that

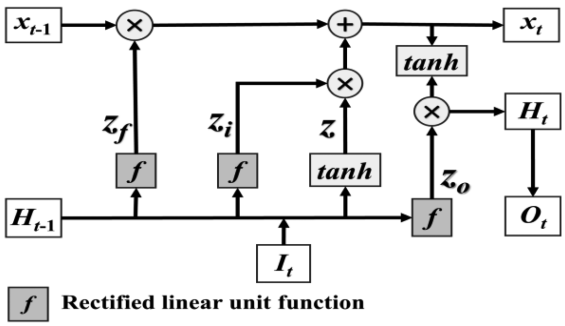

Fig. 4. The internal structure of the LSTM architecture.

controls the information over time; z_i represents the input gate that controls the streaming of information in the cell and z_o controls the extent of information in the cell that to be used for computing the output through the weight metrics, W. The x_{t-1}, H_{t-1}, and I_t denotes the cell state, previous hidden state, and current input respectively; where, x_t, H_t and O_t are three outputs that represent cell state, hidden state, and current output respectively. The internal structure of the LSTM architecture has been shown in Fig. 4.

The features extracted by the CNN have been passed through the LSTM network to extract temporal characteristics. Next, the outcome of the LSTM has been directly fed into fully connected layers having ReLU activation function. Finally, the output layer having a sigmoid function has been employed to classify the appropriate classes of seizure types. The proposed hybrid DL framework has been shown in Fig. 5.

III. Experimental Methodology

A. Data

The experimental validation has been performed by using a well-recognized epilepsy EEG dataset provided by the Temple University Hospital (TUH v1.5.2) [32]. It consists of the EEG signals of eight different seizure types, which have been recorded at different sampling frequencies by different montages. The two montages — average reference (Av) with a fixed electrode and linked ears reference (Le) have been used during EEG recordings [33]. Actually, the EEG recordings of different seizures types are not in equal proportions in the dataset. Therefore, maintaining consistency, 153 seizure sessions of 30 patients for five types of seizures have been taken into account, which is recorded based on Av hardware setup with 250Hz sampling frequency and 16-bit resolution. It includes EEG signals from common 19 channels — C3_Av, C4_Av, Cz_Av, F3_Av, F4_Av, F7_Av, F8_Av, FP1_Av, FP2_Av, Fz_Av, O1_Av, O2_Av, P3_Av, P4_Av, Pz_Av, T3_Av, T4_Av, T5_Av, and T6_Av. The five different types of seizures along with seizure-free and their durations have been listed in Table I, in which the first and second columns represent different seizure types and their corresponding duration respectively. The EEG data are already free from noise and artifact [32]–[33]. It should be noted that the recorded seizure types are not according to the International League Against Epilepsy (ILAE), but it would not have any influence to validate the proposed idea [7].

TABLE I
EEG Dataset Description

Seizure type	Duration (s)
Complex partial seizure (CPS)	1904.09
Focal non-specific seizure (FNS)	1829.19
Generalized non-specific seizure (GNS)	2060.84
Simple partial seizure (SPS)	1329.50
Tonic-clonic seizure (TCS)	517.17
Seizure-free (SF)	2000.00

B. Experiment

In the dataset, the EEG signals are already free from noise and artefacts [32], [33], so the EEG signals have been directly filtered into a frequency range of 0.5–30Hz (most of the seizures activities occur in this band) [14], [25]–[27], [28]. For this purpose, a fifth-order Butterworth band pass filter with cut-off frequencies of 0.5Hz and 30Hz has been used. Next, the whole EEG signals have been split into segments of 10s with 50% overlapping. Hereafter, the segmented signals have further been used for decomposition by the HVD method. Basically, the HVD technique, utilizes the low-pass filtering of IFs, to retrieve the slowly fluctuating frequency component at each iteration step [23], [24]. However, during filtering IFs, the cut-off frequency of the low-pass filter (LPF) should be carefully chosen, as at each iteration, the LPF keeps the unknown largest signal component by cutting down the asymmetrical oscillations [18]–[21], [23], [24]. It is suggested to keep the cut-off frequency of the LPF as low as possible, but not lower than the lowest frequency subcomponent [23]. Certainly, during designing the LPF, the shape and stability factors have also been taken into account. Therefore, the cut-off frequency of the LPF has been chosen as 0.05Hz to isolate subcomponents.

Further, among all the decomposed subcomponents by HVD, the first three subcomponents jointly preserve almost 80% energy of the main signal, therefore they have been taken into consideration for further processing. Besides, the correlation of subcomponents with the main signal has been also measured to check their suitability. Next, a separate image has been generated for the three subcomponents of an EEG segment by the CWT. Hereafter, for each segment, a single image has been constructed by horizontal and vertical stacking of generated individual images from each subcomponent (see Fig. 3). The number of generated images for different types of seizures is 7082, 6872, 7733, 5016, 1938, and 7524, for CPS, FNS, GNS, SPS, TCS, and SF respectively. All images have been resized into 32×32 to maintain uniformity. Next, the generated image data have been split into 80:20 for training and testing of the DL model. The proposed hybrid DL model has been trained with stochastic gradient descent and categorical cross-entropy as loss function. The learning rate, batch size, and the number of epochs have been regulated to 0.0001, 128, and 100 respectively. Besides, the model training validation has also been performed by considering 20% of training samples.

The classification performance has been evaluated by calculating accuracy, sensitivity, specificity, and weighted $F1$-score. The weighted $F1$-score is very much important when all sets

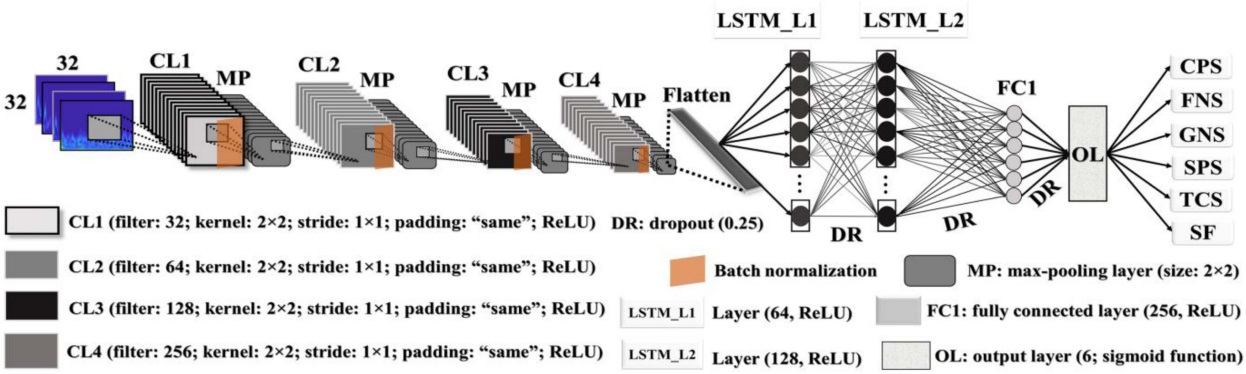

Fig. 5. The proposed hybrid deep learning model to classify different seizure types and seizure-free.

of data are not proportionally available [1], [2], [6]. Besides, it provides a more accurate assessment of incorrectly categorized instances. For performance evaluation of the proposed model, accuracy, A_{cc} (10), sensitivity, S_e (11), and specificity, S_p (12), and weighted $F1$-score, $F1$ (13) have been measured.

$$A_{cc} = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

$$S_e = \frac{TP}{TP + FN} \tag{11}$$

$$S_p = \frac{TN}{TN + FP} \tag{12}$$

$$F1 = \frac{2TP}{2TP + FP + FN} \tag{13}$$

where, *TP* and *TN* depict the true positive and negative respectively, while *FP* and *FN* represent false positive and negative respectively. In addition, the receiver operating characteristics (ROC) analysis has been performed to observe the robustness and methodical capabilities of the proposed model [1], [2], [10], [31], [37]. Actually, the ROC curve shows the trade-off between *TP* and *FP* rates of a model in terms of correct and incorrect identification of positive class respectively [10], [31].

IV. Results

A. Decomposition by HVD

The EEG signals have been segmented into several pieces of 10s length, which have been further decomposed into seven subcomponents by the HVD. For example, the main signal (an EEG segment of Cz_Av channel of CPS) and its decomposed subcomponents — C1, C2, C3, C4, C5, C6, and C7 have been displayed in Fig. 6, in which the horizontal and vertical axes indicate time (s) and amplitude (μV) respectively.

Further, the fraction of energy and correlation (14) of each decomposed subcomponents with their corresponding original signal has been measured separately and the results have been summarized in Table II, in which the first, second, and third columns depict the subcomponents, and their respective energy

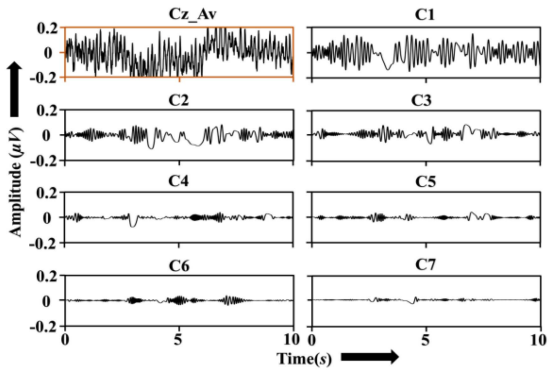

Fig. 6. The seven decomposed subcomponents by the HVD form of an EEG segment of the Cz_Av channel of complex-partial seizure.

TABLE II
Energy and Correlation of Subcomponents

Subcomponents	Energy (%)	C_c
C1	67.25	0.854
C2	11.33	0.386
C3	2.74	0.178
C4	0.55	0.120
C5	0.53	0.083
C6	0.18	0.064
C7	0.15	0.065

(%) and correlation, C_c (14) respectively.

$$C_c(E, F) = \frac{\text{cov}(E, F)}{\sigma_E \, \sigma_F} \tag{14}$$

where, $cov(E, F)$ indicates the covariance; σ_E, and σ_F refer to the standard deviation of E and F signals respectively.

It is seen that the energy and correlation of the decomposed subcomponents with the parent signal are in decreasing order, *i.e.*, the first component preserves the highest energy and has a stronger correlation with the parent signal. In further investigation, the correlation among decomposed subcomponents has been found ≤ 0.02, which is very small. It signifies that the subcomponents have a better correlation with the parent signal than among themselves. Now, from Table II, the first three subcomponents preserve almost 80% energy of the parent signal and have a high correlation. Therefore, the first three

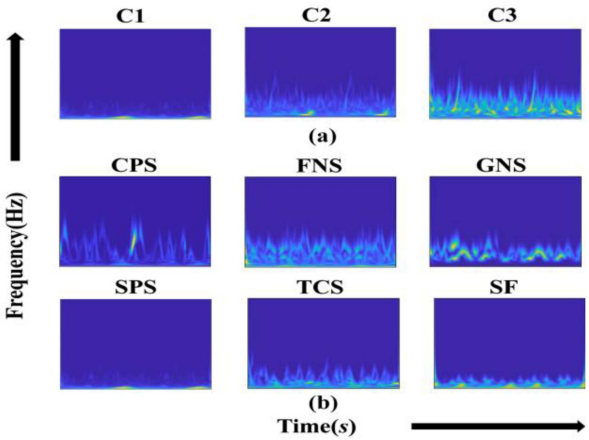

Fig. 7. The 2D images generated from decomposed subcomponents (C1, C2, and C3) of an EEG segment of (a) SPS have been displayed. (b) The images encoded from the C1 subcomponent of an EEG segment of different seizure types (CPS, FNS, GNS, SPS, and TCS) along with SF have been shown.

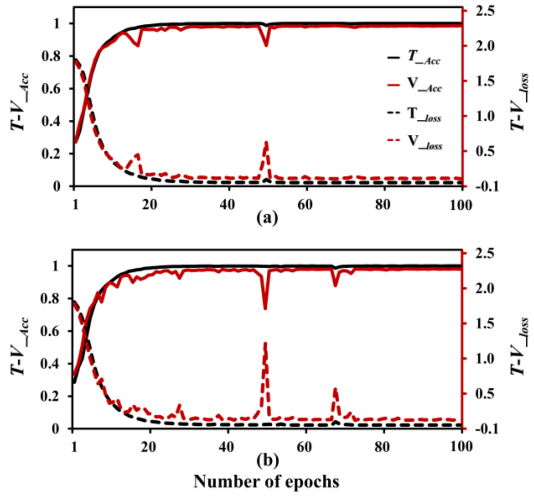

Fig. 8. The recorded $T\text{-}V\_{Acc}$ and $T\text{-}V\_{loss}$ by the proposed model with the number of epochs, when (a) Vertically and (b) Horizontally, stacked images of first three decomposed subcomponents have been used as input.

subcomponents, C1, C2, and C3 have been considered for 2D image generation.

B. Image Construction

Now, the three generated images for each EEG segment have been stacked horizontally and vertically to construct the single image. For example, 2D images generated by CWT of C1, C2, and C3 of an EEG segment of SPS have been displayed in Fig. 7(a). A significant difference among these three subcomponents can be observed, which advocates for the extraction of in-depth features through signal decomposition. Further, 2D images encoded from C1 subcomponents of all five different types of seizure along with seizure-free have also been displayed in Fig. 7(b). Next, images have been stacked and resized (32×32) before being considered as input for the proposed hybrid DL pipeline.

C. Training and Validation

Before classification, the training performance of the model has been verified by measuring training and validation accuracy ($T\text{-}V\_{Acc}$) and loss ($T\text{-}V\_{loss}$) for the different number of epochs and the results have been displayed in Fig. 8(a)(vertically stacked images) and (b) (horizontal stacked images), in which horizontal line refers to the number of epochs and along left and right side vertical lines refer to $T\text{-}V\_{Acc}$ and $T\text{-}V\_{loss}$ respectively. The accuracy and loss have been distinguished by continuous and dotted lines respectively. As seen, overall training performance improves gradually with an increase in the number of epochs. Indeed, the choice of the appropriate number of epochs is very important as it is directly related to computational burden along with overfitting and underfitting issues of the DL model. [2], [8], [9]. Empirically, the number of epochs has been determined as 100 when the $T\text{-}V\_{Acc}$ and $T\text{-}V\_{loss}$ measures become steady, as seen in Fig. 8. Therefore, for all classification tasks, the number of epochs has been set to 100.

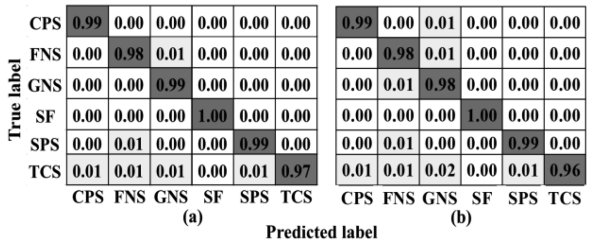

Fig. 9. The recorded normalized confusion matrices when (a) Vertically and (b) Horizontally stacked images, have been used as input for the proposed hybrid DL pipeline.

D. Classification

1) Combining First Three Subcomponents: The proposed idea has achieved the classification accuracy up to 98.82%, when the input images have been structured by vertically stacking, whereas, the accuracy reached up to 98.18% for horizontally arranged images. The results demonstrate that for both cases the classification accuracies are almost the same, but for vertically stacked images, the model has displayed a little better performance. Further, it is important to check sensitivity, specificity, and $F1$-score, as all the samples of different types of seizures are not in the same proportion in this work (see Table I). In addition, the normalized confusion matrix recorded for both the cases has been displayed in Fig. 9(a) and (b) for vertically and horizontally stacked images respectively. In the confusion matrix, the diagonal values depict the number of instances for which the predicted label is the same as the true label, whereas off-diagonal values represent misidentified instances [6]. As seen, the inputs with vertical and horizontal stacked images achieved $F1$-score up to 98.8% and 98.1% respectively, which is very high.

2) Combining Different Subcomponents: Now, it is also important to investigate the classification performance when the different combinations of input images (among C1, C2, and C3) are taken into account. For this purpose, the classification

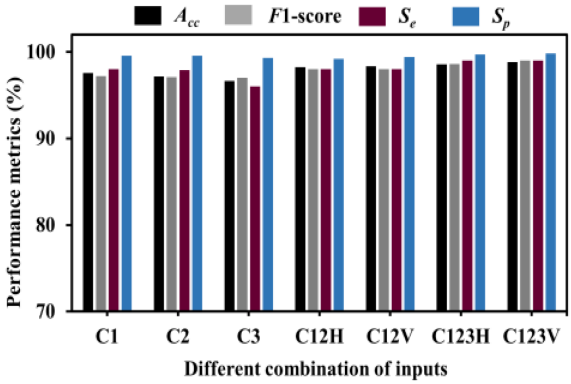

Fig. 10. The A_{cc}, S_e, S_p, and $F1$-score have been recorded by the proposed hybrid DL model when the different combinations of images have been used as inputs.

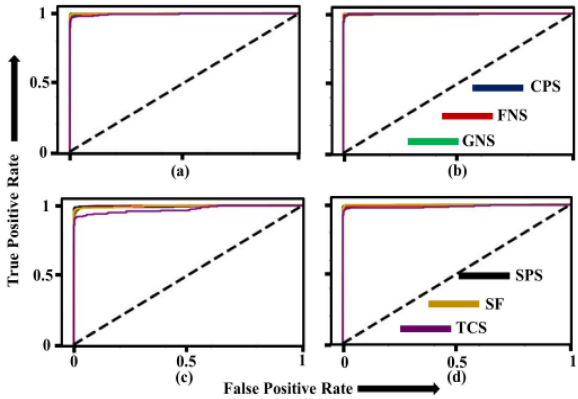

Fig. 11. The ROC curves, when different combinations of images (a) C123V, (b) C123H, (c) C3, and (d) C2 have been used as input to the proposed hybrid DL pipeline.

tasks have been performed by using the C1, C2, and C3 images separately as input. In addition, horizontally (C12H) and vertically (C12V) stacked images of C1 and C2, and horizontally (C123H) and vertically (C123V) stacked images of C1, C2, and C3 have also been used as input. The different performances metrics including A_{cc}, S_e, S_p, and $F1$-score achieved by the proposed idea have been shown in Fig. 10, in which the vertical axis represents performance metrics (%), while the horizontal axis displays different sets of input images.

As seen, the performance metrics have achieved maximum value when images of three subcomponents have been vertically stacked. In addition, the classification accuracy for C1 is better than that of C2 and C3, which reveals that C1 has carried more information. Further, the results have improved significantly when the combination of C1, C2, and C3 has been considered as input. In further analysis, the performance of the proposed model has been examined by observing ROC curves. The ROC curves for different combinations of input images have been plotted in Fig. 11, in which the vertical and horizontal axes display true positive and false positive rates respectively. It is observed that the area under the curves for all classes of seizure types is maximal, which demonstrates the efficacy of the proposed model in seizure types discrimination.

TABLE III
COMPARATIVE STUDY

Works	Input to Classifier		Classifier	ST	PM (%)	
	Type	IGM			A_{cc}	$F1$
Liu *et al.*[5]	2D image	STFT	CNN	8	-	95.5
			RNN		-	95.8
			Hybrid		-	97.4
Asif, *et al.*, [6]	Image	SES	CNN	8	-	94.0
Sriraam, *et al* [11]	2D image	STFT	CNN	8	84.1	-
Ahmedt., *et al.*, [12]	EEG	FFT	P–NMN	7	-	94.5
Wijayanto, *et al* [16]	SF	-	SVM	4	95.0	-
Kassahun *et al.*[17]	MD	-	GB-ML	2	-	77.8
			k–NN		-	90.1
Roy *et al.*, [13]	EEG	FFT	XGB	8	-	86.6
			CNN		-	72.2
Shankar *et al.*,[14]	2D image	GAF	CNN	5	84.2	84.0
This work	**2D image**	**HVD, CWT**	**CNN+ LSTM**	**6**	**98.8**	**98.8**

Note: IGM: input generating methods; PM: performance metrics; ST: number of seizure types, SF: statistical feature, SES: saliency-encoded spectrogram, P-NMN: plastic neural memory network, GB-ML: genetic-based machine learning; MD: medical data. GAF: gramian angular field.

E. Comparative Study

A comparative study has been conducted with the related works, and results have been summarized in Table III, in which the first column indicates the related works followed by their respective methods, machine learning models, number of seizure types, and classification performance metrics. As seen, the proposed idea has been achieved the best performance by achieving the highest classification accuracy along with an $F1$–score.

V. DISCUSSION

The main hypothesis of this work is to decompose the EEG signals into several subcomponents to expose the underlying features which can be classified by DL pipeline efficiently. In this direction, 2D input images for the DL were generated by segmenting the EEG signals, which could benefit in two ways — offering detailed information present in different segments and providing diverse data for successful DL-based classification. In this work, a segment length of 10s has been considered; however, the optimal length of the segment needs to be investigated. Certainly, the length should not be too small to miscue features, on the contrary, not be that long as well to meet the actual purpose of segmentation. Further, the HVD decomposition allows mining the underlying features through its subcomponents. Here, one of the advantages of HVD-based decomposition is its preservation of phase information in all the subcomponents, which plays a very significant role in discriminating different types of seizures. Next, 2D image generation by CWT reliably identifies the time and frequency information of small segments with great details, especially for nonlinear and non-stationary signals like EEG. In this regard, other time-frequency representation techniques can be considered. The classification performed by the DL model demands a large number of diverse data from all types of seizure data, which is always challenging to acquire. Although, it has been dealt with by segmenting the long EEG data as mentioned before, but for proper verification, more diverse data should be incorporated. In the case of the DL pipeline, in this work, CNN followed by LSTM block has been used. Certainly, other DL frameworks can be considered by using the same proposed input image generation mechanism.

VI. Conclusion

In this work, the EEG segments have been decomposed into several subcomponents by HVD. Next, the first three subcomponents have been used for 2D image generation by using CWT. Further, three images have been combined into single image for DL inputs. A hybrid DL pipeline combining CNN and LSTM has been used for feature extraction and classification. The EEG data from the Temple University EEG dataset (TUH) has been used for experimental validation. The proposed method has achieved the classification accuracy up to 98.82% along with the weighted $F1$-score of 98.80%, which is very high. Further, the compatibility of an in-depth feature extraction strategy based on signal decomposition and classification by DL pipeline has been verified by generating 2D input images considering different combination subcomponents. The results reveal that the proposed idea can efficiently extract the underlining features and achieve the highest classification performance. In a comparative study, the proposed method demonstrates its superiority by displaying the highest classification performance.

References

[1] A. Shoeibi *et al.*, "Epileptic seizures detection using deep learning techniques: A review," *Int. J. Environ. Res. Public Health,* vol. 18, no. 11, May 2021, Art. no. 5780.

[2] Y. Roy *et al.*, "Deep learning-based electroencephalography analysis: A systematic review," *J. Neural Eng.*, vol. 16, no. 5, Aug. 2019, Art. no. 051001.

[3] K. Rasheed *et al.*, "Machine learning for predicting epileptic seizures using EEG signals: A review," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 139–155, 2021.

[4] Y. Paul, "Various epileptic seizure detection techniques using biomedical signals: A review," *Brain Informat.*, vol. 5, no. 2, pp. 1–9, Dec. 2018.

[5] T. Liu, N. D. Truong, A. Nikpour, L. Zhou, and O. Kavehei, "Epileptic seizure classification with symmetric and hybrid bilinear models," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 10, pp. 2844–2851, Oct. 2020.

[6] U. Asif, S. Roy, J. Tang, and S. Harrer, "SeizureNet: Multi-spectral deep feature learning for seizure type classification," in *Machine Learning Clinical Neuroimaging Radiogenomics Neuro-Oncology (MLCN,RNO-AI)*, Cham, Lima, Peru: Springer, Oct. 2020, pp. 77–87.

[7] R. S. Fisher *et al.*, "Operational classification of seizure types by the international league against epilepsy: Position paper of the ILAE commission for classification and terminology," *Epilepsia*, vol. 58, no. 4, pp. 522–530, Mar. 2017.

[8] Y. Gao, B. Gao, Q. Chen, J. Liu, and Y. Zhang, "Deep convolutional neural network-based epileptic electroencephalogram signal classification," *Front. Neurol.*, vol. 11, pp. 375–385, May 2020.

[9] C. Gómez*et al.*, "Automatic seizure detection based on imaged-EEG signals through fully convolutional networks," *Sci. Rep.,* vol. 10, no. 1, pp. 1–13, Dec. 2020.

[10] K. O. Cho and H. J. Jang, "Comparison of different input modalities and network structures for deep learning-based seizure detection," *Sci. Rep.*, vol. 10, no. 1, pp. 1–11, Jan. 2020.

[11] N. Sriraam, Y. Temel, S. V. Rao, and P. L. Kubben, "A convolutional neural network-based framework for classification of seizure types," in *Proc. IEEE 41st Annu. Int. Conf. Eng. Med. Biol. Soc.*, Berlin, Germany, Jul. 2019, pp. 2547–2550.

[12] D. Ahmedt-Aristizabal, T. Fernando, S. Denman, L. Petersson, M. J. Aburn, and C. Fookes, "Neural memory networks for seizure type classification," in *Proc. IEEE 42nd Annu. Int. Conf. Eng. Med. Biol. Soc.*, Montreal, QC, Canada, Jan. 2020, pp. 569–575.

[13] S. Roy, U. Asif, J. Tang, and S. Harrer, "Seizure type classification using EEG signals and machine learning: Setting a benchmark," in *Proc. IEEE Signal Process. Med. Biol. Symp.*, PA, USA, Dec. 2020, pp. 1–6.

[14] A. Shankar, S. Dandapat, and S. Barma, "Seizure type classification using EEG based on Gramian angular field transformation and deep learning," in *Proc. IEEE 43rd Ann. Int. Conf. Eng. Med. Biol. Soc.*, Mexico, Nov. 2021, pp. 3340–3343.

[15] I. R. D. Saputro *et al.*, "Seizure type classification on EEG signal using support vector machine," *J. Phys., Conf. Ser.*, vol. 1201, no. 1, May 2019, Art. no. 012065.

[16] I. Wijayanto, R. Hartanto, H. A. Nugroho, and B. Winduratna, "Seizure type detection in epileptic EEG signal using empirical mode decomposition and support vector machine," in *Proc. IEEE Int. Seminar Intell. Tech. Appl.*, Surabaya, Indonesia, Jan. 2019, pp. 314–319.

[17] Y. Kassahun *et al.*, "Automatic classification of epilepsy types using ontology-based and genetics-based machine learning," *Artif. Intell. Med.*, vol. 61, no. 2, pp. 79–88, Mar. 2014.

[18] M. Civera and S. Cecilia, "A comparative analysis of signal decomposition techniques for structural health monitoring on an experimental benchmark," *Sensors*, vol. 21, no. 5, Mar. 2021, Art. no. 1825.

[19] A. Ullal and R. B. Pachori, "EEG signal classification using variational mode decomposition," Mar. 2020, *arXiv:2003.12690*.

[20] P. Khan, Y. Khan, S. Kumar, M. S. Khan, and A. H. Gandomi, "HVD-LSTM based recognition of epileptic seizures and normal human activity," *Comput. Biol. Med.*, vol. 136, Sep. 2021, Art. no. 104684.

[21] B. Büyükçakır, F. Elmaz, and A. Y. Mutlu, "Hilbert vibration decomposition-based epileptic seizure prediction with neural network," *Comput. Biol. Med.*, vol. 119, Apr. 2020, Art. no. 103665.

[22] W. Bomela, S. Wang, C. A. Chou, and J. S. Li, "Real-time inference and detection of disruptive EEG networks for epileptic seizures," *Sci. Rep.*, vol. 10, no. 1, May 2020.

[23] M. Feldman, "Time-varying vibration decomposition and analysis based on the Hilbert transform," *J. Sound Vib.*, vol. 295, no. 3–5, pp. 518–530, Aug. 2006.

[24] S. Barma, B. -W. Chen, W. Ji, S. Rho, C. -H. Chou, and J. -F. Wang, "Detection of the third heart sound based on nonlinear signal decomposition and time–frequency localization," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 8, pp. 1718–1727, Aug. 2016.

[25] A. Emami, N. Kunii, T. Matsuo, T. Shinozaki, K. Kawai, and H. Takahashi, "Seizure detection by convolutional neural network-based analysis of scalp electroencephalography plot images," *NeuroImage: Clin.*, vol. 22, Jan. 2019, Art. no. 101684.

[26] A. Shankar, S. Dandapat, and S. Barma, "Epileptic seizure classification based on Gramian angular field transformation and deep learning," in *Proc. IEEE Appl. Signal Process. Conf.*, Kolkata, India, Oct. 2020, pp. 147–151.

[27] A. Shankar, H. K. Khaing, S. Dandapat, and S. Barma, "Analysis of epileptic seizures based on EEG using recurrence plot images and deep learning," *Biomed. Signal Process. Control*, vol. 69, Aug. 2021, Art. no. 102854.

[28] G. A. Worrell, L. Parish, S. D. Cranstoun, R. Jonas, G. Baltuch, and B. Litt, "High-frequency oscillations and seizure generation in neocortical epilepsy," *Brain*, vol. 127, no. 7, pp. 1496–1506, Jul. 2004.

[29] O. Turk and M. S. Ozerdem, "Epilepsy detection by using scalogram based convolutional neural network from EEG signals," *Brain Sci.,* vol. 9, no. 5, pp. 115–130, May 2019.

[30] A. L. Halberstadt, "Automated detection of the head-twitch response using wavelet scalograms and a deep convolutional neural network," *Sci. Rep.*, vol. 10, no. 1, pp. 1–12, May 2020.

[31] Y. H. Byeon, S. B. Pan, and K. C. Kwak, "Intelligent deep models based on scalograms of electrocardiogram signals for biometrics," *Sensors*, vol. 19, no. 4, pp. 935–960, Feb. 2019.

[32] I. Obeid and P. Joseph, "The Temple University hospital EEG data corpus," *Front. Neurosci.*, vol. 10, pp. 196–200, May 2016.

[33] S. Ferrell *et al.*, "The temple university hospital EEG corpus: Electrode location and channel labels," *Inst. Signal Inf. Proc. Rep.*, vol. 1, no. 1, pp. 1–9, Apr. 2020.

[34] G. Xu, T. Ren, Y. Chen, and W. Che, "A one-dimensional CNN-LSTM model for epileptic seizure recognition using EEG signal analysis," *Front. Neurosci.*, vol. 14, pp. 1253–1261, Dec. 2020.

[35] A. V. Medvedev, G. I. Agoureeva, and A. M. Murro, "A long short-term memory neural network for the detection of epileptiform spikes and high frequency oscillations," *Sci. Rep.*, vol. 9, no. 1, pp. 1–10, Dec. 2019.

[36] X. Liu, J. Jian, and Z. Rui, "Automatic detection of epilepsy EEG based on CNN-LSTM network combination model," in *Proc. 4th Int. Conf. Comput. Sci. Artif. Intell.*, Zhuhai, China, Dec. 2020, pp. 225–232.

[37] Y. Yang, R. A. Sarkis, R. El Atrache, T. Loddenkemper, and C. Meisel, "Video-based detection of generalized tonic-clonic seizures using deep learning," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 8, pp. 2997–3008, Aug. 2021.

[38] M. Shahbazi and A. Hamid, "A generalizable model for seizure prediction based on deep learning using CNN-LSTM architecture," in *Proc. IEEE Glob. Conf. Signal Informat. Process.*, Anaheim, CA, USA, Nov. 2018, pp. 469–473.