IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 24 January 2023, accepted 16 February 2023, date of publication 1 March 2023, date of current version 30 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3251105

TOPICAL REVIEW

AI-Based Epileptic Seizure Detection and Prediction in Internet of Healthcare Things: A Systematic Review

SOBHANA JAHAN[1], FARHANA NOWSHEEN[2], MAHATHIR MAHMUD ANTIK[2], MD. SAZZADUR RAHMAN[3], M. SHAMIM KAISER[3], (Senior Member, IEEE), A. S. M. SANWAR HOSEN[4], (Member, IEEE), AND IN-HO RA[5], (Member, IEEE)

[1]Department of Computer Science and Engineering, Bangladesh University of Professionals, Dhaka 1216, Bangladesh
[2]Department of Information and Communication Technology, Bangladesh University of Professionals, Dhaka 1216, Bangladesh
[3]Institute of Information Technology, Jahangirnagar University, Savar, Dhaka 1342, Bangladesh
[4]Department of Artificial Intelligence and Big Data, Woosong University, Daejeon 34606, South Korea
[5]School of Computer, Information and Communication Engineering, Kunsan National University, Gunsan 54150, South Korea

Corresponding authors: MD. Sazzadur Rahman (sazzad@juniv.edu) and In-Ho Ra (ihra@kunsan.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) Grant by the Korean Government through the Ministry of Science and ICT (MSIT) under Grant 2021R1A2C2014333, and in part by the Woosong University Academic Research in 2022.

ABSTRACT Epilepsy is a neurological condition affecting around 50 million individuals worldwide, reported by the World Health Organization. This is identified as a hypersensitive disease by clinical associations. The unique characteristics of Electroencephalography have proven to be stable and universal; therefore, researchers have a lot of credibilities. So, it is the most used test for Epileptic Seizure detection and prediction. This study examines the contributions that have so far been made utilizing Electroencephalography technology to detect, predict, and monitor Epileptic Seizures. We have reviewed around 56 research articles, and those papers are selected from different academic databases. The studies explored various approaches, including Machine Learning, Deep Learning, and the Internet of Things framework. A comprehensive discussion of different classification algorithms is analyzed, and their performances are explored. Furthermore, various open issues of the stated approach are discussed, and potential future works are addressed.

INDEX TERMS Electroencephalography, brain-computer interface, artificial intelligence, Internet of Things, deep learning.

I. INTRODUCTION

The word "Epilepsy" originated from the Greek verb "epilambanein," which means to seize, or attack. Epilepsy, also known as Epileptic Seizure (ES), is a common neural disorder in which brain activities become abnormal due to sudden changes of electrical impulses in the brain. An uncontrollable seizure is the result of this sudden change in brain activity. Some common symptoms of ES are: jerking at an uncontainable rate, dizziness, tingling, seeing flashing lights, loss of awareness, changes in taste, hearing, smell, and touch. ES patients frequently experience psychic illnesses such as terror, stress, and deja vu. ES may be caused by a brain tumor, a stroke, an acute head injury, a brain infection, genetics, drug toxicity, injury before birth, lack of sound sleep, and a lack of oxygen during birth. It is reported that epilepsy can harm one's quality of life [1]. It is found that it can affect anyone at any age, but it most commonly begins in childhood or over 65 years [2]. These patients may suffer memory loss and emotional problems such as depression and anxiety. They often face disgrace and discrimination from society. People with ES are up to three times more likely to die at a premature age than person without disability. From Harris et al. [3], it is clear that there is also the risk of sudden unexpected death in epileptic seizures (SUDEP). Though the reason behind SUDEP is still unclear, in most cases, the reason behind the death is a sudden

The associate editor coordinating the review of this manuscript and approving it for publication was Md Arafatur Rahman.

fall and hurting the head. The Centers for Disease Control and Prevention (U.S.A) reported yearly 1.16 SUDEP cases per 1,000 ES patients. Patients having uncontrolled seizures have a higher death rate than people who have controlled seizures [4].

Though there is no permanent cure for ES, this disease can be managed by proper medication and treatment. According to the WHO, almost 70% of epilepsy patients can live a seizure-free life if those patients are treated and diagnosed correctly [5]. Early diagnosis and continuous monitoring of seizures can decrease the life risks and ensure a better quality of life. For this diagnosis and monitoring, experts need brain activity signals. Electroencephalography (EEG) [6], [7], Positron Emission Tomography (PET) [8], Electrocorticography (ECoG) [9], Functional MRI (fMRI) [10], [11], and Magnetoencephalography (MEG) [8] are the most essential and common neuroimaging methods. According to research on the diagnostic testing of ES, EEG data-oriented methods are popular by physicians. MRI gives information about brain activity over a longer period of time than standard EEG, which only records problematic patterns such lateralized periodic discharges (LPDs) [12]. EEG uses two non-invasive recording techniques: intracranial (IEEG) and scalp (sEEG). Due to its lower risks and easier recording, the sEEG approach is more frequently employed by neurologists and specialized doctors than IEEG. In addition, the frequency and rhythm of brain activity change during seizures, and these signal recordings are affordable. EEG is therefore frequently employed as the primary signal to identify epileptic seizures [13]. So, considering the huge popularity and the extensive use of EEG, this paper only focuses on the EEG technique for gathering brain signals. EEG signals can be collected with the help of a Head-mounted headset where electrodes are responsible for sensing the brainwaves. A computer-oriented system named Brain-Computer Interface (BCI) is used to analyze and convert EEG data into commands related to an output device for the desired action [14]. By collecting EEG data and converting them into commands, we can easily detect and predict seizures.

Multiple researchers have used ML based prediction and detection methods to address seizures over the years. DL is a cutting-edge technology that can learn patterns more exactly from a large volume of data by classifying them through multiple layered hierarchical structures. For ensuring personalized healthcare, continuous remote monitoring, and emergency treatment, the Internet of Things (IoT) is a wonderful technology. IoT works as a middleware that connects many devices or components and enables communication among them. Researchers are eager to work on IoT-based seizure prediction and monitoring systems these days [15].

This article surveys the literature over the period 2010-2021 on ES prediction, detection, and monitoring-oriented research which included ML, DL, and IoT frameworks. This study mainly focused on narrative systematic review approach. Besides, for better understanding of the outcomes of these articles, some Pie and Bar charts are also presented here. The papers were searched in databases like IEEE Xplore Digital Library, Google Scholar, and Science Direct. For searching articles, the main Key words were EEG or Electroencephalography or ML or DL or IOT or Epileptic Seizure or ES or Detection or Classification or Neural Network or SVM or Random Forest or K-Nearest Neighbor or Prediction. 159 papers were primarily discovered. Once the duplicates were omitted and screening and reviewing the abstract of these papers, 62 papers were eligible and assessed. Here, no use of EEG, no ML based ES detection, Neoro-Fuzzy, Adversarial Network, and cross-bispectrum EEG were the exclusion criteria. After full text review, Thesis and review papers are also excluded. Eventually, 6 research articles were removed after full text assessment. Total of 56 papers were reviewed for this study. Because the conference papers also contain important content and researches, this review work covers both conference and journal literature. Figure 1 represents the paper selection process for this review work through a Prisma diagram.

A. EXISTING REVIEW WORKS

ES has vast literature and to date, several surveys addressing various areas of ES have been published. Song et al. [16] reviewed advancements in automatic medical assistance systems used for EEG-based ES identification. He also addressed several methodologies employed in this sector of research and described their key qualities. Acharya et al. [17] thoroughly discussed several feature extraction approaches as well as the results of various automated epilepsy stage identification techniques. Torse et al. [18] evaluated the results of feature extraction techniques and classification algorithms. The comparison investigates the practical feasibility of implementing a seizure detection technique. Iftekhar et al. [19] presented a comparative review of many new classification and feature selection strategies used in BCI at each level. The application of DL approaches to classify brain signal is also investigated. Abualsaud et al. [20] reviewed numerous classification algorithms and demonstrated the effect of ambiguity in EEG data on classifier accuracy. They constructed a model for dividing the EEG into numerous sub-bands utilizing various transformations such as Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT).

Abbasi and Goldenholz [21] focused on some of the most popular ML methods for Epilepsy prediction. The authors have also discussed some of the challenges that ML techniques face in the field of epilepsy. Researchers and medical experts will be benefited from the knowledge of these methods. Siddiqui et al. [22] reviewed a wide range of these ML classification techniques (non-black-box and black-box) for detecting ES. The authors also analyzed various statistical features. Rasheed et al. [23] conducted a thorough review of the existing literature, highlighting why early identification of ES is necessary and how DL and ML algorithms are employed for ES prediction. Raut and Rathee [24] discussed various ML classification techniques for ES detection. After experimenting with various classification techniques, they

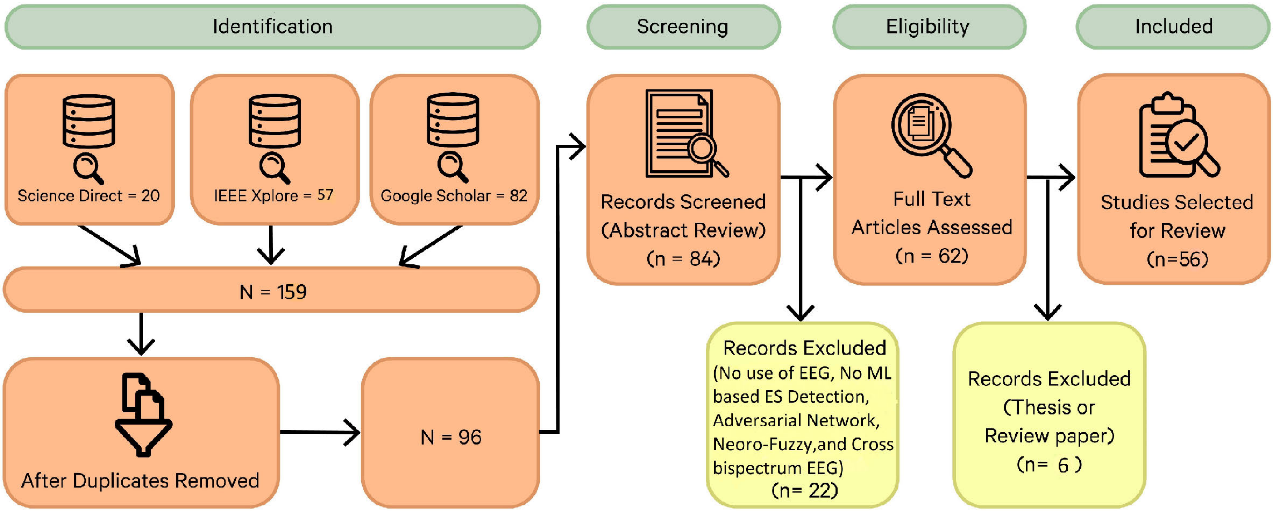

FIGURE 1. Prisma diagram showing the paper selection process for this work. Ultimately, 56 papers were reviewed for this study.

determined that Random Forest (RF) was the best classifier. Shoeibi et al. [25] reviewed various DL techniques for detecting ES. They also presented the pros and cons of each DL models and found the most promising DL model for ES detection. The authors also described challenges and future work. Table 1 presents a clear comparison between our paper and other mentioned review papers. We compared these papers based on four different criteria: feature extraction method, ML classifier, DL classifier, and IoT monitoring.

Feature extraction methods were described by [17], [18], [20], [22], and [23]. Reference [19] partially described feature extraction methods in their work. References [17], [18], [20], [22], [21], [23], and [24] completely dicussed different ML classifiers. Various DL classifiers are partially described by [17]. On the other hand, [19], [23], and [25] presented various DL classifiers completely. Reference [25] partially presented IoT-based ES patient monitoring schemes. After analyzing Table 1 it is clear that there is no review paper that covers all four criteria.

B. MAIN CONTRIBUTIONS OF THIS WORK

To the best of our knowledge, no review paper has thoroughly discussed IoT frameworks in conjunction with ML and DL classifiers for ES detection, prediction, and monitoring. So, here's the unique aspect of our work. The significant contributions of this paper can be summarized as follows:

- The working principles and application areas of the EEG techniques are analyzed.
- A detailed description of the steps and procedures of predicting ES are analyzed along with the common feature extraction methods.
- Comparison between various common ML and DL models based on adaptability, scalability, and interpretability. Besides limitations of every algorithms are also stated.
- Applications of the IoT in e-Healthcare, along with some recent IoT-based models for predicting and monitoring ES are also discussed.
- Detailed analysis of recently published ML and DL models for ES prediction and detection are presented.
- Evaluation metrics and performance analysis of those recently published models are broadly discussed.
- Detailed analysis and comparison of commonly used EEG datasets are presented.
- In-depth analysis of different challenges and future directions in this particular field are addressed.

The remainder of the paper is arranged as follows: Section II contains the EEG technique's working principles and application domains. Section III comes with a ES detection and prediction steps. Section IV contains a discussion about commonly used ML and DL classification algorithms. Section V provides an overview of the IoT in e-Healthcare. Some recently published EEG IoT-based models for ES predicting and patient monitoring are also reviewed. An elaborative discussion and comparison among recently published ML and DL models for detecting and predicting ES are presented in Section VI. Evaluation metrics and relative performance comparisons are depicted in the Section VII. Existing EEG datasets are presented in Section VIII. Section IX contains challenges and future directions in this research field. Finally, Section X contains concluding remarks. The pictorial illustration of the survey structure is shown in Figure 2.

II. ELECTROENCEPHALOGRAPHY (EEG)

The method of EEG, is used to calculate the electrical activity of the brain. The electrical impulses inside the brain and of neurons in the cerebral cortex are captured using tiny metal disks (called electrodes) implanted inside the scalp. Through adding advanced algorithms to the recorded EEG, EEG devices are capable of providing an amount of knowledge that can describe a human's general state. It is used to quantify brain activity that occurs after a particular event, such as the completion of a task or the appearance of a stimulus, or it can be used to measure random brain activity that occurs while there is no specific event. Postsynaptic potentials, or changes

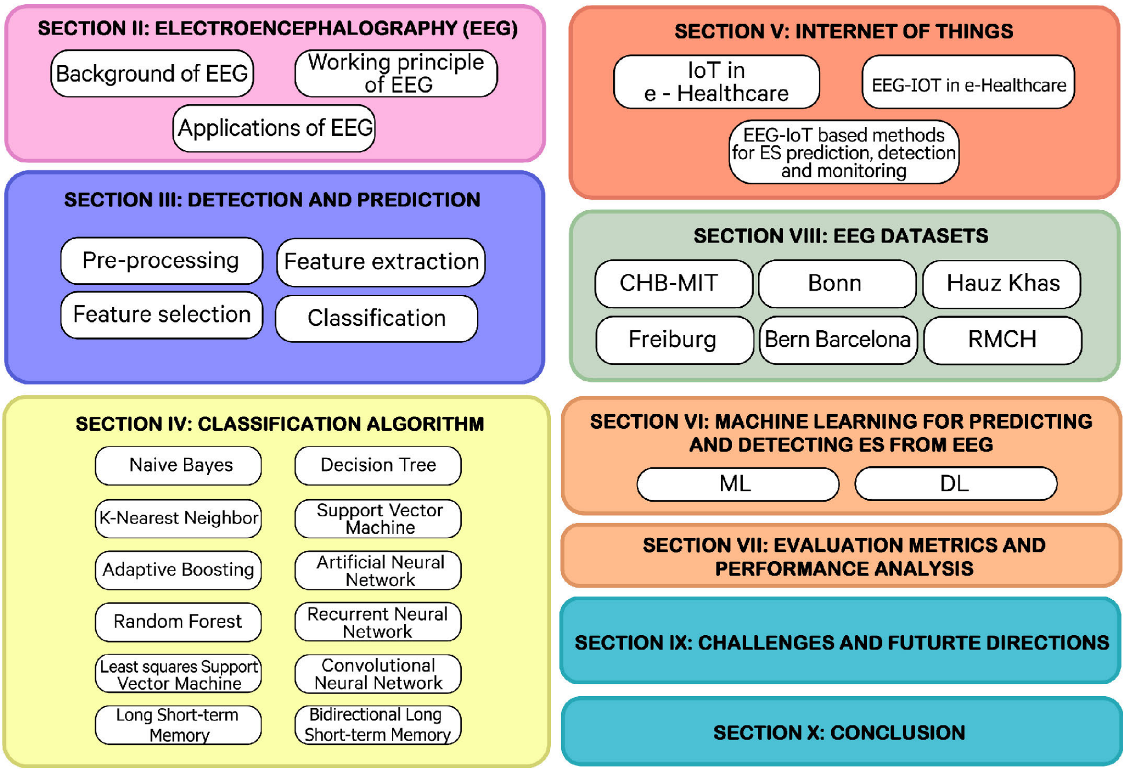

FIGURE 2. Taxonomy of this paper.

TABLE 1. The following table compares the focused area of our paper with other existing review works. legends: ✓ = discussed, X = not discussed, P̃ = partially discussed.

Ref.	Review Topic	Feature Extraction Methods	ML Classifier	DL Classifier	IoT for Monitoring
Acharya et al. [17]	Reviewed several feature extraction approaches as well as the results of various automated epilepsy stage identification techniques.	✓	✓	P̃	X
Torse et al. [18]	Evaluated the results of various feature extraction techniques and classification algorithms.	✓	✓	X	X
Iftikhar et al. [19]	Review of many new DL classification and feature selection strategies used in BCI at each level.	P̃	X	✓	X
Abu al saud et al. [20]	Reviewed numerous classification algorithms and demonstrated the effect of ambiguity in EEG data on classifier accuracy.	✓	✓	X	X
Abbasi et al. [21]	Focused on some of the most popular ML methods. The authors have also discussed some of the challenges that ML techniques face in the field of epilepsy.	X	✓	X	X
Siddiqui et al. [22]	Reviewed a wide range of ML classification techniques (non-black-box and black-box) for detecting ES.	✓	✓	X	X
Rasheed et al. [23]	Reviewed existing literature and highlighted why early identification of ES is necessary.	✓	✓	✓	X
Raut et al. [24]	Reviewed various ML classification techniques for ES detection.	X	✓	X	X
Shoeibi et al. [25]	Reviewed various DL techniques for detecting ES.	X	X	✓	P̃
This work	Reviewed ML, DL, and IoT framework for ES prediction, detection, and patient monitoring. We have also described commonly used feature extraction methods and presented a clear comparison between them.	✓	✓	✓	✓

in membrane potential elicited by neurotransmitters bind to receptors on the postsynaptic membrane, are mainly measured by EEG [26]. EEG signals cater a noninvasive and susceptible measure of brain function throughout the cognitive operations. Some of the notable benefits of EEG are:

- Exemplary fast in analyzing neural activity and data manipulation.
- Commanding reliability.
- Cost effective.
- Practical and convenient to use.

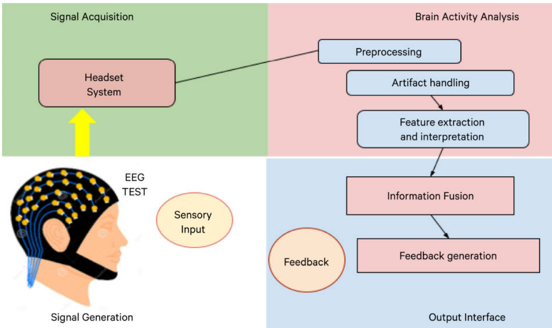

FIGURE 3. **Working principle of electroencephalography (EEG).**

- Qualified in analysing the extensive area of cognitive functions.

A. BACKGROUND OF EEG

The revelation of recordable brain impulses in animals' stimulated nerves and muscles, and subsequently, in their cerebral cortex, gave rise to the field of EEG in the last quarter of the nineteenth century. Hans Berger, a German neuropsychiatrist who is credited with discovering the human EEG, devised the procedure in the late 1920, with the driving philosophy of ''window into the brain''. EEG was originally dismissed by the scientific community, Eventually following the replication of Hans Berger's results by British physiologists Edgar Adrian and Bryan Matthews in the year 1934, EEG became recognized as a non-invasive test of the neuronal activities [27].

B. WORKING PRINCIPLE OF EEG

The functionality of EEG and its application follows a set of steps as shown in Figure 3.

1) SIGNAL GENERATION

An EEG signal is constructed as the various neural activities occur in the brain. With the help of an EEG device that is small metal electrodes connected to the scalp of the head, neuronal activities are collected and hence an EEG signal is generated. The electrical activities of the brain differ from person to person. Consequently, EEG test can distinguish the distinctive cognitive functions. EEG is the optimal choice for the study of brain activities.

2) SIGNAL ACQUISITION

The necessary EEG signal is acquired from the brain. Various methods are recently introduced for this sake. The traditional electrode on the scalp is one of the options. The electrodes are contacted to the scalp using electrolyte gel. This gel is useful as it makes the attachment to the scalp adhesive and eventually improves the signal quality. Another method of signal acquisition is the headset system. There are a number of commercially available EEG headsets. Each system has upper hand in a particular application. The main task is to collect brain electrical signals.

3) BRAIN ACTIVITY ANALYSIS

The first step of analyzing the acquired signals is preprocessing. This includes filtering noise and artifacts. Preprocessing step prepare the signal for further analysis. EEG signals can be re-referenced as the need of application. Next step is artifact handling. In case of EEG, artifacts signify the other signals acquired by the system than brain signals. EEG tests can obtain signals from the atmosphere, equipment, and physiological activities which need to be reduced. Proper analysis is not achieved if these artifacts are associated with the brain signal. Artifact rejection method is facilitated to handle the interfering artifacts. Separating EEG artifacts from the neural signal is the main focus of this method. For EEG feature extraction and implementation, a diverse group of algorithms are introduced. The optimal algorithm depends on the goal of the application. Supervised-unsupervised classification, graph theory, and many other algorithms are used. The algorithms generated are application specific which consequently provides the best result for analysis.

4) OUTPUT INTERFACE

Information fusion results in achievable refined characteristics of the observed system. In case of EEG, this method is necessary but the available algorithms and techniques are not capable of providing sufficient results. The proper implementation of information fusion techniques are application specific. A time and memory efficient algorithm for information fusion is necessary in EEG. The current feedback generation models include large screens, and provide a comprehensive feedback of the study. Smartphones and other mobile devices are potential feedback devices which can display the signal activities in a more articulate and appealing way [28].

C. APPLICATIONS OF EEG

The advancement of technology has opened up a potential new era for researchers around the world to extend the study on human emotion and complex brain processes. Introduction of EEG technology has been manifested in a variety of fields [29]. Some of the application categories are described bellow:

1) NEUROMARKETING

The concept of neuromarketing focuses on the decision making process of consumers. The measurement of consumer's response towards products, advertisement is the base of neuromarketing. Using brain signal activities during the subject's store visit or product purchase, an image of consumer behavior can be drawn [30]. This helps to gain insights of customer motivation and preferances.

2) HUMAN FACTOR

Study of the brain during human-machine interaction is one of the significant EEG applications. The mental state of individuals during activities is evaluated with this. This allows the researchers to recognize degrees of mental workload, stress and emotion during tasks. EEG can evaluate the anxiety level or characteristics of people by analyzing the brain activities while social interaction is occured [28].

3) NEUROSCIENCE AND PSYCHOLOGY

Neuroscience is the study of the nervous system. This focuses on effects of the brain on characteristics, traits and psychological functions. Actions and reactions of the brain as the human attempts activities and mental state during experiencing emotions are studied in neuroscience [31].

4) BCI

A BCI is a digital process that collects impulses from the brains, and contemplate on them. The process further converts them to instructions which are then beamed into an external device in order to perform an intended operation. It's a computer based mechanism as the name suggests. EEG is an optimal technique for recording neural activity signals of the BCI system [32]. BCI uses real time EEG data to implement applications [33]. These data are used to control and regulate devices. It is an emergent domain of EEG which allows operating devices with the help of brain activity. This enables extraordinary and tailored EEG applications [34]. Integration of BCI and EEG elevates a wide variety of applications conducted through neural activity. Application of EEG is seen in the military where personnel can lift and carry heavy weight while equipped with EEG technology. Assistance to disabled people, measuring mental disorders are some of many applications of BCI and EEG amalgamation. With BCI, neural activities are translated into command signals which can steer devices as needed. Some standard BCI applications are as follows:

- Sensor monitoring and management of smart homes.
- Controlling vehicles and other automated devices.
- Supervision of a prosthetic body part.
- Guiding the mobility of an electric wheelchair.
- Mobile applications control through Eyewinks.
- Cursor movement.
- Speech recognition.

III. DETECTION AND PREDICTION OF EPILEPTIC SEIZURE

A standard scheme for predicting and classifying ES using EEG data allows the following steps:

- Pre-processing of EEG data.
- Extraction of features that define the pattern of seizure.
- Feature selection.
- Classification of the selected features to classify the EEG signal (i.e., normal or seizure).

A. PRE-PROCESSING

The pre-processing process extracts noise and undesired artifacts from the acquired EEG signal (e.g., eye movement and muscle artifacts) and prepares it for future signal analysis. Noise and artifacts in raw EEG signals must be identified in the preprocessing step to minimize their impact on the feature extraction process. As EEG is produced by numerous electrodes, it is also critical to determine frequency and channel in this step. Although EEG is designed to record cerebral movement, it also records electrical activity from locations other than the cerebrum. Artifacts are recorded movements that are not of cerebral origin and can be divided into physiologic and extra-physiologic artifacts. Extra physiologic artifacts caused by equipment, the type of EEG (scalp or intracranial), cutoff frequencies, notch filter characteristics, and sampling rate. Body-generated physiological artifacts such as ocular (electrooculogram: EOG), muscle (electromyogram: EMG), and heart rate (electrocardiogram: ECG). Differential window (DW) is applied to EEG signals to deliver more audible signals, making it easier to distinguish seizure from interictal signals. Regression method, Wavelet transform, Principle Component Analysis (PCA), Independent Component Analysis (ICA), EEG source imaging (ESI), EMD, Adaptive, and Wiener filtering techniques are typically used to remove noise and other artifacts from the EEG signals [35]. Therefore, noise removal, filtering, re-sampling are some common pre-processing techniques.

There is a connection between the four stages of a seizure: interictal, preictal, ictal, and postictal. Preictal is a time period that precedes ictal, which is before the seizure prediction horizon and within the seizure occurs period. Postictal describes the state following seizures, while interictal refers to normal stages. Preictal and interictal distinctions are crucial for data segmentation and sampling. One of the popular approaches for data segmentation and sampling is discussed here. Preictal and interictal segments must first be located on the long-term EEG signal in order to be distinguished from one another. Preictal is a time frame prior to an ictal seizure that is influenced by the start time, the occurring period (SOP), and the seizure prediction horizon (SPH). Moving the window sampling method, which has a window length of 30 s and an 8 s overlap to increase the sample set, further divides each preictal or interictal segment into smaller samples. Here, the window size is the amount of time over which a waveform is sampled. So, sample rate and window size are closely related to each other. If the sample rate is 'p' samples per second, then a window size of 'q' samples is 'q × (1/p)' second.

B. FEATURE EXTRACTION

EEG signals can be analyzed in frequency, time or frequency-time domains. Each domain offers an EEG representation, which is needed to analyze and evaluate the EEG data and characterize the identified seizure activity. The feature extraction process intends to extract distinguishing features from the EEG representation. The motive of this step is to characterize the various patterns of seizure activity. With the help of signal decomposition techniques, feature extraction can be easily performed from the EEG input signals. Some common signal decomposition techniques which are used in seizure prediction are described below:

1) TIME DOMAIN

a: STANDARD STATISTICS

The objective of the mathematical transformations or Standard statistics is to gather additional data that the initial signal does not provide. Standard statistics for feature extraction include mean, standard deviation, kurtosis, and skewness.

The mean is the average of the values of the signal's multiple data points. The standard deviation is used to calculate the spread of data values around the mean. Kurtosis is a measure of the number of outliers in a probability distribution. Skewness is a measure of the amount of asymmetry in a probability distribution [36].

b: HJORT PARAMETER

The Hjorth parameter, which has three different types of parameters including Activity, Mobility, and Complexity, is one method for indicating a signal's statistical properties in the time domain. The surface of the power spectrum in the frequency domain can be determined by the activity parameter, which is the variance of the time function. In other words, if there are many or few high-frequency components in the signal, the value of Activity will return a large/small value. The ratio of the variances of the first derivative of the signal and the signal itself is used to define the mobility parameter. This parameter's power spectrum standard deviation is a percentage. The complexity parameter describes how closely a signal's shape resembles a pure sine wave. As the signal's shape resembles a pure sine wave more and more, complexity converges to 1. These three parameters not only aid in the analysis of signals in the time domain but also provide information about the frequency spectrum of a signal. Additionally, by using them, lower computational complexity can be attained [37].

c: HIGHER-ORDER CROSSINGS (HOC)

As time progresses, most of the time series display local and global up and down movements. This can be expressed by calculating the number of zero crossings of the time series x(t). When we apply a filter to any time series, its a number of oscillations change [38]. This in turn changes the zero crossing count. Thus, the technique that can be applied can be stated as: a filter is applied to the time series, and then count the number of zero-crossings; then apply another filter to the original time series, and count the number of zero-crossings, and so on. The Higher-Order Crossings (HOC) are nothing but the resulting zero-crossing counts. The HOC sequence is just the sequence of zero-crossing counts [39].

d: ENTROPY-BASED FEATURE EXTRACTION

Entropy-based feature extraction methods can discriminate between distinct communication signals by defining the distribution state features of the signals. The entropy-based feature extraction is very popular because the internal details of the signals are unimportant and the calculation is quite easy. Some commonly used feature extraction algorithms are discussed below.

Shannon entropy: Shannon entropy was named after Claude Shannon and was first presented in 1948. Since then, it has been most widely used in the information sciences. The Shannon entropy of a random variable is a measure of its uncertainty. Shannon entropy, in particular, quantifies the expected value of the information contained in a message [40].

Rényi entropy: Shannon entropy is the most commonly used method of quantifying information. However, there are others. Alfréd Rényi invented Rényi entropy, which generalizes Shannon entropy and incorporates other entropy metrics as special instances. The Rényi entropy is used as an indicator of diversity in ecology, medicine, and statistics. The Rényi entropy is also useful in quantum information since it may be used to calculate entanglement. The collision entropy is defined as the Rényi entropy for the case α=2.

Fuzzy entropy: Fuzzy entropy is used to calculate subjective data in the context of uncertainty. In order to recognize the EEG signal pattern, two methods, successive and direct segmentation, are introduced.

Transfer entropy: Transfer entropy is a model-free statistic and non-parametric approach that measures the amount of directed transfer of an EEG signal between stochastic variables. As a result, it provides an asymmetric technique of measuring information flow.

Tsallis entropy: The Tsallis entropy distribution is a probability distribution that is produced from maximizing the Tsallis entropy under proper restrictions. Tsallis statistics have been used to study a wide range of phenomena in fields as diverse as physics, chemistry, biology, medicine, economics, and geophysics.

e: INDEPENDENT COMPONENT ANALYSIS (ICA)

ICA is frequently used in EEG analysis at the signal preprocessing stage because of its capability to filter out artifacts from the signal. When recording multi-channel signals, the advantages of using ICA becomes most apparent. By assuming non-Gaussian signal distribution, ICA allows the separation of a mixture of signals into their various sources. The sources are extracted by the ICA by investigating the independence underlying the measured data [41].

2) FREQUENCY DOMAIN

a: FAST FOURIER TRANSFORM (FFT)

J. Fourier discovered the Fast Fourier Transform (FFT) in 1965, following the development of the Discrete Fourier Transform (DFT) algorithm in 1822. Because of the reduction in the looping process, the FFT algorithm calculates transformations faster than DFT. Such transformations are quickly computed by factorizing the DFT matrix into a product of sparse (mostly zero) factors. As a result, it handles to decrease the complexity of computing the DFT from O(n2), which arises if the definition of DFT is simply applied, to O(N log N), at which N is the data size. The speed difference can be huge, especially for large data sets with N in the thousands or millions. To filter signals from the time domain to the frequency domain, the system employs FFT [42]. FFT is truly a promising feature extraction technique. For example, A person's heart rate can be measured using FFT. Photoplethysmography (PPG) signals are required for this. PPG signals are captured using the flashlight camera on an Android phone. When the heart rates calculated by the Android app were compared to the digital blood pressure and heart-rate monitor medical equipment, Sharma et al. [43] achieved an optimistic accuracy of nearly 98%.

b: SHORT TIME FOURIER TRANSFORM (STFT)

When a signal's frequency components change over time, the STFT provides time-localized frequency information, whereas the conventional Fourier transform offers frequency information averaged over the whole signal time interval. Therefore, the STFT method could be a good choice for non-stationary signal analysis. It uses frequency domain windowing to analyze localized signals. The Gaussian window has the best resolution for time and frequency. The STFT's time resolution and frequency resolution are also impacted by the window length [44], [45].

c: AUTOREGRESSIVE MODEL

The Autoregressive model is extensively used in signal processing and system identification. In an AR model, the variable of interest is typically forecasted using a linear combination of the variable's past values. In other words, it's used for predicting when there's a relationship between the values in a time series and the values that come before and after them. The AR model coefficients are used as feature vectors in the brain-computer interface system [46].

d: EIGENVECTOR

Eigenvector techniques are used to estimate signal frequencies and powers from noisy measurements. The eigenvector approaches are based on an eigendecomposition of the noisy signal's correlation matrix. Even when the signal-to-noise ratio (SNR) is low, the eigenvector strategy generates a high-resolution frequency spectrum. These techniques provide enough resolution to estimate sinusoids from data. As a result, in order to gain some noise immunity, it is reasonable to estimate the autocorrelation matrix using only the principal eigenvector components [47].

e: PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA is an unsupervised linear transformation technique that is primarily used for dimensionality reduction and feature extraction. PCA is a technique that tries to find the eigenvectors of the a covariance matrix with the maximum eigenvalues and then uses them to project the information into a new subspace with equal or fewer dimensions [48].

3) TIME-FREQUENCY DOMAIN

a: WAVELET TRANSFORMATION (WT)

The length of the window in the Fourier Transform (FT) limits the frequency resolution. To resolve this issue, WT is invented. Wavelets are explained as ''small'' waves with a shorter period and 0 mean values. Those are mathematical functions that can localize a set or a function in terms of both frequency and time. Because of its appealing properties, such as frequency or time localization and extracting features, the WT is a powerful technique in signal processing. Analyzing Wavelet is classified into two types: Continuous Wavelet Transform (CWT) and DWT. The Wavelet basis function is matched and convolved with the signal in CWT principles refers, the original signal is matched and convolved with Wavelet basis function at a continuous frequency and time increment. Reference [49] stated that a DWT is a transform method that separates the signal into several sets, each of which is a time series coefficients explaining the signal's time evolution in the respective frequency range.

b: EMPIRICAL MODE DECOMPOSITION (EMD)

Huang et al. established a new method for evaluating non-linear and non-stationary data called ''empirical mode decomposition'' in 1998. Any complicated and complex collection of data can be reduced using this model into a limited and often small amount of ''intrinsic mode functions'' that permits well-behaved Hilbert transforms. EMD has limitations in the resolution of the frequency as well as problem in mode mixing. According to [50], if the amplitude ratio exceeds a certain threshold, the frequency result may fall.

c: ENSEMBLE EMD (EEMD)

Wu and Huang [51] proposed a noise-assisted model for analyzing data named EEMD. The EEMD method entails ''sifting'' an ensemble of white noise-added signals. EEMD can naturally divide scales without requiring the selection of previous subjective criterion, as in the original EMD algorithm's intermittence test. White noise is required to drive the ensemble for exhausting all feasible solutions during the sifting phase, causing the various scale signals to collate in the right IMF mandated by the dyadic filter banks. Although EEMD method overcomes the problem in mode mixing, it has certain downsides, including residual noise in the reconstructed EEMD signal and various realizations of the similar input signal yielding a varied number of modes [52].

d: COMPLETE EMPIRICAL ENSEMBLE MODE DECOMPOSITION WITH ADAPTIVE NOISE (CEEMDAN)

To overcome the problems of EED, Torres et al. [53] proposed CEEMDAN in 2011. CEEMDAN is an EEMD and EMD algorithm variant that gives a precise reconstruction of the main input signal. It provides improved separation of the IMFs in the spectral domain.

e: EMPIRICAL WAVELET TRANSFORMATION (EWT)

Gilles et al. [54] developed a decomposition model which removed few limitations of the EMD and named the model EWT. This model is also a successor of classical WT. The EWT approach seeks to identify the oscillatory AM and frequency FM components of a signal, both of which have compact Fourier support. By segmenting the Fourier spectrum, empirical wavelets are created. If various com-

TABLE 2. Advantages and disadvantages of popular feature extraction methods.

Domain	Feature Extraction Method	Advantage	Limitation
Time	Hjort	Removed the risk of over-fitting, lowered the computational cost, and achieved simplicity. The time domain orientation of the Hjorth representation may be appropriate for situations requiring ongoing EEG analysis.	Not appropriate for stationary signal.
	ICA	Useful for removing artifacts from short EEG samples and improving higher-order statistics like kurtosis.	ICA is based on the assumption of non-Gaussianity. If more than one underlying signal property is Gaussian, ICA will not separate these properties
Frequency	FFT	Standard for processing stationary signals, narrowband signal (e.g., sine wave).	Surfers from large noise sensitivity and poor spatial estimation.
	STFT	Standard for processing stationary signals	Based on the Hisenbarg principle, time-frequency resolution is limited.
	Autoregressive	Improved frequency resolution, reduces loss of spectral problems, and never depends on the length of signal.	Tough to select spectral estimation and if the estimation model is inappropriate and model's order is incorrect, it gives bad special estimation.
	Eigenvector	For signals such as sine waves, it provides preferable resolution.	False zero can be generated due to low eigenvalue.
Time-Frequency	WT	The length of the window in the fourier transform limits the frequency resolution. To remove this issue, wavelet transformation is invented. Because of its appealing properties, such as frequency or time localization and extracting features, the WT is a powerful technique in signal processing.	Three significant drawbacks of DWT are poor directionality, less information about phase, and shifting sensitivity.
	EMD	EMD evaluates the nonlinear and non-stationary data. Typically used in natural signals.	EMD has limitations in the resolution of the frequency as well as problems in mode mixing.
	EEMD	EEMD entails sifting an ensemble of white noise-added signals and calculated the mean of the noise signal.	EEMD has certain downsides, including residual noise in the reconstructed EEMD signal and various realizations of the similar input signal yielding various modes.
	CEEMDAN	CEEMDAN provides greater spectrum separation of modes and requires fewer sifting rounds, lowering costs for computation.	The authors did not completely determine the amplitude of the introduced noise and the appropriate ensemble size.
	EWT	EWT overcomes few limitations of EMD. Whenever a low to high-frequency ratio is greater than 0.75, the two elements of a signal could not be distinguished. EWT can overcome this constraint.	By segmenting the Fourier spectrum, empirical wavelets are created. If various components of a signal could not be separated in the Fourier spectrum, The decomposition findings of EWT will be incorrect.
	WPD	The benefit of WPD analysis is that it allows you to combine several levels of decomposition to create the best time-frequency representation.	-
	HHT	The Hilbert transform is useful for computing instantaneous time series properties such as amplitude and frequency. It is a simple and helpful approach for determining the frequency of a signal in real-time.	The excessive envelope will diverge at the endpoints when using HHT, causing substantial error.
	VMD	VMD decomposes the original input signal into distinct band-limited IMFs. VMD has advantages over HHT and WT, such as minimal modal aliasing and sensitivity to noise.	VMD divides the Fourier spectrum into segments to separate various elements of a signal so this approach faces the limitation of the Fourier spectrum; when various elements cannot be distinguished in the Fourier spectrum, they may not be distinguished by VMD.

ponents of a signal could not be separated in the Fourier spectrum, the decomposition findings of EWT will be incorrect [55].

f: WAVELET PACKET DECOMPOSITION (WPD)

Wavelet Packet Decomposition (WPD), often known as wavelet packets, is a wavelet transform that uses more filters than the DWT. Wavelet packets are a linear combination of wavelets. They form bases that retain many of their parent wavelets' orthogonality, smoothness, and localization features. The coefficients in the linear combinations are produced by a recursive technique, with each newly computed wavelet packet coefficient, resulting in minimal computing cost for expansions in wavelet packet bases. Each level of the DWT is computed by passing the preceding approximation coefficients through high and low pass filters. The WPD, on the other hand, decomposes both the detail and approximation coefficients.

g: HILBERT-HUANG TRANSFORMATION (HHT)

The Hilbert-Huang transformation (HHT) is a method that combines empirical mode decomposition with Hilbert spectrum analysis. The empirical mode decomposition adaptively decomposes signals based on their properties into various intrinsic mode functions. The intrinsic mode functions are then transformed into instantaneous frequencies using Hilbert transforms to provide the signal's time-frequency-energy distributions and characteristics. Biological signals, mechanical diagnosis signals, natural physical signals such as earthquake waves, ocean acoustic signals, and winds can all benefit from HHT-based time-frequency analysis.

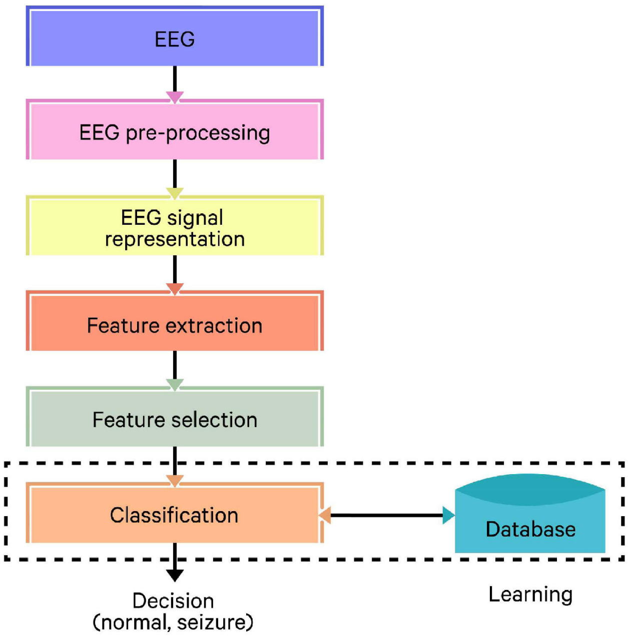

FIGURE 4. Flow chart of epileptic seizure (ES) classification.

h: VARIATIONAL MODE DECOMPOSITION (VMD)

Dragomiretskiy and Zosso [56] suggested a VMD model in 2014. The approach, which is based on the EMD, presupposes that the main signal F is made up of several IMFs [57]. VMD outperforms WT in terms of noise sensitivity and eliminates modal aliasing [58]. VMD can be used combined with other optimization approaches to provide exact separation results in the actual air quality forecasting models [59]. Table 2 clearly says the limitations and main contributions of popular feature extraction methods.

C. FEATURE SELECTION

The features selection is an essential step in lowering classification error rate and computation complexity. The feature selection process refers to removing redundant and irrelevant features or optimizing the feature's small subsection for the specific classifiers. Wrapper, filter, and embedded methods are the three types of feature selection techniques. The classification method is treated as a black box by the wrapper. Filtering methods are used in the pre-processing step to perform a preliminary evaluation of the feature's importance. While wrapper methods select features based on interaction with a classifier, i.e., an underlying model, filter methods are model-independent. An advantage of filters is that they usually require less computational power than wrappers and are more suitable for big data sets [38].

D. CLASSIFICATION

In the classification phase, the classifier assumes the appropriate class (seizure or normal). This assumption takes place based on selected features that came from the feature selection phase. The section VI of this paper contains a detailed description of various classification models. Figure 4 depicts the process of ES classification.

IV. CLASSIFICATION ALGORITHM

Popular classification algorithms can be divided into two major sub parts. One is ML and another is DL. ML is a type of AI that can adjust automatically with little assistance from humans. On the other hand, a Artificial neural networks are used in DL which is a subset of ML and used to simulate the way the human brain learns. Some popular ML and DL algorithms are discussed bellow:

A. MACHINE LEARNING (ML)

1) NAIVE BAYES (NB)

NB is a ML approach used in probabilistic classification. This algorithmic technique predicts varied class probabilities based on distinct criteria. The core of the classifier is based on the Bayesian theory with a substantial implication that attributes are conditionally independent given the class. This classifier has the advantage as it needs a lesser amount of training data. Multinomial, Bernoulli, and Gaussian Naive Bayes are some of the popular types of NB classifiers. NB operates by assuming that the inclusion of some class features has no relevance on the presence of other features. It evaluates each feature separately to compute the feature qualities which influence the classification outcome of any given class. This is a speedier way and, even though it seems to be over-simplified, in many real-world scenarios NB classifiers have been effectively working. One limitation of NB is it makes the assumption that all features are independent, which is rarely the case in practice.

2) DECISION TREE (DT)

DT is a supervised learning algorithm with the added benefit of solving regression and classification problems. The objective is to develop a model which predicts the value of a target variable, in order to resolve the problem in which the leaf node correlates with a class label, the decision tree employs the tree representative. Attributes are represented on the tree's internal node. This is also known as a statistical classifier based on information gain, where the best-standardized profit is the criteria for splitting by producing DTs.

DTs apply several strategies to divide a node into two or more sub-nodes. The emergence of sub-nodes improves the uniformity of the sub-nodes. That is, with relation to the target variable, we may say that the pure node rises. The decision tree divides the nodes on all relevant factors and then picks the split that leads to the most uniform subnodes. C4.5, ID3, CART, and MARS are some of the algorithms to formulate DTs [60]. DT's relative instability in comparison to other decision predictors is one of their drawbacks. A minor change in the data can have a significant impact on the decision tree's structure, which can convey a different outcome than what users would typically receive.

3) RANDOM FOREST (RF)

RF is a supervised learning algorithm. The forest, which is a combination of a set of DTs, is generally trained in "bagging" methods. The main principle of this method is that the total outcome is increased by a mixture of learning models. Every individual tree disseminates a class prediction in the RF, and the class with the most votes becomes the forecast for the model. Instead of looking for the most essential feature when a node is scattered, a random subset of features seeks for the best feature suggested by Breiman in 1999. This covers a wide range which usually leads to a superior model [61]. With a view to increase prediction accuracy while preventing overfitting issues, this classifier flexibly selects the optimizing parameters (i.e., number of DTs formed and greatest depth of the decision tree) [62]. This algorithm may become too slow and ineffective for real-time predictions if there are a lot of trees. But lots of work is going on to remove this barrier.

4) K-NEAREST NEIGHBOR (KNN)

KNN algorithm is a nonparametric, nonlinear and reasonably easy classification method [63]. It works effectively with the bigger training dataset. For the bigger training dataset, it works great. In this approach, data object categorization is done by computing the majority of neighbors' votes, and the object is the class that is most prevalent among the neighbors. KNN is mostly based on similarities such as Euclidean Distance (ED), Manhattan Distance and the other metrics between training and test data sets. The latest samples are allocated to the class for training on the basis of similarity measures based on close K dataset thus the majority vote in the case is determined to classify the case. For K, the ideal value is from 3 to 10. Based on neighboring K training sets, a test data set is allocated to the class [64]. KNN is proved to be slower in execution when the number of predicates increases.

5) SUPPORT VECTOR MACHINE (SVM)

SVM is one of the most commending supervised ML algorithms. Its popularity remains as being used in classifications and regressions problems. This algorithm is useful in solving big data classification problems. The SVM algorithm is designed to build the optimal line or decision boundary to separate n-dimensional space in classes. As a result, the new data item may readily be assigned to the appropriate section. This optimal decision limit is known as a hyperplane [65].

Linear and Non-Linear SVMs are the categories of this algorithm. SVM selects the extreme vectors to assist in building the hyperplane. These extreme vectors are called support vectors and thus the algorithm is named SVM. Whenever two classes cannot be linearly distinguished, the SVMs attempt to determine the hyperplane maximizing the margin while decreasing the amount of misclassification errors proportionate to them. A favorable user-defined parameter controls the balance between margin and misclassification error [66].

6) LEAST SQUARES SUPPORT VECTOR MACHINE (LS-SVM)

LS-SVM algorithm exists as a modification of the SVM algorithm. This can be utilized in classifications as well as in function estimation. Application of equality constraints and least squares are seen to disentangle problems. Utilization of target values is considered contrary to threshold values [67].

7) ADAPTIVE BOOSTING (AdaBoost)

The AdaBoost algorithm is a strategy used as an ensemble method used in ML elongated as adaptive boosting. It is termed as the weights are allocated to each case, with bigger weights for examples which have been incorrectly classified. In supervised learning, boosting is used to minimize bias and change. It is centered on the gradual growth of the learner. Each successor is produced from a formerly developed learner, with the exception of the initial one. In other words, weak learners are transformed into strong ones [68]. As a weak classifier, a decision stump is typically employed in the AdaBoost method as a sub-classifier and these weak sub-classifiers are combined together to build a stronger final classifier.

B. DEEP LEARNING (DL)

1) ARTIFICIAL NEURAL NETWORK (ANN)

The theme of ANN is based upon the computer simulation of the human brain as a response to any incident. It consists of hundreds or thousands of artificial neurons interlinked by processing units. Processing units are the weights and biases which interconnect neurons. The weights and biases are thus modified so that by training an ANN, it creates an outcome significantly analogous to the original result [69]. In general, an ANN has an input layer, a single or hidden layer, for the production of a feature, and a classification output layer.

2) CONVOLUTIONAL NEURAL NETWORK (CNN)

CNNs can successfully capture spatial and temporal dependencies. These use local spatial correlation by implementing a pattern of local connections among adjoining layer neurons. CNN classifies images by detecting lower-level characteristics (such as borders and curves) and then building up a number of convolutional layers into more abstract representations [70]. In comparison to other algorithms, pre-processing is substantially reduced in CNN. While hand-made filters are made in a simple manner, CNNs are able to learn these filters/functions with adequate training. At least four separate layers, including convolutinal, pooling/subsampling layers, completely connected layers and the output layers are included in a standard CNN design. Generally CNN is referred to as Two Dimensional CNN (2D-CNN) although there are two other variations of CNN i.e., One Dimensional

CNN (1D-CNN) and Three Dimensional CNN (3D-CNN) exist in real life applications. 1D-CNN is used for time series data such as data obtained from accelerometer to recognize activity. 3D-CNN is utilized with 3D images for-instance Computerized Tomography and Magnetic Resonance Imaging in order to classify them or for feature extraction.

3) RECURRENT NEURAL NETWORK (RNN)

The NNs with memories that collect all the information saved in succession from the preceding element are RNNs. This means that, because the RNNs carry out the identical duties for each element in the series, with the result being reliant upon all previous calculations, the data is used in a somewhat lengthy sequence. RNN-based algorithms can generally track and store relationships of long-term reliance. In RNN-based learning techniques, the Long Short-Term Memory model (LSTM) as well as the Gated Recurrent Unit (GRU) are common variables. RNN has a "memory," which retains all of the computed information. For each input, it utilizes the same settings as it does for all the inputs or hidden layers to create the output. In contrast to other ML approaches, this decreases parameter complexity. It is often used to handle mundane issues such as language translation, processing, speech recognition, and captioning [71].

Commonly used mobile featured applications such as Siri, google translate use RNNs. By supplying the identical weights and partitions to all layers, RNN turns the independent activations into dependent activations, decreasing the complexity of growing parameters and studies each previous output by supplying an input to the next hidden level for every output.

4) LONG SHORT-TERM MEMORY (LSTM)

Long short-term memory networks, or LSTMs, are employed in deep learning. A variety of RNNs, particularly in problems involving sequence prediction, are capable of learning long-term dependencies. A series of "gates" used by LSTMs regulate how data in a sequence enters, is stored in, and leaves the network. A typical LSTM has three gates: an output gate, an input gate, and a forget gate. Each of these gates is a separate neural network and can be thought of as a filter.

5) BIDIRECTIONAL LONG SHORT-TERM MEMORY (BI-LSTM)

Attempting to make any neural network have the sequence information in both directions—backwards (future to past) or forward—is known as bidirectional long-short term memory (BI-LSTM) (past to future). An extension of conventional LSTMs that can enhance model performance is called a bidirectional LSTM. A bidirectional LSTM differs from a conventional LSTM in that our input flows in two directions. Bidirectional LSTMs train two LSTMs rather than one on the input sequence where all timesteps of the input sequence are available.

Adaptability and scalability are two major concerns of all ML algorithms. IoT-based applications are dealing with a large amount of real-time data. Adaptive machine learning is a more sophisticated solution that prioritizes real-time data collection and analysis. It readily adapts to new information and offers insights almost immediately, as its name would imply. In other words, it is capable of quickly adjusting to new information and understanding its significance. Scalability in machine learning refers to the capacity of ML applications to handle large amounts of data and carry out complex computations quickly and affordably for millions of users worldwide. The scope to which a model can be understood in terms of people is its interpretability. The following Table 3 and 4 shows a comparison between the above-mentioned algorithms in terms of adaptability and scalability and interpretability. The limitations of these algorithms are also mentioned in this table. After reviewing different research articles on ML algorithms it is observed that in the case of adaptability and scalability, very limited works are done. Most of the researchers didn't focus on adaptable and scalable ML algorithms for EEG-based ES detection. Besides, maximum ML algorithms don't have any proper interpretable form. For EEG-based ES detection, no such interpretable form of ML algorithm is developed. Though explainable AI algorithms such as LIME and Shap can be used to interpret the decisions of these ML algorithms. But more emphasis should be given to creating interpretable ML Algorithms.

V. INTERNET OF THINGS (IoT)

The IoT is synonymous with the revolution of technology. It depicts the evolution of information and communication technologies, and its growth is dependent on rapid technological advancement in a variety of fields. The phrase Internet of Things which is abbreviated as IoT is coined by Kevin Ashton, cofounder of the Auto-ID Center in MIT, consists of the terms "Internet" and "Things". The creation of the World Wide Web which is referred to as the internet has connected people worldwide to exchange information, news and views. As per Internet Live Stats, the total number of Internet users worldwide was projected to be 4,990,597,893 as of December 13, 2021. The number of internet users accounts for roughly 40% of the global population. When it comes to the Things, this can be any entity or individual that can be identified in the real world. The term 'Things' is not only electronic devices or technologically advanced tools but "Things" which are around us and not thought to be technological [94]. By 2025, the total number of estimated IoT devices will be 30.9 billion units. There have beed a significant increase from approximately 13.8 billion units in the year 2021 [95].

The unique benefit of IoT and a wide variety of IoT applications enables users to operate in all fields, such as healthcare, home automation, industrial automation, business automation, urbanization, smart agriculture, emergency alerts and disaster recovery. New IoT technology applications allow firms to develop and deploy sophisticated risk management methods, in the form of enhanced operational performance, companies leverage technology.

TABLE 3. Comparison between various ML algorithms.

Model	Adaptability	Scalability	Interpretability	Limitation
NB	NB is quick and can create real-time predictions. Tan et al. [72] proposed Adapted Naive Bayes (ANB) for sentiment analysis which is a weighted transfer edition of the Naive Bayes Classifier, to acquire knowledge from new-domain data.	With a large number of predictors and data points, NB is highly scalable	-	Naive Bayes makes the assumption that all features are independent, which is rarely the case in practice.
RF	Gomes et al. presented the adaptive random forest (ARF) algorithm for categorizing evolving data streams. ARF includes an efficient resampling technique and adaptive operators that can cope with various types of concept drifts without requiring complex optimizations for various data sets, in contrast to earlier attempts to replicate random forests for data stream learning.	To enhance the scalability of RF, Scalable Random Forest (SRF) was proposed by Li et al in 2012 [73]. The SRF algorithm can generate a random forest of 100 trees in less than an hour from 110 Gigabytes of data with 1000 attributes and 10 million records.	It becomes impossible for RF to interpret the non-parametric and the high-dimensional model that is produced.	The algorithm may become too slow and ineffective for real-time predictions if there are a lot of trees. But lots of work is going on to remove this barrier.
DT	Imperfect data streams increase the size of trees and have negative effects on accuracy. The original decision tree algorithm for stream mining performs worse because of the overfitting issue and the unbalanced class distribution. To solve these issues In data stream mining, Yang et al. [74] proposed the Optimized Very Fast Decision Tree (OVFDT) algorithm. It is a significantly improved algorithm because it has a control mechanism for node splitting that is optimized.	Usually, DT doesn't perform well on the large dataset or the scalability property of DT is poor.	-	In comparison with other decision predictors, DT is relatively unstable. A minor shift in the data can result in a significant change in the DT structure.
KNN	Sun et al. [75] proposed an adaptive k-nearest neighbor algorithm (AdaNN) that focused on determining the appropriate number of k for each test example. This proposed algorithm determines the optimal k, or the amount of nearest neighbors that each training example can use to determine its correct class label.	Pablo et al. [76] proposed GPU-SME-kNN, scalable and memory-efficient architecture for a GPU-based kNN that works with a large amount of data. It eliminates the relationship between data size and memory usage while maintaining high performance.	-	As KNN is a distance-based algorithm, the cost of determining the distance between a new point and each original point is very high. It reduces the algorithm's performance. The quality of the data determines the accuracy.
SVM	Murugavel et al. [77] proposed a unique wavelet-based Combined Seizure Index (CSI) features and an innovative Adaptive Multi-Class SVM for the classification of multi-class EEG signals, with a focus on ES detection. The goal was to find the best classification scheme for this problem while also inferring information about the extracted features.	Attempting to solve optimization models for SVM (including parameter fitting) on large-scale training data is a heavy computational task. According to Razzaghi et al., [78] a multilevel algorithmic framework can efficiently deal with very large data sets. Rather than solving the entire training set in a single optimization process, support vectors are acquired and gradually refined at multiple levels of data coarseness. This scalable multilevel framework significantly reduces computational time while maintaining classifier quality.	Barragan et al. [79] presented an Interpretable SVM for Functional Data, which provides a classification result with high predictive power that can also be interpreted. The proposed method's utility is demonstrated in real-world applications by producing interpretable classifiers with comparable, if not better, predictive ability than classical SVM.	SVM doesn't perform well when the data set contains more noise such as the target classes being overlapped and when the amount of features for each data point surpasses the amount of training data samples.
LS-SVM	For trustable communication and interaction, BCI systems need accurate and fast identification of brain activity patterns. Due to the low signal-to-noise ratio in EEG signals and the wide variation of sensorimotor rhythms, achieving this accuracy is difficult. Li et al. [80] proposed a novel self-adaptive least squares twin SVM classifier that integrates a frequency band selection common spatial feature algorithm and a particle swarm optimization technique to address this need.	Least Square Twin Support Vector Machine (LSTSVM) is a speedier variant of SVM, but it has scalability issues and computational storage limitations on large datasets. Prasad et al. [81] proposed a Distributed LSTSVM (DLSTSVM), a scalable solution to LSTSVM. DLSTSVM is built on a cluster of various machines using distributed parallel processing. DLSTSVM trains in distributed parallel after horizontal segmentation on massive datasets. The proposed method achieves storage and computational scalability while maintaining prediction accuracy.	-	LS-SVM is sensitive to noises or outliers and their ability to create an interpretable model is severely limited.

TABLE 4. Comparison between various DL algorithms.

Model	Adaptability	Scalability	Interpretability	Limitation
ANN	Chakrabarti et al. [82] proposed an adaptable method for ES detection that is rooted in ANN for classification and wavelet for feature extraction. This complete process is discussed in Section 6.2.	Saric et al. [83] used a feed-forward multi-layer neural network architecture (MLP ANN) to establish a Field Programmable Gate Array (FPGA)-based approach for the classification of focal and generalized ES types. FPGA solutions based on MLP ANN are portable and scalable, making them extremely useful in the real-time diagnosis of ES in both medical and non-clinical settings. The entire model is discussed in Section 6.2.	ANNs are commonly labeled as black-boxes means lacking in interpretability.	Some limitations of ANN are: training takes a long time, solutions are not guaranteed, and models are prone to be overfitted.
CNN	Due to the scarcity of subject-specific data, the performance boost for DL models has been minimal. To address this, numerous transfer-based approaches have been developed, in which deep networks are trained on pre-existing data from previous individuals and evaluated on new subjects. This technique of transfer learning, however, is hampered by significant inter-subject variability in brain data. Zhang et al. [84] presented five approaches for adapting a deep CNN-based EEG-BCI system for interpreting hand motor imagery. Each strategy fine-tunes and adapts an extensively trained, pre-trained model to improve evaluation performance on a certain subject.	Placidi et al. [85] proposed a fully automatic, effective, rapid, and scalable approach for detecting artifacts in EEG signals recorded as IC Topoplots for use in online BCI. This technique splits Topoplots into four classes: three forms of artifacts and relevant brain signals.	Borra et al. [86] created a lightweight and interpretable shallow CNN (Sinc-ShallowNet) through stacking a temporal sinc-convolutional layer (planned to learn band-pass filters with only the two cut-off frequency bands as trainable sample), a spatial depthwise convolutional layer (minimizing channel interconnection and learning spatial filters attached to each band-pass filter), and a fully-connected layer to complete the classification. This convolutional module reduces the number of trainable parameters while still allowing direct interpretation of the learned spectral-spatial properties via simple kernel visualizations.	CNN fails to encode object location and orientation. CNN requires a large amount of training data to be effective. Because of procedures such as max-pooling, it is substantially slower.
RNN	In order to efficiently address data dependencies, Yazdani et al. [87] presented an intelligent tiled-based dispatching strategy for boosting the adaptiveness of RNN computation. Sharp is a hardware accelerator that streams RNN computation using an effective scheduling strategy to mask the majority of the dependent serialization. Sharp also uses dynamic reconfigurable architecture to adapt to the model's properties.	Paulin et al. [88] developed an RNN accelerator structure for LSTM inference in UMC 65nm technology with a silicon-measured energy-efficiency of 3.25 TOP/s/W and performance of 30.53 GOP/s. Muntaniala's scalable design enables the execution of huge RNN models by mixing several tiles in a systolic array.	Wisdom et al. [89]introduced an interpretable RNN for tackling the sequential sparse recovery issue, which models a sequence of correlated observations with a succession of sparse latent vectors, based on the sequential iterative soft-thresholding algorithm (SISTA).	When applying the activation functions, processing large sequences becomes exceedingly tiresome. It has difficulties such as Exploding and Gradient Vanishing.
LSTM	When working with the sequence elements, the general LSTM models cannot leverage the property of sequential data and ends up with poor performance. In order to create a new LSTM unit, Niu et al. [90] introduced a unique adaptive LSTM for continuous sequential data that takes advantage of the temporal continuity of the input data.	M. Alonso et al. [91] presented a basic framework for processing and forecasting a large number of smart meter time series. Rather than utilizing traditional and univariate methodologies, authors created a single, complicated recurrent neural-network model with long short-term memory that can capture individual consumption patterns as well as consumptions from diverse families.	-	LSTMs take longer to train, need more memory, and are susceptible to overfitting. Furthermore, dropout is far more difficult to apply in LSTMs.
Bi-LSTM	Samavat et al. [92] suggested a hybrid multi-input deep model based on convolution neural networks (CNNs) and bidirectional Long Short-term Memory (LSTM) (Bi-LSTM). CNNs collect time-invariant attributes from raw EEG data, while Bi-LSTM allows for long-range interactions. The authors also used the adaptive regularization method upon every parallel CNN layer to take into account the spatial features of EEG acquisition electrodes.	To investigate the sentiment trend of Chinese texts, Gan et al. [93] developed a scalable multi-channel expanded joint architecture of CNN-BiLSTM model with an attention mechanism. This model can collect both the original context characteristics and the multiscale high-level context features due to its multichannel nature.	-	BiLSTM is a substantially slower model that takes longer to train.

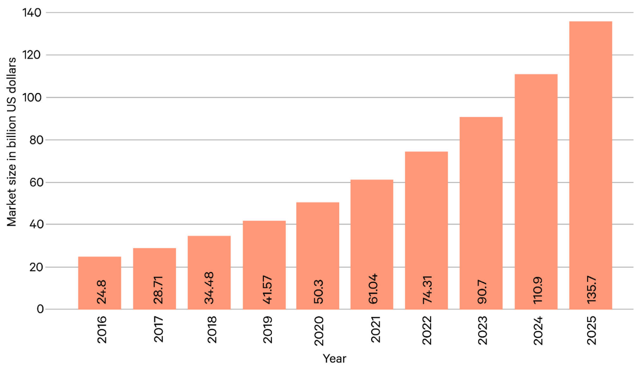

FIGURE 5. Projected market share of IoT in healthcare industry (year 2016 to 2025) [96].

A. IoT IN E-HEalthcarE

To minimize the insufficiency of advanced patient care, a remarkable incorporation of IoT with the healthcare industry is necessitated. The application of IoT and healthcare is immense and plays significant roles. Internet of Healthcare Things (IoHT) is a subset of IoT. IoHT is a connected infrastructure of software applications and medical devices. It allows the exchange of healthcare data with the help of wireless technology. IoHT opens a new dimension providing various possibilities and advantages. It plays an essential role to make the modern healthcare system efficient and robust. IoT adoption in healthcare has substantially transformed the healthcare sector by facilitating interconnectivity among medical devices, patients, and medical professionals. Statistic confirms the healthcare currently accounts for 70% of best-selling IoT devices mostly wearables and smartwears, and by 2026, 40% of all IoT-enabled products will be accounted by healthcare. By 2025, the healthcare industry is expecting to generate more than $135 billion in revenue [96]. Estimated universal market share of IoT healthcare industry is shown in Figure 5. Some of the notable benefits of IoT in healthcare are:

- IoT allows patient supervision on a real-time basis and thereby reduces needless trips to doctors, hospitals, and readmission rates substantially.
- It permits doctors to reach intelligent conclusions based on evidence and offers total clarity.
- Constant patient surveillance and real-time data aid in the early detection of illnesses or even before the symptoms appear.
- Medicines and healthcare device management in the healthcare sector is a major issue. These are efficiently maintained and used at lower costs via IoHT.
- IoT data not only serves to make an appropriate decision but is also responsible for seamless health operations with fewer failures, wastage, and operational expenses.

B. EEG-IoT IN E-HEalthcarE

1) PSYCHOLOGICAL STATE OR EMOTION ANALYSIS

Soundarya [97] proposed a new work that uses EEG to measure psychological state that helps to initialize the parameters of the NN as emotion is hard to understand and measure, but still, suicide is a huge issue on a global scale. The system creates a synchronized wave containing values and details of the patient's emotions. Which finally helps to understand the state of the emotion of an individual. It gives an accuracy of 75.42% (got from softmax classifier), which gives better results when compared to other ML classification algorithms.

Kaur et al. [98] made an investigation where EEG is used to analyze all sorts of impacts of positive and negative emotion. They worked on three types of happiness, emotion, calm, and anger. Real-time EEG signals are recorded from 10 subjects and simultaneously different emotions video clips of 2 minutes each are being watched. Later they extract fractal dimension features from raw EEG. It gets an average accuracy of 60%. The proposed methodology proves it is possible to recognize emotions from EEG signals.

2) PATHOLOGY DETECTION

Muhammad et al. [99] proposed a remote pathology detection system that is based on EEG. A Deep Convolutional Neural Network (DCNN) with 1D and 2D convolutions are used, and fusion is done with features from different convolutional layers. Various networks are investigated, and a publicly available EEG signal is used to experiment. The proposed system achieved accuracy greater than 89% using the CNN and later the (Multi Layer Perceptron) MLP with two hidden layers. It has been evaluated in a cloud-based framework also, and its performance is found comparable with the performance gathered from only a local server. The accuracy rate of this model is satisfactory but not excellent.

3) SLEEPINESS DETECTION

In 2018, a sleepiness detection method based on a BCI headset with three electrodes was proposed by borulkar et al. [100]. The frequency of brain waves is computed using a headset worn on the person's head. Unwanted noises are removed from the received signals. The calculated frequencies are then compared to the brain state's threshold frequencies, allowing a specific judgment such as if an individual is awake or asleep to be made. If a person is drowsy, a specific alert is produced on his or her Android smartphone to wake him or her up.

4) FATIGUE DETECTION

Zhang et al. [101] presented a system that detects fatigue based on monitoring train driver vigilance for high-speed train safety using a wireless wearable EEG. It detects the drowsiness of the driver. The three parts of this system record data, transmit it to pc using Bluetooth. Then it implements a SVM classification algorithm, which determines the level of vigilance. Also, if the fatigue is detected then an early alarm system starts working. The accuracy rate of this model is 90.70%, which is satisfactory but not excellent.

5) DEPTH OF ANESTHESIA MONITORING

For monitoring the depth of Anesthesia intelligently, proactively, and wirelessly, Fernandes et al. [102] proposed an electric Depth of Anesthesia (DoA) monitoring activity in the course of an intraoperative period in a completely simulated environment. The authors had used a micro-controller to collect EEG raw data from the ''Mindwave Mobile Headset'' during the process of sensing. They combined the Arduino with a Bluetooth module because the headset sends EEG signals via Bluetooth and the Arduino did not have communications infrastructure in this method. A three-layered architecture has been proposed by the authors where layer 1 consists of data collection, monitoring, storage, visualization, and monitoring. REST API stays in the second layer and the final layer consists of the cloud platform. To determine the depth of anesthesia, EEG attention levels are measured. When anomalies in attention values are detected, the agents respond by sending text messages to care-providers. The ''SmartDoA-Monitoring'' app takes an average of 3.375 seconds to detect the anomaly and notify the agent if found.

6) PHYSICALLY DISABLED OR IMMOBILE PATIENTS MONITORING

Carrasquilla-Batista et al. [103] proposed an EEG data-dependent wheelchair that is specially customized for physically immobile people. Here, small voluntary eye blinking brain wave data are captured by the sensor-based headset prepared by NEUROSKY, and that raw information is processed and computed by Raspberry Pi single-board computer. The authors claimed that their system can perfectly detect the difference between voluntary and involuntary eye blinks. So, based on those voluntary eye blinks of patients', a wheelchair can do its movements.

Soman and Murthy [104] designed and implemented a BCI-based system for synthesized speech generation that is based on the user's EEG signals. This type of system is especially beneficial for patients with locomotive disorders like locked-in disease, who can share information with their caregivers via this functionality. This solution utilizes the lightweight and easily wearable EMOTIV headset, is designed on an open-source application, and no individual training is needed for the users. By collecting the eye blinking data of five volunteers, the authors trained the system and achieved an offline accuracy of approximately 95%.

Kurapa et al. [105] proposed a home appliance automation system where EEG-based BCI signal is used, those signals are filtered by a hybrid filter and finally can extract the Electromyography (EMG) signals of a physically disabled person. This system collects brain waves from an electrode cap then those data are being filtered and processed in mat data. Further processing is done with the help of a microprocessor and with the help of Bluetooth those data are sent to the home appliances. The EEG signals were obtained through an experiment on eight healthy male subjects aged 18 to 22 years old who were free of any abnormalities. In this experiment, a 29-second video is shown to perform a specific task at a specific time, such as hand movement. The power and time peak generated from the hand movement data of three subjects exist between 21.4 Hz to 21.8 Hz, and 14.1 seconds to 14.69 seconds, respectively. One possible drawback of this model is here, EEG data are collected and analyzed based on the movement of the right hand. This would create a problem for those persons who have difficulties in right-hand movement or may not have the right hand.

C. EEG-IoT BASED EXISTING METHODS FOR ES PREDICTION, DETECTION, AND MONITORING

Hosseini et al. [106] proposed a BCI framework for predicting seizure where DL and cloud computing techniques are used. An optimization of general CNN and Stacked Auto-encoder (SAE) is implemented in this model where, Differential Search Algorithm (DSA), PCA, and Independent Component Analysis (ICA). Sayeed et al. [107] proposed a three-staged seizure detection technique where in the first stage EEG data are decomposed by DWT method. After that, different Hjorth Parameters (HPs) are retrieved from the decomposed data and prominent changes are observed between ictal and interictal data. Finally, KNN classification is performed, final predictions are collected and saved in the cloud. This framework supports the seizure detection technique and offers the communication of acquired results. Data security issues and a complete patient monitoring scheme were not focused on here. A complete seizure prediction and IoT-based remote monitoring framework is proposed by Sayeed et al. [108]. In the proposed framework, DWT, statistical extraction of features, and a NB classifier are used to detect seizures. The patient's medical data can be viewed at any time with the help of the edge-IoT framework. Security of edge devices was not focused on this proposed work.

Alhussein et al. [109] proposed a cognitive IoT-based smart health care framework that can provide timely, assessable, modern health care services at a very minimal cost. In this framework, various physical data are collected from patients body with the help of sensors like EEG, ECG, EMG, accelerometer, oximeter, thermometer etc. With the help of Bluetooth, Zigbee, RFID those collected data are passed to the hosting layer and a Wireless Area Network (WAN) interface will send those data to the cloud. The data is sent to the cognitive engine, which analyzes multimodal sensor data to determine if the patient requires emergency treatment. If the cognitive engine predicts the patient may have a seizure, all those physical data are passed to the feature extraction module and also inform patients seizure tendency to all possible stakeholders. In the feature extraction module, DL techniques are used to predict and detect seizure and sends those data to the cloud database. Based on those results doctors and caregivers can provide necessary healthcare. With the help of smart communication technique, all healthcare service details are shared with all city stakeholders so that patients reports and details can be used in future. This model is not compatible with very large EEG dataset.

TABLE 5. Comparison between various EEG-IoT based models for ES prediction, detection, and monitoring.

Ref.	Model	IoT based detection, prediction, and monitoring	Limitation
Hosseini et al. [106]	SAE, CNN, PCA, ICA, and DSA	A BCI framework is presented using DL and real-time cloud-based IoT techniques.	-
Sayeed et al. [107]	DWT, HPs, and KNN	This framework supports the seizure detection technique and offers the communication of acquired results.	Data security issues and a complete patient monitoring scheme were not focused on here.
Sayeed et al. [108]	DWT and NB	The patient's medical data can be viewed at any time with the help of the edge-IoT framework.	Security of edge devices was not focused on this proposed work.
Alhussein et al. [109]	Cognitive IoT and DL	In this framework, various physical data are collected from patients body, passed to the hosting layer and a WAN interface will send those data to the cloud.	This model is not compatible with very large EEG data.
Sayeed at al. [110]	SRA and VLD	IoMT-based seizure detection. By continuously extracting the hyper-synchronous pulses seizures were detected.	Despite being an IoMT framework, the monitoring of seizure patient is not described in this paper.
Gupta et al. [111]	DWT-SVD and STFT	Here, EEG data were classified in the cloud and necessary information was passed to Doctor. Emergency services can also be offered by this technique.	This paper only focused on security using watermarking. Feature extraction and Classification parts were not focused on here.
Sayeed et al. [112]	DWT, HPs, and DNN	The Proposed model only detects seizure, send EEG data and seizure state in the cloud.	A complete monitoring scheme is not focused on here.
Singh et al. [113]	HOS and RF	Patient's EEG data were passed to the cloud through 4G or WiFi, classification operation is performed and necessary notifications were passed to hospital and family members.	Data security issues were not focused on here.
Daoud et al. [114]	DCNN	An instant alarm is created in the event of an impending seizure, and it is communicated to the doctor and any emergency services that have been selected.	In spite of being a real-time cloud-based system, the data security aspects of this model were not described.

Sayeed et al. [110] proposed a system that continually analyzes brain inputs and extracts hyper-synchronous impulses to detect the starting of seizure. If the number of pulses reaches a predetermined threshold level within a particular time window, a seizure is reported. First, the brain's EEG data are collected; after that, with the help of Signal Rejection Algorithm (SRA) and Voltage Level Detection (VLD) of the signal, a seizure can easily be detected. Despite being an Internet of Medical Things (IoMT) framework, the monitoring of seizure patients is not described in this paper.

IoT in health opens the way for emergency treatment for epileptic patients. Gupta et al. [111] suggested a cloud health IoT system based monitoring paradigm for epileptic patients. To assure data security a DWTs Singular Value Decomposition (SVD) and the Short-time Fourier Transform (STFT) technique can be used for watermarking. The suggested watermarking system, which is founded on DWT-SVD and STFT, was tested on class S and class Z EEG signal. Here, the authors only focused on security using watermarking. Feature extraction and classification parts are not focused on here.

Sayeed et al. [112] proposed an IoT framework to detect the seizure using EEG data. In this approach, HPs and DWT techniques are used. DWT technique will help to decompose the EEG signals and signals will be converted to subbands and features are extracted. In the last stage, DL is used for performing the classification. An approach named hardware-in-the-loop is used in this framework. The proposed model only detects a seizure, send EEG data and seizure state in the cloud. So, a complete monitoring scheme is not presented here.

Singh and Malhotra [113] introduced an autonomous ES detection method and layered architecture for the early identification of ES by combining existing modern communications techniques with cloud computing and ML. This system sends detected EEG data from the patient to the cloud via a WiFi or 4G network. Fast Walsh Hadamard Transform algorithm is employed to handle EEG data in the cloud, and Higher Order Spectra (HOS) are employed to extract statistical and entropy-based characteristics. The RF technique was used to categorize EEG data into seizure phases: ictal, preictal, and normal. Because of the numerous hand-crafted features that must be extracted, existing seizure prediction algorithms are computationally intensive and require a lot of memory to store their parameters. For seizure prediction, Daoud et al. [114] suggested a DL-based IoT platform. Because the classification and feature extraction stages are combined, the computational complexity will be minimized. The Spatio-temporal properties of nonlinear and non-stationary EEG data are extracted using a DCNN model. Table 5 shows recently proposed EEG-IoT based models for ES prediction, detection, and monitoring.

VI. MACHINE LEARNING FOR DETECTING EPILEPTIC SEIZURE FROM EEG

A. MACHINE LEARNING (ML)

1) SUPPORT VECTOR MACHINE (SVM)

In 2016, a multiclass SVM-based seizure prediction method was proposed by Direito et al. [115]. The feature sets, when integrated with multiclass classification as well as post-processing strategies, aim to create warnings while reducing the impact of false positives. Using SVM and Multi-Features of EEG data, Sriraam and Raghu [116] proposed a model to classify seizure data. To understand non-focal and focal ES, multi-features obtained from various domains are used. Subasi et al. [117] proposed a two hybrid SVM approach to detect an ES. Here, Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) helped to detect the optimum parameters for SVM. Those two new approaches open a new door of research on hybrid SVM. But, the high time complexity is observed by those hybrid approaches.

Hamad et al. [118] proposed a hybrid SVM method, where Grasshopper Optimization Algorithm (GOA) was combined with SVM. To ensure a good EEG classification, GOA was used to choose the impactful set of features as well as the suitable SVMs variable settings. This model to detect seizure was a real promising approach that provides good performance in classification. A high-quality ES prediction model using the CNN-SVM model was proposed by Agarwal et al. [119] in the year 2018. CNN has the ability to modify various requirements from different applications. SVM is a computational method described by a segregating ideal hyperplane that is used to categorize new samples. The integration of SVM and CNN has been shown to be an effective method for epileptic prediction. Raghu et al. [120] proposed a novel method to detect an ES. The authors presented a sigmoid entropy-based DWT approach for extracting features. SVM was used here as the classifier. The wavelet coefficients in each sub-band were used to calculate the sigmoid entropy. From this work, it is clear that sigmoid entropy can be a potential biomarker to detect a seizure.

A combination of SAE and SVM is proposed by Siddharth et al. [121] to detect seizures' focal area. The authors used a filter bank named Fourier-Bessel series expansion domain empirical wavelet transform (FBSE-EWT). From multi-channel EEG signals, EEG data are passed to the FBSE-EWT filter, which can separate rhythms. After that, filtrated data are passed to the SAE-SVM network and the classification between focal and non-focal data are observed. Specific rhythms were extracted and selected from each channel EEG data using the FBSE-EWT filter bank with a predetermined sequence range. After the experiment, the authors observed that the FBSE-EWT based rhythms (theta band) coupled with SAE-SVM outperformed other existing models.

2) RANDOM FOREST (RF)

Mursalin et al. [122] presented an ES detection model where the feature selection process is performed by improved correlation method. RF was used as a classifier in this model. To begin the analysis, Improved Correlation-based Feature Selection (ICFS) method was used to identify the most important features from the frequency-time domain and entropy-based attributes. This model outperforms various conventional correlation-based methods. A seizure detection method was proposed by Zhang et al. [123] that combines the SVD, Generalized Stockwell Transform (GST), and a classifier called RF. The unprocessed EEG was converted into a matrix (frequency-time) using GST, and the singular values were retrieved using SVD. After those steps, classification operation took place using RF. The time complexity of this method was not specified. RF typically takes longer to generate the tree.

Wang et al. [124] proposed a model that employs a unique RF-based model in conjunction with Grid Search Optimization (GSO). After normalization Seizure features were visualized using the short-time Fourier transformation. Prior to actually feeding the features into the classifiers, the dimensionality of the attributes was reduced using PCA. Wang et al. [125] also proposed a novel ES state prediction method where Wavelet Packet Features (WPFs) was used to extract features from data and the RF was used as a classifier. For the multi-channel EEG signal, WPFs such as the sub-band ratio (energy) and three wavelet coefficients are retrieved. For preictal phase prediction, the RF is used. Sameer and Gupta [126] proposed a novel automated seizures detection technique using the alpha band (8 Hz-12 Hz). The authors used STFT to convert the frequency domain signal into the time domain. From this time-domain signal, four statistical features were extracted. Mean, variance, skewness, and kurtosis are used as the features of this signal. A total of six classifiers were used in this model and their result comparison shows that the RF classifier provides better results than others. Chakraborty et al. [127] presented an ES detection method using VMD as the feature extraction technique and RF as the classifier. The Kruskal-Wallis test is used to identify significant features, which are then fed into the classifier. Four different classifiers are tested in this experiment and the RF classifier outperformed all others.

3) K-NEAREST NEIGHBOR (KNN)

Rajaguru and Prabhakar [128] presented a modified KNN-based classifier to classify epilepsy. The dimensions of the collected EEG signals were reduced using Power Spectral Density (PSD). The dimensionally decreased values were classified using KNN Dependent AdaBoost Classifier for classifying epilepsy. Ibrahim et al. [129] presented KNN-based classifier to predict epilepsy. In this model feature, extraction was performed by Shannon entropy. The distribution and the complexity of the EEG signals may reveal information about the brain's features and condition. These findings compel us to conduct additional research into the use of entropy as a tool for seizure prediction. With KNN approach, this seizure predictor compares the entropy of every iteration throughout the moving window to analyze the pre-seizure and normal baselines.

For detecting epilepsy using different EEG data, Choubey and Pandey [130] presented a method where some statistical features are compared and for classification comparison, KNN and ANN classifiers were used. Here, Higuchi Fractal Dimension (HFD), Expected Activity Measurement (EAM), and sample entropy those statistical features were used and the best suited statistical feature for the classification was obtained. Akbari et al. [131] presented a seizures detection method where rhythms' phase space is reconstructed in the EWT domain. In this method, rhythmic separation and noise cancellation are performed on raw EEG data. After that, feature extraction has occurred and necessary features were selected with the help of the GA. Finally, classification is performed using the KNN algorithm.

TABLE 6. Comparison between various ML-based models for ES prediction and detection.

Ref.	Model	Main contribution	Limitation
Direito et al. [115]	Multiclass SVM	A novel method to predict seizure.	The sensitivity rate is very poor.
Sriraam et al. [116]	Multi Featured SVM	A novel multi-feature based SVM classifier to classify seizure.	The accuracy of the model was satisfactory but not very high.
Subasi et al. [117]	GA-SVM and PSO-SVM	2 hybrid SVM approach to detect an ES.	High time complexity is observed.
Hamad et al. [118]	GOA-SVM and DWT	A hybrid SVM based model to detect a seizure.	-
Agarwal et al. [119]	CNN-SVM	A high-quality ES prediction using CNN-SVM.	Suffers from time complexity.
Raghu et al. [120]	DWT, Sigmoid entropy, and SVM	A novel approach for ES detection where sigmoid entropy-based DWT was used with an SVM classifier.	-
Siddharth et al. [121]	FBSE-EWT and SAE-SVM	Seizures' focal area detection.	-
Mursalin et al. [122]	ICFS-RF	An ES detection model where the feature selection process is performed by improved correlation method and the random forest was used as classifier.	-
Zhang et al. [123]	GST, SVD, and RF	A seizure detection approach by combining three algorithms.	The time complexity of this model was not clearly mentioned. RF usually takes a longer time to generate the tree.
Wang et al. [124]	RF-GSO	A unique RF-GSO model to detect a seizure.	Highly affected by noise.
Wang et al. [125]	WPFs-RF	Epileptic state prediction model.	Poor accuracy.
Sameer et al. [126]	STFT and RF	The first successful approach to work with the alpha band.	Hauz khans dataset was smaller than Bonn dataset. For conducting such experiments, large datasets are preferred.
Chakraborty et al. [127]	VMD and RF	ES detection using VMD and RF.	-
Rajaguru et al. [128]	PSD, AdaBoost, and KNN	Epilepsy classification by combining 3 algorithms.	The accuracy value was satisfactory not very high.
Ibrahim et al. [129]	KNN	Epilepsy prediction.	Poor accuracy.
Choubey et al. [130]	KNN and HFD	For epilepsy detection, three statistical features and two classifiers were examined, and the best one was found.	-
Akbari et al. [131]	EWT, GA, and KNN	A seizures detection method where rhythms' phase space is reconstructed in the EWT domain.	-
Sharmila et al. [132]	NB and DWT	Comparison between KNN and NB for detecting epilepsy.	-
Xiao et al. [133]	ANBP	Online-based personalized seizure prediction.	The model's accuracy is not very high.

4) NAIVE BAYES (NB)

To detect epilepsy, Sharmila and Geethanjali [132] presented a DWT feature extraction-based model KNN and NB classifiers are used. The result of 14 various epilepsy detection mixtures was investigated using KNN and NB classifiers for the generated statistical features from the DWT. For online-based personalized seizure prediction, Xiao et al. [133] proposed three adaptive predictors. Those are: Adaptive Naive Bayes Predictor (ANBP), and Adaptive Probabilistic Prediction (APP). From long-term EEG recordings, the online pattern learning and forecasting model produced highly impressive prediction accuracy for ten epileptic patients.

5) DECISION TREE (DT)

Martis et al. [134] presented an EMD based automated ES detection method where the C4.5 algorithm is used as the classifier. EMD works as the feature extraction method that decomposes inter-ictal and ictal EEG data. Few IMFs are produced by EMD, which are frequency and amplitude modulated waves. After that, data are passed into the C4.5 that is a DT algorithm. For epilepsy classification, Rajaguru and Prabhakar [135] presented a Soft Decision Tree (SDT) method where Sparse PCA (SPCA) was considered as a dimensionality reduction factor. In this approach first, raw EEG data were sampled and passed to the sparse PCA for reducing the dimensionality then the outputs were sent to SDT for classification. Wu et al. [136] proposed an intelligent classifier for epilepsy detection where a performance comparison between three types of DT algorithm and Multi Layer Perceptron (MLP) algorithm was performed. The three types of DT algorithms are C4.5, CHAID, and CART. Table 6 discusses recently proposed ML-based models for ES prediction and detection.

6) ADAPTIVE BOOSTING (ADABOOST)

Rajaguru and Prabhakar [137] proposed an AdaBoost-based ES detection method in which AdaBoost are used as post-classifiers to improve the model's accuracy. This proposed model is built with the code converter and the AdaBoost classifier, and the main reason for using the AdaBoost classifier is the code converter's poor performance. First, the EEG datasets are collected, followed by the samples. The samples are then encoded and processed using a code converter. Third, an AdaBoost classifier is used as the post classifier, which helps to ensure that a specific weight is maintained throughout the entire training set. Finally, the model's performance is evaluated. One limitation of this model is the accuracy of the model is good but not the best.

There are some models which provide more accurate results than this work.

Al-Hadeethi et al. [138] used AdaBoost LS-SVM to present a two-phased seizure detection technique. To reduce the dimensionality, covariance matrix in the first phase. Here, a feature extraction operation was also performed. In the second step, the AdaBoost LS-SVM classifier can deal with unbalanced data. Hassan et al. [139] proposed a method for resolving the issue of automatic ES detection from single-channeled EEG signals using AdaBoost classifier where CEEMDAN is used for extracting the features. Firstly, segmentation of EEG data takes place and after that, those segmented data are decomposed by the CEEMDAN method. Secondly, from the output of CEEMDAN, inverse Gaussian parameters are extracted carefully. Now, data are ready to be divided into training ad testing sets. Lastly, with the help of the AdaBoost classifier, a classification operation was performed.

B. DEEP LEARNING (DL)

1) ARTIFICIAL NEURAL NETWORK (ANN)

Chakrabarti et al. [82] proposed an adaptable method for ES detection that is rooted in ANN for classification and wavelet for feature extraction. EEG signals were used as the input for the DWT, which was processed using multi-resolution analysis. For feature extraction, four different wavelets, including Daubechies, Symlet, Bi-orthogonal, and Coiflet, were used. The method was validated using a prototype microcontroller-based model to enhance classification accuracy in the shortest amount of time. The prototype circuit is not complex and challenging, and the microcontroller is well-equipped and suitable for working in various environments with various sensors.

For the categorization of focal and generalized ES types that used a feed-forward multi-layer neural network architecture (MLP ANN), Saric et al. [83] constructed a Field Programmable Gate Array (FPGA)-based solution. Scalable and portable FPGA systems based on MLP ANN are very useful for real-time ES diagnosis in both biomedical and non-clinical settings. 822 recorded signals from the Temple University Hospital Seizure Detection Corpus (TUH EEG Corpus) dataset are used to train, verify, and test the ANN algorithm. The system took five main features as inputs, which were extracted from EEG data using time-frequency assessment, Continuous Wavelet Transform (CWT), and statistical methods. 583 which means that 70% of the entire sample was used for developing the system in MATLAB and TensorFlow, and 30% of the samples which is 239 samples were also used for later testing of the model's result on the FPGA. After that, k-Fold Cross-Validation was used to find the ANN model's appropriate parameters. Ultimately, the highest-performing ANN model for real-time seizure classification was applied to the FPGA in terms of average validation data accuracy obtained during cross-validation.

Guo et al. [140] proposed an automated ES detection technique where ANN is used with WT technique for decomposition. The abilities to combine line length features with an ANN to identify ES in EEGs were investigated. Since this dataset was pre-processed by erasing artifacts through visual observation, extensive testing in real-world clinical settings is required. Juárez-Guerra et al. [141] presented a novel method to identity epileptic patient and the normal person using DWT and ANN. First, noisy data were removed then data were passed to the DWT. Here, DWT is used for extracting the features of the EEG data. After the feature extraction, data were passed to the Feed-forward ANN (FFANN). Ambulkar and Sharma [142] proposed a five-stage ES detection method. First, the data optimization was performed using S-transform then time-frequency representation of EEG segments using window width occurred, after that the PSD was calculated next data features were extracted, and finally, classification using ANN is performed.

2) CONVOLUTIONAL NEURAL NETWORK (CNN)

A DCNN of 13 layers was proposed by Acharya et al. [143] to detect seizure in an efficient way. Feature selection and extraction steps are not mandatory in this method. One limitation of this work is, it requires a lot of datasets as it is a DL technique. Wei et al. [144] proposed a 3D-CNN algorithm to classify seizure from multi-channel input signals. In the first step, EEG data are converted into 2D images then all images are converted into 3D images. Lastly, a CNN model is responsible to classify those images into different seizure stages.

Zhou et al. [145] proposed a CNN-based ES detection method where the authors avoided manual feature extraction. Huang et al. [146] proposed a novel seizure detection method with the help of attention-based CNN-BiRNN. This model is constructed using three steps. the first step consists of a multi-scale convolution model, an attention-based model constricts the second step, and a multi-stream recurrent bidirectional algorithm completes the final step. This model can also deal with EEG signals that have missing or different channels.

Jana et al. [147] proposed a combined method to detect the seizure. Here spectrogram and 1D CNN are combined together to get a better and efficient result. One drawback of this model is the accuracy value of this model is poor. To detect ES offset and onset, Boonyakitanont et al. [148] proposed a CNN-based model where EEG signals act as the input signal. To capture Spatio-temporal patterns separately a filter is factorized here. This model can successfully detect the onset and offset seizure. For classifying seven different variants of seizure, Raghu et al. [149] proposed a unique transfer learning and CNN-based method. The method has been evaluated by using the transfer learning algorithm, and the features of images were extracted by utilizing ten pre-trained networks.

To decrease false alarms during seizure detection, Takahashi et al. [150] merged a CNN that prepared images of EEG plots with patient-specific autoencoders (AE) of

TABLE 7. Comparison between various DL-based models for ES prediction and detection.

Ref.	Model	Main contribution	Limitation
Chakrabarti et al. [82]	ANN and DWT	An adaptable method for ES detection that is rooted in ANN for classification and wavelet for feature extraction.	-
Saric et al. [83]	MLP ANN	A Field Programmable Gate Array (FPGA)-based solution. Scalable and portable FPGA systems based on MLP ANN are very useful for real-time ES diagnosis in both biomedical and non-clinical settings.	-
Guo et al. [140]	ANN and WT	The abilities to combine line length features with an artificial neural network to identify ES in EEGs were investigated.	Since this dataset was pre-processed by erasing artifacts through visual observation, extensive testing in real-world clinical settings is required.
Juarez et al. [141]	DWT and FFANN	Identification of epileptic patient and the normal person.	The final weight and bias values of the neural network were not mentioned.
Ambulkar et al. [142]	ANN and Optimized S-Transform	Five-stage ES detection.	-
Acharya et al. [143]	Deep-CNN	Automated Seizure detection.	Model accuracy was poor. Requires a lot of datasets as it was a DL technique.
Wei et al. [144]	3D-CNN	Automatic Seizure stage detection from multi-channel input signals.	Modal validation analysis was missing.
Zhou et al. [145]	CNN	Seizure Detection.	For DL-based classification, a huge dataset was needed.
Huang et al. [146]	CNN-BiRNN	Automation seizure detection method.	Because of the imbalanced number of negative and positive samples, the accuracy score was unsuitable for testing the model.
Jana et al. [147]	Spectrogram and 1D CNN	A new approach to detect seizure.	Poor accuracy rate.
Boonyakitanont et al. [148]	CNN	Onset and offset seizure detection.	Less emphasization on epoch-based classification method.
Raghu et al. [149]	Transfer learning and CNN	Classifying seven different variants of seizure.	The accuracy rate was not very high.
Takahashi et al. [150]	AE-CNN	Patient-specific AE of EEG signals to decrease the false alarms at the time of seizure detection.	-
Wen et al. [151]	AE-CDNN	Performs unsupervised feature learning from EEG in epilepsy.	The proposed model performed badly with a lower feature dimension.
Vidyaratne et al. [152]	DRNN	ES detection.	No clear information about the used activation function.
Hussein et al. [153]	ESD-LSTM	Detection of ES with High Accuracy.	This model is not suitable for multi-channel EEG data.
Abbasi et al. [154]	LSTM	Epilepsy detection.	Performance comparison with multiple ML and DL methods were missing.
Aliyu et al. [155]	RNN	Epileptic EEG classification.	Performance comparison with other Deep learning approaches was missing.
Daoud et al. [156]	Bi-LSTM and DCAE	Early seizure prediction.	Gathering the full data sequence before starting the prediction can create a huge concern, as this model was suitable for a real-time application.
Baskar et al. [157]	EEMD and LSTM	An ES detection technique using Akima Spline Interpolation oriented EEMD where LSTM is used as classifier.	Sensitivity and specificity scores are satisfactory but not very high.
Liu et al. [158]	Deep C-LSTM	Tumor and ES detection.	A larger dataset is needed for training the model.
Hu et al. [159]	Deep Bi-LSTM	Seizure detection.	Requires huge chunk of data.

EEG signals. Seizures and artifacts were both automatically recorded by the AE. They built a CNN with three output classes (seizure, non-seizure-but-abnormal, and non-seizure) based on seizure logs gathered by expert epileptologists and AE errors. A seizure alarm was issued based on the total number of consecutive seizure labels. The deep convolution network and autoencoder-based model, known as the AE-CDNN, was proposed by Wen and Zhang [151]. To do unsupervised feature learning from EEG in epilepsy, this was built. Based on two publicly available EEG data sets, the authors extracted features using the AE-CDNN model and classified the features (Bonn and Boston). According to experimental findings, features derived by AE-CDNN perform better in classification than features obtained by principal component analysis and sparse random projection.

3) RECURRENT NEURAL NETWORK (RNN)

A deep RNN (DRNN) based epilepsy detection method is proposed by Vidyaratne et al. [152]. In addition, one mapping technique for efficient signal processing is also presented by the authors. This mapping helps to learn the Spatio-temporal features from raw EEG signals. Hussein et al. [153] proposed an ESD-LSTM method for detecting ES with high accuracy. This method can understand high-level EEG depictions and differentiate between seizure and normal EEG activity. An LSTM-based epilepsy detection approach was proposed by Abbasi et al. [154]. This classifier classifies interictal, preictal, and ictal data very efficiently. Even Binary classification was also performed by this model. Single-layered and double-layered memory-based LSTM architecture was developed by authors and they found double-layered memory-based LSTM gives better performance than single-layered one.

Aliyu et al. [155] proposed an RNN model to classify epileptic EEG data. Here DWT is used for extracting the features and after that those data are sent to the classifier. The authors compared their RNN with other ML model and found their model outperforms all the others. Daoud and Bayoumi [156] proposed a combined model to predict seizure where Bi-LSTM, and Deep Convolutional Autoencoder (DCAE). The authors compared Bi-LSTM and DCAE with other 4 approaches and stated, Bi-LSTM and DCAE outperformed other methods. The Bi-LSTM extracts temporal information from raw EEG data and DCCN is responsible for learning spatial data. The transfer learning method was used to explore DCAE-based semi-supervised learning strategies, which resulted in a reduction in the training period. The authors reported that this model is suitable for real-time applications. But, gathering the full data sequence before starting the prediction can create a huge concern, as it is a real-time application.

Baskar and Karthikeyan [157] proposed an ES detection technique using EEMD as the feature extraction method and LSTM as the classifier. Using the Akima Spline Interpolation technique to decompose the signal, they have successfully found the intrinsic mode function. After that signal was decomposed by EEMD and Kalman filter helps to remove the Gaussian noise. Then this signal is feed to an LSTM classifier. Liu et al. [158] proposed a tumor and ES detection technique using the Deep C-LSTM method. First, the input data are passed through the DCNN layer after that an LSTM layer got those signals and an Adam optimizer is used to optimize those signals. Data will pass to the dropout layer and lastly, the softmax activation function was used. A larger dataset is needed for training the model. Hu et al. [159] proposed a Bi-LSTM network to detect seizure where the local mean decomposition technique is used to reduce the computational complexity. Here two opposite directional LSTM networks are combined together. On major benefit of this model is it can utilize the benefits of both after and before stages. Table 7 discusses recently published DL-based models for ES prediction and detection.

VII. EVALUATION METRICS AND PERFORMANCE ANALYSIS

A. EVALUATION METRICS

Classification models are notorious techniques and widely used. Evaluating a model is a core part of building a practical ML model. The performance of a model is described by evaluation metrics. The capacity of the metrics to discern between model outcomes is a key feature. The goal is not only to create a prediction model but also developing and selecting a model that provides high accuracy on out-of-sample data. As a result, it is critical to validate the model prior to computing anticipated values. Out of several metrics, confusion matrix is a prominent example. A confusion matrix generates a matrix that describes the model's overall performance. It is an $N \times N$ matrix, where N represents the total of classes being predicted. Some of the essential definitions related to confusion matrix are:

1) **Accuracy:** Accuracy is the most prevalent metric for evaluating classification models. This metric computes the proportion of correct predictions to total instances analyzed. Best use case is when the data are balanced.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

2) **Precision:** The precision metric is used for measuring the positive patterns accurately predicted in positive classes from the total predicted patterns. It is the percentage of positive instances detected accurately.

$$Precision = \frac{TP}{TP + FP}$$

3) **Recall or Sensitivity:** The sensitivity metric is used to calculate the percentage of correctly categorized positive patterns.This is the percentage of true positive instances that are appropriately detected. Best use case is when the data are imbalanced like the disease detection data-set.

$$Sensitivity = Recall = \frac{TP}{TP + FN}$$

4) **Specificity:** In contradiction to sensitivity metric, specificity quantifies the percentage of negative instances that are appropriately detected. The fraction of true negative cases detected accurately has its best use case when data are imbalanced as well as the minority class is a negative class [160].

$$Specificity = \frac{TN}{FP + TN}$$

Here, True Positive (TP) means the observation is true so is the prediction. True Negative (TN) means observation and prediction both are negative. False Negative (FN) stands for prediction being negative while the observation was a positive case. In the case of False Positive (FP) which is, the observation is negative but the prediction was positive Lever et al. [161].

B. PERFORMANCE ANALYSIS

1) INTERNET OF THINGS (IOT)

Hosseini et al. [106] developed a DL and cloud computing enabled ES prediction technique. The accuracy of optimized CNN and optimized SAE were 96% and 94%, respectively. Sayeed et al. [107] developed a three-staged seizure detection technique using Bonn datasets. Dataset 'A' contains EEG recordings from five healthy volunteers with their eyes wide open. Intracranial EEG data were recorded from the patients during ictal and interictal stages, in datasets 'E' and 'D', respectively. The authors claimed a classification accuracy of 100% for ictal vs. normal EEG; and 97.9% for ictal vs. normal and interictal EEG data. Authors had also developed an IoMT-based platform for effectively predicting seizures. ThingSpeak, SimulinkR, and off-the-shelf microcontrollers

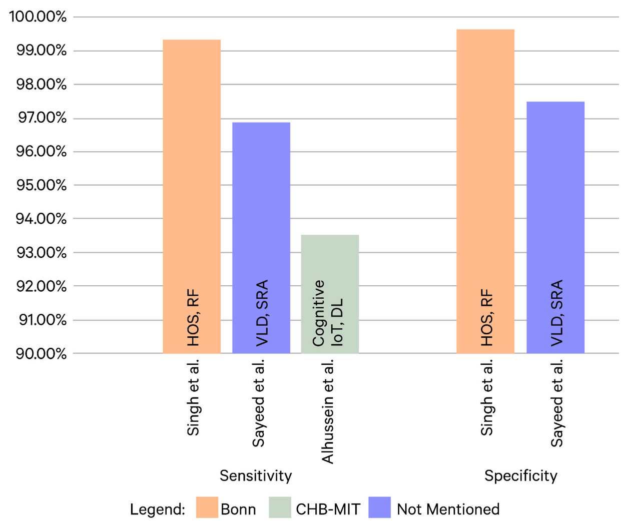

FIGURE 6. Comparison between various IoT models for epileptic seizure (ES) detection, prediction, and epileptic patient monitoring based on sensitivity and specificity.

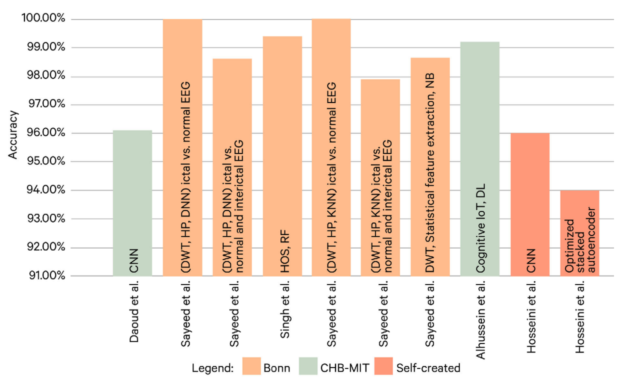

FIGURE 7. Comparison between various internet of things (IoT) models for Epileptic Seizure (ES) detection, prediction, and epileptic patient monitoring based on accuracy.

were used by Sayeed et al. [108] to develop and validate a complete seizure prediction and IoT-based remote monitoring framework. The proposed solution reduces time latency by 44% compared to typical cloud IoT systems and has an accuracy of 98.65% while performing the classification.

Alhussein et al. [109] developed a cognitive IoT-enabled smart health care framework for epileptic patients and used the CHB-MIT dataset that is collected from the Children's Hospital situated in Boston. This dataset is made with the help of 23 volunteers. All of them are epileptic patients, and their ages are between 10 to 22 years old. The sensitivity and accuracy rates of this framework are 93.5% and 99.2%, respectively. Sayeed et al. [110] proposed an IoMT-based seizure detector that has a specificity of 97.5% and sensitivity of 96.9%. Gupta et al. [111] developed a cloud-enabled IoT system for epileptic patients and used a publicly available dataset given by the University of Bonn. The findings reveal that this method performed well in watermarking, with 35.25 Peak Signal to Noise Ratio (PSNR) and 31.32 Signal to Noise Ratio (SNR).

Sayeed et al. [112] developed an IoT framework to detect the seizure using EEG data. Using the Bonn dataset, the authors found a 100% accuracy for ictal vs. normal EEG classification and 98.6% accuracy for ictal vs. interictal and normal EEG classification. Singh and Malhotra [113] reported 99.40% accuracy, 99.66% specificity, and 99.40% sensitivity rate using the RF classifier in an autonomous ES detection and early identification method. Daoud et al. [114] developed a DL-based IoT platform for ES prediction and used the CHB-MIT dataset in this experiment. EEG signals from 22 pediatric epileptic patients were gathered for this dataset. This system's accuracy, sensitivity, and specificity rates were 96.1%, 97.41%, and 94.8%, respectively. This model can be implemented in a wearable device.

Performance analysis of various IoT-based models for ES prediction, detection, and epileptic patient monitoring are presented in Table 8. Figure 6 and 7 represents the performance comparison between various IoT-based models for ES prediction, detection, and epileptic patient monitoring based on sensitivity, specificity, and accuracy, respectively. From Figure 6 and 7 it is clear that [107] and [112] presented two models that achieved 100% accuracy. Reference [113] reported the highest specificity and sensitivity score, which was 99.66% and 99.4%, respectively.

2) MACHINE LEARNING (ML)

Agarwal et al. [119] used the American Epilepsy Society (Kaggle) EEG dataset for establishing a high-quality epileptic seizure prediction model using the CNN-SVM. The model achieved an accuracy of (97.86±10.5)%, specificity (98.81±10.5)%, and sensitivity (96.47±10.5)%. RMCH, Bonn, and CHB-MIT datasets were used by Raghu et al. [120] for developing a novel method to detect an epileptic seizure using DWT and SVM. From the RMCH dataset, the authors found an accuracy of 96.34%, a detection delay (mean) of 1.2s, and a false detection rate of 0.5 per hour. The CHB-MIT, and Bonn and databases scored the highest sensitivity of 94.21%, and 100%, respectively. Siddharth et al. [121] presented a hybrid model to classify focal and non-focal seizure areas. The authors used SAE-SVM and FBSE-EWT filter bank and achieved a 100% accuracy, sensitivity, and specificity score. This experiment is performed on Bern-Barcelona dataset. For developing an ES detection model using ICFS, the Bonn dataset was used by Mursalin et al. [122]. They found 100% (case A-E) accuracy from the ICFS-RF model.

A combination of SVD, GST, and RF was presented by Zhang et al. [123] for ES prediction. Using the Bonn dataset, this method successfully predicts ES with an accuracy of 99.63%. This method outperformed various typical methods. Wang et al. [124] used the Bonn dataset, and the RFGSO classifier has an AUC of 99.0%, signifying near-perfect output. An accuracy score of 84.8% was reported by Wang et al. [125] using WPFs-RF to predicts seizure states. Here a benchmark dataset named CBH-MIT was used. Sameer and Gupta [126] presented a new approach for detecting seizures

TABLE 8. Performance comparison between various internet of things (IoT) based models for epileptic seizure (ES) detection, prediction, and epileptic patient monitoring.

Ref.	Model	Dataset	Performance
Hosseini et al. [106]	SAE, CNN, PCA, ICA, and DSA	Self-made dataset (11 volunteers participated)	Accuracy 96% (Optimized CNN) and 94% (Optimized SAE).
Sayeed et al. [107]	DWT, HP, and KNN	Bonn	Accuracy 100% (ictal vs. normal) and 97.9% (ictal vs. normal and interictal).
Sayeed et al. [108]	DWT and NB	Bonn	Accuracy 98.65%.
Alhussein et al. [109]	Cognitive IoT and DL	CHB-MIT	Accuracy 99.2% and sensitivity 93.5%.
Sayeed et al. [110]	SRA and VLD	-	Specificity 97.5% and sensitivity 96.9%.
Gupta et al. [111]	DWT-SVD and STFT	Bonn	PSNR= 5.25 and SNR=31.32.
Sayeed et al. [112]	DWT, HP, and DNN	Bonn	Accuracy 100% (ictal vs. normal) and 98.6% (ictal vs. normal and interictal).
Singh et al. [113]	HOS and RF	Bonn	Accuracy 99.40%, specificity 99.66%, and sensitivity 99.40%.
Daoud et al. [114]	DCNN	CHB-MIT	Accuracy 96.1%, sensitivity 97.41%, and specificity 94.8%.

TABLE 9. Performance Comparison between various ML based models for ES prediction and detection.

Ref.	Model	Dataset	Performance
Direito et al. [115]	Multiclass SVM	European Epilepsy	Sensitivity 38.47% and FP rate 0.20 per hour.
Sriram et al. [116]	Multi Featured SVM	Bern Barcelona	Accuracy 92.15%, sensitivity 94.56%, and specificity 89.74%.
Subasi et al. [117]	Hybrid SVM	Bonn	Accuracy 98.75% (GA-SVM) and 99.38% (PSO-SVM).
Hamad et al. [118]	GOA-SVM and DWT	Bonn	Accuracy 100%.
Agarwal et al. [119]	CNN-SVM	American Epilepsy Society (Kaggle)	Accuracy (97.86±0.5)%, specificity (98.81±0.5)%, and sensitivity (96.47±0.5)%.
Raghu et al. [120]	DWT, Sigmoid entropy, and SVM	RMCH, Bonn, and CHB-MIT	Accuracy 96.34% (RMCH), detection delay (mean) of 1.2s, and false detection rate of 0.5, Sensitivity 94.21% (CHB-MIT) and 100% (Bonn).
Siddharth et al. [121]	FBSE-EWT and SAE-SVM	Bern Barcelona	Accuracy 100%, specificity 100%, and sensitivity 100%.
Mursalin et al. [122]	ICFS-RF	Bonn	Accuracy 100% (case A-E).
Zhang et al. [123]	GST, SVD, and RF	Bonn	Accuracy 99.63%.
Wang et al. [124]	RF-GSO	Bonn	AUC 99.0.
Wang et al. [125]	WPFs-RF	CHB-MIT	Accuracy 84.8%.
Sameer et al. [126]	STFT and RF	Bonn and Hauz Khans	Acuracy 98%.
Chakraborty et al. [127]	VMD and RF	Bonn	Accuracy 98.7%.
Rajaguru et al. [128]	PSD, AdaBoost, and KNN	20 epileptic patients of Sri Ramakrishna Hospital, India	Accuracy 97.53%, sensitivity 95.27% , and specificity 99.79%.
Ibrahim et al. [129]	KNN	CHB-MIT	Accuracy 76.4%.
Choubey et al. [130]	KNN and HFD	Bonn	Accuracy 98%.
Akbari et al. [131]	EWT, GA, and KNN	Bonn	Accuracy 98.33%, sensitivity 96%, and specificity 99.50%.
Sharmila et al. [132]	NB and DWT	Bonn	Accuracy 100%.
Xiao et al. [133]	ANBP	Self-created	Accuracy 82%.
Martis et al. [134]	EMD and C4.5	Bonn	Accuracy 95.33%, sensitivity 98%, and specificity 97%.
Rajaguru et al. [135]	SPCA and SDT	Self-created	Accuracy 96.83%, specificity 98.54%, and sensitivity 95.13%.
Wu et al. [136]	CHAID and MLP	CHB-MIT	Accuracy 99%.
Rajaguru et al. [137]	Code Converter and AdaBoost	20 epileptic patients of Sri Ramakrishna Hospital, India	Accuracy, sensitivity, and specificity are 97.29%, 96.45%, and 98.12%.
Al et al. [138]	AdaBoost LS-SVM	Bonn	Accuracy 99% and sensitivity 99%.
Hassan et al. [139]	CEEMDAN and AdaBoost	Bonn	Accuracy 100%, specificity 100%, and sensitivity 100% (class A,E).

using the alpha band (8 Hz-12 Hz). Using the RF classifier, they obtained an accuracy of 98% and an AUC score of 1 in distinguishing healthy and seizure patients. For conducting this study, the authors used the Bonn and Hauz Khans EEG dataset. Chakraborty et al. [127] used VMD and RF to detect ES and obtained 98.7% accuracy using the Bonn dataset.

Rajaguru and Prabhakar [128] reported an average classification accuracy of 97.53% using a modified KNN-based classifier to classify epilepsy. The sensitivity and specificity values were 95.27% and 99.79%, respectively. The average obtained time delay was 2.18 seconds. Here, the dataset was collected from 20 epileptic patients in Sri Ramakrishna Hospital, India. Ibrahim et al. [129] presented a KNN-based

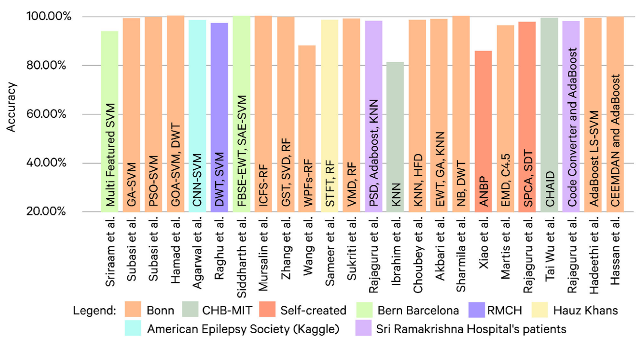

FIGURE 8. Comparison between various machine learning (ML) models for epileptic seizure (ES) prediction, and detection based on accuracy.

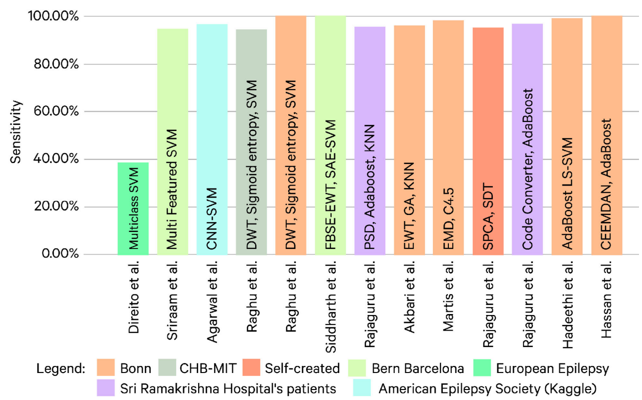

FIGURE 9. Comparison between various machine learning (ML) models for epileptic seizure (ES) prediction and detection based on sensitivity.

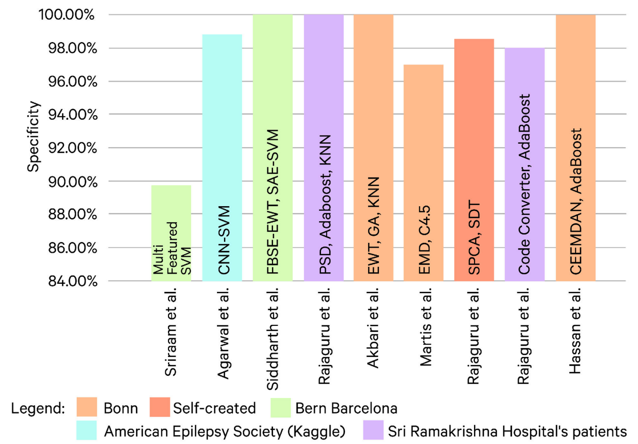

FIGURE 10. Comparison between various machine learning (ML) models for epileptic seizure (ES) prediction and detection based on specificity.

classifier to predict epilepsy and achieved 76.4% accuracy while working with the CBH-MIT dataset. This model successfully detected 44 patients among 55. Choubey and Pandey [130] conducted an experiment for detecting epilepsy using different EEG data. Using the Bonn dataset, the HFD model yields ANN classifier accuracy of around 94%, and the KNN accuracy rate of around 98%.

Akbari et al. [131] presented a seizure detection approach using EWT, GA, and KNN. The accuracy of this model was 98.33% using the Bonn dataset. Sensitivity and specificity scores are 96% and 99.50%, respectively. For personalized online seizure prediction, Xiao et al. [133] created a model and evaluated it using a self-created dataset. The average accuracy of ANBP was 82%, which was reported by the authors. One drawback of this model was that this model's performance was not compared to other well-known classifiers. Sharmila and Geethanjali [132] presented an ES classifier and reported that using the Bonn dataset NB classifier provides 100% accuracy, and the computational time using NB was very low than KNN. So, in this case, NB outperformed KNN classifier.

Martis et al. [134] presented a new EMD based ES detection approach and achieved an accuracy of 95.33%, sensitivity of 98%, and specificity of 97%. Here, the Bonn dataset was used for conducting this work. For ES classification, Rajaguru and Prabhakar [135] presented a model where they used a self-created dataset. Specificity 98.54%, sensitivity 95.13%, and accuracy 96.83% were reported by the authors. The model's accuracy was not very high. Wu et al. [136] presented an intelligent classifier for epilepsy detection. It was reported by the authors that both MLP and CART models had an accuracy of more than 99%.

Rajaguru and Prabhakar [137] proposed an ES detection method with the help of an AdaBoost classifier and code converter. The authors reported that this model's specificity, sensitivity, and accuracy are 98.12% and 96.45%, 97.29%, respectively. EEG data were collected from Twenty patients' of Sri Ramakrishna Hospital. The mean perfect classification rate of this model is 94.58%. Using the Bonn dataset, Al-Hadeethi et al. [138] reported 99% accuracy and sensitivity in his two-phased seizure detection technique. Using Bonn's dataset, Hassan et al. [139] successfully detected ES. The combination of CEEMDAN and Adaboost provides a very good detection accuracy of 100% (class A and E). The sensitivity and specificity scores are 100% and 100%, respectively.

Performance analysis of various ML-based models for ES prediction and detection are presented in Table 9. Figure 8, 9, and 10, and represents the performance comparison between various ML models for ES prediction and detection based on accuracy, sensitivity, and specificity, respectively. ML-based models presented by Hamad et al. [118], Sharmila and Geethanjali [132], Mursalin et al. [122], Siddharth et al. [121], and Hassan et al. [139] achieved an accuracy of 100%. The ES prediction model presented by Siddharth et al. [121], Raghu et al. [120], and Hassan et al. [139] achieved the highest sensitivity and specificity score, which is 100%.

3) DEEP LEARNING (DL)

The work of Chakrabarti et al. [82] has been validated using the pediatric patient EEG database from CHB-MIT. Records from 10 patients were chosen, and the performance metrics revealed that the sym4 (a version of Symlet) wavelet had the highest sensitivity, specificity, and accuracy at 97.2%, 93.5%, and 95.3% respectively. Furthermore, the prototype model consumes very little power. As a result, the proposed method has a good chance of working as a good seizure detection

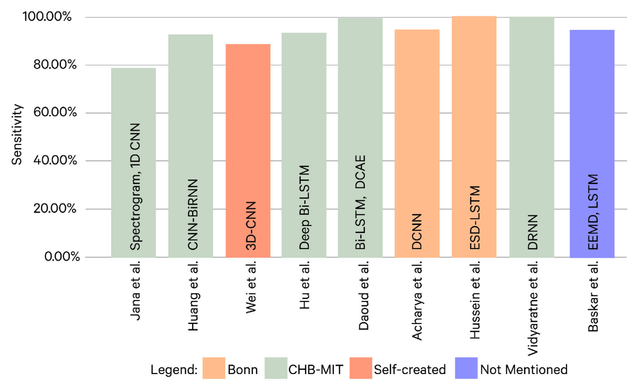

FIGURE 11. Comparison between various deep learning (DL) models for epileptic seizure (ES) prediction and detection based on sensitivity.

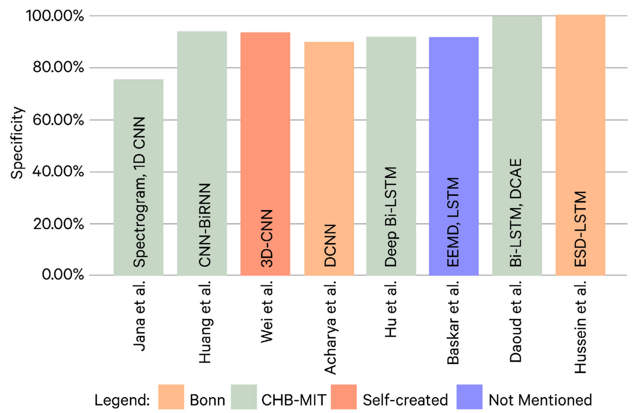

FIGURE 12. Comparison between various deep learning (DL) models for epileptic seizure (ES) prediction and detection based on specificity.

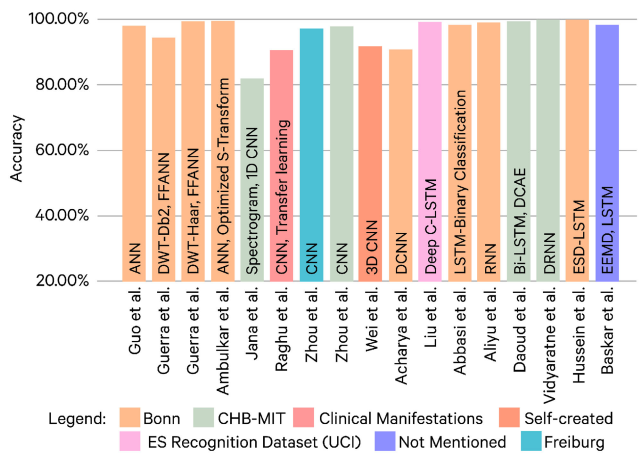

FIGURE 13. Comparison between various deep learning (DL) models for epileptic seizure (ES) prediction and detection based on accuracy.

device in real-life scenarios. The findings of Saric et al. [83] showed that an FPGA-based MLP ANN can diagnose ES with a rate of accuracy of 95.14%. The results also demonstrate the steps needed to properly implement ANN on the FPGA.

Using the Bonn dataset, Guo et al. [140] developed an automated ES detection technique and achieved a classification accuracy of 97.75%.

Juárez-Guerra et al. [141] developed a novel method to identity epileptic patients and person without disability using DWT and ANN. The authors reported 93.23% (DWT-Db2) and 99.26% (DWT-Haar) accuracy when using the Bonn dataset. Ambulkar and Sharma [142] proposed a five-stage epileptic seizure detection method, and by working on the Bonn dataset, the authors achieved 99.5% accuracy. Acharya et al. [143] developed an effective way for detecting seizures. Using the Bonn dataset, the authors reported specificity, sensitivity, and accuracy of 90.00%, 95.00%, and 88.67%, respectively.

Wei et al. [144] presented a 3D-CNN algorithm to classify seizures from multi-channel input signals, which outperformed 2D CNN models. Here accuracy rate was more than 90%. 88.90% and 93.78% are the value of sensitivity and specificity, respectively. Zhou et al. [145] developed a CNN-based ES detection method where the authors avoided manual feature extraction. The authors reported accuracies of 96.7% (preictal vs. interictal), 95.4% (ictal vs. interictal), and 94.3% (interictal vs. preictal vs. ictal) by using the Freiburg dataset. Accuracies of 95.6% (preictal vs. interictal), 97.5% ictal vs. interictal), and 93% (interictal vs. preictal vs. ictal) were observed by using the CBH-MIT dataset. A novel seizure detection method with the help of attention-based CNN-BiRNN was developed by Huang et al. [146]. Using the CBH-MIT dataset, the authors reported 94% specificity and 93% sensitivity. For ES detection, an average accuracy noted by Jana et al. [147] is 77.57%. Using the CBH-MIT dataset, the specificity and sensitivity value of this model is 75.59% and 79.54%, respectively. For offset and onset ES detection, the CHB-MIT dataset was used by Boonyakitanont et al. [148], and the authors got an accuracy of over 90%. Here, the recorded F1 score is 64.40%. For classifying seven different variants of seizures, 82.85% (Adam Solver and LR) and 88.30%(Inceptionv3) classification accuracies were reported by Raghu et al. [149]. In the case of computation time and accuracy, the extracted image features approach significantly outperformed the transfer learning approach.

The proposed AE-median CNN's [150] false alarm rate was lowered, from 0.17 h^{-1} to 0.034 h^{-1}, compared to the original CNN's rate of 0.17 h^{-1}.Wen and Zhang [151] proposed an AE-CDNN model. Whenever the feature dimension was more more than 16, the average classification accuracy of the features without parameter adjustment, reach more than 92%.

Vidyaratne et al. [152] developed a DRNN-based epilepsy detection method and found 100% accuracy and sensitivity with a 7 (mean) second of detection delay. By using data from Bonn's dataset, Hussein et al. [153] achieved 100% accuracy, specificity, and sensitivity. This ES prediction model outperforms general ML-based models. For binary classification (ictal and non-ictal) of ES, Abbasi et al. [154] presented an LSTM model. This model gives 98% accuracy. One limitation of this model is the performance of this model is compared with only the SVM. For classifying ES, 99% accuracy is achieved by the proposed model of Aliyu et al. [155]. One of the popular datasets called ''Bonn'' was used by authors

TABLE 10. Performance comparison between various DL based models for ES prediction and detection.

Ref.	Model	Dataset	Performance
Chakrabarti et al. [82]	ANN DWT	CHB-MIT	Accuracy, sensitivity, and specificity 95.3%, 97.2%, and 93.5% respectively.
Saric et al. [83]	MLP ANN	TUH EEG Corpus	Accuracy 95.14%.
Guo et al. [140]	ANN	Bonn	Accuracy 97.75%.
Juarez et al. [141]	DWT and FFANN	Bonn	Accuracy 93.23% (DWT-Db2) and 99.26% (DWT-Haar).
Ambulkar et al. [142]	ANN and Optimized S-Transform	Bonn	Accuracy 99.5%.
Acharya et al. [143]	DCNN	Bonn	Accuracy 88.67%, specificity 90.00%, and sensitivity 95.00%.
Wei et al. [144]	3D-CNN	Self-made dataset (13 volunteers participated)	Accuracy 90%, sensitivity 88.90%, and specificity 93.78%.
Zhou et al. [145]	CNN	Freiburg and CHB-MIT	Accuracy 96.7% (preictal vs. interictal) using the Freiburg dataset.
Huang et al. [146]	CNN-BiRNN	CHB-MIT	Specificity 94% and sensitivity 93%.
Jana et al. [147]	Spectrogram and 1D CNN	CHB-MIT	Accuracy 77.57%, specificity 75.59%, and sensitivity 79.54%.
Boonyakitanont et al. [148]	CNN	CHB-MIT	Accuracy over 90% and F1 score 64.4%.
Raghu et al. [149]	CNN, transfer learning	Different electroclinical, electrographic, and clinical manifestations.	Accuracy 88.30% (Inceptionv3) and 82.85% (Adam Solver and LR).
Takahashi et al. [150]	AE-CNN	-	False alarm reduced to 0.034 h^{-1}.
Wen et al. [151]	AE-CDNN	Bonn and Boston	Accuracy more than 92% when features are more than 16.
Vidyaratne et al. [152]	DRNN	CHB-MIT	Accuracy 100% and sensitivity 100% .
Hussein et al. [153]	ESD-LSTM	Bonn	Accuracy 100%, specificity 100%, and sensitivity 100%.
Abbasi et al. [154]	LSTM	Bonn	Accuracy 95% (trinary classification) and 98% (binary classification).
Aliyu et al. [155]	RNN	Bonn	Accuracy 99%.
Daoud et al. [156]	Bi-LSTM and DCAE	CHB-MIT	Accuracy 99.6%, specificity 99.60%, sensitivity 99.72%, and false alarm rate 0.004/h.
Baskar et al. [157]	EEMD and LSTM	-	Accuracy 98.2%, sensitivity 94.96%, and specificity 93.72%.
Liu et al. [158]	Deep C-LSTM	UCI	Accuracy 99.38%.
Hu et al. [159]	Deep Bi-LSTM	CHB-MIT	Sensitivity 93.61% and specificity 91.85%.

to conduct this work. Daoud and Bayoumi [156] used the CBH-MIT dataset for predicting seizures and found an accuracy of 99.6%. Here, the specificity rate is 99.60%, sensitivity is 99.72%, and false alarm rate is 0.004 per hour. Baskar and Karthikeyan [157] proposed an LSTM based ES detection technique that achieves 98.2% accuracy, 94.96% sensitivity, and 93.72% specificity. Liu et al. [158] developed a tumor and epileptic seizure detection technique and achieved 99:38% accuracy. Here, the dataset is collected from the UCI data repository. The duration between each detection was very short, and that value was 0.006 seconds. A Bi-LSTM network to detect seizures was developed by Hu et al. [159]. Using the CHB-MIT dataset, this model achieved 93.61% sensitivity and 91.85% specificity, respectively.

Performance analysis of various DL-based ES detection and prediction models are presented in Table 10. Figure 11, 12, and 13 represents the performance comparison between various DL models for ES detection and prediction based on sensitivity, specificity, and accuracy, respectively. After analyzing all the DL models, we found that the model of Hussein et al. [153] and Vidyaratne et al. [152] achieved an outstanding accuracy score, and that is 100%. Hussein et al. [153] reported the highest specificity and sensitivity score, which was 100%.

VIII. EEG DATASETS

A. CHB-MIT

Shoeb et al. [162] created CHB-MIT dataset by capturing EEG signals from 23 individuals with 163 seizures. 844 hours were needed to create the full dataset. Here, the sample rate per second was 256 and 10 to 20 standard electrodes are used in this case. This dataset contains two kinds of seizure, one is the main seizure, and another is a combined seizure. The amount of data in this dataset is satisfactory.

B. BONN

The data were sampled at a rate of 173.61 Hz. The time series have the acquisition system's bandwidth (0.5 - 85) Hz. The use of a low-pass filter with a frequency of 40 Hz, as specified in the manuscript, is considered the The initial stage of analysis. There are sets of data in this dataset. The Bonn database is made up of five sets of data: A, B, C, D, and E, each of which contains 100 separate channels of EEG signals. Here data recording took 23.6 seconds. Datasets B and A contain normal signals from five participants with their eyes closed and open, respectively. EEG data from database C and D are associated with the pre-ictal area. The EEG data in the E dataset are associated with the ictal area. 10–20 scalp EEG standard electrodes were used in the A and the B dataset.

TABLE 11. Comparison between various datasets.

Dataset	Sample frequency (Hz)	Number of subjects	Data collection duration	Total electrodes	Limitation
CHB-MIT	256 Hz	23	844 hours	10-20 standard electrodes	-
Bonn	173.61Hz	5	23.6 seconds	10-20 standard electrodes	As dataset A contains EEG signal of only five participants, there is a possibility of less diversity in the dataset. The number of participants for constructing C, D and E datasets is not clearly mentioned.
EEG Epilepsy Datasets (Hauz Khas)	200Hz	10	-	10-20 standard electrodes	The total duration for collecting EEG data is not mentioned.
Freiburg	256 Hz	21	Around 24 hours	6 electrodes	-
Bern Barcelona	512 Hz	5	82 hours	10-20 standard electrodes	-
Ramaiah Medical College and Hospital (RMCH)	128 Hz	115	-	19 electrodes	To the best of our knowledge, the data collection duration is not mentioned clearly.

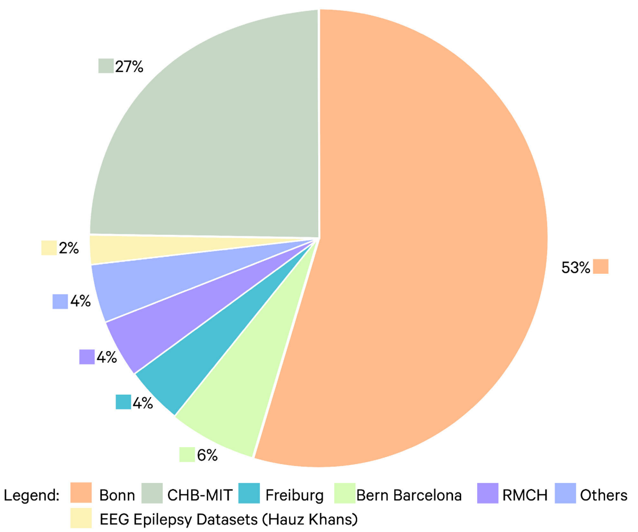

FIGURE 14. The prevalence of various popular EEG datasets for epileptic seizure (ES) prediction and detection in the literature.

Depth electrodes were used for constructing the C and the D dataset and using both strip and depth electrodes, the E dataset was created by [163].

C. EEG EPILEPSY DATASET (Hauz Khas)

These are segmented EEG time series data collected from ten epilepsy patients obtained from the Neurology Sleep Centre in Hauz Khas in New Delhi. This dataset was created by Swami et al. [164]. The data was collected using the AS40 Amplifier at a sample rate of 200 Hz. Gold-plated scalp EEG electrodes were put according to the 10-20 electrode placement protocol. Filtered between (0.5-70) Hz. Ictal, interictal, and preictal stages are the three stages of the signal. Each downloading folder consists of fifty EEG time-series signal MAT files. The folder's name correlates to the stage of an ES.

D. FREIBURG

Klatt et al. [165] constructed this dataset from EEG recordings of 21 participants with medically uncontrollable focal epilepsy. It was captured during an intrusive presurgical epilepsy scanning at the University Hospital of Freiburg in Germany. The recording time for eight human subjects was pretty less than 24 hours; on the other hand, this time for 13 subjects was 24 hours. This data set contains 88 seizures data altogether. Here the sampling frequency was 256 Hz, and six electrodes were used to capture those data.

E. BERN BARCELONA

This dataset provided intracranial EEG of focal epileptic patients and was obtained from Bern Hospital in Barcelona (brain department). Sriraam and Raghu [116] recorded signals using AD-Tech electrodes. 10–20 standards were used here with one additional electrode. There were two kinds of EEG data in the database: one is focal data, and another one is extra focal data. Each database comprises 3750 pairs of concurrent observed data with a 20-second duration and a 512-Hz sampling frequency. The database contains 83 hours of EEG data which are collected from five human subjects of various ages.

F. RAMAIAH MEDICAL COLLEGE AND HOSPITAL (RMCH)

This dataset was created by Raghu et al. [120].They generated using a sampling frequency of 128 Hz from the Galileo Suite-EB neuro EEG system. The unipolar scalp EEG signal was used to record the signal using 19 electrodes, according to the International 10–20 configuration system. The database includes 115 individuals, 48 female and 67 male, ranging from 2.5 to 75. 38 of the 115 people had epilepsy, and 77 had healthy individuals. Table 11 represents comparison between various mentioned datasets. Figure 14 shows the prevalance of various popular EEG datasets for ES.

IX. CHALLENGES AND FUTURE DIRECTIONS

A. UNAVAILABILITY OF OPEN ACCESS LONG TERM EEG DATA

Long-term EEG data is vital for the accurate prediction and detection of ES. The most unfortunate fact is the scarcity of open access long-term EEG datasets [166]. Only a chunk of the data may be publicly accessible, as the entire dataset is not shared. As a result, real-time detection of ES remains difficult. However, the clinical dataset has been used in research into real-time ES prediction. But the collection process of the clinical datasets is very slow, tedious, and complex. As a result, there is an emergency need for open access to EEG databases containing long-term recordings. Wireless EEG headsets can be helpful to collect long term data rather than wired headsets.

B. MISSING DATA

As an EEG headset is implanted on the volunteer's head and the communication between the headset and the data storing and converting device is completely wireless, there is a clear possibility of data dropout (missing data). Due to missing data, the quality of data decreases rapidly and also decreases prediction quality. Adding a missingness indicator and predictor before the converter will help remove the missing data problems. The Missingness indicator will warn about missing data. Based on the previous data, an auto predictor will generate predicted brain impulse data.

C. TIME REQUIRED TO PROCESS EEG DATA

It is essential to keep the time complexity of a real-time ES prediction and monitoring system as low as possible. Using the raw data, one can decrease the time complexity of a real-time system. But the upsetting fact is that one can not work with the raw data; rather, one has to operate the feature extraction method to extract features from raw data. DL can solve this problem because it automatically extracts features from raw data.

D. BLACK-BOX NATURE OF ML

Due to the ML technique's black-box nature, doctors and patients are less likely to trust it. The reliability of ES prediction using ML and DL is always a concern to general users. So, a technique is needed that will predict like ML, but the prediction's reasons will be very clear and well explainable to the general user. Explainable AI (XAI) can serve this need. With the help of XAI, users can analyze the reasons behind any specific decision.

E. HETEROGENEOUS NATURE OF SEIZURES

The variability in the causes of ES, trouble in finding the area of incidence, and a poor understanding of how and why the seizures propagated are the significant reasons for the failure of the massive usage of ML in ES prediction [167]. To create a strong solution, researchers must first understand the various potential causes of ES. Researchers should avoid binary classifiers and use multimodal classification algorithms to predict ES.

F. EXPENSE AND USABILITY OF EEG HEADSET

EEG headsets are costly because of the price and a large number of electrodes. EEG testing can take a reasonable period, and the subject must remain calm during the procedure. It is difficult to implant electrodes in the correct location; even a conducting gel is required before placing such electrodes in the scalp. Though dry electrodes are invented, they are really expensive. Usually, wet Ag/AgCl is used to manufacture electrodes, but active dry electrodes are a great solution for a less expensive EEG headset. Researchers should focus on inventing more cost-effective materials for the electrodes.

G. DATA DIMENSIONALITY

Brain signals are complex and non-linear. To capture this non-linearity, multiple channeled EEG headsets are required. For reducing computational complexity, these multi-channeled data should be converted into a single channel. But the matter of fact is all channels don't contain useful data. So, finding the important data channel is necessary, and Signal Quality Index (SQI) is an adaptive algorithm to choose the best data channel. Significant effort has been put into creating adaptive algorithms for selecting the best EEG channel using various dimensionality reduction techniques such as PCA, ICA, etc. Moreover, developing an optimal technique for reducing the dimensionality of EEG signals and selecting the most important channel in a multi-channel EEG system could be a potential work field.

H. THE TRADE-OFF BETWEEN TECHNICAL COMPLEXITY AND DATA VOLUME

Smaller headsets with fewer electrodes are more convenient to use. It will also reduce the technical difficulty. However, because there are fewer electrodes, the volume of gathered data or training sets is reduced. However, heavy training sets can yield more accurate outcomes, but they are also computationally expensive. A major difficulty in creating a BCI is a trade-off between the technical complexity of understanding the participant's brain activity signals and the volume of training required for successful functioning. More and more researches can be performed to create a benchmark to solve this specific challenge.

I. REAL-TIME IoT-BASED ES PREDICTION, DETECTION, AND MONITORING

IoT allows continuous monitoring of a patient as well as early seizure prediction [168]. IoT-based systems must be low-powered and have low computational complexity. Brokers in the Cloud, Fog, and Mist work on EEG data based on its type and necessity. In IoMT systems, edge fog computing can be used to detect epileptic seizures using lossless EEG data [169]. On-site patient monitoring and real-time brain signal processing are extremely difficult tasks. As instant EEG data processing is a real concern for IoT-based models, more and more research is on this particular topic. The researcher should strike a good balance between data volume, model memory, and processing speed to ensure real-time data processing. Because patients' data are stored and transmitted via the internet in an IoT framework, ensuring online data security for IoT systems is also a hot topic for research. Even,

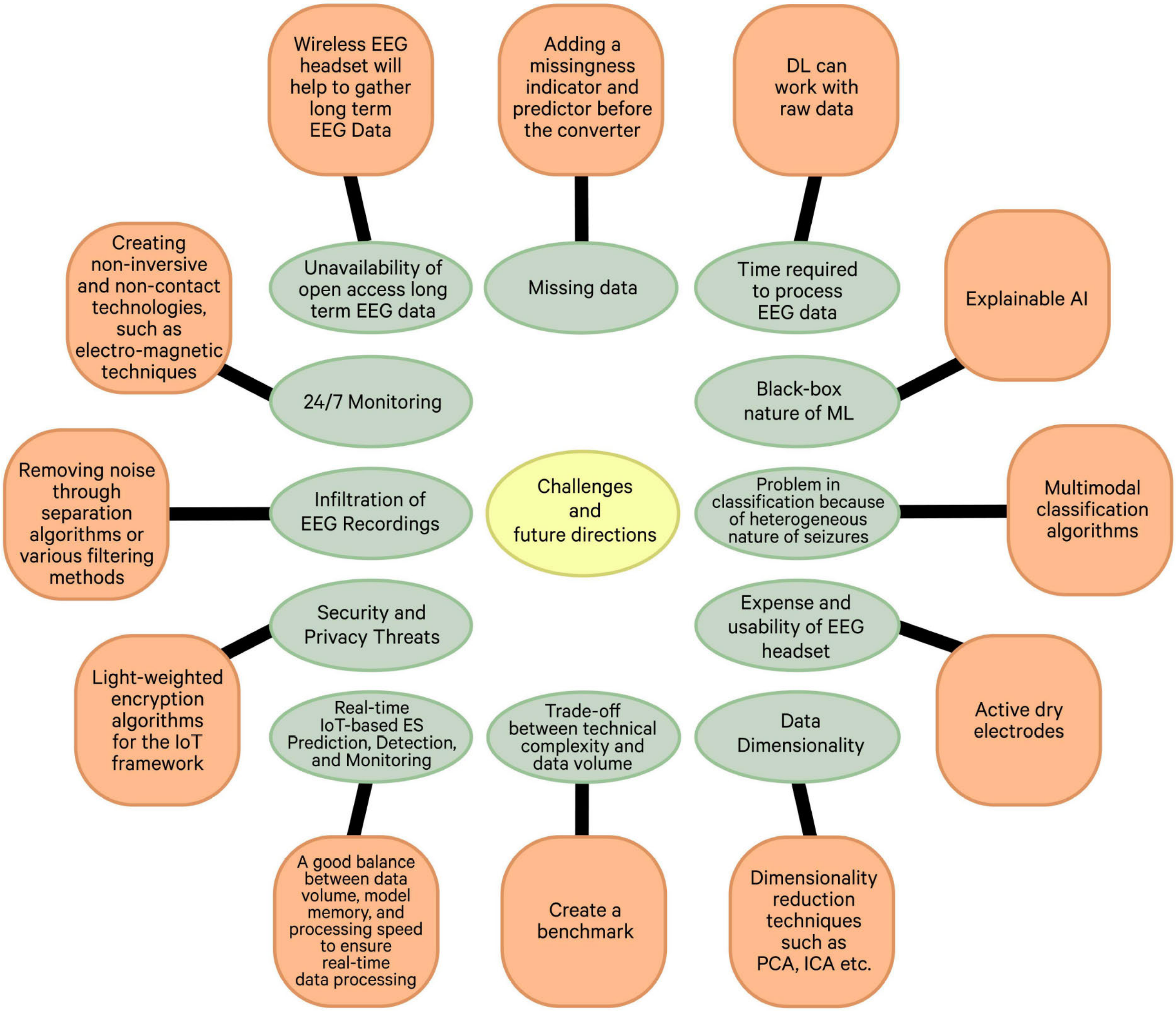

FIGURE 15. **Challenges and future directions.**

for developing IoMT systems, digital twin and metaverse can be a good option for understanding the characteristics of the physical devices [169], [170].

J. SECURITY AND PRIVACY THREATS

Serious security breaching problems may arise in the field of EEG data-oriented IoT frameworks. As IoT frameworks are entirely dependent on internet connectivity, attackers and intruders are often more privileged to hack sensitive brain data and manipulate that information [171]. Typical BCI authentication methods are 2-factor authentication, patterns, PIN, passwords, and fingerprints. Denial of action or service, data theft, faking patients' identities, and illegally tracking patients' geolocation are some of the common security and privacy concerns in IoT-based e-Healthcare models. Some standard solutions to remove these problems are: controlling data access, implementing a firewall, providing data integrity, ensuring authentication, and Encrypting all information. Message Authentication Code (MAC), Hash algorithm, Message Digest version 5 (MD5), Triple Data Encryption Standard (DES), and Advanced Encryption Standard (AES) are some popular and powerful encryption algorithms. Researchers should develop novel, light-weighted encryption algorithms for the IoT framework based on the techniques mentioned above.

K. INFILTRATION OF EEG RECORDINGS

Artifacts, which are especially evident on scalp recordings of EEG, are noises generated by activities that do not arise in the brain yet unfeasible to be eliminated at the moment of recording. Artifacts can emerge as a result of eye blinks, chewing, muscle movement of scalp musculature or less prevalently because of heartbeat. Incorrect way of EEG headset usage or environmental interference can cause noise in EEG recordings. These artifacts are relatively easy to detect and categorize for an expert. However, their removal in digital analysis is more complex, since it is difficult to remove the artifact without also compromising the measurement of actual brain activity. Training the system to distinguish and manage artifacts, removing them through separation algorithms or various filtering methods, alternatively excluding the signals

TABLE 12. Challenges and potential solutions of EEG-IoT based epileptic seizure Detection.

Challenges	Potential Solutions
The lack of long-term EEG datasets with open access is the most regrettable reality. Since the complete dataset is not disclosed, just a portion of it might be available to the general public. Real-time ES detection is therefore still challenging.	Wireless EEG headsets can be helpful to collect long term data rather than wired headsets.
Data quality and prediction quality can rapidly deteriorate due to missing data.	To solve the missing data issues, add a missingness indicator and predictor before the converter. The Missingness indicator will issue a warning when data is missing. An automatic predictor will produce anticipated brain impulse data based on historical data.
One can make a real-time system less time-complex by using the raw data. Unfortunately, working with raw data is not an option; one must instead use the feature extraction approach to extract features from raw data.	Because DL automatically pulls features from unprocessed data, it can address this issue.
Doctors and patients are less inclined to trust the ML methodology since it is a black-box method.	With the help of Explainable AI(XAI), general users can analyze the reasons behind any specific decision.
The failure of the widespread application of ML in ES prediction is mostly due to the diversity in the etiology of ES, difficulty locating the area of occurrence, and poor knowledge of how and why the seizures propagated.	To begin with, researchers must comprehend the many ES-related probable causes. To anticipate ES, researchers should stay away from binary classifiers and instead employ multimodal classification systems.
Due to their high cost and extensive use of electrodes, EEG headsets are expensive. Electrodes are challenging to implant in the proper spot; even a conducting gel is necessary before putting such electrodes in the scalp.	An excellent option for a less-priced EEG headset is active dry electrodes. The focus should be placed on developing more affordable electrode materials.
EEG headsets with many channels are necessary to record this non-linearity. These multi-channeled data should be combined into a single channel in order to simplify the calculation. However, the truth is that not all channels offer meaningful information.	A potential area of effort would be to create the best method for decreasing the dimensionality of EEG signals and choosing the most crucial channel in a multi-channel EEG system.
The use of smaller headsets with fewer electrodes is more practical. However, because there are fewer electrodes, there are fewer training sets or data sets collected.	More and more studies can be conducted to establish a standard for resolving this particular problem.
The monitoring of ES patients in real-time is crucial. Real-time brain signal processing and on-site patient monitoring are exceedingly challenging undertakings.	Models for IoT-based monitoring can address this problem. To ensure real-time data processing, the researcher should find a fair balance between data volume, model memory, and processing speed.
Because IoT frameworks are entirely reliant on internet access, attackers and intruders are frequently more privileged to hack and modify sensitive brain data.	Creating a lightweight encryption algorithm could be a possible solution to this problem.
Raw EEG data as well as infiltrated data can give misleading information.	Some possible strategies for managing artifacts include training the system to recognize and manage artifacts, removing them using separation algorithms or various filtering methods, or excluding contaminated signals or channels from the analysis.

or channels which are contaminated from the analysis are some of the probable strategies to manage artifacts.

L. 24/7 MONITORING

Commonly used EEG headsets can collect data when they stay in contact with the patient's body. For creating a large dataset, it is essential to continue data recording for around 24 hours. To create an ES patients monitoring system, patients have to wear the EEG headset continuously, which is very tedious and inconvenient. To resolve this problem, researchers should create cost-effective wireless EEG headsets. Creating non-invasive and non-contact technologies, such as electromagnetic techniques, may provide a great solution to this problem. A summarized version of challenges and potential solutions are presented in Table 12. A visual representation of all challenges and future directions are depicted in Figure 15.

X. CONCLUSION

EEG has been introduced approximately a hundred years ago, remains the chosen way for much relevant literature, from neuroscience study to real-time applications, from functional brain imagery, through movement to interface brain-computer applications. While EEG has historically played a focal role in assessing neural function, this paper explores the contributions made to seizure detection and prediction techniques, where different ML and DL approaches are successfully used. This review focuses on recent ML, DL, and IoT-based ES detection, prediction, and patient monitoring approach. After reviewing 56 research articles, it is very clear that SVM, RF, and KNN are the most powerful ML-based classifiers. Nowadays, more and more research is going on hybrid ML-based classification schemes, and those research approaches provide better outcomes than traditional ML. Hybrid RNN, Hybrid LSTM, and Deep RNN are the most promising DL-based classifiers and contain a lot of research opportunities. Real-time detection, prediction, and monitoring of ES can provide a better life to the patients. With the help of IoT-based prediction and monitoring, real-time services like early prediction, and emergency medical support can easily be provided. There are a lot of research opportunities on early seizure prediction and IoT-based real-time monitoring approaches. For removing data limitations, various research work should be done on unsupervised and semi-supervised methods. This review highlights the performance, limitation, and usage of various ML, DL, and IoT-based seizure classification and monitoring approaches. Researchers should focus on the adaptability, scalability and interpretability of ML algorithms. This work will also provide a clear direction to future researchers for developing more efficient and feasible research works in this field.

REFERENCES

[1] L. De Clerck, A. Nica, and A. Biraben, "Clinical aspects of seizures in the elderly," *Geriatrie et Psychologie Neuropsychiatrie du Vieillissement*, vol. 17, no. 1, pp. 7–12, 2019.

[2] A. Daoud, A. Batieha, M. Bashtawi, and H. El-Shanti, "Risk factors for childhood epilepsy: A case-control study from Irbid, Jordan," *Seizure*, vol. 12, no. 3, pp. 171–174, Apr. 2003.

[3] K. Harris. (2018). *The Dangers of Seizures: Why You Need Immediate Treatment*. [Online]. Available: https://www.osfhealthcare.org/blog/dangers-of-seizures/

[4] N. H. V. Higuera. (May 2019). *Can you Die From a Seizure? What to Know If You Have Seizures*. [Online]. Available: https://www.healthline.com/health/epilepsy/can-you-die-from-a-seizure

[5] (Jun. 2019). *Epilepsy*. Accessed: Dec. 13, 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/epilepsy

[6] G. Zazzaro, S. Cuomo, A. Martone, R. V. Montaquila, G. Toraldo, and L. Pavone, "EEG signal analysis for epileptic seizures detection by applying data mining techniques," *Internet Things*, vol. 14, Jun. 2021, Art. no. 100048. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660519300204

[7] N. van Klink, A. Mooij, G. Huiskamp, C. Ferrier, K. Braun, A. Hillebrand, and M. Zijlmans, "Simultaneous MEG and EEG to detect ripples in people with focal epilepsy," *Clin. Neurophysiol.*, vol. 130, no. 7, pp. 1175–1183, Jul. 2019.

[8] N. Pianou and S. Chatziioannou, "Imaging with PET/CT in patients with epilepsy," in *Epilepsy Surgery Intrinsic Brain Tumor Surgery*. Berlin, Germany: Springer, 2019, pp. 45–50.

[9] L. M. M. Chacón, J. G. González, M. R. Castillo, S. B. Batista, K. B. García-Ramo, A. S. Santos, N. Q. Cordero, M. Z. Bermudez, R. G. Fernández, B. E. Díaz, Z. H. Díaz, J. E. B. Del Busto, A. S. Coroneux, M. M. B. Martin, and L. L. Pedre, "Surgical outcome in extratemporal epilepsies based on multimodal pre-surgical evaluation and sequential intraoperative electrocorticography," *Behav. Sci.*, vol. 11, no. 3, p. 30, Mar. 2021.

[10] S. Kulaseharan, A. Aminpour, M. Ebrahimi, and E. Widjaja, "Identifying lesions in paediatric epilepsy using morphometric and textural analysis of magnetic resonance images," *NeuroImage, Clin.*, vol. 21, Jan. 2019, Art. no. 101663.

[11] M. Carbó-Carreté, C. Cañete-Massé, M. D. Figueroa-Jiménez, M. Peró-Cebollero, and J. G. Rdia-Olmos, "Relationship between quality of life and the complexity of default mode network in resting state functional magnetic resonance image in down syndrome," *Int. J. Environ. Res. Public Health*, vol. 17, no. 19, p. 7127, Sep. 2020.

[12] P. Gelisse, P. Genton, A. Crespel, and P. H. Lefevre, "Will MRI replace the EEG for the diagnosis of nonconvulsive status epilepticus, especially focal?" *Revue Neurologique*, vol. 177, no. 4, pp. 359–369, Apr. 2021.

[13] A. Biondi, V. Santoro, P. F. Viana, P. Laiou, D. K. Pal, E. Bruno, and M. P. Richardson, "Noninvasive mobile EEG as a tool for seizure monitoring and management: A systematic review," *Epilepsia*, vol. 63, no. 5, pp. 1041–1063, May 2022.

[14] J. J. Shih, D. J. Krusienski, and J. R. Wolpaw, "Brain-computer interfaces in medicine," *Mayo Clinic Proc.*, vol. 87, no. 3, pp. 268–279, Mar. 2012. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3497935/

[15] S. Cheruvu, A. Kumar, N. Smith, and D. M. Wheeler, "IoT frameworks and complexity," in *Demystifying Internet of Things Security*. Berlin, Germany: Springer, 2020, pp. 23–148.

[16] Y. Song, "A review of developments of EEG-based automatic medical support systems for epilepsy diagnosis and seizure detection," *J. Biomed. Sci. Eng.*, vol. 4, no. 12, pp. 788–796, 2011.

[17] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, and J. S. Suri, "Automated EEG analysis of epilepsy: A review," *Knowl.-Based Syst.*, vol. 45, pp. 147–165, Jun. 2013.

[18] D. Torse, V. Desai, and R. Khanai, "A review on seizure detection systems with emphasis on multi-domain feature extraction and classification using machine learning," *BRAIN. Broad Res. Artif. Intell. Neurosci.*, vol. 8, no. 4, pp. 109–129, 2017.

[19] M. Iftikhar, S. A. Khan, and A. Hassan, "A survey of deep learning and traditional approaches for EEG signal processing and classification," in *Proc. IEEE 9th Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON)*, Nov. 2018, pp. 395–400.

[20] K. Abualsaud, A. Mohamed, T. Khattab, E. Yaacoub, M. Hasna, and M. Guizani, "Classification for imperfect EEG epileptic seizure in IoT applications: A comparative study," in *Proc. 14th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2018, pp. 364–369.

[21] B. Abbasi and D. M. Goldenholz, "Machine learning applications in epilepsy," *Epilepsia*, vol. 60, no. 10, pp. 2037–2047, Oct. 2019.

[22] M. K. Siddiqui, R. Morales-Menendez, X. Huang, and N. Hussain, "A review of epileptic seizure detection using machine learning classifiers," *Brain Informat.*, vol. 7, no. 1, p. 5, May 2020.

[23] K. Rasheed, A. Qayyum, J. Qadir, S. Sivathamboo, P. Kwan, L. Kuhlmann, T. O'Brien, and A. Razi, "Machine learning for predicting epileptic seizures using EEG signals: A review," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 139–155, 2020.

[24] S. Raut and N. Rathee, "Comparative study on machine learning classifiers for epileptic seizure detection in reference to EEG signals," in *Proc. Int. Conf. Artif. Intell. Appl.* Cham, Switzerland: Springer, 2021, pp. 185–194.

[25] A. Shoeibi, M. Khodatars, N. Ghassemi, M. Jafari, P. Moridian, R. Alizadehsani, M. Panahiazar, F. Khozeimeh, A. Zare, H. Hosseini-Nejad, A. Khosravi, A. F. Atiya, D. Aminshahidi, S. Hussain, M. Rouhani, S. Nahavandi, and U. R. Acharya, "Epileptic seizures detection using deep learning techniques: A review," *Int. J. Environ. Res. Public Health*, vol. 18, no. 11, p. 5780, May 2021.

[26] A. Biasiucci, B. Franceschiello, and M. M. Murray, "Electroencephalography," *Current Biol.*, vol. 29, no. 3, pp. 80–85, 2019.

[27] J. L. Stone and J. R. Hughes, "Early history of electroencephalography and establishment of the American clinical neurophysiology society," *J. Clin. Neurophysiology*, vol. 30, no. 1, pp. 28–44, Feb. 2013. [Online]. Available: https://journals.lww.com/00004691-201302000-00002

[28] V. Mihajlovic, B. Grundlehner, R. Vullers, and J. Penders, "Wearable, wireless EEG solutions in daily life applications: What are we missing?" *IEEE J. Biomed. Health Inform.*, vol. 19, no. 1, pp. 6–21, Jan. 2015.

[29] M. Soufineyestani, D. Dowling, and A. Khan, "Electroencephalography (EEG) technology applications and available devices," *Appl. Sci.*, vol. 10, no. 21, p. 7453, Oct. 2020.

[30] W. Lin, Q. Chen, M. Jiang, J. Tao, Z. Liu, X. Zhang, L. Wu, S. Xu, Y. Kang, and Q. Zeng, "Sitting or walking? Analyzing the neural emotional indicators of urban green space behavior with mobile EEG," *J. Urban Health*, vol. 97, no. 2, pp. 191–203, Apr. 2020.

[31] S. Machado, F. Araujo, F. Paes, B. Velasques, M. Cunha, H. Budde, L. F. Basile, R. Anghinah, O. Arias-Carrion, M. Cagy, and R. Piedade, "EEG-based brain-computer interfaces: An overview of basic concepts and clinical applications in neurorehabilitation," *Rev. Neurosci.*, vol. 21, no. 6, pp. 451–468, Jan. 2010.

[32] C. Kamarajan, "Chapter 13—Brain electrophysiological signatures in human alcoholism and risk," in *Neuroscience of Alcohol*, V. R. Preedy, Ed. Cambridge, MA, USA: Academic Press, Jan. 2019, pp. 119–130. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128131251000131

[33] E. Q. Wu, "Detecting fatigue status of pilots based on deep learning network using EEG signals," *IEEE Trans. Cogn. Developmental Syst.*, vol. 13, no. 3, pp. 575–585, Sep. 2021.

[34] M. S. Kalas and B. F. Momin, "Stress detection and reduction using EEG signals," in *Proc. Int. Conf. Electr., Electron., Optim. Techn. (ICEEOT)*, Mar. 2016, pp. 471–475.

[35] A. Shoka, M. Dessouky, A. El-Sherbeny, and A. El-Sayed, "Literature review on EEG preprocessing, feature extraction, and classifications techniques," *Menoufia J. Electron. Eng. Res.*, vol. 28, no. 1, pp. 292–299, Dec. 2019.

[36] R. Haloi, J. Hazarika, and D. Chanda, "Selection of appropriate statistical features of EEG signals for detection of Parkinson's disease," in *Proc. Int. Conf. Comput. Perform. Eval. (ComPE)*, Jul. 2020, pp. 761–764.

[37] S.-H. Oh, Y.-R. Lee, and H.-N. Kim, "A novel EEG feature extraction method using Hjorth parameter," *Int. J. Electron. Elect. Eng.*, vol. 2, no. 2, pp. 106–110, 2014.

[38] R. Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from EEG," *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, pp. 327–339, Jul. 2014.

[39] A. Patil, C. Deshmukh, and A. R. Panat, "Feature extraction of EEG for emotion recognition using Hjorth features and higher order crossings," in *Proc. Conf. Adv. Signal Process. (CASP)*, Jun. 2016, pp. 429–434.

[40] L. Hussain, "Detecting epileptic seizure with different feature extracting strategies using robust machine learning classification techniques by applying advance parameter optimization approach," *Cogn. Neurodyn.*, vol. 12, no. 3, pp. 271–294, Jun. 2018.

[41] I. Rejer and P. Gorski, "Benefits of ICA in the case of a few channel EEG," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 7434–7437.

[42] S. Suwanto, M. H. Bisri, D. C. R. Novitasari, and A. H. Asyhar, "Classification of EEG signals using fast Fourier transform (FFT) and adaptive neuro fuzzy inference system (ANFIS)," *J. Matematika*, vol. 5, no. 1, pp. 35–44, May 2019.

[43] A. Sharma, R. S. Tanwar, Y. Singh, A. Sharma, S. Daudra, G. Singal, T. R. Gadekallu, and S. Pancholi, "Heart rate and blood pressure measurement based on photoplethysmogram signal using fast Fourier transform," *Comput. Electr. Eng.*, vol. 101, Jul. 2022, Art. no. 108057.

[44] K. Samiee, P. Kovács, and M. Gabbouj, "Epileptic seizure classification of EEG time-series using rational discrete short-time Fourier transform," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 541–552, Feb. 2014.

[45] M. Sameer, A. K. Gupta, C. Chakraborty, and B. Gupta, "Epileptical seizure detection: Performance analysis of gamma band in EEG signal using short-time Fourier transform," in *Proc. 22nd Int. Symp. Wireless Pers. Multimedia Commun. (WPMC)*, Nov. 2019, pp. 1–6.

[46] Y. Zhang, B. Liu, X. Ji, and D. Huang, "Classification of EEG signals based on autoregressive model and wavelet packet decomposition," *Neural Process. Lett.*, vol. 45, no. 2, pp. 365–378, Apr. 2017.

[47] S. A. Awang, M. P. Paulraj, and S. Yaacob, "Analysis of EEG signals by eigenvector methods," in *Proc. IEEE-EMBS Conf. Biomed. Eng. Sci.*, Dec. 2012, pp. 778–783.

[48] L. Wang, W. Xue, Y. Li, M. Luo, J. Huang, W. Cui, and C. Huang, "Automatic epileptic seizure detection in EEG signals using multi-domain feature extraction and nonlinear analysis," *Entropy*, vol. 19, no. 6, p. 222, Jun. 2017.

[49] H. Adeli, Z. Zhou, and N. Dadmehr, "Analysis of EEG records in an epileptic patient using wavelet transform," *J. Neurosci. Methods*, vol. 123, no. 1, pp. 69–87, Feb. 2003.

[50] G. Rilling and P. Flandrin, "One or two frequencies? The empirical mode decomposition answers," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 85–95, Jan. 2008.

[51] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, 2009.

[52] V. R. Carvalho, M. F. D. Moraes, A. P. Braga, and E. M. A. M. Mendes, "Evaluating five different adaptive decomposition methods for EEG signal seizure detection and classification," *Biomed. Signal Process. Control*, vol. 62, Sep. 2020, Art. no. 102073.

[53] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2011, pp. 4144–4147. [Online]. Available: http://ieeexplore.ieee.org/document/5947265/

[54] J. Gilles, "Empirical wavelet transform," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 3999–4010, Aug. 2013.

[55] T. Liu, Z. Luo, J. Huang, and S. Yan, "A comparative study of four kinds of adaptive decomposition algorithms and their applications," *Sensors*, vol. 18, no. 7, p. 2120, Jul. 2018.

[56] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.

[57] Y. Huang, J. Lin, Z. Liu, and W. Wu, "A modified scale-space guiding variational mode decomposition for high-speed railway bearing fault diagnosis," *J. Sound Vibrat.*, vol. 444, pp. 216–234, Mar. 2019.

[58] B. Chaitanya, A. Yadav, M. Pazoki, and A. Y. Abdelaziz, "A comprehensive review of islanding detection methods," in *Uncertainties in Modern Power Systems*. Amsterdam, The Netherlands: Elsevier, 2021, pp. 211–256.

[59] H. Liu, G. Yan, Z. Duan, and C. Chen, "Intelligent modeling strategies for forecasting air quality time series: A review," *Appl. Soft Comput.*, vol. 102, Apr. 2021, Art. no. 106957.

[60] F.-J. Yang, "An extended idea about decision trees," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Dec. 2019, pp. 349–354.

[61] S. Bose, V. Rama, N. Warangal, and C. B. R. Rao, "EEG signal analysis for seizure detection using discrete wavelet transform and random forest," in *Proc. Int. Conf. Comput. Appl. (ICCA)*, Sep. 2017, pp. 369–378.

[62] J. S. Park, H. S. Cho, J. S. Lee, K. I. Chung, J. M. Kim, and D. J. Kim, "Forecasting daily stock trends using random forest optimization," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2019, pp. 1152–1155.

[63] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 1, pp. 21–27, Jan. 1967.

[64] K. Fukunaga and P. M. Narendra, "A branch and bound algorithm for computing K-nearest neighbors," *IEEE Trans. Comput.*, vol. 100, no. 7, pp. 750–753, Jul. 1975.

[65] S. Suthaharan, "Machine learning models and algorithms for big data classification," *Integr. Ser. Inf. Syst.*, vol. 36, pp. 1–12, Feb. 2016.

[66] M. Pal, "Random forest classifier for remote sensing classification," *Int. J. Remote Sens.*, vol. 26, no. 1, pp. 217–222, 2007.

[67] J. Chai, J. Du, K. K. Lai, and Y. P. Lee, "A hybrid least square support vector machine model with parameters optimization for stock forecasting," *Math. Problems Eng.*, vol. 2015, pp. 1–7, Jan. 2015.

[68] S. Chen, B. Shen, X. Wang, and S.-J. Yoo, "A strong machine learning classifier and decision stumps based hybrid AdaBoost classification algorithm for cognitive radios," *Sensors*, vol. 19, no. 23, p. 5077, Nov. 2019.

[69] M. Chen, X. Shi, Y. Zhang, D. Wu, and M. Guizani, "Deep feature learning for medical image analysis with convolutional autoencoder neural network," *IEEE Trans. Big Data*, vol. 7, no. 4, pp. 750–758, Oct. 2021.

[70] A. D. Dongare, R. R. Kharde, and A. D. Kachare, "Introduction to artificial neural network," *Int. J. Eng. Innov. Technol.*, vol. 2, no. 1, pp. 189–194, 2012.

[71] R. Zhu, X. Tu, and J. X. Huang, "Deep learning on information retrieval and its applications," in *Deep Learning for Data Analytics*. Amsterdam, The Netherlands: Elsevier, 2020, pp. 125–153.

[72] S. Tan, X. Cheng, Y. Wang, and H. Xu, "Adapting naive Bayes to domain adaptation for sentiment analysis," in *Proc. Eur. Conf. Inf. Retr.* Cham, Switzerland: Springer, 2009, pp. 337–349.

[73] B. Li, X. Chen, M. J. Li, J. Z. Huang, and S. Feng, "Scalable random forests for massive data," in *Proc. Pacific-Asia Conf. Knowl. Discovery Data Mining*. Cham, Switzerland: Springer, 2012, pp. 135–146.

[74] H. Yang and S. Fong, "Improving adaptability of decision tree for mining big data," *New Math. Natural Comput.*, vol. 9, no. 1, pp. 77–95, Mar. 2013.

[75] S. Sun and R. Huang, "An adaptive K-nearest neighbor algorithm," in *Proc. FSKD*, vol. 1, 2010, pp. 91–94.

[76] P. D. Gutiérrez, M. Lastra, J. Bacardit, J. M. Benítez, and F. Herrera, "GPU-SME-kNN: Scalable and memory efficient kNN and lazy learning using GPUs," *Inf. Sci.*, vol. 373, pp. 165–182, Dec. 2016.

[77] A. S. M. Murugavel, S. Ramakrishnan, U. Maheswari, and B. S. Sabetha, "Combined seizure index with adaptive multi-class SVM for epileptic EEG classification," in *Proc. Int. Conf. Emerg. Trends VLSI, Embedded Syst., Nano Electron. Telecommun. Syst. (ICEVENT)*, Jan. 2013, pp. 1–5.

[78] T. Razzaghi and I. Safro, "Scalable multilevel support vector machines," *Proc. Comput. Sci.*, vol. 51, pp. 2683–2687, Jan. 2015.

[79] B. Martin-Barragan, R. Lillo, and J. Romo, "Interpretable support vector machines for functional data," *Eur. J. Oper. Res.*, vol. 232, no. 1, pp. 146–155, Jan. 2014.

[80] D. Li, H. Zhang, M. S. Khan, and F. Mi, "A self-adaptive frequency selection common spatial pattern and least squares twin support vector machine for motor imagery electroencephalography recognition," *Biomed. Signal Process. Control*, vol. 41, pp. 222–232, Mar. 2018.

[81] B. R. Prasad and S. Agarwal, "Scalable least square twin support vector machine learning," in *Proc. Int. Conf. Big Data Analytics Knowl. Discovery*. Cham, Switzerland: Springer, 2019, pp. 239–249.

[82] S. Chakrabarti, A. Swetapadma, A. Ranjan, and P. K. Pattnaik, "Time domain implementation of pediatric epileptic seizure detection system for enhancing the performance of detection and easy monitoring of pediatric patients," *Biomed. Signal Process. Control*, vol. 59, May 2020, Art. no. 101930.

[83] R. Sarić, D. Jokić, N. Beganović, L. G. Pokvić, and A. Badnjević, "FPGA-based real-time epileptic seizure classification using artificial neural network," *Biomed. Signal Process. Control*, vol. 62, Sep. 2020, Art. no. 102106.

[84] K. Zhang, N. Robinson, S.-W. Lee, and C. Guan, "Adaptive transfer learning for EEG motor imagery classification with deep convolutional neural network," *Neural Netw.*, vol. 136, pp. 1–10, Apr. 2021.

[85] G. Placidi, L. Cinque, and M. Polsinelli, "A fast and scalable framework for automated artifact recognition from EEG signals represented in scalp topographies of independent components," *Comput. Biol. Med.*, vol. 132, May 2021, Art. no. 104347.

[86] D. Borra, S. Fantozzi, and E. Magosso, "Interpretable and lightweight convolutional neural network for EEG decoding: Application to movement execution and imagination," *Neural Netw.*, vol. 129, pp. 55–74, Sep. 2020.

[87] R. Y. Aminabadi, O. Ruwase, M. Zhang, Y. He, J.-M. Arnau, and A. Gonazalez, "SHARP: An adaptable, energy-efficient accelerator for recurrent neural networks," *ACM Trans. Embedded Comput. Syst.*, vol. 22, no. 2, pp. 1–23, Mar. 2023.

[88] G. Paulin, F. Conti, L. Cavigelli, and L. Benini, "Vau da muntanialas: Energy-efficient multi-die scalable acceleration of RNN inference," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 1, pp. 244–257, Jan. 2022.

[89] S. Wisdom, T. Powers, J. Pitton, and L. Atlas, "Building recurrent networks by unfolding iterative thresholding for sequential sparse recovery," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 4346–4350.

[90] D. Niu, Z. Xia, Y. Liu, T. Cai, T. Liu, and Y. Zhan, "ALSTM: Adaptive LSTM for durative sequential data," in *Proc. IEEE 30th Int. Conf. Tools Artif. Intell. (ICTAI)*, Nov. 2018, pp. 151–157.

[91] A. M. Alonso, F. J. Nogales, and C. Ruiz, "A single scalable LSTM model for short-term forecasting of disaggregated electricity loads," 2019, *arXiv:1910.06640*.

[92] A. Samavat, E. Khalili, B. Ayati, and M. Ayati, "Deep learning model with adaptive regularization for EEG-based emotion recognition using temporal and frequency features," *IEEE Access*, vol. 10, pp. 24520–24527, 2022.

[93] C. Gan, Q. Feng, and Z. Zhang, "Scalable multi-channel dilated CNN–BiLSTM model with attention mechanism for Chinese textual sentiment analysis," *Future Gener. Comput. Syst.*, vol. 118, pp. 297–309, May 2021.

[94] E. A. Kosmatos, N. D. Tselikas, and A. C. Boucouvalas, "Integrating RFIDs and smart objects into a unified Internet of Things architecture," *Adv. Internet Things*, vol. 1, no. 1, pp. 5–12, 2011, doi: 10.4236/ait.2011.11002.

[95] Statista. (2021). *Global IoT and Non-IoT Connections 2010–2025*. [Online]. Available: https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/

[96] C. Stewart. (Dec. 2016). *IoT in Healthcare Market SIZE Worldwide 2016–2025 Forecast*. [Online]. Available: https://www.statista.com/statistics/997959/worldwide-internet-of-things-in-healthcare-market-size/

[97] S. Soundarya, "An EEG based emotion recognition and classification using machine learning techniques, I," *J. Emerg. Technol. Innov. Eng.*, vol. 5, pp. 744–750, Apr. 2019.

[98] K. Barjinder, S. Dinesh, and P. R. Partha, "EEG based emotion classification mechanism in BCI," *Proc. Comput. Sci.*, vol. 132, pp. 752–758, Jan. 2018.

[99] G. Muhammad, M. S. Hossain, and N. Kumar, "EEG-based pathology detection for home health monitoring," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 2, pp. 603–610, Feb. 2021.

[100] N. Borulkar, P. Pandey, C. Davda, and J. Chettiar, "Drowsiness detection and monitoring the sleeping pattern using brainwaves technology and IoT," in *Proc. 2nd Int. Conf. I-SMAC (IoT Social, Mobile, Analytics Cloud) (I-SMAC)I-SMAC (IoT Social, Mobile, Analytics Cloud) (I-SMAC), 2nd Int. Conf.*, Aug. 2018, pp. 703–706.

[101] X. Zhang, J. Li, Y. Liu, Z. Zhang, Z. Wang, D. Luo, X. Zhou, M. Zhu, W. Salman, G. Hu, and C. Wang, "Design of a fatigue detection system for high-speed trains based on driver vigilance using a wireless wearable EEG," *Sensors*, vol. 17, no. 3, p. 486, 2017.

[102] C. O. Fernandes, C. J. P. de Lucena, and D. D. S. E. Silva, "Smart depth of anesthesia monitoring with EEG sensors and agent-based technology," in *Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Adv. Trusted Comput., Scalable Comput. Commun., Cloud Big Data Comput., Internet People Smart City Innov. (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Aug. 2017, pp. 1–8.

[103] A. Carrasquilla-Batista, K. Quirós-Espinoza, and C. Gómez-Carrasquilla, "An Internet of Things (IoT) application to control a wheelchair through EEG signal processing," in *Proc. Int. Symp. Wearable Robot. Rehabil. (WeRob)*, p. 1, 2017, doi: 10.1109/WEROB.2017.8383877.

[104] S. Soman and B. K. Murthy, "Using brain computer interface for synthesized speech communication for the physically disabled," *Proc. Comput. Sci.*, vol. 46, pp. 292–298, Mar. 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050915000873

[105] A. Kurapa, D. Rathore, D. R. Edla, A. Bablani, and V. Kuppili, "A hybrid approach for extracting EMG signals by filtering EEG data for IoT applications for immobile persons," *Wireless Pers. Commun.*, vol. 114, no. 4, pp. 3081–3101, Oct. 2020, doi: 10.1007/s11277-020-07518-5.

[106] M.-P. Hosseini, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, "Optimized deep learning for EEG big data and seizure prediction BCI via Internet of Things," *IEEE Trans. Big Data*, vol. 3, no. 4, pp. 392–404, Dec. 2017.

[107] M. A. Sayeed, S. P. Mohanty, E. Kougianos, and H. Zaveri, "A fast and accurate approach for real-time seizure detection in the IoMT," in *Proc. IEEE Int. Smart Cities Conf. (ISC)*, Sep. 2018, pp. 1–5.

[108] A. Sayeed, S. P. Mohanty, E. Kougianos, V. P. Yanambaka, and H. Zaveri, "A robust and fast seizure detector for IoT edge," in *Proc. IEEE Int. Symp. Smart Electron. Syst. (iSES) (Formerly iNiS)*, Dec. 2018, pp. 156–160.

[109] M. Alhussein, G. Muhammad, M. S. Hossain, and S. U. Amin, "Cognitive IoT-cloud integration for smart healthcare: Case study for epileptic seizure detection and monitoring," *Mobile Netw. Appl.*, vol. 23, pp. 1624–1635, Dec. 2018.

[110] M. A. Sayeed, S. P. Mohanty, E. Kougianos, and H. P. Zaveri, "ESeiz: An edge-device for accurate seizure detection for smart healthcare," *IEEE Trans. Consum. Electron.*, vol. 65, no. 3, pp. 379–387, Aug. 2019.

[111] A. Gupta, C. Chakraborty, and B. Gupta, "Monitoring of epileptical patients using cloud-enabled health-IoT system," *Traitement du Signal*, vol. 36, no. 5, pp. 425–431, Nov. 2019.

[112] M. A. Sayeed, S. P. Mohanty, E. Kougianos, and H. P. Zaveri, "Neurodetect: A machine learning-based fast and accurate seizure detection system in the IoMT," *IEEE Trans. Consum. Electron.*, vol. 65, no. 3, pp. 359–368, Aug. 2019.

[113] K. Singh and J. Malhotra, "IoT and cloud computing based automatic epileptic seizure detection using HOS features based random forest classification," *J. Ambient Intell. Humanized Comput.*, vol. 10, pp. 1–16, Dec. 2019, doi: 10.1007/s12652-019-01613-7.

[114] H. Daoud, P. Williams, and M. Bayoumi, "IoT based efficient epileptic seizure prediction system using deep learning," in *Proc. IEEE 6th World Forum Internet Things (WF-IoT)*, Jun. 2020, pp. 1–6.

[115] B. Direito, C. A. Teixeira, F. Sales, M. Castelo-Branco, and A. Dourado, "A realistic seizure prediction study based on multiclass SVM," *Int. J. Neural Syst.*, vol. 27, no. 3, May 2017, Art. no. 1750006. [Online]. Available: https://www.worldscientific.com/doi/abs/10.1142/S012906571750006X

[116] N. Sriraam and S. Raghu, "Classification of focal and non focal epileptic seizures using multi-features and SVM classifier," *J. Med. Syst.*, vol. 41, no. 10, pp. 1–14, Oct. 2017.

[117] A. Subasi, J. Kevric, and M. A. Canbaz, "Epileptic seizure detection using hybrid machine learning methods," *Neural Comput. Appl.*, vol. 31, no. 1, pp. 317–325, Jan. 2019, doi: 10.1007/s00521-017-3003-y.

[118] A. Hamad, E. H. Houssein, A. E. Hassanien, and A. A. Fahmy, "Hybrid grasshopper optimization algorithm and support vector machines for automatic seizure detection in EEG signals," in *Proc. Int. Conf. Adv. Mach. Learn. Technol. Appl.* Cham, Switzerland: Springer, 2018, pp. 82–91.

[119] P. Agarwal, H.-C. Wang, and K. Srinivasan, "Epileptic seizure prediction over EEG data using hybrid CNN-SVM model with edge computing services," in *Proc. MATEC Web Conf.*, vol. 210, 2018, p. 03016.

[120] S. Raghu, N. Sriraam, Y. Temel, S. V. Rao, A. S. Hegde, and P. L. Kubben, "Performance evaluation of DWT based sigmoid entropy in time and frequency domains for automated detection of epileptic seizures using SVM classifier," *Comput. Biol. Med.*, vol. 110, pp. 127–143, Jul. 2019.

[121] T. Siddharth, P. Gajbhiye, R. K. Tripathy, and R. B. Pachori, "EEG-based detection of focal seizure area using FBSE-EWT rhythm and SAE-SVM network," *IEEE Sensors J.*, vol. 20, no. 19, pp. 11421–11428, Oct. 2020.

[122] M. Mursalin, Y. Zhang, Y. Chen, and N. V. Chawla, "Automated epileptic seizure detection using improved correlation-based feature selection with random forest classifier," *Neurocomputing*, vol. 241, pp. 204–214, Jun. 2017.

[123] T. Zhang, W. Chen, and M. Li, "Generalized stockwell transform and SVD-based epileptic seizure detection in EEG using random forest," *Biocybernetics Biomed. Eng.*, vol. 38, no. 3, pp. 519–534, 2018.

[124] X. Wang, G. Gong, N. Li, and S. Qiu, "Detection analysis of epileptic EEG using a novel random forest model combined with grid search optimization," *Frontiers Hum. Neurosci.*, vol. 13, p. 52, Feb. 2019.

[125] Y. Wang, J. Cao, X. Lai, and D. Hu, "Epileptic state classification for seizure prediction with wavelet packet features and random forest," in *Proc. Chinese Control Decis. Conf.*, Jun. 2019, pp. 3983–3987.

[126] M. Sameer and B. Gupta, "Detection of epileptical seizures based on alpha band statistical features," *Wireless Pers. Commun.*, vol. 115, no. 2, pp. 909–925, Nov. 2020.

[127] M. Chakraborty and D. Mitra, "Epilepsy seizure detection using kurtosis based VMD's parameters selection and bandwidth features," *Biomed. Signal Process. Control*, vol. 64, Feb. 2021, Art. no. 102255.

[128] H. Rajaguru and S. K. Prabhakar, "Power spectral density and KNN based AdaBoost classifier for epilepsy classification from EEG," in *Proc. Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Apr. 2017, pp. 441–444.

[129] S. W. Ibrahim, R. Djemal, A. Alsuwailem, and S. Gannouni, "Electroencephalography (EEG)-based epileptic seizure prediction using entropy and K-nearest neighbor (KNN)," *Commun. Sci. Technol.*, vol. 2, no. 1, pp. 6–10, May 2017.

[130] H. Choubey and A. Pandey, "A combination of statistical parameters for the detection of epilepsy and EEG classification using ANN and KNN classifier," *Signal, Image Video Process.*, vol. 15, no. 3, pp. 475–483, Apr. 2021.

[131] H. Akbari, S. S. Esmaili, and S. F. Zadeh, "Detection of seizure EEG signals based on reconstructed phase space of rhythms in EWT domain and genetic algorithm," *Signal Process. Renew. Energy*, vol. 4, no. 2, pp. 23–36, 2020.

[132] A. Sharmila and P. Geethanjali, "DWT based detection of epileptic seizure from EEG signals using naive Bayes and k-NN classifiers," *IEEE Access*, vol. 4, pp. 7716–7727, 2016.

[133] C. Xiao, S. Wang, L. Iasemidis, S. Wong, and W. A. Chaovalitwongse, "An adaptive pattern learning framework to personalize online seizure prediction," *IEEE Trans. Big Data*, vol. 7, no. 5, pp. 819–831, Nov. 2021.

[134] R. J. Martis, U. R. Acharya, J. H. Tan, A. Petznick, R. Yanti, C. K. Chua, E. Y. K. Ng, and L. Tong, "Application of empirical mode decomposition (EMD) for automated detection of epilepsy using EEG signals," *Int. J. Neural Syst.*, vol. 22, no. 6, Dec. 2012, Art. no. 1250027.

[135] H. Rajaguru and S. K. Prabhakar, "Sparse PCA and soft decision tree classifiers for epilepsy classification from EEG signals," in *Proc. Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Apr. 2017, pp. 581–584.

[136] J. M.-T. Wu, M.-H. Tsai, C.-T. Hsu, H.-C. Huang, and H.-C. Chen, "Intelligent signal classifier for brain epileptic EEG based on decision tree, multilayer perceptron and over-sampling approach," in *Advances in Information and Communication* (Lecture Notes in Networks and Systems), K. Arai and R. Bhatia, Eds. Cham, Switzerland: Springer, 2020, pp. 11–24.

[137] H. Rajaguru and S. K. Prabhakar, "Analysis of AdaBoost classifier from compressed EEG features for epilepsy detection," in *Proc. Int. Conf. Comput. Methodologies Commun. (ICCMC)*, Jul. 2017, pp. 981–984.

[138] H. Al-Hadeethi, S. Abdulla, M. Diykh, R. C. Deo, and J. H. Green, "Adaptive boost LS-SVM classification approach for time-series signal classification in epileptic seizure diagnosis applications," *Exp. Syst. Appl.*, vol. 161, Dec. 2020, Art. no. 113676.

[139] A. R. Hassan, A. Subasi, and Y. Zhang, "Epilepsy seizure detection using complete ensemble empirical mode decomposition with adaptive noise," *Knowledge-Based Syst.*, vol. 191, Mar. 2020, Art. no. 105333.

[140] L. Guo, D. Rivero, J. Dorado, J. R. Rabuñal, and A. Pazos, "Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks," *J. Neurosci. Methods*, vol. 191, no. 1, pp. 101–109, 2010.

[141] E. Juarez-Guerra, V. Alarcon-Aquino, and P. Gomez-Gil, "Epilepsy seizure detection in EEG signals using wavelet transforms and neural networks," in *New Trends in Networking, Computing, E-Learning, Systems Sciences, and Engineering*. Berlin, Germany: Springer, 2015, pp. 261–269.

[142] N. K. Ambulkar and S. N. Sharma, "Detection of epileptic seizure in EEG signals using window width optimized S-transform and artificial neural networks," in *Proc. IEEE Bombay Sect. Symp. (IBSS)*, Sep. 2015, pp. 1–6.

[143] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Comput. Biol. Med.*, vol. 100, pp. 270–278, Sep. 2017.

[144] X. Wei, L. Zhou, Z. Chen, L. Zhang, and Y. Zhou, "Automatic seizure detection using three-dimensional CNN based on multi-channel EEG," *BMC Med. Informat. Decis. Making*, vol. 18, no. 5, pp. 71–80, Dec. 2018.

[145] M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo, and J. Xiang, "Epileptic seizure detection based on EEG signals and CNN," *Frontiers Neuroinform.*, vol. 12, p. 95, Dec. 2018. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fninf.2018.00095/full

[146] C. Huang, W. Chen, and G. Cao, "Automatic epileptic seizure detection via attention-based CNN-BiRNN," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Nov. 2019, pp. 660–663.

[147] G. C. Jana, R. Sharma, and A. Agrawal, "A 1D-CNN-spectrogram based approach for seizure detection from EEG signal," *Proc. Comput. Sci.*, vol. 167, pp. 403–412, Jan. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050920307146

[148] P. Boonyakitanont, A. Lek-Uthai, and J. Songsiri, "Automatic epileptic seizure onset-offset detection based on CNN in scalp EEG," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 1225–1229.

[149] S. Raghu, N. Sriraam, Y. Temel, S. V. Rao, and P. L. Kubben, "EEG based multi-class seizure type classification using convolutional neural network and transfer learning," *Neural Netw.*, vol. 124, pp. 202–212, Apr. 2020.

[150] H. Takahashi, A. Emami, T. Shinozaki, N. Kunii, T. Matsuo, and K. Kawai, "Convolutional neural network with autoencoder-assisted multiclass labelling for seizure detection based on scalp electroencephalography," *Comput. Biol. Med.*, vol. 125, Oct. 2020, Art. no. 104016.

[151] T. Wen and Z. Zhang, "Deep convolution neural network and autoencoders-based unsupervised feature learning of EEG signals," *IEEE Access*, vol. 6, pp. 25399–25410, 2018.

[152] L. Vidyaratne, A. Glandon, M. Alam, and K. M. Iftekharuddin, "Deep recurrent neural network for seizure detection," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 1202–1207.

[153] R. Hussein, H. Palangi, R. K. Ward, and Z. J. Wang, "Optimized deep neural network architecture for robust detection of epileptic seizures using EEG signals," *Clin. Neurophysiol.*, vol. 130, no. 1, pp. 25–37, 2019.

[154] M. U. Abbasi, A. Rashad, A. Basalamah, and M. Tariq, "Detection of epilepsy seizures in neo-natal EEG using LSTM architecture," *IEEE Access*, vol. 7, pp. 179074–179085, 2019.

[155] I. Aliyu, Y. B. Lim, and C. G. Lim, "Epilepsy detection in EEG signal using recurrent neural network," in *Proc. 3rd Int. Conf. Intell. Syst., Metaheuristics Swarm Intell.*, Mar. 2019, pp. 50–53.

[156] H. Daoud and M. A. Bayoumi, "Efficient epileptic seizure prediction based on deep learning," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 5, pp. 804–813, Oct. 2019.

[157] K. Baskar and C. Karthikeyan, "Epilepsy seizure detection using Akima spline interpolation based ensemble empirical mode Kalman filter decomposition by EEG signals," *J. Med. Imag. Health Informat.*, vol. 9, no. 6, pp. 1320–1328, Aug. 2019.

[158] Y. Liu, Y.-X. Huang, X. Zhang, W. Qi, J. Guo, Y. Hu, L. Zhang, and H. Su, "Deep C-LSTM neural network for epileptic seizure and tumor detection using high-dimension EEG signals," *IEEE Access*, vol. 8, pp. 37495–37504, 2020.

[159] X. Hu, S. Yuan, F. Xu, Y. Leng, K. Yuan, and Q. Yuan, "Scalp EEG classification using deep bi-LSTM network for seizure detection," *Comput. Biol. Med.*, vol. 124, Sep. 2020, Art. no. 103919.

[160] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *Int. J. Data Mining Knowl. Manag. Process*, vol. 5, no. 2, pp. 1–11, Mar. 2015.

[161] J. Lever, M. Krzywinski, and N. Altman, "Classification evaluation," *Nature Methods*, vol. 13, no. 8, pp. 603–604, 2016, doi: 10.1038/nmeth.3945.

[162] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. thesis, Massachusetts Inst. Technol., Cambridge, MA, USA, 2009. [Online]. Available: https://dspace.mit.edu/handle/1721.1/54669

[163] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 64, no. 6, Nov. 2001, Art. no. 061907.

[164] P. Swami, B. Panigrahi, S. Nara, M. Bhatia, and T. Gandhi, "EEG epilepsy datasets," 2016, doi: 10.13140/RG.2.2.14280.32006.

[165] J. Klatt, H. Feldwisch-Drentrup, M. Ihle, V. Navarro, M. Neufang, C. Teixeira, C. Adam, M. Valderrama, C. Alvarado-Rojas, A. Witon, M. Le Van Quyen, F. Sales, A. Dourado, J. Timmer, A. Schulze-Bonhage, and B. Schelter, "The EPILEPSIAE database: An extensive electroencephalography database of epilepsy patients," *Epilepsia*, vol. 53, no. 9, pp. 1669–1676, Sep. 2012.

[166] I. Ahmad, X. Wang, M. Zhu, C. Wang, Y. Pi, J. A. Khan, S. Khan, O. W. Samuel, S. Chen, and G. Li, "EEG-based epileptic seizure detection via machine/deep learning approaches: A systematic review," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–20, Jun. 2022.

[167] A. Das, S. S. Cash, and T. J. Sejnowski, "Heterogeneity of preictal dynamics in human epileptic seizures," *IEEE Access*, vol. 8, pp. 52738–52748, 2020.

[168] S. Hassan, E. Mwangi, and P. K. Kihato, "IoT based monitoring system for epileptic patients," *Heliyon*, vol. 8, no. 6, Jun. 2022, Art. no. e09618.

[169] R. Chengoden, N. Victor, T. Huynh-The, G. Yenduri, R. H. Jhaveri, M. Alazab, S. Bhattacharya, P. Hegde, P. K. R. Maddikunta, and T. R. Gadekallu, "Metaverse for healthcare: A survey on potential applications, challenges and future directions," 2022, *arXiv:2209.04160*.

[170] M. Alazab, L. U. Khan, S. Koppu, S. P. Ramu, P. Boobalan, T. Baker, P. K. R. Maddikunta, T. R. Gadekallu, and A. Aljuhani, "Digital twins for healthcare 4.0—Recent advances, architecture, and open challenges," *IEEE Consum. Electron. Mag.*, early access, Sep. 23, 2022, doi: 10.1109/MCE.2022.3208986.

[171] E. Bruno, P. F. Viana, M. R. Sperling, and M. P. Richardson, "Seizure detection at home: Do devices on the market match the needs of people living with epilepsy and their caregivers?" *Epilepsia*, vol. 61, no. 1, pp. 11–24, Nov. 2020.

SOBHANA JAHAN received the B.Sc. and M.Sc. degrees in information and communication engineering from the Bangladesh University of Professionals (BUP), in 2020 and 2022, respectively. She was a Lecturer with the Green University of Bangladesh (GUB) and a Teaching Assistant with BUP. She is currently a Lecturer with the Department of Computer Science and Engineering, BUP. She has published several articles at international conferences. Her research interests include artificial intelligence, machine learning, and the IoT. In 2021, she was a recipient of the Research Fellowship from the ICT Division of the Bangladesh Government.

FARHANA NOWSHEEN received the B.Sc. degree in information and communication engineering from the Bangladesh University of Professionals, in 2020, where she is currently pursuing the M.Sc. degree. Her research interests include artificial intelligence, machine learning, and the IoT.

MAHATHIR MAHMUD ANTIK received the B.Sc. degree in information and communication technology from the Bangladesh University of Professionals, in 2020, where he is currently pursuing the M.Sc. degree. He has previously published articles in the field of artificial intelligence. His areas of research interests include machine learning, the Internet of Things, material science for computer applications, and artificial intelligence.

MD. SAZZADUR RAHMAN received the B.Sc. and M.S. degrees in applied physics, electronics, and communication engineering from the University of Dhaka, Dhaka, Bangladesh, in 2005 and 2006, respectively, and the Ph.D. degree in material science from Kyushu University, Japan, in 2015. He was a Faculty Member with the Faculty of Computer Science and Engineering, Hajee Mohammad Danesh Science and Technology University, from May 2009 to November 2018. Since November 2018, he has been with the Institute of Information Technology, Jahangirnagar University, Savar, Dhaka, where he is currently an Associate Professor. His research interests include material science for computer applications, surface science, ubiquitous computing, WSNs, machine learning, and the IoT.

M. SHAMIM KAISER (Senior Member, IEEE) received the bachelor's and master's degrees in applied physics, electronics, and communication engineering from the University of Dhaka, Bangladesh, in 2002 and 2004, respectively, and the Ph.D. degree in telecommunication engineering from the Asian Institute of Technology (AIT), Pathum Thani, Thailand, in 2010. In 2005, he joined the Department of ETE, Daffodil International University, as a Lecturer. In 2010, he was an Assistant Professor with the Department of EEE, Eastern University, Bangladesh, and the Department of MNS, BRAC University, Dhaka. Since 2011, he has been an Assistant Professor with the Institute of Information Technology, Jahangirnagar University, Dhaka, where he became an Associate Professor, in 2015, and a Full Professor, in 2019. He is currently the Director of the Institute of Information Technology, Jahangirnagar University. He has authored more than 100 papers in different peer-reviewed journals and conferences. His current research interests include data analytics, machine learning, wireless networks and signal processing, cognitive radio networks, big data and cyber security, and renewable energy. He is the Founding Chapter Chair of the IEEE Bangladesh Section Computer Society Chapter.

A. S. M. SANWAR HOSEN (Member, IEEE) received the M.S. and Ph.D. degrees in computer science and engineering from Jeonbuk National University, Jeonju, South Korea. He is currently an Assistant Professor with the Department of Artificial Intelligence and Big Data, Woosong University, Daejeon, South Korea. He has published several articles in journals and international conferences. His recent research interests include wireless sensor networks, the Internet of Things, fog-cloud computing, cyber security, artificial intelligence, and blockchain. He was an expert Reviewer of IEEE TRANSACTIONS, Elsevier, Springer, and MDPI journals and magazines. He was invited to serve as a guest editor and the technical programme committee member in several reputed international conferences such as IEEE and ACM.

IN-HO RA (Member, IEEE) received the Ph.D. degree in computer engineering from Chung-Ang University, Seoul, South Korea, in 1995. He was with the School of Computer, Information and Communication, Kunsan National University, where he is currently a Professor. From February 2007 to August 2008, he was a Visiting Scholar with the University of South Florida, Tampa, FL, USA. His major research interests include wireless ad hoc and sensor networks, blockchain, the IoT, PS-LTE, and microgrids.