From Seizure Detection to Smart and Fully Embedded Seizure Prediction Engine: A Review

Article *in* IEEE Transactions on Biomedical Circuits and Systems · August 2020

DOI: 10.1109/TBCAS.2020.3018465

2 authors:

Jie Yang
Westlake University
37 PUBLICATIONS **199** CITATIONS
SEE PROFILE

Mohamad Sawan
Polytechnique Montréal
798 PUBLICATIONS **7,862** CITATIONS
SEE PROFILE

From Seizure Detection to Smart and Fully Embedded Seizure Prediction Engine: A Review

Jie Yang and Mohamad Sawan, *Fellow, IEEE*

***Abstract*—Recent review papers have investigated seizure prediction, creating the possibility of preempting epileptic seizures. Correct seizure prediction can significantly improve the standard of living for the majority of epileptic patients, as the unpredictability of seizures is a major concern for them. Today, the development of algorithms, particularly in the field of machine learning, enables reliable and accurate seizure prediction using desktop computers. However, despite extensive research effort being devoted to developing seizure detection integrated circuits (ICs), dedicated seizure prediction ICs have not been developed yet. We believe that interdisciplinary study of system architecture, analog and digital ICs, and machine learning algorithms can promote the translation of scientific theory to a more realistic intelligent, integrated, and low-power system that can truly improve the standard of living for epileptic patients. This review explores topics ranging from signal acquisition analog circuits to classification algorithms and dedicated digital signal processing circuits for detection and prediction purposes, to provide a comprehensive and useful guideline for the construction, implementation and optimization of wearable and integrated smart seizure prediction systems.**

***Index Terms*—Seizure prediction, hardware system, signal processing, wearable devices, processor, electrodes, analog-to-digital converter.**

I. Introduction

EPILEPSY is a chronic disease typically characterized by recurrent seizures that are triggered by the misfiring of neurons and create an uncontrolled electrical disturbance in the brain. A seizure can cause severe disorders in the behavior, movements, emotions, and consciousness of patients, and can lead to injury or even death. Unfortunately, 1% of the world population is affected by epilepsy, and 30% of epileptic sufferers encounter drug-resistance problems during their medical treatment [1]. Epilepsy surgery is frequently conducted when patients encounter medication failure. The goal of epilepsy surgery is to remove the lesion from which the misfiring neuron originates. However, owing to reasons such as the inaccurate localization of the lesion or limitation of various medical equipment, locating the lesion is sometimes difficult [2]. Statistics have indicated that the probabilities of becoming seizure-free are 75% and 50% for lesional and nonlesional patients, respectively, after temporal lobe epilepsy surgeries [3], while approximately 60% of lesional patients and merely 35% of nonlesional patients are cured of seizures after frontal lobe epilepsy surgeries. For patients that cannot be completely cured through medicine or surgery, if seizure onsets occur in some unexpected scenarios without medical staff or family members present, they could suffer from severe comorbidity injury, sometimes even death [4]. Seizure detection enables monitoring of ongoing seizures and provide medical teams with useful seizure data as well as reduce the response time for patients in need of medical attention. Seizure prediction is expected to warn the patients at least few minutes before the seizure onset which gives chance to patients to find a safe position [5]. Therefore, an accurate seizure detection or even more advanced seizure prediction system would be significant to intractable patients as it can significantly improve the quality of their lives [6], [7].

Seizure detection and prediction rely on the sensing and processing of electrophysiological signals, such as through electroencephalography (EEG), electrocorticography (ECoG) and local field potential (LFP). These signals are the direct cause of neuron activities. Hence, they are logically used in applications for monitoring and classifying various brain states, including seizure detection and prediction. Because epilepsy is chronic, systems that acquire and process these electrophysiological signals should be compact such that they can be implanted or carried by patients to provide real-time prediction. Wearable and, in particular, implantable devices require low-power budgets to achieve long-term operation. Moreover, a reliable prediction with high sensitivity and specificity is mandatory to reduce the number of missed onset and false alarms to ensure the effectiveness and minimal disturbance of the system to patients' daily lives. Advanced semiconductor technology can integrate billions of transistors on a die of a few square millimeters with milliwatt power consumption. This is the best method to fulfill the above-mentioned features. Much research has been devoted to developing systems-on-chips that can acquire neural signals and detect seizure, and they satisfy the key features mentioned above. For example, highly integrated application-specific integrated circuits (ASIC) for EEG signal acquisition [8]–[11], ECoG signal acquisition [12]–[14], and seizure detection [15]–[17] have been reported. Recently, with the expansion of artificial intelligence, particularly with the remarkable performance that state-of-the-art networks have enabled [18]–[20], the deep learning approach is used to seek solutions in many medical applications [21]–[23]. Deep neural networks trained with numerous parameters can automatically determine a pattern through data and make highly accurate decisions. This trend has logically migrated to be applicable to seizure prediction and has spurred a number of deep learning-based seizure prediction methods [24]–[26]. Deep learning-

Jie Yang, and Mohamad Sawan are with Cutting-Edge Net of Biomedical Research and INnovation (CenBRAIN), School of Engineering, Westlake University, and with Institute of Advanced Technology, Westlake Institute for Advanced Study, 18 Shilongshan Road, Hangzhou 310024, Zhejiang Province, China. E-mail: yangjie@westlake.edu.cn

Manuscript submitted on ...

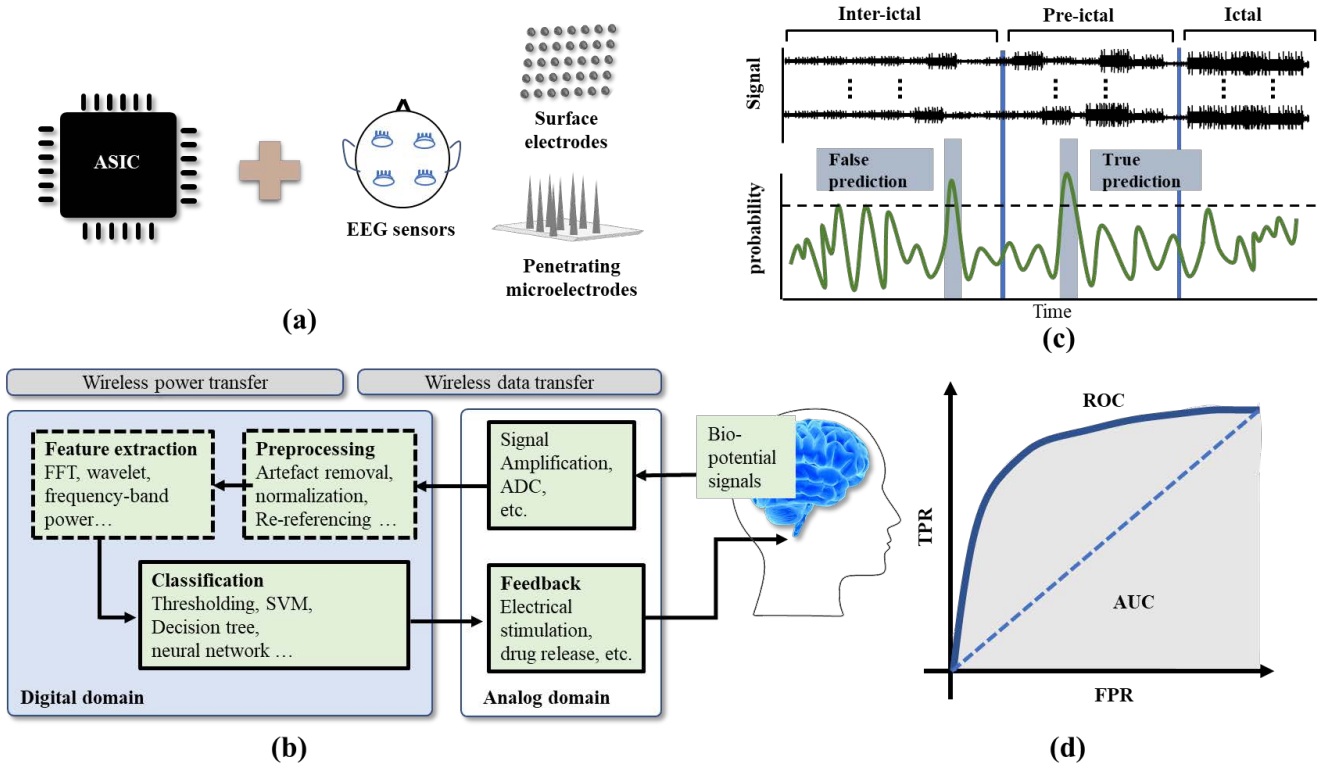


Fig. 1. Typical seizure prediction system: (a) Highly integrated seizure prediction system including a bio-interface or microelectrode array to collect data such as EEG, ECoG, and an integrated circuit for analog and digital processing. (b) Hardware building blocks of the integrated circuit. (c) The function of the circuits is to classify different seizure stages based on the captured bio-signal. (d) Performance of the system characterized by various metrics such as ROC, AUC, which are characterized by TPR, FPR of the system.

based approaches have become the most promising solutions for realistic seizure prediction owning to their excellent classification performance. However, the study above focused only focus on the algorithmic perspective. Some review papers summarized the state-of-the-art algorithms for seizure prediction [27], [28], however, these papers emphasized algorithmic approaches without considering the actual physical hardware implementations required to translate these technologies to clinical usage. Moreover, much important content was neglected such as methods of acquiring high-quality data for prediction, the type of computing platform required to execute the algorithms, and what the critical metrics of the system are. To date, dedicated seizure prediction chips have not been developed yet.

Existing research efforts in signal acquisition, seizure detection algorithms, and system integrations have provided a solid foundation toward the future development of smart and fully embedded seizure prediction integrated circuits (ICs). In this review, we first outline the building blocks of a seizure detection and prediction system, subsequently, we provide detailed review of each functional block of the system, including analog front-end (AFE) circuits, classification algorithms, and digital signal processing circuits, with special emphasis on ASIC integration. Finally, future prospects and challenges to building smart and fully embedded seizure prediction systems are discussed.

II. Systems Overview

Clinic trials have proved the efficacy of seizure detection and treatment. Bergy *et al.* reported the long-term treatment effects of the responsive neurostimulation (RNS) system [30]. The RNS device has been approved by US Food and Drug Administration (FDA), it operates using strip leads with electrodes that are surgically placed at seizure foci. The RNS feedback system includes a sensor to monitor seizures and a programmable neurostimulator to deliver stimuli. When the RNS system detects abnormal ECoG signals at the onset of a seizure, it delivers a set of stimuli to the foci. A long-term study indicated that the device reduces the frequency of seizure and improves the quality of patients' lives [30]. Despite its significant impact, RNS can only detect the onset of seizure, it cannot perform seizure prediction. The world's first seizure prediction system that underwent a clinical trial was reported in 2013 [31]. It is a seizure advisory system comprises intracranially implanted electrodes, a subdermal telemetry unit and a handheld unit. The electrodes and telemetry unit are connected using subdermal wires, and the obtained signals are transmitted wirelessly to the handheld device. The handheld device executes an algorithm based on the received signals and can provide three levels of warning of the likelihood of a seizure onset to the patients, namely, high, moderate, or low. The system reported above the chance prediction accuracy for nine individuals that participated in the clinical trial. It is

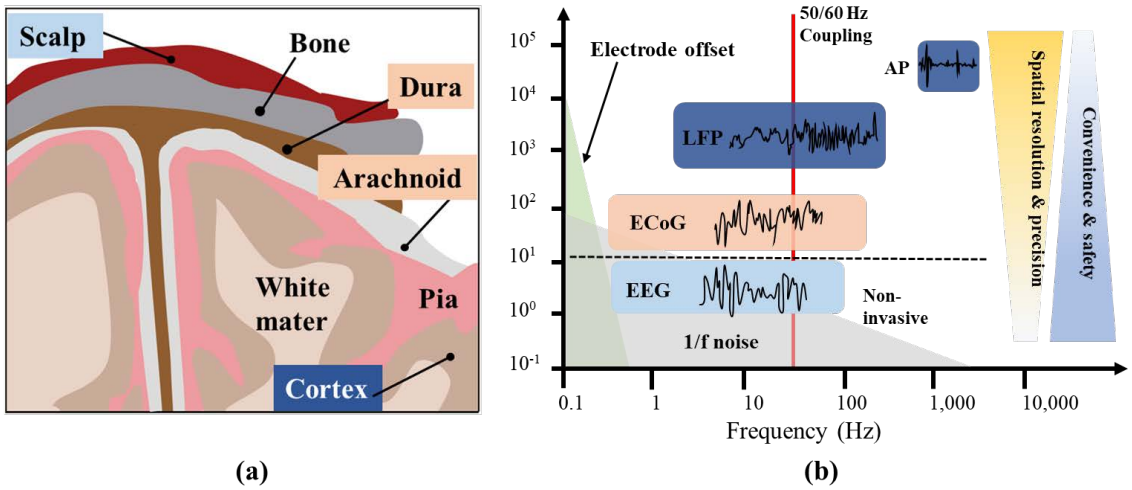


Fig. 2. Different signal acquisition methods and their characteristics: (a) brain anatomy [29], (b) neural signal characteristics.

worth noting that 86%, 100% and 100% seizures are correctly predicted for three patients. Nine individuals participated in the clinical trial of this system. Note that 86%, 100% and 100% seizures were predicted for three patients. The sensitivities of the other six individuals ranged from 54% to 71%. This system combines all necessary components and provides a good blueprint for a seizure prediction system, however, it requires improvement. For example, the handheld device can be replaced with a dedicated processor to increase system compactness. Wireless data transfer can directly transfer acquired data to an off-body device [32], [33]. Improvement in machine learning algorithms [24] can be employed to provide a better prediction performance. Furthermore, lower power consumption and an easy-to-use system are worth implementing.

Fig. 1 (a) shows a general wearable or implant seizure detection or prediction scheme. It contains a certain type of electrodes and a dedicated processing IC. The electrodes can be dry or gel-based for EEG, invasive non-penetrating electrodes for ECoG, or invasive penetrating microelectrodes for LFP or action potential (AP) recording. Signals from various types of electrodes are continuously sampled and processed by the chip. Without constraining the actual implementation method, Fig. 1(b) lists all possible building blocks. They comprise analog circuits such as filters, instrumentation amplifier to de-noise and amplify the signals of interest, analog to digital converter (ADC) to bridges the analog and digital domain to enable further algorithmic processing, and digital processing circuits that analyze and classify the stream of digital signals using the pre-defined algorithms. Moreover, feedback circuits implemented for further actions such as applying electrical stimulation and drug release. Wireless power and data transfer blocks are commonly used in implantable devices to facilitate long-term operation and mitigate discomfort caused by skin penetration. The sampled signals contain typical brain states that can be used for classification. Fig. 1(c) shows different brainwave states. The prediction and detection of seizures can be considered as classification of preictal and ictal signals. The successful recognition of preictal brainwave signals allows intervention before seizure starts. Preictal duration range from 30 to few more minutes has been reported [25], [34]–[36]. A short-term preictal duration is considered enough for patients to take some basic precautions [5]. The goal of the classification algorithm is to distinguish different epileptic brainwave states accurately. The performance of the classification algorithm can be characterized by true positive and false positive which provide further metrics such as receiver operating characteristic (ROC) curve and area under the curve (AUC). The details of different epileptic brainwave states and the figures of merit are provided in Section IV.

Reflecting Fig. 1(b), we analyze the AFE for non-invasive and invasive brainwave signal acquisition in Section III, and signal processing algorithm and dedicated digital processing circuits in Section IV.

III. Signal Acquisition Circuits

Neuron activities generate AP signals that propagate from intracellular to extracellular sites and finally reach the scalp. The signals can be captured at various stages (Fig. 2(a)). EEG consists of electrical signals measured from the scalp using noninvasive electrodes. ECoG reads signals from sensors placed either above or below the dura mater. Microelectrodes implanted at the cortical layers are used to record LFP or AP. Among all methods shown in Fig. 2(a), ECoG and microelectrodes require neurosurgical implantation. Microelectrodes can cause damage to cells, hence EEG is the most dominant method to date owing to its safety and easy installation. In contrast, EEG signals are less informative than others because measured EEG signals are spatially filtered by the dura mater and skull and have a smaller spatial resolution. Fig. 2(b) shows the characteristics, and the changing trends of the spatial resolution, signal quality, safety and convenience of different type of neural signals. These signals have limited amplitude of a few microvolts to millivolts and commonly affected by various types of noise and interference. In the remainder of this section, we review dedicated AFE used to acquire these signals.

TABLE I
NON-INVASIVE EEG SIGNAL ACQUISITION SYSTEMS

Work	JSSC2007 Yazicioglu *et al.* [37]	JSSC2008 Yazicioglu *et al.* [9]	JSSC2010 Verma *et al.* [38]	TBioCAS2011 Xu *et al.* [39]	JSSC2014 Xu *et al.* [11]	JSSC2015 Xu *et al.* [40]	JSSC2018 Ha *et al.* [41]	ISSCC2020 Tang *et al.* [42]
Noise	60 nV/$\sqrt{}$Hz	0.59 μVrms	1.3 μVrms	0.8 μVrms	1.75μVrms	0.65μVrms	0.48 μVrms	0.27μVrms
CMRR (dB)	120	120	>60	82	84	102	110	> 100
Elec. offset	± 50 mV	± 45 mV	rail-to-rail	rail-to-rail	± 250 mV	± 350 mV	± 300 mv	rail-to-rail
Impedance	-	>1$G\Omega$	> 700 $M\Omega$	> 2$G\Omega$	1.2 $G\Omega$	100$M\Omega$@50Hz 1$G\Omega$@1Hz	1 $G\Omega$	-
Techonlogy (μm)	0.5	0.5	0.18	0.18	0.18	0.18	0.065	0.18
Area (mm^2)	0.65 × 3	3.9 × 4.5	0.3	1.8 × 3.6	3.9 × 4.5	3.5 × 4.5	-	-
No. of channels	-	8	-	8	8	16	2	1-16
ADC	-	11-bit SAR	12-bit SAR	-	12-bit SAR	12-bit SAR	12-bit SAR	10-bit SAR
Power μW	60	200*	3.5	160*	<700*	∼ 105	∼ 42.9*	-

* Total power consumption of all channels.

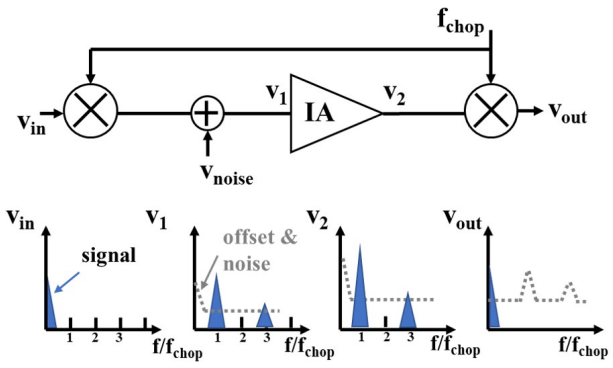


Fig. 3. Chopping modulation to reduce 1/f noise and offset.

A. EEG Acquisition

As Fig. 2(b) shows, the amplitude of an EEG signals is significantly weak. As a consequence, suitable electrodes and circuits are needed to acquire high-quality signals. In noninvasive applications, wet (gel-based) electrodes are the most commonly used in laboratory and clinical environments. Conductive gels make capturing signals easier. Nevertheless, obtaining high-quality EEG signals is challenging. Fig. 2(b) indicates that the EEG signal is susceptible to 1/f noise, electrode offset, and interference from a mains supply. Hence, dedicated circuits are necessary to clear and amplify the signals. A conventional approach is to use a low-noise amplifier after the electrodes to condition the signal. The circuits are specially designed to optimize key criteria such as input-referred noise, input impedance, common-mode rejection ratio (CMRR), electrode offset rejection, and power consumption.

Table I summaries the results of recent studies that attempted to overcome the mentioned challenges and optimize the corresponding criterion. To reduce the 1/f noise and reject the electrode offset, Yazicioglu *et al.* proposed an AC-coupled chopped instrumentation amplifier (ACCIA) in 2007 [37]. As shown in Fig. 3, chopping works by alternating the input signal at the input stage and then chopping it again at the output stage. It can effectively reduce 1/f noise and amplifier offset voltage by modulating them to the chopping frequency. AC-coupling is introduced using a DC servo loop which rejects DC signals, thereby filtering out the electrode offset. Fig. 4 illustrates the concept of blocking DC signal using a low-pass filter in the feedback. The trade-off chopping spikes introduced by the input chopper are filtered using a chopping spike filter (CSF) stage after the ACCIA in the design. The reported chip achieved a CMRR of over 120 dB, an input-referred noise density of 57 nV$\sqrt{}$Hz, and power consumption of 60 μW. However, the chopping scheme used in Fig. 5(a) achieves an equivalent DC-coupled amplifier, limits the electrode offset to only ±50 mV. The ACCIA was further improved later in 2008 by dividing the single servo loop with a coarse-fine servo loop for power-noise performance [9]. Coarse and the fine servo had discrete and continuous output levels, respectively. The superimposed output mimicked continuous output range of a single servo loop but minimized the output range for the fine servo, decreasing the power dissipation and increasing the power–noise performance. Furthermore, the study integrated eight readout front-end channels that could acquire 8-channel EEG signals simultaneously. Each readout front-end contained an ACCIA, a CSF and a programmable gain stage (PGS). An 11-bit successive-approximation-register (SAR) ADC was employed to digitize the outputs of the eight readout front-end channels. The chip fabricated under 0.5 μm technology had a 3.9×4.5 mm^2 die area and consumed power of 200 μW. The power consumption and noise performance of the amplifier were significantly improved with the aid of the coarse-fine servo loop. The active offset cancellation using the servo loop limited the minimum supply voltage of the amplifier which prohibited further power reduction through V_{DD} scaling. In 2010, Verma *et al.* proposed to chop signals at virtual ground node of the amplifier (shown in Fig. 5(b)) to perform passive offset cancellation [38]. The amplifier supply voltage can be reduced to improve the power efficiency. Additionally, a large DC offset can be rejected because it is decoupled by biasing resistors. However, CMRR is compromised because the differential noise introduced by the input capacitor mismatch cannot be mitigated, as the chopping is after the input capacitors. The AFE of [38] was 0.3 mm^2 with a power consumption of 3.5

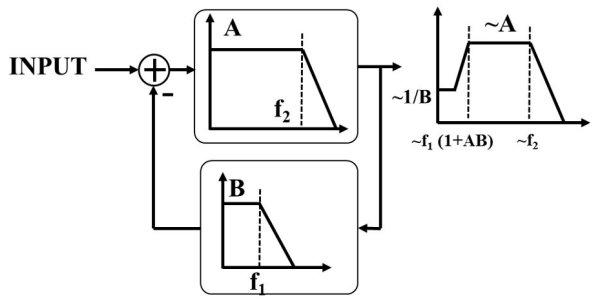


Fig. 4. Block DC offset using a low-pass filter in a feedback configuration.

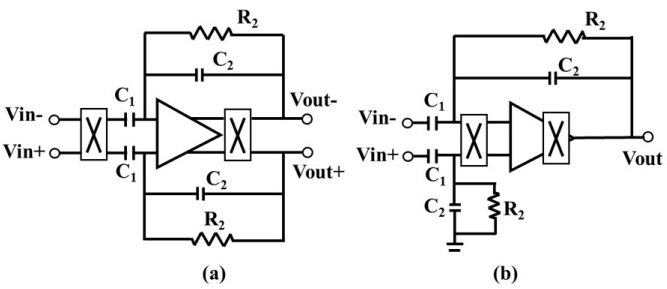


Fig. 5. Two types of commonly used chopping schemes: (a) chopping before the input capacitors, (b) chopping at the virtual ground.

μW. Although the achieved electrode offset voltage was larger than 1.0 V, the CMRR of this AFE was only 60 dB.

Despite the wide use of the wet electrodes in clinics, they have many obvious drawbacks [43], [44]. The signal integrity and system performance decrease if the gel dries out, they require time for setup, and they create discomfort and inconvenience for patients. These drawbacks make gel-based electrodes systems cumbersome and irritating for daily and long-term usage. Dry electrodes are more comfortable and suitable for daily seizure detection and prediction applications. However, the high electrode-tissue contact impedance increases significantly without the aid of gel, thereby increasing the severity of cable motion artifact and mains interference. A voltage buffer can be co-integrated with the electrodes to lower the impedance to eliminate the cable motion artifacts and mains interference. However, without signal amplification, a power-consuming amplifier or high-resolution ADC is required to maintain the total input referred noise at acceptable levels. An alternative method is to replaceA the voltage buffer with an amplifier, which leads to higher noise-power efficiency and lower power consumption [39]. The first low-power active electrode (AE) ASIC design that co-integrated the electrode and amplifier was presented in [39]. Similar to [38], the chopping modulation occurred at the virtual ground of an AC-coupled capacitive feedback amplifier to suppress 1/f noise. To satisfy the high input impedance requirement of dry electrodes, an additional input impedance boosting loop was integrated. Moreover, a ripple reduction loop and DC servo loop were utilized to compensate for the nonidealities of the amplifier. The CMRR degradation caused by relocating the input chopper to the virtual ground was mitigated by adopting an extra common-mode feedback (CMFB) circuit to feed the common-mode signal back to the input of each amplifier to compensate for the CMRR. The reported ASIC chip integrated with eight such amplifiers was fabricated using 0.18 μm CMOS technology, the chip size was 1.8×3.6 mm^2 and the power consume was 160 μW power. A high input impedance, rail-to-rail electrode offset rejection, and low-noise performance were achieved. In this design, the extra CMFB circuit for the feedback loop increased the area and power budget. A common-mode feed forward (CMFF) scheme was proposed in [11] to boost CMRR as opposed to using the CMFB scheme. The CMFF scheme provided an average node for all common-mode input signals and could effectively reduce the common-mode gain and increase the CMRR. A back-end chip was co-designed to process and digitize as many as eight channels of the front-end output. In [11], nine front-end chips and one back-end chip operated together to record and output digital signals of an 8-channel EEG. The system achieved CMRR, electrode offset rejection and noise level of 84 dB, $\pm 250mV$, and 1.75 μV_{rms}, respectively. Generally, when the number of channels increases, more wires are required to connect the front-end chip to the back-end chip. These wires make the system bulky, and introduce noise to the analog signals. Moreover, the gain mismatch of different channels limits the CMRR of the system. A digital active electrode (DAE) which integrates amplifier, filters, ADC and digital interface can deliver the digital output of each electrode's site, eliminating the noise introduced by the wires [40]. The CMRR of the system can be improved using a new CMFF which involves buffering the CM signal before applying it to the inverting inputs of all DAEs. The fabricated chip occupied an area of 3.5×4.5 mm^2 when fabricated using 0.18 μm technology, it could output digital signals directly to a microcontroller via the integrated I^2C interface. The power consumption of the chip was approximately around 105 μW, CMRR was 102 dB and the electrode offset tolerance was ± 350 mV. Up to 16 chips could be connected via an I^2C bus to form a multi-channel system. In 2018, Ha *et al.* proposed a mixed-mode DC servo loop solution that could tolerate DC offset up to ± 300 mV [41]. The servo loop was composed of a wide range digital servo loop and a analog fine resolution servo loop. This design methodology is similar to the course-fine servo loop proposed in [9]. With impedance boosting loop and ripple reduction loop, the capacitive coupled amplifier based front-end achieved 0.44 μV_{rms} input referred noise, 1 $G\Omega$ input impedance, and over 110 dB CMRR. Recently, Tang *et al.* presented EEG dust, a body channel communication (BCC) based EEG recording and transmitting AFE [42]. Each AFE chip could record, digitize, and wirelessly transmit one channel of EEG signals. This eliminated the use of wires. Signals were transmitted or received from other AFEs to obtain a common reference for all electrodes, thus eliminating the need for a reference node. A common-mode averaging unit implemented using two unity gain buffers and two average capacitors was adopted in the chip to dynamically suppress and adjust the common mode interference.

TABLE II
IMPLANTABLE SIGNAL ACQUISITION SYSTEMS

Work	JSSC2007 Denison *et al.* [45]	JSSC2012 Muller *et al.* [46]	JSSC2015 Muller *et al.* [12]	JSSC2016 Kassiri *et al.* [13]	JSSC2017 Kassiri *et al.* [14]	JSSC2018 Kim *et al.* [47]	TBioCAS2019 Lee *et al.* [48]	JSSC2019 Ghanbari *et al.* [49]
Noise	0.98 μVrms	4.9 μVrms	1.43 μVrms	4.2 μVrms	1.13 μVrms	$<$ 1 μVrms	-	5.3 μVrms
CMRR(dB)	$>$ 80	75	88	70	90	81	-	-
Elec. offset	-	$\pm$50 mV	$\pm$50 mV	$\pm$50 mV	rail-to-rail	$\pm$260 mV	300 mV	$\pm$10 mV
Technology	0.8	0.065	0.065	0.13	0.13	0.065	0.13	0.065
Area (mm^2)	1.7	0.013	2.4 $\times$ 2.4	4.8 $\times$ 3.3	2.6 $\times$ 2.3	1.0 $\times$ 1.0	-	0.5 $\times$ 0.5
Power	2 μW	5 μW	225 μW	2.17 mW -5.8 mW	1.07 mW - 5.44 mW	12.8 μW	375 μW	37.7 μW
No. of channels	-	2	64	64	64	16	32	1
Band-width	0.5-100 Hz	1-10 kHz	1-500 Hz	1-5 kHz	(0.1-500 Hz)	1-500 Hz	-	5 kHz
Power harvesting	-	-	300 MHz, 13 mW, 10 mm	1.5 MHz, 30 mW, 15 cm	1.5 MHz, 10 mW, 15 cm	-	13.56 MHz, 35mW, 18cm	Ultrasound, 2 MHz, 50 mm
Wireless comm.	-	-	300 MHz, Back-scattering, 1 Mb/s, 12.5 mm	3.1-10.6GHz 45 Mbps, 10 cm, $<$1GHz, 10 Mbps, 1 m 916.4 Mhz, 1.2 Mbps, 10 m	UWB 3.1-10.6GHz 10-46 Mbps, 10cm-2m	-	OOK 434 MHz 9 Mbps	AM back-scattering, 1.78 MHz, 35kbps

B. Invasive Neural Signal Acquisition

Compared with the EEG method, signals acquired close to the neuron cells have higher spatial resolution and precision. A few hundred neural microelectrodes can be placed sufficiently close to the neural cell to obtain more accurate recordings of neural activities [50], [51]. However, the circuits listed in Table I cannot be used for such purposes owing to their different signal characteristics such as their bandwidth and amplitude level. As Fig. 2(b) shows, ECoG, LFP, and AP signals have higher amplitudes than EEG signals, hence they are less sensitive to noise. However, when placed in the body, power and circuit area become critical concerns. Furthermore, wireless data transmission and power transfer become important to avoid the use of tissue-damaging wires. Some ICs targeted for implantable neural signals acquisition were developed. Table II lists and reviews some representative studies.

Similar to EEG acquisition, ACCIA was successfully adopted for neural signal acquisition as it can eliminate 1/f noise and electrode offset. For example, in 2007, Denison *et al.* proposed an AC coupled chopper-stabilized amplifier to record neural signals [45]. The amplifier achieved 2 μW power consumption and 100 nV/rtHz noise performance. The input AC-coupling capacitor occupies a large area, making the implementation of more channels in implantable chips difficult. To decrease the chip area, Muller *et al.* removed the AC coupling capacitor and adopted a dual mixed-signal servo loop architecture [46]. In this architecture, the forward pass was a DC-coupled amplifier followed by an ADC. Although removing the input decoupling capacitor reduced the chip area, additional circuits were required to address the DC offset that is directly connected to the amplifier. In this design, the fully digital servo loop which was composed of a digital low-pass filter (LPF) and a digital-analog converter (DAC), served this purpose. The digital LPF delivered the low-frequency signals which were then converted by the DAC to suppress the offset and reduce the dynamic range requirement of the input amplifier. The advantage of this design is that the digital feedback occupies small additional area and has low power consumption, and the LFPs and Spikes can be digitized simultaneously with the ADC and digital LPF. The chip area under 65 nm technology was 0.013 mm^2 with 5 μW power consumption, and DC offset up to $\pm 50mV$ can be eliminated. In 2015, Muller *et al.* proposed a chopper-stabilized, open-loop amplifier with mixed-signal feedback [12]. Chopper switch was placed between the electrode and input capacitors to avoid the high-pass pole setting introduced large area issue. Servo loop circuits similar to those in [46] were used to suppress the offset. The chip integrated 64 channels of such front-end circuits and a wireless subsystem in 2.4 $\times$ 2.4 mm^2 chip. The wireless subsystem used electromagnetic field backscattering to transmit data. To achieve a continuous wireless power transfer, the system was designed to modulate the impedance of the matching network between a matched condition and finite high impedance. The incident RF could always be received on-chip and rectified, enabling continuous-wave power transfer with continuous data modulation. The chip could harvest 12 mW of power at 300 MHz while transmitting data at 1 Mb/s. In 2016, Kassiri *et al.* proposed a DC-coupled front-end with digitally assisted feedback [13]. The design idea was similar to that in [46] utilizing DC-coupling to decrease chip area, and small area digital feedback for DC offset cancellation. The output of the amplifier was converted by an ADC and then compared with a reference number that represented the midrange voltage at the ADC input. The difference was integrated using a digital LPF and the input offset was canceled by biasing the amplifier through a current-steering DAC. DC offset up to ± 50 mW could be removed. The gain mismatch of different channels was calibrated using an additional calibration feedback loop. An inductive link operating at 1.5 MHz and provide up to 30 mW

power at a 15 cm range was integrated in the system. Ultra-wideband (UWB), frequency-shift keying (FSK) transmitters were integrated into the system to provide centimeter-to-meter-level communication. Analysis indicated that a discrete-time front-end with a high oversampling rate and small switch capacitor can satisfy the noise and impedance requirements for ECoG bandwidth [14]. Hence, in 2017, Kassiri et al. used a $\Delta\Sigma$ based architecture to leverage the low-cost low-complexity switched-capacitor techniques for input DC offset removal, and to make the channel circuits active-component-dominate [14]. This provides greater technological scalability. The fabricated chip under 0.13 μm was significantly smaller than that in [13] and achieved rail-to-rail max offset. The use of $\Delta\Sigma$ ADC for direct digitization without an amplification stage has recently become popular. For example, Kim *et al.* proposed a hybrid analog-digital second-order $\Delta\Sigma$ modulator oversampling ADC for neural recording [47]. The reported neural interface achieved 0.8 μW power consumption, 81 dB CMRR, and <1 μV_{rms} input referred noise with only 0.024-mm^2 area budget. $\Delta\Sigma$ ADC with oversampling and noise shaping features inherently allow to have low-noise performance, thus eliminating the need of an amplification stage. The scalablity of $\Delta\Sigma$ ADCs enable reduce the area significantly. Lee *et al.* proposed an ASIC chip with 32 neural recording channels. Conventional LNA and PGA architecture was adopted for signal acquisition. 35 mW power was continuously delivered to power up the chip through 13.56 Mhz frequency. Acquired neural can be upload to external devices through 434 MHz on-off keying RF transmitter. The proposed ASIC chip was system integrated with other devices such as MCU, bluetooth modules, demonstrated a complete system for wireless neural recording and transmission. Magnetic field attenuate rapidly when penetrate tissue, it is challenging to reduce the size of implant and extend power delivery range. Recently, taking advantage of low tissue propagation loss of acoustic waves, wireless powered neural implant through ultrasound has been proposed [49]. External transducer was able to power up the 0.8 mm^3 implant in a 50 mm range, neural signals were acquired by a LNA and transmitted to external device through amplitude modulation of the ultrasound echo.

IV. Signal Processing VLSI Implementations

Fig. 6 shows a typical 5-channel neural recording and the signal training and inference steps. The illustrated brainwave signal can be segmented into five different classes. Among them, preictal state presents the changeover from interictal (in-between) to ictal (during) seizure states. Many publications have demonstrated that the preictal state can be used as an indication of an imminent seizure [52]. Hence, seizure detection and prediction can be regarded as the recognition of the ictal and preictal states, respectively, from various epileptic brain states. The purpose of the signal processing VLSI is to perform certain classification algorithms that can recognize the ictal or preictal states from neural recordings. Time windows in different classes are selected to sample corresponding signals. Each window is then transferred to feature vectors using raw data or through various signal pre-processing methods, including fast Fourier transform (FFT),

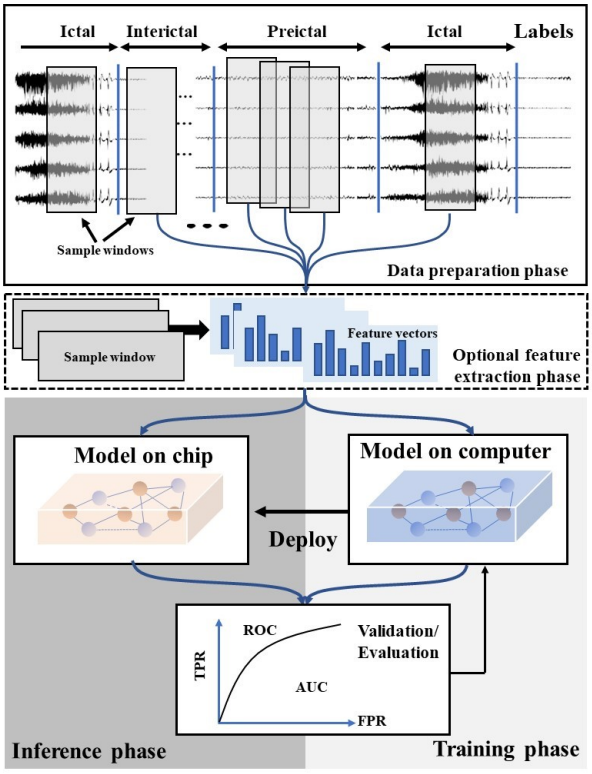


Fig. 6. Typical training and inference process.Training phase: brainwave signals database are labeled with different epileptic states and passed on to a deep learning network on computer to find a set of optimized parameters. The performance of the predictive model can be evaluated with RoC and AUC metrics. Inference phase: training obtained parameters will be deployed to a SoC, brainwave signals obtained from users will be passed to the SoC.

short time Fourier transform etc. The feature vectors are then fed to the classification algorithm to train a predictive model. Limited by the number of training samples, many research works utilized cross-validation to select model and reduce the chance of overfitting [34], [35], [53]–[55]. When training is completed, the model can be deployed on chip for inference. In the inference phase, brainwave data will be sampled using the same time window, and the same feature are used for the inference model. Seizure detection is considered correct if a correct ictal signal is recognized during inference. A seizure prediction is considered correct if a seizure follows the inferred alarm within the seizure prediction horizon (SPH). The SPH is the defined time between alarm and seizure onset and can range from minutes to an hour. Currently, there is no precise conclusion on how long the SPH should be. For example, SPHs from 30 to 5 minutes have been reported [25], [34]–[36]. Generally, an SPH between 5 and 30 minutes is considered suitable. If a seizure does not occur within the SPH after an alarm, then the alarm is regard as false prediction. The performance of a model can be characterized by metrics such as sensitivity, specificity, ROC curves, and AUC. Moreover, a low-power digital processing circuit is required to satisfy the $<$ mJ level energy budget for wearable and implantable devices. In the remainder of this section, we review the feature extraction process, different classification algorithms, and the

corresponding processing circuits for seizure detection and prediction.

A. Feature extraction

The goal of the feature extraction process is to derive a biomarker from physiological signals that are unique during the defined state and does not occur at other states. Frequency domain features are the most commonly used features in various detection and prediction studies. In the frequency domain, features such as spectral energy [15], [17], [59], [63], approximated entropy [16], discrete wavelet transform (DWT) [64], spectrograms [65], power spectral density [34], [67], cross-frequency coupling [55] and a combination of multiple features [66], [68], [70] are used as biomarkers to create multidimensional vectors for the following classification algorithms to differentiate various brain states.

Fig. 7(a) shows the simplified diagrams of frequency transformations. Time domain neural recordings are transformed into frequency domain using FFT or DWT algorithms with hardware accelerators. FFT or DWT hardware accelerators can be implemented with a butterfly architecture (shown in Fig. 7(a)) or low-, band- and high-pass filters. The details of time-frequency transformations are not discussed in this paper since it is a widely studied topic, and only a small part of processing systems.

B. SVM

Fig. 7(b) shows a simplified VLSI architecture and algorithmic illustration of the support vector machine (SVM) classifier. The decision boundary (dashed line) of the SVM classifier defined by the support vectors separates the two classes of feature vectors. The SVM is one of the most popular classifiers used in seizure detection and prediction applications [5], [34], [53], [54]. Eq. (1) provides a general form of the SVM classification function [71]:

$$f(x) = \sum_{i=1}^{N} \alpha_i K(v_i, x) + b \tag{1}$$

where $K(v_i, x)$ is the kernel function, and x is the feature vector that represents the acquired signal, v_i is one of the N support vectors and α_i and b are the parameters. The sign of the computed $f(x)$ determines the class of the feature vector. The most commonly used kernel functions are radial basis function (RBF), polynomial, and linear kernels. The RBF kernel [71] expressed as follows:

$$K(v_i, x) = exp(-\frac{||v_i - x||^2}{2\sigma}) \tag{2}$$

Fig. 7(b) provides a simplified VLSI architecture of SVM classifier. It contains parallel units, weights and vector memories, and a controller. The main computation components are the processing units for the kernel processing and summation. The memories store the support vectors and weights. The control module coordinates and reuses the processing units according to Eq. (1). The actual hardware cost of the system shown in Fig. 7(b) can depend on the number and actual design of the parallel unit, and the memory size required for the weights and support vectors. Some general SVM implementations are available [72], [73], however, these designs are either power or area consuming. Moreover, without dedicated feature extractor, they are not suitable for neural signal classification. The design of most SVM VLSIs for biomedical applications are optimized to satisfy certain specifications, such as area and power. For instance, radial basis function (RBF) kernel shown in Eq. (2) requires subtraction, squaring, division, and complex exponentiation operations which results in high hardware resource costs. The direct implementation of Eqs. (1) and (2) will consume considerable hardware resources.

Approximation methods such as the coordinate rotation digital computer (CORDIC) CORDIC function, can be used to implement the RBF kernel. The CORDIC approximation can complete complex functions by recurrently look up and summing of some pre-stored data. The SVM implementations shown in Table III have utilized various methods. In [58], features were first identified from the compressed EEG signals through eight compressed-domain band-pass filters. Subsequently, the extracted features were classified using an SVM classifier in which an RBF kernel was performed by an embedded CORDIC engine. The chip could compress EEG data by a factor of 2-24, and the total processor power was in the range 0.6-107 μW for all SVM kernels including linear, polynomial and linear kernels. Seizure detection experiments conducted on the CHB-MIT dataset indicated the performance for sensitivity, latency and specificity was 96-91%, 4.7-5.3 s, and 0.17-0.30 false-alarms/h. Lee *et al.* proposed a custom processor that integrated a low-power microcontroller unit (MCU) with a configurable SVM accelerator [59]. The dedicated data-path unit (DPU) and CORDIC engine of the SVM accelerator could support linear, polynomial and RBF kernels based on different applications. The fabricated chip could conduct seizure detection on the CHB-MIT dataset with 274 μJ per detection. Yoo *et al.* [15] used a linear SVM (LSVM) that involved only vector multiplication and addition to reduce the hardware cost. Despite the simplicity of the LSVM, it degraded the prediction performance because some features were not linearly separable. The reported average detection rate on the CHB-MIT database was 82.7% while it had 4.5% false positive rate. As a supplement to a single LSVM, Altaf *et al.* [17] proposed a dual detector (D^2A) architecture with one LSVM detector trained for sensitivity optimization and another for specificity optimization. The D^2A architecture achieved an average sensitivity and specificity of 95.7%, 98% on the CHB-MIT dataset, respectively. Altaf *et al.* modified Eq. (2) by first converting the computation into log (ln) domain and then converting it back [62]. This optimization reduced 61.2% of gate count for the exponent computation unit and replaced a complex float point multiplexer with a float point adder. In [64], a table-driven algorithm employing a predetermined polynomial approximation was used to calculate the exponent operation of an RBF function. It achieved both high accuracy and high speed with a limited hardware cost. O'Leary *et al.* adopted an exponentially decay memory (EDM) SVM to enable low-power time series classification [68]. The EDM approach defines a continuous sampling recursive window that

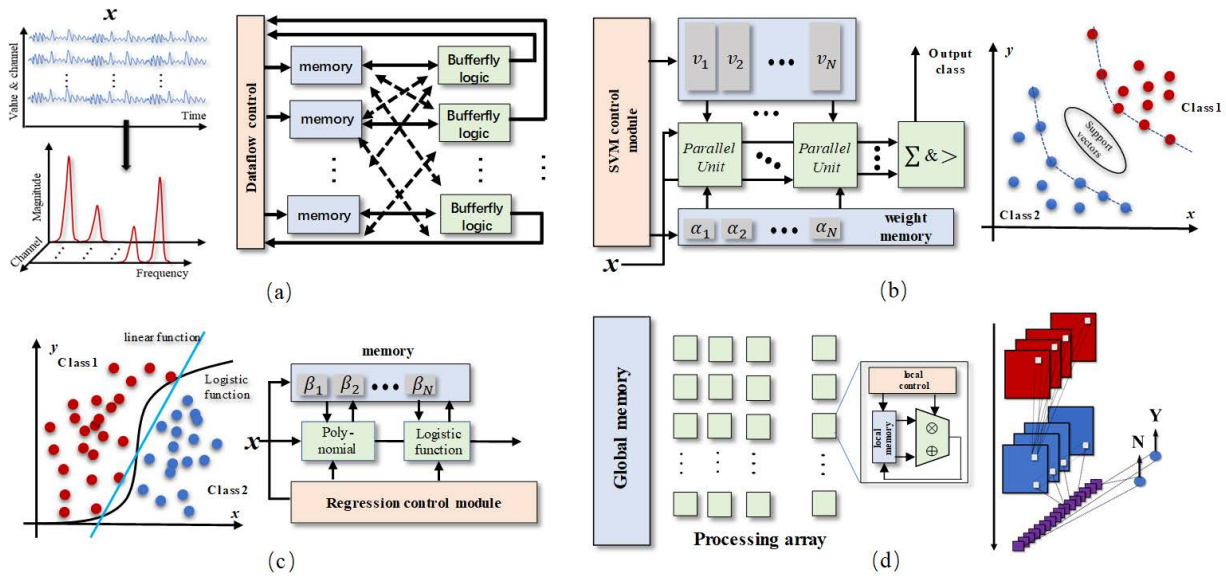


Fig. 7. Simplified schematic of common models used in seizure prediction VLSI: (a) feature extraction model, (b) support vector machine, (c) regression, (d) convolutional neural network.

considers feature's history. As the EDM updates every sample, the SVM classification is performed constantly rather than only when a window has been processed. The system-on-a-chip fabricated under 0.13 μm technology achieved sensitivity of 97.7% and a false detection rate of 0.185/h at cost of 169 μJ per classification. The EDMSVM implementation occupied a silicon area of 1.3 $\times$ 2.5 mm^2. In 2019, Huang *et al.* proposed an SVM processor capable of detecting seizure and dynamically adjusting the model [69]. The actual architecture of the SVM processor is shown in Fig. 8(a). It consisted of a feature extractor that performed real-valued FFT, an SVM processing engine that comprised an 8 $\times$ 4 CORDIC-based processing element (PE) array, and a memory bank. The 8 $\times$ 4 PE array provided all the required computation for the inference and model adaptation. The CORDIC-based PE first converted numerical values to the convergence range and then performed the supported functions using the custom floating-point CORDIC. The output values were finally scaled back to the correct numerical range. The chip was implemented using 40 nm technology and the chip area was 2.1 $\times$ 2.1 mm^2. The power consumption of the chip was 1.9 mW and 2.9 mW when performing seizure detection and adaptation, respectively.

C. Logistic Regression

Logistic regression (LR) provides the probability of a data point being the positive class based on Eq. (3):

$$P(y = 1|x) = \frac{1}{1 + exp(\beta \cdot x + b)} \tag{3}$$

where x is the input vector, β is the calculated coefficients vector and b is offset. Additionally, regarding binary classification, if $\beta \cdot x + b$ is larger than zero, then the probability obtained using Eq. (3) will be greater than 0.5 and the input sample can be considered positive. It can eliminate unnecessary computation and use only multiplication and addition to minimize computation.

Fig. 7(c) shows a simplified LR classification example and hardware diagram. One memory is used to store the coefficients β, the operation of computational core is controlled by the controller to perform the required polynomial or logarithmic operations. Coefficients β usually contain few to ten scalar values, thereby requiring a small memory space. The coefficients associated computation are also light-weighted, involving mostly multiplication and addition, making it suitable for low-power and small-area, hardware implementation. Some results are also shown in Table III. Samie *et al.* implemented logistic regression on an MCU [70]. The low-power MCU consumed 13.13 mJ power to perform LR for one segment of approximately 5 seconds. This implementation achieved a 0.79 AUC score on the Kaggle dataset. However, a smartphone was still needed for further processing. Chen *et al.* proposed the use of a least linear square (LLS) classifier for seizure detection [16]. The Approximated entropy (ApEn) and frequency spectrum were used as the epileptic features in the LLS model. The classifier achieved 92% detection accuracy inn the reported experiment. However, comparing this accuracy with that of other studies was difficult because the data was not available. Further improvement was done by replacing the LLS classier with a ridge regression classifier [67]. Owing to the simplicity of regression classifiers, they only occupy small silicon area. For example, a logistic regression classifier implemented in 65 nm technology was reported by Page *et al.* [60], the estimated layout was only 0.008 mm^2 and the power consumption could be as low as 37 nW when operate under low frequency.

TABLE III
SEIZURE DETECTION AND PREDICTION SYSTEMS

Work	TBioCAS2011 Abdelhalim *et al.* [56]	JSSC2011 Sridhara *et al.* [57]	TCAS-I2014 Shoaib *et al.* [58]	JSSC2013 Lee et al. [59]	JSSC2013 Yoo *et al.* [15]	JSSC2014 Chen *et al.* [16]	JSSC2015 Altaf *et al.* [17]	TCASII2015 Page *et al.* [60]
Func.	Seizure prediction	General for biosignal	Seizure detection	General for biosignal	Seizure detection	Seizure detection & stimulate	Seizure detection & stimulate	Seizure detection
Implementation	ASIC	ASIC	ASIC	ASIC	ASIC	ASIC	ASIC	FPGA
Algorithm	Phase -synchronization	[61]	SVM	SVM	BPF+ LSVM	FFT + LLS	FTDM + LSVM	LR
Area (mm^2)	0.178 @0.13 μm	2.98 $\times$ 2.98 @0.13 μm	1.5 $\times$ 1.0 @0.13 μm	2.7 $\times$1.9 @0.13 μm	5 $\times$ 5 @0.18 μm	13.47 mm^2 @0.18 μm	5 $\times$ 5 @0.18 μm	0.008 @65 nm (simulation)
Power	3.6 μW per input pair	1.6-2.9 μW	70.8 nJ feature ext. 53.2 μJ - 209 nJ /feature	273 μJ /classification	2.03μJ /classification	77.91 μJ /feature ext. + detection	2.73 μJ /classification	19nJ /classification 37nW (simulation)
Performance	TPR = 66%, FPR = 1/hr. (Freiburg)	-	Sensitivity 96-91% specificity 0.17-0.30/hr, latency = 4.7-5.3s. (CHB-MIT)	sensitivity = 100%, false alarm = 1.2/day, latency = 4.8s (simulated)	TPR = 82.7%, FPR = 4.5%, latency $<$ 2s. (CHB-MIT)	accuracy = 92% latency $<$ 0.8s	sensitivity = 95.7% specificity = 98% latency = 1s. (CHB-MIT)	onset sensitivity = 95.24%, accuracy $>$ 80%. (CHB-MIT)
Other features	Three CORDIC cores, 10-bit precision.	Fast FFT + MCU + subthreshold 6T SRAM	CPF + CORDIC + operate in compressed domain	Support RBF, linear, poly-nomial kernels. CORDIC core	DQ-LUT + AFE	wireless power-data transmission + AFE + stimulator	AFE + 2 LSVMs + stimulator	-

Work	TBioCAS-2016 Altaf *et al.* [62]	TBioCAS2016 Valle *et al.* [63]	TBioCAS2018 Feng *et al.* [64]	EBioMedicine-2018 Kiral *et al.* [65]	JETCAS 2018 Shoaran *et al.* [66]	JSSC2018 Cheng *et al.* [67]	JSSC2018 O' Leary *et al.* [68]	JSSC2019 Huang *et al.* [69]
Func.	Neural decoding	Seizure detection	Detection	Prediction	Seizure detection	Seizure detection & simulation	General EEG classification	Seizure detection & adaptation
Implementation	ASIC	ASIC	FPGA	IBM TrueNorth	ASIC	ASIC	ASIC	ASIC
Algorithm	NLSVM	Custom	DWT + SVM	CNN	XGBoost	Ridge regression	EDM SVM	SVM
Area mm^2	5 $\times$ 5 @0.18 μm	5 $\times$ 5 @0.18 μm	-	4.3 cm^2 @28nm	1.85 $\times$ 0.54 @65 nm	5 $\times$5 @0.18 μm	1.3 $\times$ 2.55 @ 0.13 μm	2.1 $\times$ 2.1 @40 nm
Power	1.83 μJ /classification	0.45 μW /channel	45 mW	4.3-7.5 mW	41.2nJ /class	62.5 μJ / feature ext. + classification	169 μJ /classification	1.9 mW (detection) 2.9 mW (adaptation)
Performance	Detection rate = 95.1%, FPR = 0.94%, latency = 2 s (CHB-MIT)	TPR = 98.5%, false alarm = 4.4/hr (CHB-MIT)	DR = 96.8, FR = 0.91%	mean sensitivity 69%, mean time in warning 27%	Sensitivity = 83.7%, specificity = 88.10%, latency = 1.79s (iEEG)	Sensitivity = 96%, speci-ficity = 100% late-ncy = 0.76s	Sensitivity = 97.7% false alarm = 0.185/h (iEEG)	sensitivity = 96.6% false alarm rate = 0.48%
Other features	AFE + TDM-BPF + LL-GBF	AFE + BPF + probability LUT	On-chip learning	-	12 features + programmable FIR filters + memory control	AFE + ApEn + FFT + wireless power-data transmission + stimulator	On-chip autoencoder, neuro-stimulation.	Custom ALU, model adaptation

D. Other Conventional Methods

In addition to SVM and regression classifier, other implementations that use less popular algorithms such as phase synchronization [56], extreme gradient boosting (XGBoost) [5], [66], [70] and customized algorithms [57], [63] are also listed in Table III. Abdelhalim *et al.* computed phase locking value (PLV) that quantified the synchronic level of two neural signals to predict the onset of seizures [56]. The complicated arctan and sine cosine operations that were required to compute PLV were performed using a CORDIC algorithm. The processor fabricated using 130 nm technology only occupied an area of 0.178 mm^2 and dissipated 3.6 μW of power. However, the prediction performance was low. XGBoost combines multiple base classifiers to form one strong classifier. Each base classifier can be built with simple comparators. This can significantly reduce system complexity. Shoaran *et al.* implemented this algorithm on a chip with 65 nm technology that achieved 41.2 nJ/class in a total area of 540 $\times$ 1850 μm^2 [66]. The hardware architecture for the proposed ensemble of

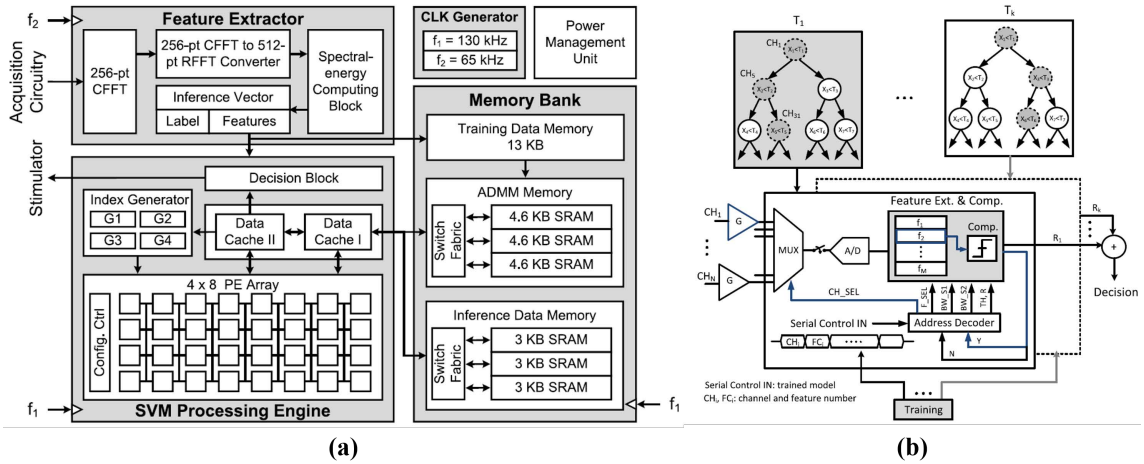


Fig. 8. Recent processor architectures that are used for seizure prediction: (a) processor architecture of [69], (b) hardware architecture for an ensemble of gradient boosted decision trees [66].

gradient boosted decision trees is shown in Fig. 8(b). Each tree could perform feature extraction and comparison, and the architecture was an ensemble of up to eight trees. Since only one path existed between the root and the leaf, each tree compared only the feature of the selected channel with a threshold. All the answers of each tree were summed to form a final answer at the end. Due to the simple comparsion operation decision tree, the fabricated chip under 65 nm technology achieved energy efficiency of 41.2nJ/class on a 1.8 $\times$ 0.5 mm^2 die. Seizure detection was evaluated with the iEEG database, 83.7% and 88.1% average sensitivity and specificity were achieved.

E. Convolutional Neural Network

A convolutional neural network (CNN) is composed of convolution, pooling and fully connected layers. The convolutional layers are the most computationally expensive part of a CNN. The convolution layer shown in Fig. 7(d) could be expressed as

$$c_{xyz} = \phi((\sum_{j=0}^{l-1}\sum_{i=0}^{l-1}\sum_{k=0}^{d-1} w_{ijkz} \times f_{(sx+i)(sy+j)k}) + b_z \qquad (4)$$

where f, c, w, s and b are the input feature, output feature, kernel, stride size, and offset value, respectively, and ϕ is an activation function. The subscript of a parameter indicates its index. International collaboration and standard databases [75]–[77] have provided a large amount of research data that can be used for deep learning-based seizure prediction research. For example, Korshunova *et al.* proposed a five layer CNN architecture with two convolutional layers, one pooling layer and two dense layers. Although the model was simple, its prediction performance was still comparable with state-of-the-arts [36]. Khan *et al.* described a CNN model which consists of six convolutional layers, two dense layers and one output layer, sensitivity of 87.8% and false prediction rate of 0.142/h were achieved [35]. Currently, owing to the complexity of CNN algorithms, most CNN algorithms are executed using CPU or GPUs, which cannot be used for wearable applications because of the significant power consumption. Hence, implementing the algorithm on a low-power processing platform is very important. Hugle *et al.* implemented a network architecture called SeizureNet on a low-power MCU to predict seizure onset [78]. SeizureNet was evaluated on the EPILEPSIAE database [75], and it achieved sensitivity of 0.96, false detection rate of 10.1/h, and detection delay of 3.7 seconds. The reported power consumption was 850 μW. Moving power consuming processing to the cloud is an alternative to reduce power consumption. Hosseini *et al.* proposed a cloud-based system for seizure prediction [79]. This system first reduced the EEG data dimension to decrease the communication bandwidth, and, the EEG data was then transmitted to a cloud platform where data was stored and a stacked autoencoder was used for processing. In addition to the low-power MCU and cloud-based system, a more direct approach is to use a special designed processor for low-power CNN processing. A fundamental structure employed in CNN processor is illustrated in Fig. 7(d). It is generally composed of a 2D PE array. Each PE includes local memory, an ALU and a control unit that controls the inter-PE communication independently. This architecture is adopted by many processors including TrueNorth [74], a dedicated chip developed by IBM, and has been used for seizure prediction [65]. The chip and its conceptual architecture is shown in Fig. 9. It consists of a two-dimentional array of neurosynaptic cores that are connected with a network on chip. Each neurosynaptic core has axons acting as inputs, neurons as outputs, and synapses as directed connections between the axons and neurons. A four-layer CNN network is mapped to this chip, a mean sensitivity of 69% and mean time in warning of 27% have been achieved. This chip cumsumes approximately 4.3-7.5 mW power and occupies 4.3 cm^2 area under 28 nm technology. A CNN architecture that consists of three alternating convolutional layers and three max-pooling layers were used in [65] for patient-specific seizure prediction. The implementation achieves mean sensitivity of 69% and mean time in warning of 27%. Recently, Kueh proposed a

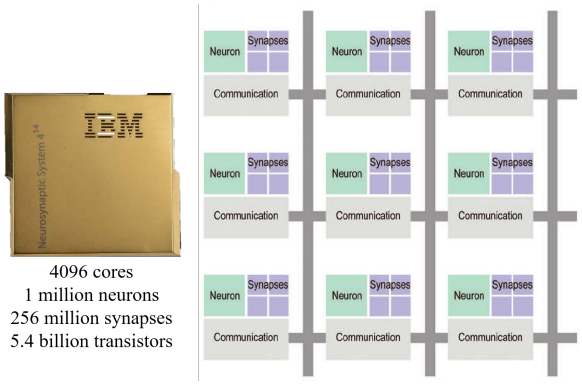


Fig. 9. IBM neuromorphic chip [74] and its conceptual architecture. The chip is used for seizure prediction in [65]. The chip contains is a two-dimensional array of neurosynaptic cores, each core integrates memory, computation, and communication modules. Cores operate in parallel and communicate via an event-driven network.

FPGA implementation of a multi-layer perceptron that was realized with bit-serial operations to achieve low-power seizure prediction [80]. The bit-serial data-path can significantly reduce the cost of hardware resources, because when process EEG signal in a bit by bit fashion, complex ALUs are no longer required.

V. Discussion and Future Prospects

With numerous circuits and systems that can record EEG, ECoG or other neural activities, increasingly reliable and accurate detection and prediction algorithms and dedicated processing circuits, we can observe a clear migration trend from seizure detection to seizure prediction. In this section, we discuss the key challenges and major progress of existing ICs and systems. Furthermore, ongoing bottlenecks for future fully embedded and smart seizure prediction systems are analyzed.

AFE design requires low input referred noise, high electrode offset tolerance, high input impedance and high CMRR performance to obtain high-quality EEG signals. In addition to these, low power budget is needed to allow long-term operation. Chopping is the most commonly used technique to reduce 1/f noise which is the dominant type of noises in different recorded biomedical signals. Also, as reviewed in Section III, two common chopping practices can be used. One is to chop before the input coupling capacitor of capacitively coupled amplifier like Fig. 5(a). However, this method can lead to a switched-capacitor conductance which lowers the input impedance to a value that inverse proportional to the input capacitor [81]. On the other hand, large input capacitor is needed in this scheme to mitigate the introduced thermal noise. Additionally, this chopper scheme achieves an equivalent DC-coupled amplifier, hence limits the electrode offset tolerance to few tens of mV. As shown in Fig. 5(b), another scheme is chopping at the amplifier's virtual ground, this scheme ensures rail-to-rail electrode offset tolerance but suffer from a reduced input impedance. Chopping introduces ripples, requiring extra circuits to eliminate them. Electrode offset can be directly rejected using AC-coupling, but come with the cost of a large capacitor and blocking of low frequency signals. DC-coupling can preserve some low frequency signal but electrode offset and EEG signals would affect the gain of amplifier. A DC servo loop, which can provide corrective current or voltage as feedback to compensate the offset is commonly used method to nullifying electrode offset. Limited by the maximum feedback current, a current feedback DC servo loop can only compensate for tens of millivolts. Transconductor introduced noise also contributes to the overall amplifier noise. Voltage feedback DC servo loop can provide higher tolerance to electrode offset at lower power budget. Impedance boosting feedback circuits can be used to increase the input impedance, thereby reducing the input signal division and mitigating the electrode-tissue impedance mismatch thus improve CMRR of system.

Based on the IEC [82] and IFCN standards [83], the critical metrics for acquiring high-quality EEG signals are satisfied with the above discussed design techniques. State-of-the-art technologies can achieve $< 1\ \mu V_{rms}$ input referred noise, ± 300 mV electrode offset tolerance, $> G\Omega$ level input impedance and > 100 dB CMRR with only tens of microwatts power consumption. However, the spatial resolution of EEG signals is still limited and the current EEG systems require further improvement before being truly wearable. Further decreasing the size of the electrodes so that more electrodes can be placed on top of the scalp is critical to increasing the resolution of EEG signals. This requires improving the functionality of each electrode and lowering the power consumption per electrode. Advancement and collaboration may be required in material science, packaging, and process technology to develop CMOS-compatible, smaller and comfortable electrodes. The number of connecting wires for the electrodes is another major challenge that causes EEG systems to be bulky. These wires include power supply and data transport lines. The wireless distributed EEG system proposed in [42] exhibits a promising direction for solving the problem of connecting wires. Individual AEs that can acquire, amplify, digitize, and transmit one channel of EEG signal can make the EEG system more comfortable. Additionally, a combination of multimodal bio-information has shown the potential to increase the spatial resolution of EEG signals. For example, functional near-infrared spectroscopy (fNIRS) signals can be considered as a complementary additions to EEG measurement [84], noninvasive neural imaging with improved EEG resolutions can boost performance classification significantly [85].

ICs for invasive neural signal acquisition have higher constrains on area and power and relaxed constraint on noise and impedance performance owing to their invasive features. Early implant ICs adopted AC coupled closed-loop amplifier topology similar to the ones used in EEG sensors. However, as the number of channels increases, the area occupied by the decoupling capacitor can become significant. Removing the capacitors mitigates this problem, however, additional circuits are required to remove offset. A functional DC-coupled amplifier utilizing DC Servo loop is commonly used to cancel the offset by biasing the amplifier input. More importantly, having more passive components means that the design can be scaled

when using advanced fabrication process. Recently, recording architecture that can directly convert neural signals to the digital domain without neural amplifier have been reported. $\Delta\Sigma$ modulators with scalable feature, can significantly reduce area, but come with the cost of higher noise and challenging ADC design. Wireless power and data transfer are two important features in implant neural signal acquisition. Currently, inductive links operating at low frequency and millimeter ranges are commonly used to provide power supply and low data rate communication to the implants. UWB, owing to its high data rate that can be over 100 Mbps, low-power, and small antenna features, is implemented for centimeter range multichannel neural signal transmission purposes. Recently, ultrasound provides an alternative for wireless powering and data transmission. State-of-the-art neural recording front-ends consume less than 10 μW power per channel. The power budget for wireless data transmissions can range from a dozen μW to few mW based on the transmission distance. Inductive links can usually transfer dozens of mWs which are sufficient for a limited number of channels. However, with the increasing number of channels, more power and higher data rate communication are required. This is can perhaps be solved jointly by compressing or preprocessing the neural signals locally and increasing the communication bandwidth. Improving safety and eliminating the need for surgery are also important for invasive neural signal acquisition systems. Recently, Neuralink reported a robotic inserter machine and probe that can be safely insert tiny electrode into the brain without damaging the blood vessels. The surgical robot was designed to avoid blood vessels when inserting μm level electrodes. With minimal invasive, the entire operation can be expected to complete without any hospital stay [86]. New technology can make invasive methods more acceptable. One goal for signal acquisitions is to find a good balance between high-resolution, high-precision bio-signals and user safety and comfortability.

In Section IV, algorithm and circuits implementations capable of seizure detection are reviewed. Feature extraction and classification based on the extracted features are two essential steps for all studies in Table III. The feature extraction methods are relatively simple and mainly involving FFT and DWT for transforming time domain raw signals to the frequency domain to construct features such as spectral energy, spectrogram and entropy. However, the features are still manually selected. Recently, many biomedical applications that only utilize raw bio-signals demonstrated clinician-level performance. For example, Hannun *et al.* proposed a CNN that accepts raw ECG signals for the detection and classification of arrhythmia [21]. Supratak *et al.* proposed DeepSleepNet which can automatically score sleep stages based on raw EEG signals [87]. Recorded ECoG signals are used to synthesize speech using bidirectional long short-term memory (bLSTM) [23]. We postulate that future smart and fully embedded seizure prediction systems will employ deep-learning methods that utilize only raw signals.

For detection purposes, SVM, LR, and other customized algorithms are implemented. Logistic regression usually requires the least number of parameters, and the computation is simple,

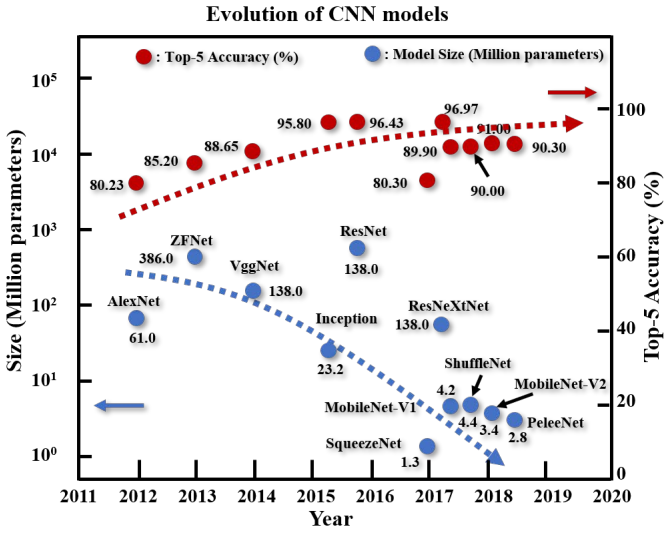


Fig. 10. Top-5 accuracy and model size of different CNN models.With suitable optimization techniques, CNN models can shrink rapidly without affecting the classification performance.

involving only multiplication and addition. Hence, LR based hardware is easy to implement with low silicon area and power budget. However, LR-based classification has limitations in detection accuracy owing to its linear characteristic. In contrast, SVM-based classification gains non-linearity from the kernel functions, achieving a good detection accuracy. Moreover, it only needs to store a few numbers of support vectors, hence small memory could be enough. Kernel functions involve complex operations, but can be bypassed using CORDIC to reduce hardware cost. State-of-the-art seizure detection hardware can achieve energy efficiency from a few to $<$200 μJ per class based on different classification methods.

To date, reliable and accurate detection has been achieved with $<$95% sensitivity, $< 0.2/h$ false detection rate. However, existing seizure detection systems cannot be used in seizure prediction application directly, because of two reasons. First, the majority of seizure detection hardware relies on the SVM classification algorithm, which is not suitable for prediction purposes. Preictal signals are more difficult to distinguish than onset ictal signals. SVM classification with frequency features can seldom provide favorable result. The comparison in [26] indicated that CNNs generally provide better results than SVM does. Second, these chip architectures are not optimized for CNN computation and lack support for specific operations such as multiply–accumulate operation. For future implementation of reliable seizure prediction, some challenges remain. Current deep-learning based models still require a large number of parameters, necessitating a large on-chip memory. Implementing MB level memory in area-critical implant chips is impractical. Therefore, further efforts are required to develop compact deep-learning seizure prediction models with acceptably satisfactory prediction performance. Fortunately, several technologies have been developed in recent years to address these challenges. Fig. 10 shows various efforts to reduce the number of network parameters while

maintaining acceptable classification accuracy. The blue and red dots with the same horizontal coordinates represent a certain model's number of parameters and top-5 accuracy over ImageNet database, respectively. ImageNet is a standard image database for measuring classification performance [88]. The top-5 accuracy measures the precision of any of a model's five highest probability answers that match the expected label. With technologies such as depth-wise separable convolution [89], [90], pointwise group convolution, channel shuffle [91] and other methods [92], [93], a rapid reduction in the number of parameters is clearly observable in Fig. 10. Some EEG classifications already adopted abovementioned technology to reduce the size of the classification model to 100 kB level [94]. Hence, we consider that the model for seizure prediction can be compressed to a smaller scale using similar techniques. Another major challenge is implementing deep learning related computation in a low-power topology. This should be performed at both the circuit and system levels. At the system level, the computation should be performed locally within the chip or remotely with less power critical devices, through wireless communication, based on different scenarios. Seizure onset only accounts for a small portion of time, this character can be used to address the low-power design challenge. For example, two-stage classification can be used to reduce system power [17], [95]. At the circuit level, dedicated integer multipliers and accumulators should be designed to accelerate multiply-accumulate operations, variable width bit operation [96] and binary operation [97] designs are popular methods to reduce the power consumption.

We have focused on chip level implementation for practical seizure prediction, several interdisciplinary collaborations have been conducted to achieve progress in this domain. International collaboration has given birth to many databases and dedicated seizure prediction workshops [77], [98]. For clinical translation, closer cooperation with clinicians is required to improve the systems from several aspects, such as easy of use, automated operation, intervention and patient personalization.

VI. Conclusion

Progress in circuit design and algorithm development have contributed considerably toward seizure prediction. Dedicated analog ICs enable the acquisition of relevant signals with low-noise interference, low-power consumption and high CMRR from increasing number of electrodes. The technology transfer of machine learning algorithms from the computer vision society to bio-medical applications are continue to increase the reliability and precision of the seizure prediction performance. However, this impact has not yet reflected in the seizure prediction system designs. Most digital processing circuits are still built upon more traditional seizure detection frameworks which include feature extraction and a classifier such as SVM or LR.

The buildings blocks of prediction systems analyzed in this paper can become the focus of future research to boost individual performance to improve overall system performance. Based on our comprehensive review, it is apparent that, together with the bloom in machine learning and silicon technology, the prospect of a seizure prediction system that is smart, full embedded, accurate and reliable is promising. A wide collaboration of neural scientists, material scientists, analog and digital designers, data scientists can propagate the process. We believe that the lives of epileptic patients will be significantly improved in the near future when the remaining challenges are addressed.

Acknowledgment

The authors would like to acknowledge start-up funds from Westlake University to the Cutting-Edge Net of Biomedical Research and INnovation (CenBRAIN) for supporting this project.

References

[1] P. Kwan, S. C. Schachter, and M. J. Brodie, "Drug-resistant epilepsy," *New England Journal of Medicine*, vol. 365, no. 10, pp. 919–926, 2011.

[2] A. Harroud, A. Bouthillier, A. G. Weil, and D. K. Nguyen, "Temporal lobe epilepsy surgery failures: a review," *Epilepsy research and treatment*, vol. 2012, 2012.

[3] J. F. Téllez-Zenteno, L. H. Ronquillo, F. Moien-Afshari, and S. Wiebe, "Surgical outcomes in lesional and non-lesional epilepsy: a systematic review and meta-analysis," *Epilepsy research*, vol. 89, no. 2-3, pp. 310–318, 2010.

[4] L. Ridsdale, J. Charlton, M. Ashworth, M. P. Richardson, and M. C. Gulliford, "Epilepsy mortality and risk factors for death in epilepsy: a population-based study," *Br J Gen Pract*, vol. 61, no. 586, pp. e271–e278, 2011.

[5] P. Detti, G. Z. M. de Lara, R. Bruni, M. Pranzo, F. Sarnari, and G. Vatti, "A patient-specific approach for short-term epileptic seizures prediction through the analysis of eeg synchronization," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 6, pp. 1494–1504, 2018.

[6] R. Nickel, C. E. Silvado, F. M. B. Germiniani, L. d. Paola, N. L. d. Silveira, J. R. B. d. Souza, C. Robert, A. P. Lima, and L. M. Pinto, "Quality of life issues and occupational performance of persons with epilepsy," *Arquivos de neuro-psiquiatria*, vol. 70, no. 2, pp. 140–144, 2012.

[7] R. S. Fisher, B. G. Vickrey, P. Gibson, B. Hermann, P. Penovich, A. Scherer, and S. Walker, "The impact of epilepsy from the patient's perspective i. descriptions and subjective perceptions," *Epilepsy research*, vol. 41, no. 1, pp. 39–51, 2000.

[8] K. A. Ng and P. K. Chan, "A cmos analog front-end ic for portable eeg/ecg monitoring applications," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 11, pp. 2335–2347, 2005.

[9] R. F. Yazicioglu, P. Merken, R. Puers, and C. Van Hoof, "A 200 μw eight-channel eeg acquisition asic for ambulatory eeg systems," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 12, pp. 3025–3038, 2008.

[10] F. Shahrokhi, K. Abdelhalim, D. Serletis, P. L. Carlen, and R. Genov, "The 128-channel fully differential digital integrated neural recording and stimulation interface," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 4, no. 3, pp. 149–161, 2010.

[11] J. Xu, S. Mitra, A. Matsumoto, S. Patki, C. Van Hoof, K. A. Makinwa, and R. F. Yazicioglu, "A wearable 8-channel active-electrode eeg/eti acquisition system for body area networks," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 9, pp. 2005–2016, 2014.

[12] R. Muller, H. Le, W. Li, P. Ledochowitsch, S. Gambini, T. Bjorninen, A. Koralek, J. M. Carmena, M. M. Maharbiz, E. Alon, and J. M. Rabaey, "A minimally invasive 64-channel wireless μecog implant," *IEEE Journal of Solid-State Circuits*, vol. 50, pp. 344–359, Jan 2015.

[13] H. Kassiri, A. Bagheri, N. Soltani, K. Abdelhalim, H. M. Jafari, M. T. Salam, J. L. Perez Velazquez, and R. Genov, "Battery-less tri-band-radio neuro-monitor and responsive neurostimulator for diagnostics and treatment of neurological disorders," *IEEE Journal of Solid-State Circuits*, vol. 51, pp. 1274–1289, May 2016.

[14] H. Kassiri, M. T. Salam, M. R. Pazhouhandeh, N. Soltani, J. L. Perez Velazquez, P. Carlen, and R. Genov, "Rail-to-rail-input dual-radio 64-channel closed-loop neurostimulator," *IEEE Journal of Solid-State Circuits*, vol. 52, pp. 2793–2810, Nov 2017.

[15] J. Yoo, L. Yan, D. El-Damak, M. A. B. Altaf, A. H. Shoeb, and A. P. Chandrakasan, "An 8-channel scalable eeg acquisition soc with patient-specific seizure classification and recording processor," *IEEE Journal of Solid-State Circuits*, vol. 48, pp. 214–228, Jan 2013.

[16] W.-M. Chen, H. Chiueh, T.-J. Chen, C.-L. Ho, C. Jeng, M.-D. Ker, C.-Y. Lin, Y.-C. Huang, C.-W. Chou, T.-Y. Fan, *et al.*, "A fully integrated 8-channel closed-loop neural-prosthetic cmos soc for real-time epileptic seizure control," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 1, pp. 232–247, 2013.

[17] M. A. B. Altaf, C. Zhang, and J. Yoo, "A 16-channel patient-specific seizure onset and termination detection soc with impedance-adaptive transcranial electrical stimulator," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 11, pp. 2728–2740, 2015.

[18] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[21] A. Y. Hannun, P. Rajpurkar, M. Haghpanahi, G. H. Tison, C. Bourn, M. P. Turakhia, and A. Y. Ng, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature medicine*, vol. 25, no. 1, p. 65, 2019.

[22] R. Poplin, A. V. Varadarajan, K. Blumer, Y. Liu, M. V. McConnell, G. S. Corrado, L. Peng, and D. R. Webster, "Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning," *Nature Biomedical Engineering*, vol. 2, no. 3, p. 158, 2018.

[23] G. K. Anumanchipalli, J. Chartier, and E. F. Chang, "Speech synthesis from neural decoding of spoken sentences," *Nature*, vol. 568, no. 7753, p. 493, 2019.

[24] L. Kuhlmann, P. Karoly, D. R. Freestone, B. H. Brinkmann, A. Temko, A. Barachant, F. Li, G. Titericz Jr, B. W. Lang, D. Lavery, *et al.*, "Epilepsyecosystem. org: crowd-sourcing reproducible seizure prediction with long-term human intracranial eeg," *Brain*, vol. 141, no. 9, pp. 2619–2630, 2018.

[25] N. D. Truong, A. D. Nguyen, L. Kuhlmann, M. R. Bonyadi, J. Yang, S. Ippolito, and O. Kavehei, "Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram," *Neural Networks*, vol. 105, pp. 104–111, 2018.

[26] H. Daoud and M. A. Bayoumi, "Efficient epileptic seizure prediction based on deep learning," *IEEE transactions on biomedical circuits and systems*, vol. 13, no. 5, pp. 804–813, 2019.

[27] E. B. Assi, D. K. Nguyen, S. Rihana, and M. Sawan, "Towards accurate prediction of epileptic seizures: A review," *Biomedical Signal Processing and Control*, vol. 34, pp. 144–157, 2017.

[28] L. Kuhlmann, K. Lehnertz, M. P. Richardson, B. Schelter, and H. P. Zaveri, "Seizure prediction—ready for a new era," *Nature Reviews Neurology*, p. 1, 2018.

[29] E. C. Leuthardt, G. Schalk, J. Roland, A. Rouse, and D. W. Moran, "Evolution of brain-computer interfaces: going beyond classic motor physiology," *Neurosurgical focus*, vol. 27, no. 1, p. E4, 2009.

[30] G. K. Bergey, M. J. Morrell, E. M. Mizrahi, A. Goldman, D. King-Stephens, D. Nair, S. Srinivasan, B. Jobst, R. E. Gross, D. C. Shields, *et al.*, "Long-term treatment with responsive brain stimulation in adults with refractory partial seizures," *Neurology*, vol. 84, no. 8, pp. 810–817, 2015.

[31] M. J. Cook, T. J. O'Brien, S. F. Berkovic, M. Murphy, A. Morokoff, G. Fabinyi, W. D'Souza, R. Yerra, J. Archer, L. Litewka, *et al.*, "Prediction of seizure likelihood with a long-term, implanted seizure advisory system in patients with drug-resistant epilepsy: a first-in-man study," *The Lancet Neurology*, vol. 12, no. 6, pp. 563–571, 2013.

[32] R. Bashirullah, "Wireless implants," *IEEE microwave magazine*, vol. 11, no. 7, pp. S14–S23, 2010.

[33] A. Zhou, S. R. Santacruz, B. C. Johnson, G. Alexandrov, A. Moin, F. L. Burghardt, J. M. Rabaey, J. M. Carmena, and R. Muller, "A wireless and artefact-free 128-channel neuromodulation device for closed-loop stimulation and recording in non-human primates," *Nature biomedical engineering*, vol. 3, no. 1, p. 15, 2019.

[34] Y. Park, L. Luo, K. K. Parhi, and T. Netoff, "Seizure prediction with spectral power of eeg using cost-sensitive support vector machines," *Epilepsia*, vol. 52, no. 10, pp. 1761–1770, 2011.

[35] H. Khan, L. Marcuse, M. Fields, K. Swann, and B. Yener, "Focal onset seizure prediction using convolutional networks," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 9, pp. 2109–2118, 2017.

[36] I. Korshunova, P.-J. Kindermans, J. Degrave, T. Verhoeven, B. H. Brinkmann, and J. Dambre, "Towards improved design and evaluation of epileptic seizure predictors," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 3, pp. 502–510, 2017.

[37] R. F. Yazicioglu, "A 60μw 60 nv/ $\sqrt{}$hz readout front-end for portable biopotential acquisition systems," in *IEEE International Solid-State Circuits Conference Digest of Technical Papers, Feb. 2006*, 2006.

[38] N. Verma, A. Shoeb, J. Bohorquez, J. Dawson, J. Guttag, and A. P. Chandrakasan, "A micro-power eeg acquisition soc with integrated feature extraction processor for a chronic seizure detection system," *IEEE journal of solid-state circuits*, vol. 45, no. 4, pp. 804–816, 2010.

[39] J. Xu, R. F. Yazicioglu, B. Grundlehner, P. Harpe, K. A. A. Makinwa, and C. Van Hoof, "A160 μw8-channel active electrode system for eeg monitoring," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 5, pp. 555–567, Dec 2011.

[40] J. Xu, B. Büsze, C. Van Hoof, K. A. Makinwa, and R. F. Yazicioglu, "A 15-channel digital active electrode system for multi-parameter biopotential measurement," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 9, pp. 2090–2100, 2015.

[41] U. Ha, J. Lee, M. Kim, T. Roh, S. Choi, and H.-J. Yoo, "An eeg-nirs multimodal soc for accurate anesthesia depth monitoring," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 6, pp. 1830–1843, 2018.

[42] T. Tang, L. Yan, J. H. Park, H. Wu, L. Zhang, H. Y. B. Lee, and J. Yoo, "34.6 eeg dust: A bcc-based wireless concurrent recording/transmitting concentric electrode," in *2020 IEEE International Solid-State Circuits Conference-(ISSCC)*, pp. 516–518, IEEE, 2020.

[43] N. Meziane, J. Webster, M. Attari, and A. Nimunkar, "Dry electrodes for electrocardiography," *Physiological measurement*, vol. 34, no. 9, p. R47, 2013.

[44] M. Lopez-Gordo, D. Sanchez-Morillo, and F. Valle, "Dry eeg electrodes," *Sensors*, vol. 14, no. 7, pp. 12847–12870, 2014.

[45] T. Denison, K. Consoer, W. Santa, A.-T. Avestruz, J. Cooley, and A. Kelly, "A μw 100 nv/rthz chopper-stabilized instrumentation amplifier for chronic measurement of neural field potentials," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 12, pp. 2934–2945, 2007.

[46] R. Muller, S. Gambini, and J. M. Rabaey, "A 0.013 mm^2, 5μw dc-coupled neural signal acquisition ic with 0.5 v supply," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 1, pp. 232–243, 2011.

[47] C. Kim, S. Joshi, H. Courellis, J. Wang, C. Miller, and G. Cauwenberghs, "Sub-μ v rms-noise sub- μ w/channel adc-direct neural recording with 200-mv/ms transient recovery through predictive digital autoranging," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 11, pp. 3101–3110, 2018.

[48] B. Lee, Y. Jia, S. A. Mirbozorgi, M. Connolly, X. Tong, Z. Zeng, B. Mahmoudi, and M. Ghovanloo, "An inductively-powered wireless neural recording and stimulation system for freely-behaving animals," *IEEE transactions on biomedical circuits and systems*, vol. 13, no. 2, pp. 413–424, 2019.

[49] M. M. Ghanbari, D. K. Piech, K. Shen, S. F. Alamouti, C. Yalcin, B. C. Johnson, J. M. Carmena, M. M. Maharbiz, and R. Muller, "A sub-mm 3 ultrasonic free-floating implant for multi-mote neural recording," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 11, pp. 3017–3030, 2019.

[50] K. D. Wise, A. M. Sodagar, Y. Yao, M. N. Gulari, G. E. Perlin, and K. Najafi, "Microelectrodes, microelectronics, and implantable neural microsystems," *Proceedings of the IEEE*, vol. 96, no. 7, pp. 1184–1202, 2008.

[51] L. Guo, G. S. Guvanasen, X. Liu, C. Tuthill, T. R. Nichols, and S. P. DeWeerth, "A pdms-based integrated stretchable microelectrode array (ismea) for neural and muscular surface interfacing," *IEEE transactions on biomedical circuits and systems*, vol. 7, no. 1, pp. 1–10, 2012.

[52] K. Gadhoumi, J.-M. Lina, F. Mormann, and J. Gotman, "Seizure prediction for therapeutic devices: A review," *Journal of neuroscience methods*, vol. 260, pp. 270–282, 2016.

[53] H.-T. Shiao, V. Cherkassky, J. Lee, B. Veber, E. E. Patterson, B. H. Brinkmann, and G. A. Worrell, "Svm-based system for prediction of epileptic seizures from ieeg signal," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 5, pp. 1011–1022, 2016.

[54] M. Z. Parvez and M. Paul, "Seizure prediction using undulated global and local features," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 208–217, 2016.

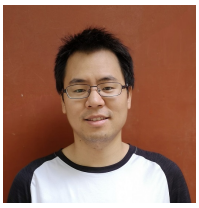
[55] D. Jacobs, T. Hilton, M. Del Campo, P. L. Carlen, and B. L. Bardakjian, "Classification of pre-clinical seizure states using scalp eeg cross-frequency coupling features," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 11, pp. 2440–2449, 2018.

[56] K. Abdelhalim, V. Smolyakov, and R. Genov, "Phase-synchronization early epileptic seizure detector vlsi architecture," *IEEE transactions on biomedical circuits and systems*, vol. 5, no. 5, pp. 430–438, 2011.

[57] S. R. Sridhara, M. DiRenzo, S. Lingam, S.-J. Lee, R. Blazquez, J. Maxey, S. Ghanem, Y.-H. Lee, R. Abdallah, P. Singh, *et al.*, "Microwatt embedded processor platform for medical system-on-chip

applications," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 4, pp. 721–730, 2011.
[58] M. Shoaib, K. H. Lee, N. K. Jha, and N. Verma, "A 0.6–107 μw energy-scalable processor for directly analyzing compressively-sensed eeg," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 4, pp. 1105–1118, 2014.
[59] K. H. Lee and N. Verma, "A low-power processor with configurable embedded machine-learning accelerators for high-order and adaptive analysis of medical-sensor signals," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 7, pp. 1625–1637, 2013.
[60] A. Page, C. Sagedy, E. Smith, N. Attaran, T. Oates, and T. Mohsenin, "A flexible multichannel eeg feature extractor and classifier for seizure detection," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 62, pp. 109–113, Feb 2015.
[61] S. Ravindran and R. Cole, "Low complexity algorithms for heart rate and epileptic seizure detection," in *2009 2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies*, pp. 1–5, IEEE, 2009.
[62] M. A. B. Altaf and J. Yoo, "A 1.83 muj/classification, 8-channel, patient-specific epileptic seizure classification soc using a non-linear support vector machine," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 10, no. 1, pp. 49–60, 2016.
[63] B. G. Do Valle, S. S. Cash, and C. G. Sodini, "Low-power, 8-channel eeg recorder and seizure detector asic for a subdermal implantable system," *IEEE transactions on biomedical circuits and systems*, vol. 10, no. 6, pp. 1058–1067, 2016.
[64] L. Feng, Z. Li, and Y. Wang, "Vlsi design of svm-based seizure detection system with on-chip learning capability," *IEEE transactions on biomedical circuits and systems*, vol. 12, no. 1, pp. 171–181, 2017.
[65] I. Kiral-Kornek, S. Roy, E. Nurse, B. Mashford, P. Karoly, T. Carroll, D. Payne, S. Saha, S. Baldassano, T. O'Brien, *et al.*, "Epileptic seizure prediction using big data and deep learning: toward a mobile system," *EBioMedicine*, vol. 27, pp. 103–111, 2018.
[66] M. Shoaran, B. A. Haghi, M. Taghavi, M. Farivar, and A. Emami-Neyestanak, "Energy-efficient classification for resource-constrained biomedical applications," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 8, no. 4, pp. 693–707, 2018.
[67] C.-H. Cheng, P.-Y. Tsai, T.-Y. Yang, W.-H. Cheng, T.-Y. Yen, Z. Luo, X.-H. Qian, Z.-X. Chen, T.-H. Lin, W.-H. Chen, *et al.*, "A fully integrated 16-channel closed-loop neural-prosthetic cmos soc with wireless power and bidirectional data telemetry for real-time efficient human epileptic seizure control," *IEEE Journal of Solid-State Circuits*, no. 99, pp. 1–13, 2018.
[68] G. O'Leary, D. M. Groppe, T. A. Valiante, N. Verma, and R. Genov, "Nurip: Neural interface processor for brain-state classification and programmable-waveform neurostimulation," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 11, pp. 3150–3162, 2018.
[69] S.-A. Huang, K.-C. Chang, H.-H. Liou, and C.-H. Yang, "A 1.9-mw svm processor with on-chip active learning for epileptic seizure control," *IEEE Journal of Solid-State Circuits*, 2019.
[70] F. Samie, S. Paul, L. Bauer, and J. Henkel, "Highly efficient and accurate seizure prediction on constrained iot devices," in *2018 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 955–960, IEEE, 2018.
[71] L. Wang, *Support vector machines: theory and applications*, vol. 177. Springer Science & Business Media, 2005.
[72] K. Kang and T. Shibata, "An on-chip-trainable gaussian-kernel analog support vector machine," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 7, pp. 1513–1524, 2009.
[73] R. Genov and G. Cauwenberghs, "Kerneltron: support vector" machine" in silicon," *IEEE Transactions on Neural Networks*, vol. 14, no. 5, pp. 1426–1434, 2003.
[74] P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza, A. S. Cassidy, J. Sawada, F. Akopyan, B. L. Jackson, N. Imam, C. Guo, Y. Nakamura, *et al.*, "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, 2014.
[75] M. Ihle, H. Feldwisch-Drentrup, C. A. Teixeira, A. Witon, B. Schelter, J. Timmer, and A. Schulze-Bonhage, "Epilepsiae–a european epilepsy database," *Computer methods and programs in biomedicine*, vol. 106, no. 3, pp. 127–138, 2012.
[76] J. Klatt, H. Feldwisch-Drentrup, M. Ihle, V. Navarro, M. Neufang, C. Teixeira, C. Adam, M. Valderrama, C. Alvarado-Rojas, A. Witon, *et al.*, "The epilepsiae database: An extensive electroencephalography database of epilepsy patients," *Epilepsia*, vol. 53, no. 9, pp. 1669–1676, 2012.
[77] J. B. Wagenaar, G. A. Worrell, Z. Ives, M. Dümpelmann, B. Litt, and A. Schulze-Bonhage, "Collaborating and sharing data in epilepsy research," *Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society*, vol. 32, no. 3, p. 235, 2015.
[78] M. Hügle, S. Heller, M. Watter, M. Blum, F. Manzouri, M. Dumpelmann, A. Schulze-Bonhage, P. Woias, and J. Boedecker, "Early seizure detection with an energy-efficient convolutional neural network on an implantable microcontroller," in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, IEEE, 2018.
[79] M.-P. Hosseini, H. Soltanian-Zadeh, K. Elisevich, and D. Pompili, "Cloud-based deep learning of big eeg data for epileptic seizure prediction," in *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 1151–1155, IEEE, 2016.
[80] S. M. Kueh and T. J. Kazmierski, "Low-power and low-cost dedicated bit-serial hardware neural network for epileptic seizure prediction system," *IEEE journal of translational engineering in health and medicine*, vol. 6, pp. 1–9, 2018.
[81] J. Xu, S. Mitra, C. Van Hoof, R. F. Yazicioglu, and K. A. Makinwa, "Active electrodes for wearable eeg acquisition: Review and electronics design methodology," *IEEE reviews in biomedical engineering*, vol. 10, pp. 187–198, 2017.
[82] "Medical electrical equipment: Particular requirements for basic safety and essential performance of electroencephalographs.,"
[83] M. R. Nuwer, G. Comi, R. Emerson, A. Fuglsang-Frederiksen, J.-M. Guérit, H. Hinrichs, A. Ikeda, F. J. C. Luccas, and P. Rappelsburger, "Ifcn standards for digital recording of clinical eeg," *Electroencephalography and clinical Neurophysiology*, vol. 106, no. 3, pp. 259–261, 1998.
[84] J. Xu, M. Konijnenburg, S. Song, H. Ha, R. Van Wegberg, M. Mazzillo, G. Fallica, C. Van Hoof, W. De Raedt, and N. Van Helleputte, "A 665 μw silicon photomultiplier-based nirs/eeg/eit monitoring asic for wearable functional brain imaging," *IEEE transactions on biomedical circuits and systems*, vol. 12, no. 6, pp. 1267–1277, 2018.
[85] B. Edelman, J. Meng, D. Suma, C. Zurn, E. Nagarajan, B. Baxter, C. Cline, and B. He, "Noninvasive neuroimaging enhances continuous neural tracking for robotic device control," *Science Robotics*, vol. 4, no. 31, p. eaaw6844, 2019.
[86] E. Musk *et al.*, "An integrated brain-machine interface platform with thousands of channels," *Journal of medical Internet research*, vol. 21, no. 10, p. e16194, 2019.
[87] A. Supratak, H. Dong, C. Wu, and Y. Guo, "Deepsleepnet: A model for automatic sleep stage scoring based on raw single-channel eeg," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 11, pp. 1998–2008, 2017.
[88] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255, Ieee, 2009.
[89] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.
[90] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4510–4520, 2018.
[91] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6848–6856, 2018.
[92] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size," *arXiv preprint arXiv:1602.07360*, 2016.
[93] R. J. Wang, X. Li, and C. X. Ling, "Pelee: A real-time object detection system on mobile devices," in *Advances in Neural Information Processing Systems*, pp. 1963–1972, 2018.
[94] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces," *Journal of neural engineering*, vol. 15, no. 5, p. 056013, 2018.
[95] Y. Wang, Q. Sun, H. Luo, X. Chen, X. Wang, and H. Zhang, "26.3 a closed-loop neuromodulation chipset with 2-level classification achieving 1.5 v pp cm interference tolerance, 35db stimulation artifact rejection in 0.5 ms and 97.8% sensitivity seizure detection," in *2020 IEEE International Solid-State Circuits Conference-(ISSCC)*, pp. 406–408, IEEE, 2020.
[96] J. Lee, C. Kim, S. Kang, D. Shin, S. Kim, and H.-J. Yoo, "Unpu: An energy-efficient deep neural network accelerator with fully variable weight bit precision," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 1, pp. 173–185, 2018.

[97] D. Bankman, L. Yang, B. Moons, M. Verhelst, and B. Murmann, "An always-on 3.8 μj 86% cifar-10 mixed-signal binary cnn processor with all memory on chip in 28-nm cmos," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 1, pp. 158–172, 2018.
[98] K. Lehnertz and B. Litt, "The first international collaborative workshop on seizure prediction: summary and data description," *Clinical neurophysiology*, vol. 116, no. 3, pp. 493–505, 2005.



Jie Yang was born in Chongqing, China, in 1987. He received the B.S. degree in electronic science and technology from Tianjin University, Tianjin, China, in 2010, and the Ph.D. degree in microelectronics from the Institute of Semiconductors, Chinese Academy of Sciences, Beijing, China, in 2015. He was a post-doctoral fellow with the I2Sense Lab, Department of Electrical and Computer Engineering, University of Calgary, Calgary, Canada from 2015 to 2019, where he developed tone mapping algorithms and circuits for wide dynamic range imaging systems. In 2019, he joined the Cutting-Edge Net of Biomedical Research and INnovation (CenBRAIN), School of Engineering, Westlake University, Hangzhou, China where he is currently a research scientist in biomedical circuits and systems.

His current research focus is on the interdisciplinary study of the energy-efficient machine learning algorithms and mixed-signal LSI for intelligent biomedical diagnosis and prediction.

Mohamad Sawan (S'88-M'89-SM'96-F'04) received the Ph.D. degree in Electrical Engineering from Universite de Sherbrooke, Sherbrooke, QC, Canada, in 1990. He is a Professor of Microelectronics and Biomedical Engineering, in leave of absence from Polytechnique Montréal, Canada. He joined Westlake University, Hangzhou, China, in 2019, where he is a Chair Professor founder and director of the Center for Biomedical Research And INnovation (CenBRAIN). He was Chair Professor awarded the Canada Research Chair in Smart Medical Devices (2001-2015), was leading the Microsystems Strategic Alliance of Quebec - ReSMiQ (1999-2018). He founded and Chaired the IEEE-Solid State Circuits Society Montreal Chapter (1999-2018) and founded the Polystim Neurotech Laboratory in Polytechnique Montréal (1994-Present), including two major research infrastructures intended to build advanced Medical devices. Dr. Sawan was Deputy Editor-in Chief of the IEEE Transactions on Circuits and Systems-II: Express Briefs (2010-2013), Co-Founder, Associate Editor and Editor-in-Chief of the IEEE Transactions on Biomedical Circuits and Systems (2016-2019), Associate Editor of the IEEE Transactions on Biomedical Engineering. He is founder of the Interregional IEEE-NEWCAS Conference, and Co-Founder of the International IEEE-BioCAS, ICECS and LSC conferences. He hosted in Montreal, as General Chair, the 2016 IEEE International Symposium on Circuits and Systems (ISCAS), and will host as General Chair, the 2020 IEEE International Engineering, Medicine, Biology Conference (EMBC). He served as member of the Board of Governors (2014-2018), and he is Vice-President Publications (2019-Present) of the IEEE CAS Society. Dr. Sawan published more than 800 peer reviewed papers, two books, 10 book chapters, and 12 patents. He received several awards, among them the Queen Elizabeth II Golden Jubilee Medal, the Barbara Turnbull 2003 Award for spinal-cord research, the Bombardier and Jacques-Rousseau Awards for academic achievements, the Shanghai International Collaboration Award, and the medal of merit from the President of Lebanon for his outstanding contributions. Dr. Sawan is Fellow of the IEEE, Fellow of the Canadian Academy of Engineering, and Fellow of the Engineering Institutes of Canada. He is also "Officer" of the National Order of Quebec.