A Low Power Multi-Class Migraine Detection Processor Based on Somatosensory Evoked Potentials

Zain Taufique, *Student Member, IEEE*, Bingzhao Zhu, *Student Member, IEEE*, Gianluca Coppola, Mahsa Shoaran, *Member, IEEE*, and Muhammad Awais Bin Altaf, *Member, IEEE*

***Abstract*—Migraine is a disabling neurological disorder that can be recurrent and persist for long durations. The continuous monitoring of the brain activities can enable the patient to respond on time before the occurrence of the approaching migraine episode to minimize the severity. Therefore, there is a need for a wearable device that can ensure the early diagnosis of a migraine attack. This brief presents a low latency, and power-efficient feature extraction and classification processor for the early detection of a migraine attack. Somatosensory Evoked Potentials (SEP) are utilized to monitor the migraine patterns in an ambulatory environment aiming to have a processor integrated on-sensor for power-efficient and timely intervention. In this work, a complete digital design of the wearable environment is proposed. It allows the extraction of multiple features including multiple power spectral bands using 256-point fast Fourier transform (FFT), root mean square of late HFO bursts and latency of N20 peak. These features are then classified using a multi-classification artificial neural network (ANN)-based classifier which is also realized on the chip. The proposed processor is placed and routed in a 180nm CMOS with an active area of 0.5mm^2. The total power consumption is 249μW while operating at a 20MHz clock with full computations completed in 1.31ms.**

***Index Terms*—Continuous health monitoring, classification processor, migraine, machine learning, Somatosensory evoked potential (SEP), artificial neural network (ANN).**

Manuscript received February 4, 2021; accepted March 2, 2021. Date of publication March 17, 2021; date of current version April 30, 2021. This work was supported by the Higher Education Commission (HEC), Pakistan under Grant 7978/Punjab/NRPU/R&D/HEC/2017. This brief was recommended by Associate Editor D. G. Muratore. *(Corresponding author: Muhammad Awais Bin Altaf.)*

Zain Taufique is with the Electrical Engineering Department, Lahore University of Management Sciences, Lahore 54792, Pakistan.

Bingzhao Zhu is with the Institute of Electrical Engineering and Center for Neuroprosthetics, EPFL, 1015 Lausanne, Switzerland, and also with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA.

Gianluca Coppola is with the Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome Polo Pontino, 04100 Latina, Italy.

Mahsa Shoaran is with the Institute of Electrical Engineering and Center for Neuroprosthetics, EPFL, 1015 Lausanne, Switzerland (e-mail: mahsa.shoaran@epfl.ch).

Muhammad Awais Bin Altaf is with the Electrical Engineering Department, Lahore University of Management Sciences, Lahore 54792, Pakistan (e-mail: awais.altaf@lums.edu.pk).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCSII.2021.3066389.

Digital Object Identifier 10.1109/TCSII.2021.3066389

I. Introduction

MIGRAINE is a type of chronic neurological disorder that causes recurrent headaches of variable duration and intensity. Migraine with aura (MA) patients faces visual changes, blind vision spots, and flashes of light before headache [1], whereas the patients without aura (MO) do not have any physical indicators. Despite being a common chronic neuro-disorder, migraine treatment follows only the conventional diagnosis mechanism of patient interviews followed by physical examinations. The early diagnosis of the upcoming attack, preferably with an automated wearable, can be vital for effective treatment [2]. Migraine patients can use a wearable device to monitor the possibility of an upcoming attack in real-time, to receive proper medications or neuromodulation therapy [4]. Wearable devices are quite popular in the early detection and handling of chronic neurological disorders [3] but are less explored for migraine treatment. Recent migraine studies have shown that somatosensory evoked potentials (SEP) can be a useful noninvasive technique for the diagnosis of migraine [2]–[5]. Here, we aim to classify migraine patients from healthy controls using a human SEP dataset on a hardware chip-based classifier. This processor can be integrated into a wearable setting in the future for the early diagnosis of migraine headaches.

Migraine patients typically show hyper-responsivity in their SEP and visual evoked potentials (VEP), which can be useful in the early diagnosis of the disease [6]. The proposed work is based on SEP as it is more suitable for a wearable environment. Both MA and MO patients show an absence of acclimatization to repetitive stimuli in the interictal state [6]. High-Frequency Oscillations (HFO) bursts are a prime indicator for the classification of migraine. Particularly, the increase in late HFOs during the attacks and the reduction of early HFOs in between attacks have been reported [6], [8]. Similarly, low-frequency components of SEP show increased N20-P25 amplitude in the ictal group which can be an important factor in migraine diagnosis. Even if the underlying causes of migraine are not completely known, the changes in the SEP signal can be used as informative features for migraine classification. Fig. 1 shows the comparative trend of LF N20-P25 peaks with early and late HFOs extracted from different groups of patients including healthy volunteers (HV), migraine interictal with aura (MA), migraine interictal without aura (MO), and migraine during ictal stage (MI).

Electroencephalogram (EEG) has been quite a popular data acquisition technique that is used in the classification of

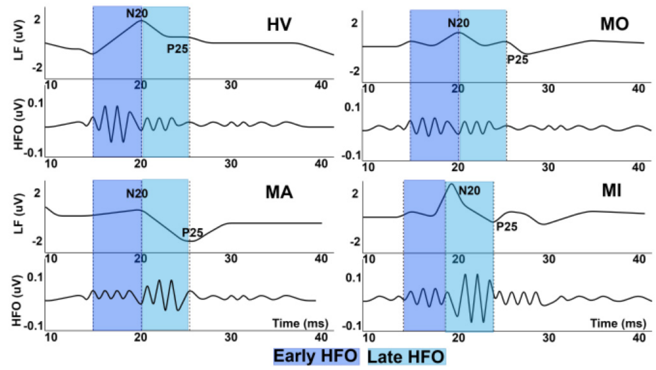


Fig. 1. The trend of early and late HFOs and N20-P25 peaks in migraine patients and healthy controls.

different neurological disorders. For the diagnosis of migraine, EEG complexity, power spectral bands, and time-domain features have been used with artificial neural network (ANN) and SVM classifiers [9], [10]. However, the higher number of EEG channels add complexity to the hardware chip realization. SEP classification based on multiple features and different classifiers was tested in [4]. They were able to successfully differentiate migraine patients from healthy volunteers using multiple classifiers such as random forest (RF), support vector machines (SVMs) having linear and non-linear kernel variants, K-nearest neighbors (KNN), extreme gradient-boosting trees (XGB), regression, and neural networks. This is a significant advancement in migraine diagnosis using SEP. The proposed work is motivated by this research and aims to realize the on-chip implementation of migraine detection using a multi-task classification processor. The previous works on migraine detection are all software-based [10]–[13], and cannot be of much assistance for MO cases. Here, the hardware implementation is optimized to ensure both low power and compact area by incorporating an area-efficient RTL, clock gating and power shutoff optimization techniques. Efficient on-chip realization of the feature extraction unit is achieved by a 256-point fast Fourier transform (FFT) implementation that exploits 2-dimensional blocks of two 16-point FFTs [14]. The implemented 256-point FFT consumes only 0.1mW at 20MHz (5pJ/FFT-operation) with the full 256-point computation completed in 1.31ms. Moreover, a patient-specific multi-class customized ANN is proposed and realized for on-chip implementation of both MA and MO detection.

The dataset was comprised of HV, MA, MO, and MI groups including male and female participants. Both MA and MO were considered as migraine interictal (MII) states. We have utilized the database recorded in [4]. The data from 57 subjects were used to validate the performance of the proposed system. To elicit SEP signals, a CED 1401 device was utilized to stimulate the median nerve electrically. A square wave of constant current was applied to have a pulse width of 0.2ms with a recurrence rate of 4.4Hz. The sampling frequency was set to 5000Hz and the data was recorded for 40ms post-stimulation. The dataset of each group is labelled as HV, MI, and MII. The MI stage is considered as the 12 hours before and after the migraine episode, while the MII stage is considered as the 72 hours before and after the MI stage.

II. Proposed Migraine Classification Algorithm

The early detection of migraine using SEP signals includes acquisition, pre-processing, feature extraction, and classification of the feature set. The pre-processing of the acquired dataset that is provided by [4], is done by artifact removal from the raw signals by averaging multiple SEP trials. Multiple features have been extracted from the pre-processed data including power spectral bands, LF features, and HFO based biomarkers. Three different classification models have been developed for classifying three different groups including HV-MI, HV-MII, and HV-MI-MII [4]. The complete system diagram is shown in Fig. 2. In this work, we have mainly implemented HV-MI-MII multi-level classification using a one-vs-one multi-classification approach along with a final decision block to enhance the overall accuracy by suppressing the false positives.

A. Feature Extraction

The most distinguishing features for the classification of SEP data for early diagnosis of migraine include N20 latency, late HFO RMS, and power spectral bands [4]. The N20 peaks are extracted by sending the pre-processed signal through a low pass 0-450 Hz filter. The HFO features are extracted by using 450-750 Hz bandpass filters (BPF). A Bartlett-Hanning window with a filter order of 50 is used to realize both BPF and low-pass filters.

Power spectral density (PSD) bands are calculated using Welch's periodogram that takes an average of weighted overlapped segments [15]. Welch's technique utilizes windowed samples from the pumped input data to calculate the FFT. For x[j] vector comprised of the discrete-time domain input data, as in eq. (1).

$$x[j] = x_0 + x_1 + x_2 + \cdots + x_j \tag{1}$$

where j is the total number of sample points. The window of samples is evaluated by taking the cross product of the window vector with the input vector, as in eq. (2)

$$X_w[j] = w[j] \times x[j] \tag{2}$$

There are numerous methods for evaluating the window vector, whereas in this work the Bartlett-Hanning window is used to ensure area-and-power efficient design. The window vector for the proposed system is designed to extract a window of 256 data samples. These data samples are sent to an FFT module with 256 points input and output length. The conventional Cooley–Tukey radix-2 256-point FFT requires excessive hardware area and computational power due to O(Nlog$_2$N) complexity and 1024 complex butterfly operations. Each butterfly operation requires two additions and one complex multiplication. The proposed system is based on split radix based FFT implementation which splits the overall architecture in a 2D structure (eq. 3) of two 16-point FFT and 256 complex multiplications while the overall energy per FFT and area requirements are reduced by 58% and 39%, respectively [16].

$$A(s+16t) = \sum_{l=0}^{15} W_{16}^{lt} \left[W_{256}^{sl} \sum_{m=0}^{15} B(l+16m) W_{16}^{sm} \right] \tag{3}$$

where s, l, m, t $\epsilon\{0, 1, \ldots, 15\}$ and s, t represents the rows and columns numbers, respectively, of 256-point FFT. The BPF is applied to the output of the FFT accelerator to evaluate predetermined frequency bands.

$$X_{bp}[n] = \forall X_W[n] \exists F_{LOW} < n < F_{HIGH} \tag{4}$$

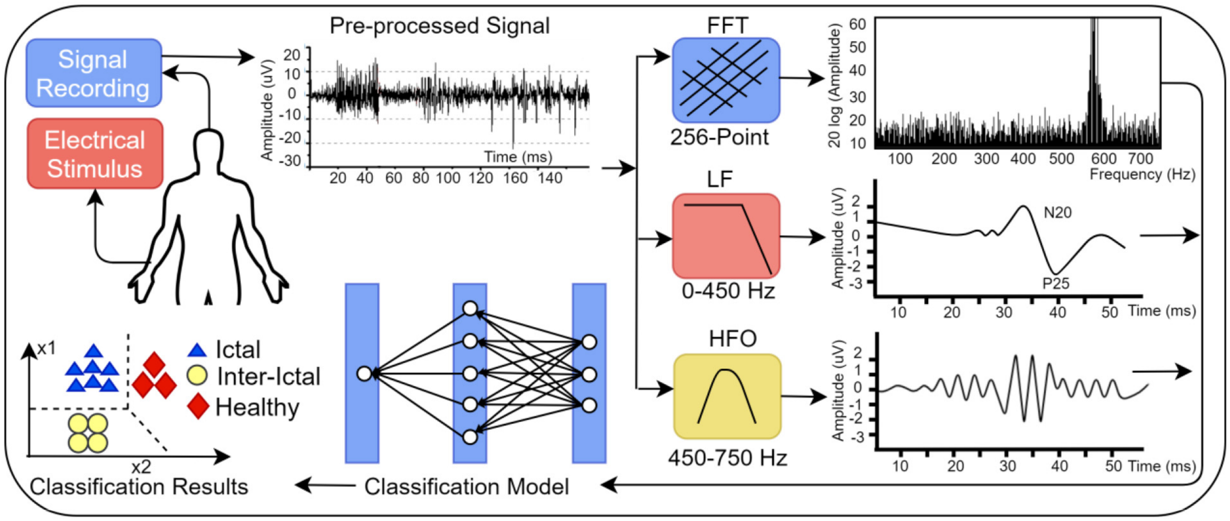


Fig. 2. The overall system diagram showing the complete migraine diagnosis process, including data acquisition, feature extraction, and classification.

TABLE I
FEATURE SETS OF THE SEP DATA THAT IS USED FOR CLASSIFICATION

Feature Number	Feature Description
Feature 1 (F1)	PSD (80-200 Hz)
Feature 2 (F2)	PSD (450-750 Hz)
Feature 3 (F3)	Late HFO RMS
Feature 4 (F4)	N20 latency

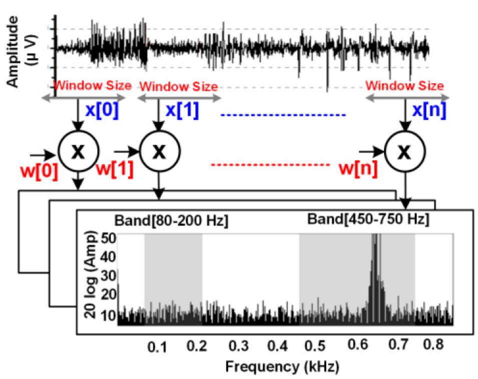


Fig. 3. Power Spectral bands extracted using welch method.

where X_{bp} is valued after applying BPF, F_{LOW} and F_{HIGH} are the lower and higher frequency values, respectively. The mean of the data sample points is evaluated to estimate the PSD and results in a single value that is utilized in the feature vector for training purposes, as in (5).

$$P[n] = \frac{1}{L}\sum_{n=1}^{L}[X_{bp}[n]]^2 \tag{5}$$

The best performing features for the classification of migraine are shown in Table I. The concept diagram of the welch method used for the extraction of power spectral bands from the SEP dataset is shown in Fig. 3.

B. Classification

An ANN is a matrix multiplication between the feature value with the layer weight, added to a bias value. The output

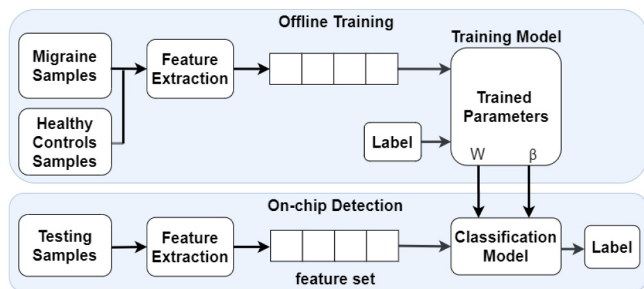


Fig. 4. ANN training and testing algorithms.

of each node is calculated by the weighted sum of its input features while progressing through different layers and operated by a non-linear activation function, as in (6).

$$z = f(\mathrm{b} + \mathrm{X} * \mathrm{W}) = f\left(b + \sum_{i=1}^{n} Xi * Wi\right) \tag{6}$$

A classification process has two major steps: The learning phase and the testing phase. The detection of migraines from healthy controls can be taken as a binary ANN classification problem. This decision is made based on the training of the model according to the training label (Fig. 4).

The dataset requires labelling in such a way that MI or MII data samples can be identified from the HV data samples. To classify three different classes HV, MI, and MII from the data in real-time, we have utilized a multi-classification approach. The multi-classification approach is based on a combination of multiple binary approaches. Our analysis favored one-vs-one compared to one-vs-all in terms of classification accuracy and training comfort. Both approaches resulted in the same hardware resources as for n=3, both require 3 binary classifiers. For the training of the ANN model, it is required that data samples with both labels are nearly equal in number, this is called data normalization. Therefore, the labelled data samples in higher numbers are down sampled in a random way to enable data normalization. The model for all three binary classifiers (HV-MI, MI-MII, and HV-MII) is based on one input layer (IL), 2 hidden layers (HL's), and one output

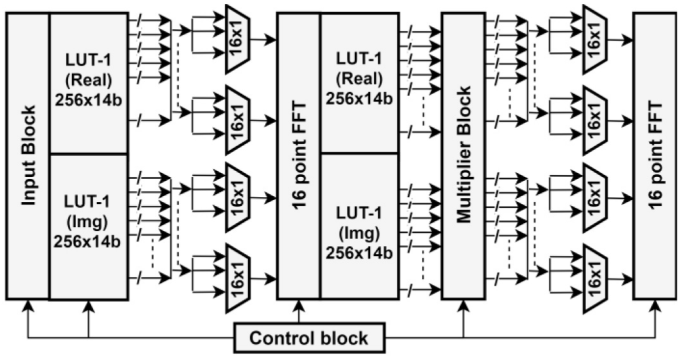


Fig. 5. 256-point FFT hardware Implementation.

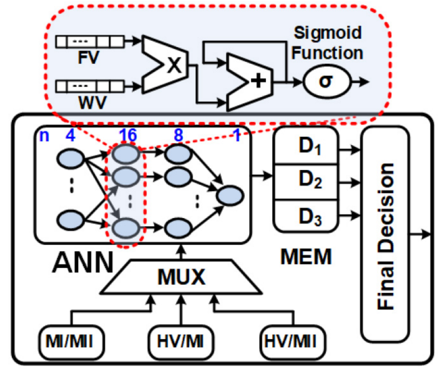
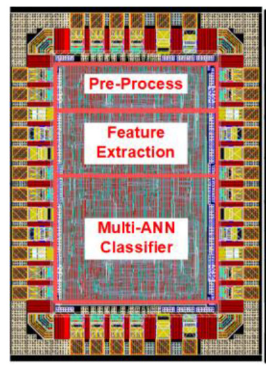


Fig. 6. Classifier patient-specific hardware implementation.



Implemented Results	
Process	180nm
Area	0.5 x 1.0 mm^2
Supply Voltage	1.0 V
Power	249µ W
Classifier	Multi-ANN
Accuracy	76%
Latency	180 CC
Precision	100%
Application	Migraine

Fig. 7. Chip layout for implemented Multi-ANN classification processor.

layer (OL). The classification model has four IL neurons, the HL-1 has 16 neurons and HL-2 has 8 neurons in both models. The OL has only one neuron with information on the binary classification results which can be either '0' or '1'. Using sigmoid activation, the parameters of the models are tuned to achieve the optimal classification performance and realize an implementation with the lowest hardware cost.

III. Migraine Classification Processor

To realize the migraine classifier on-chip, we need to implement pre-processor, feature extraction, and the multi-classifier in hardware, as shown in Fig. 2. The daunting task is feature extraction and classifier implementation with strict area and power budget constraints. F1 and F2 features are based on PSD for a specific frequency band, therefore, we are proposing and implementing an efficient 256-point FFT realization shown in Fig. 5 to achieve it. The proposed hardware architecture for the PSD calculations includes an input block, a 16-point FFT, a multiplier block, and an intermediate memory storage block. There are numerous wires and sub-modules in the overall architecture that manage the data flow from beginning to end. A control block based on a finite state machine keeps track of the overall status flags and counters. The input block takes the fixed-point input samples, converts them to a 14-bit IEEE 754 floating-point number, and sends them to the first 16-point FFT module in an appropriate order. The results of the 16-point FFT reach the multiplier block and here the twiddle factor multiplication takes place. The twiddle factors are the pre-known constants, decomposed in the power of 2, and realized by add-shift operations. The calculated results from the multiplier block get saved in intermediate memory storage and the system waits until all the input data samples are processed accordingly. The saved samples are again passed through the 16-point FFT block following the order of eq. (3) and the results are saved in a final output block. The split radix approach is not fully parallel, and the overall area of the implementation is saved on the cost of computational time. The operating frequency of the FFT architecture is 20MHz and the total time required by the complete architecture to compute 256-point FFT is 1.31ms. F3 and F4 features are realized using standard floating-point arithmetic and IP blocks provided by the cadence design ware therefore details are not provided.

The implemented classifier is based on three binary ANN classifiers, i.e., HV-MI, MI-MII, and HV-MII. The weights (W) and offset (b) are calculated in a model that is trained offline on the SEP dataset using python. These training parameters along with the feature set (x) reach the on-chip classifier as inputs. Fig. 6 shows that a single hardware architecture is used for all the three classifications to use hardware resources efficiently. The weights of the three binary classification models are uploaded in the SRAM and can be updated in real-time based on new training data. This weight vector is multiplied with the respective feature and then a bias is added to the result, which is then passed through the sigmoid function while crossing through each layer and the output is calculated.

The proposed work aims to reduce the overall power consumption of the system. Clock Gating (CG) and Power Shutoff (PSO) are two popular low power optimization techniques that are utilized in the proposed processor implementation. If the input data remains unchanged for the current clock period, clock-gating disables the clock of that Flip Flop (DFF). Hence, it disables the clock of DFF and saves power up to 15% in our implementation whereas the PSO helps to reduce power by a further 37% in our implementation at the expense of an additional area of 18% for state retention power gating cells and extra wire routings.

IV. Implementation and Measurement Results

The proposed early migraine detection classifier is implemented in a CMOS 180nm 1P6M process with a 1.0V supply voltage. Fig. 7 shows the chip layout photo and the performance summary for the processor with an active area of 0.5 mm^2. The power consumption of the multi-ANN classification processor is 249μW. We are estimating that the overall

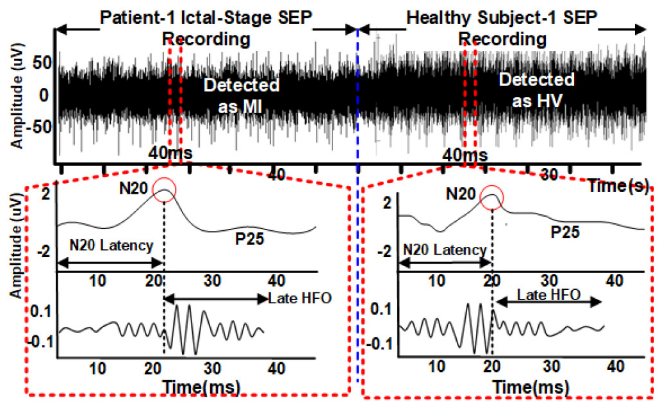


Fig. 8. Measurement results for HV and MI case from the recorded SEP data.

TABLE II
COMPARISON TO PRIOR ART ON MIGRAINE DETECTION

Parameter	ISSPA'12 [10]	CNE'13 [13]	ICCIC'18 [11]	Cephalalgia'19 [4]	TFS'20 [12]	This Work
Hardware	No	No	No	No	No	Yes
Features	7	>10	3	4	-	4
Classifier	ANN	Parzen	RF	XGB	ADA	ANN
Accuracy	90.9%	73%	88%	73.3%/ M	81%	**76%/ M**
Data Extract	EEG	SSVEP	EEG	SEP	SSVEP	SEP

SoC power will be under 300μW, including the SEP signal acquisition and digitization [18], [19].

The implemented processor is tested and validated through an FPGA Virtex5 (XC5VLX110T) real-time measurement based on SEP migraine benchmark database [4]. The accuracy of the processor is estimated using a 5-fold validation scheme and achieved an accuracy of 76% for 3-class classification. We loaded the SEP data from the database and the trained patient parameters on FPGA and tested the implemented system on a specific patient to mimic the realistic scenario. The measurement scenario is also detailed in Fig. 8. The SEP recordings of both the MI and HV are concatenated as a testing dataset for the classifier.

A comparison with state-of-the-art works is shown in Table II. References [10] and [11] used EEG based migraine detection while utilizing complex features and classifiers. They achieved fairly good accuracy (>85%) but are not practical for hardware realization due to intensive computations and multiple channels involved. Whereas [12] and [13] used steady state VEP (SSVEP) for migraine detection but both the reported accuracies and techniques are not suited for the MO detection. Moreover, all of the mentioned works were based on binary classification, classifying migraine from an HV only. The proposed processor is the first multi-classification hardware implementation for early migraine detection. It achieves an overall accuracy and precision of 76% and 100%, respectively.

V. CONCLUSION

We have implemented the first hardware-based person-specific SEP-based classification processor using a machine learning-based ANN classifier which can help the patient in the early diagnosis of migraines. The proposed processor classifies with an accuracy of 76%. The system is synthesized and placed and routed using a 180nm CMOS processor while consuming the overall power of 249μW at 20MHz with system latency of 50ms for every classification.

REFERENCES

[1] M. Garingo, F. Sahba, W. Ma, and M. Doidge, "A non-linear support vector machine approach to testing for migraine with aura using electroencephalography," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Las Vegas, NV, USA, Dec. 2018, pp. 1713–1719.

[2] K. A. Foley *et al.*, "Treating early versus treating mild: Timing of migraine prescription medications among patients with diagnosed migraine," *Headache J. Head Face Pain*, vol. 45, no. 5, pp. 538–545, May 2005.

[3] A. R. Aslam and M. A. B. Altaf, "An on-chip processor for chronic neurological disorders assistance using negative affectivity classification," *IEEE Trans. Biomed. Circuits Syst.*, vol. 14, no. 4, pp. 838–851, Aug. 2020.

[4] B. Zhu, G. Coppola, and M. Shoaran, "Migraine classification using somatosensory evoked potentials," *Cephalalgia*, vol. 39, no. 9, pp. 1143–1155, Aug. 2019.

[5] M. Valeriani, S. Rinalduzzi, and F. Vigevano, "Multilevel somatosensory system disinhibition in children with migraine," *Pain*, vol. 118, nos. 1–2, pp. 137–144, Nov. 2005.

[6] G. Coppola *et al.*, "Somatosensory evoked high-frequency oscillations reflecting thalamo-cortical activity are decreased in migraine patients between attacks," *Brain*, vol. 128, pp. 98–103, Jan. 2005.

[7] M. de Tommaso *et al.*, "Altered processing of sensory stimuli in patients with migraine," *Nat. Rev. Neurol.*, vol. 10, no. 3, pp. 144–155, Feb. 2014.

[8] G. Coppola, E. Iacovelli, M. Bracaglia, M. Serrao, C. di Lorenzo, and F. Pierelli, "Electrophysiological correlates of episodic migraine chronification: Evidence for thalamic involvement," *J. Headache Pain*, vol. 14, no. 1, p. 76, Dec. 2013.

[9] Z.-H. Cao, L.-W. Ko, K.-L. Lai, S.-B. Huang, S.-J. Wang, and C.-T. Lin, "Classification of migraine stages based on resting-state EEG power," in *Proc. Int. Joint Conf. Neural Netw.*, Killarney, Ireland, Sep. 2015, pp. 1–5.

[10] E. Sayyari, M. Farzi, R. R. Estakhrooeieh, F. Samiee, and M. B. Shamsollahi, "Migraine analysis through EEG signals with classification approach," in *Proc. 11th Int. Conf. Inf. Sci. Signal Process. Appl. (ISSPA)*, Montreal, QC, Canada, 2012, pp. 859–863.

[11] K. Jindal *et al.*, "Migraine disease diagnosis from EEG signals using non-linear feature extraction technique," in *Proc. IEEE Int. Conf. Comput. Intell. Comput.*, Madurai, India, Dec. 2018, pp. 1–4.

[12] Z. Cao *et al.*, "Extraction of SSVEPs-based inherent fuzzy entropy using a wearable headband EEG in migraine patients," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 1, pp. 14–27, Jan. 2020.

[13] L.-W. Ko, K.-L. Lai, P.-H. Huang, C.-T. Lin, and S.-J. Wang, "Steady-state visual evoked potential based classification system for detecting migraine seizures," in *Proc. IEEE/EMBS Int. Conf. Neural Eng. (NER)*, San Diego, CA, USA, 2013, pp. 1299–1302.

[14] W. Saadeh, M. A. B. Altaf, H. Alsuradi, and J. Yoo, "A pseudo OFDM with miniaturized FSK demodulation body-coupled communication transceiver for binaural hearing aids in 65 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 52, no. 3, pp. 757–768, Mar. 2017.

[15] P. D. Welch, "The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. 15, no. 2, pp. 70–73, Jun. 1967.

[16] W. Saadeh, F. H. Khan, and M. A. B. Altaf, "Design and implementation of a machine learning based EEG processor for accurate estimation of depth of anesthesia," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 4, pp. 658–669, Aug. 2019.

[17] F. H. Khan, U. Ashraf, M. A. B. Altaf, and W. Saadeh, "A patient-specific machine learning based EEG processor for accurate estimation of depth of anesthesia," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Cleveland, OH, USA, 2018, pp. 1–4.

[18] 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 J. Solid-State Circuits*, vol. 50, no. 11, pp. 2728–2740, Nov. 2015.

[19] A. R. Aslam, T. Iqbal, M. Aftab, W. Saadeh, and M. A. B. Altaf, "A10.13uJ/classification 2-channel deep neural network-based SoC for emotion detection of autistic children," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Boston, MA, USA, Mar. 2020, pp. 1–4.