eyeSay: Brain Visual Dynamics Decoding With Deep Learning & Edge Computing

Jiadao Zou and Qingxue Zhang, *Senior Member, IEEE*

***Abstract*— Brain visual dynamics encode rich functional and biological patterns of the neural system, and if decoded, are of great promise for many applications such as intention understanding, cognitive load quantization and neural disorder measurement. We here focus on the understanding of the brain visual dynamics for the Amyotrophic lateral sclerosis (ALS) population, and propose a novel system that allows these so-called 'lock-in' patients to 'speak' with their brain visual movements. More specifically, we propose an intelligent system to decode the eye bio-potential signal, Electrooculogram (EOG), thereby understanding the patients' intention. We first propose to leverage a deep learning framework for automatic feature learning and classification of the brain visual dynamics, aiming to translate the EOG to meaningful words. We afterwards design and develop an edge computing platform on the smart phone, which can execute the deep learning algorithm, visualize the brain visual dynamics, and demonstrate the edge inference results, all in real-time. Evaluated on 4,500 trials of brain visual movements performed by multiple users, our novel system has demonstrated a high eye-word recognition rate up to 90.47%. The system is demonstrated to be intelligent, effective and convenient for decoding brain visual dynamics for ALS patients. This research thus is expected to greatly advance the decoding and understanding of brain visual dynamics, by leveraging machine learning and edge computing innovations.**

***Index Terms*— Brain visual dynamics, deep learning, edge inference, amyotrophic lateral sclerosis, electrooculography, lock-in syndrome.**

I. Introduction

HEALTHCARE innovations are everlastingly attractive to the whole society [1]. Nowadays, smart health technologies are becoming a fast-evolving interdisciplinary area where scientific theories, mathematical tools, computing, informatics, and engineering are quickly fusing. Many practices are targeting long-existing healthcare associated challenges, from life-assistive technologies, intelligent home units, to vital signs tracking [2]–[5]. In this study, we focus on brain visual dynamics that encode rich functional and biological patterns of the neural system, and if decoded, is of great promise for many applications such as intention understanding, cognitive load quantization and neural disorder measurement.

Decoding brain visual dynamics is an important and promising technology for Amyotrophic lateral sclerosis (ALS) patients [6]. ALS is a severe neural disease which causes irreversible body degeneration that breaks the neurons in brain and spinal cord. As the connections are destroyed progressively, the patients gradually lose the capability of walking, grasping, eating, talking, and even breathing. Research shows than 90% of the ALS cases may not be caused by gene inheritance, so it can potentially happen on anyone. Every 90 minutes, there is one new patient diagnosed. Currently, more than 450,000 patients are suffering from this disease worldwide and their average life expectancy may even be within several years [6]. Even worse, most patients may have difficulty to afford the expensive healthcare.

Machine learning, especially deep learning, has been advancing intelligent data analytics dramatically, leveraging its capability to reveal nonlinear, complex, and time-varying patterns [7], [8]. On the other hand, new generation of easily accessible smart hardware, especially, the smart phone, allows the deep learning-empowered software to bring the true "intelligence" to the edge, bringing us the co-called real-time edge computing. With the advancement of intelligence, functionality, usability, understandability, and aesthetic of products, smart technologies are playing increasingly important roles in neural dynamics decoding.

In this study, we bridge deep learning, edge computing and brain visual dynamics decoding, aiming to enable a real-time deep learning inference system for neural dynamics understanding, as shown in Fig. 1. ALS patients, fortunately, may still be able to move their eyes, which motivates us to design a system that can understand their brain visual dynamics. Humans reflect their emotions, thoughts, and intentions in visual movements. Therefore, based on the logic of handwriting, we can let the patients to eye-write meaningful words, which, if decoded, can indicate the intension of ALS patients and reduce the suffering from the 'locked-in' condition.

There are some previous studies reported on visual movement analysis, which are usually based on cameras. The camera-based methods [9]–[11] capture the visual movement videos and then analyze the pupil trajectories visual movement understanding. Kate *et al.* reported a screen-based visual-typing system with a virtual keyboard for eye typing [12].

Manuscript received 13 April 2022; revised 26 June 2022; accepted 5 July 2022. Date of publication 25 July 2022; date of current version 11 August 2022. *(Corresponding author: Qingxue Zhang.)*

Jiadao Zou is with the Department of Electrical and Engineering, School of Engineering, Purdue University, Indianapolis, IN 46202 USA.

Qingxue Zhang is with the Department of Electrical and Computer Engineering and the Department of Biomedical Engineering, School of Engineering, Purdue University, Indianapolis, IN 46202 USA (e-mail: qxzhang@purdue.edu).

Digital Object Identifier 10.1109/TNSRE.2022.3193714

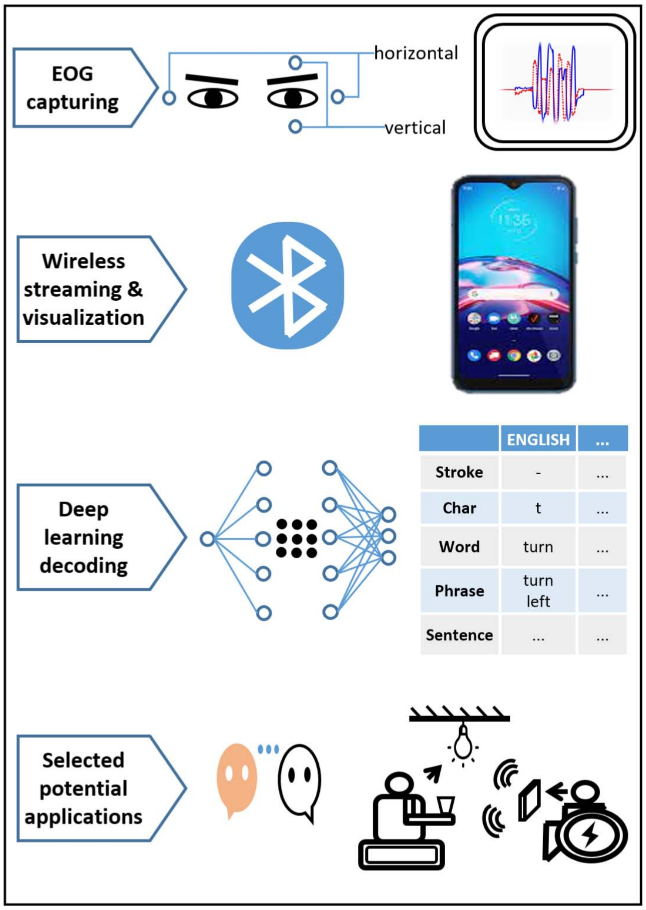

Fig. 1. The proposed eyeSay system for decoding brain visual dynamics with deep learning and edge computing. Note. EOG: Electrooculography, generated during visual movements.

Ward *et al.* designed a visual-tracker based on a screen keyboard called 'Dasher' [13]. Raudonis *et al.* dedicated a visual-tracker-based text-writing program with dimension reduction and neural networks for disabled people [14]. Ishguro *et al.* developed an image descriptor for gaze recognition [15]. Lupu *et al.* designed a camera-based visual decoding system with possible words and sentences to be chosen and displayed on the display [16]. Ogaki *et al.* proposed a method to decode the visual movements captured by a smart and convenient camera system, with both gaze signal and video frames analytics [17]. Some low-resolution cameras have also been used in replace of the expensive, high quality camera with acceptable accuracy [18]. Zdarsky *et al.* applied multi-layer perceptron with stochastic gradient descent for video-based eye tracking analysis [19]. Eye-tracking techniques have also been applied to fatigue measurement [20], mental tracking [21], [22], attention monitoring [23], [24]. These studies have advanced the field forward, nevertheless, camera-based methods are usually constrained by the environmental light intensities, and privacy concerns.

Some other studies focus on the physiological signal-based methods. The common modalities include Electrooculogram (EOG), Electroretinogram (ERG) and Electroencephalogram (EEG). EOG is the bio-potential difference between cornea and ocular fundus, as shown in Fig. 1, which changes during visual movements [25]. ERG usually needs the electrodes to be directly on eyes or very close [26], [27]. Another modality is EEG, which directly reflects the brain's intention [28], [29]. Paul *et al.* tested a wireless EEG monitor in several cognitive tasks and signal processing methods used include discrete wavelet transform, dimension reduction and support vector machine [30]. This study is focusing on the eye EOG dynamics, targeting the long-term real-time application scenarios.

Barea *et al*. developed an EOG-controlled wheelchair [31]. Xiao *et al.* proposed a similar work using a screen for displaying an a keyboard to be selected by blink [32]. Huang *et al.* have reported blink and head rotation-controlled robot system [33]. Another eye selecting/typing system was developed by Heo *et al*., which included two virtual eye-typing keyboards and one virtual wheelchair controller [34]. These systems usually need an input screen that could limit the input speed and deployment. Another category of eye-writing is based on handwritten-style writing. Different algorithms have been proposed to decode eye signals, such as Hidden Markov Model, and Dynamic Time Warping (DTW) method by Fang *et al.* [35], [36]. Ding *et al.* applied also DTW for character recognition [37]. Chang *et al.* proposed to use an ensemble deep neural network (DNN) with inception modules for eye-writing [38]. Kang *et al.* applied an ensemble network with attention mechanisms [39]. Pérez-Reynoso *et al.* developed an EOG-controlled robotic system [40]. However, gap still exists in how deep learning can further decode complex patterns and how real-time inference can be deployed.

Targeting the challenges of brain visual dynamics decoding, we propose to design and develop a novel system by leveraging both deep learning and real-time edge inference, for intelligent, real-time, and unobtrusive decoding intensions of ALS patients. More specifically, we first propose to leverage a deep learning framework for automatic feature learning and classification of the brain visual dynamics, aiming to translate the eye EOG to meaningful words. We afterwards design and develop an edge computing platform on the smart phone, which can execute the deep learning algorithm, visualize the brain visual dynamics, and demonstrate the edge inference results, all in real-time. Evaluated on 4,500 trials of brain visual movements performed by six users, our system has been demonstrated to be intelligent, easy-to-use, user-friendly, and effective for decoding brain visual dynamics for ALS patients.

Our contributions are summarized below:

1) Design and develop a novel system, eyeSay, which can, in real-time, stream, visualize, and decode the brain visual dynamics for neural dynamics understanding. To the best of our knowledge, this is the first system that leverages both deep learning and edge inference for seamless brain visual dynamics understanding and demonstration, targeting the ALS patient population.
2) Propose a multi-stage deep learning approach that can learn complex and highly non-linear dynamics hidden in the brain visual dynamics, thereby enabling efficient learning from scarce data.
3) Design and develop an edge computing platform, with end-to-end data streaming, visualization and deep

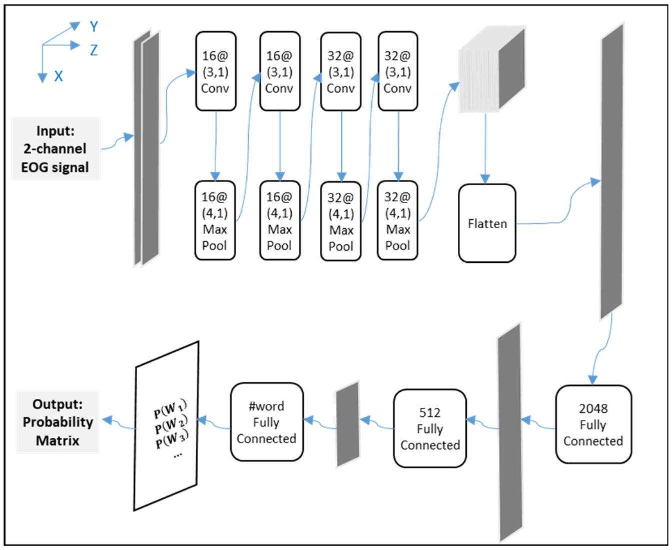

Fig. 2. The deep convolutional neural network (CNN) architecture for brain visual dynamics decoding, which includes convolutional layers for hierarchical feature abstraction, max pooling layers for dimension reduction, and fully connected layers for final inference. Notes. COV: convolution; Max Pool: max-pooling; P(·⊲: probability of the eye-written word.

learning-based inference, thereby enabling a real-time and user-friendly interaction method for ALS patients.

4) Evaluate the eyeSay system on real-world experiments and demonstrate its feasibility and potential, which indicates how machine learning and edge computing can advance neural dynamics decoding.

This research thus is expected to greatly advance the decoding and understanding of brain visual dynamics, by leveraging machine learning and edge computing innovations. We below detail the approaches, results and finally conclude the study.

II. Approaches

A. System Architecture

As shown in Fig. 1, the eyeSay system can, not only wirelessly stream and visualize the eye EOG signal, but also perform deep learning inference on the smart phone for real-time brain visual dynamics decoding. We here will detail the deep learning approach, the edge inference, wireless streaming, and real-time visualization, respectively, to demonstrate the design principles and considerations of the proposed system.

B. Deep Learning of Brain Visual Dynamics Behind EOG

We have designed a convolutional neural network (CNN), as shown in Fig. 2, for brain visual dynamics decoding. Comparing with traditional multilayer perceptron, CNN has fewer parameters to be learned through shared convolutional filters for efficient and effective learning. We here treat the 2-channel EOG signal, both horizontal and vertical, as an image and feed it into CNN for hierarchical pattern extraction.

The proposed multi-stage CNN contains four convolutional layers (COV) and four max pooling layers (MP) for spatial motif learning and dimension reduction, respectively [41]. We choose a small COV kernel size considering that, a model with multiple lighter COV kernels can achieve similar efficiency as the one with fewer but heavier COV kernels, while the earlier method has fewer parameters [42]. In the fully connected neural layers, every unit in one layer connects with each unit in its previous layer, thereby summarizing previous extracted partial features and generating higher-level abstraction. In the end, the model outputs a probability matrix for each eye-written word with the sum to one, and the output node with the highest probability is determined to be the final recognition result.

The non-linearity in our CNN architecture is introduced by an activation function called Rectifier Linear Unit (*ReLU*), which is a non-constant, monotone-increasing continuous function whose gradient is bounded within 0 and 1 [43]. *ReLU* sets the gradient no greater than zero to be zero, to effectively depress gradient vanishing by simplifying the computation of backpropagation. *ReLU* is used for all COV and fully connected layers, except the output layer in which the *Softmax* activation function is used for class probability generation.

C. Multi-Class Deep Learning Process

Since the eye-writing task will have only one high-confident output, we transfer the sample's labels into the on-hot encodings. The *Softmax* activation function in the last dense layer is given in (1), where z_i is the weighted sum of neurons in the previous layer, and the denominator is a normalization factor to make sure the sum of all output neural nodes is one. The CNN optimization object is defined as the categorical cross-entropy as (2), where P_i is the ground truth distribution of a given class i among all C classes, and y_i is the generated probability of the output neuron node i, i.e., $Softmax\,(z_i)$. With one-hot encoding, (2) is reduced to (3), where all other classes except the correct class p are neglected.

$$Softmax\,(z_i) = \frac{e^{z_i}}{\sum_{j \in allclasses} e^{z_j}} \tag{1}$$

$$CrossEntropy = -\sum_{i=1}^{C} P_i \log\left(Softmax\,(y_i)\right) \tag{2}$$

$$CrossEntropy_{one-hot} = -log\left(Softmax\,(y_p)\right) \tag{3}$$

During the backpropagation-based neural parameter updating, the gradient based on the cross-entropy loss is determined by (4) and (5), for the ground truth output node p and any other output node n, respectively. Further, the loss is backpropagated to all previous layers, to adjust the neural parameters for better pattern abstraction.

$$\frac{\partial}{\partial y_p}\left(-log\left(Softmax\left(y_p\right)\right)\right) = Softmax\left(y_p\right) - 1 \tag{4}$$

$$\frac{\partial}{\partial y_n}\left(-log\left(Softmax\left(y_p\right)\right)\right) = Softmax\left(y_n\right) \tag{5}$$

D. Edge Inference

The proposed edge computing architecture is given in Fig. 3. The edge inference means letting peripheral device(s)

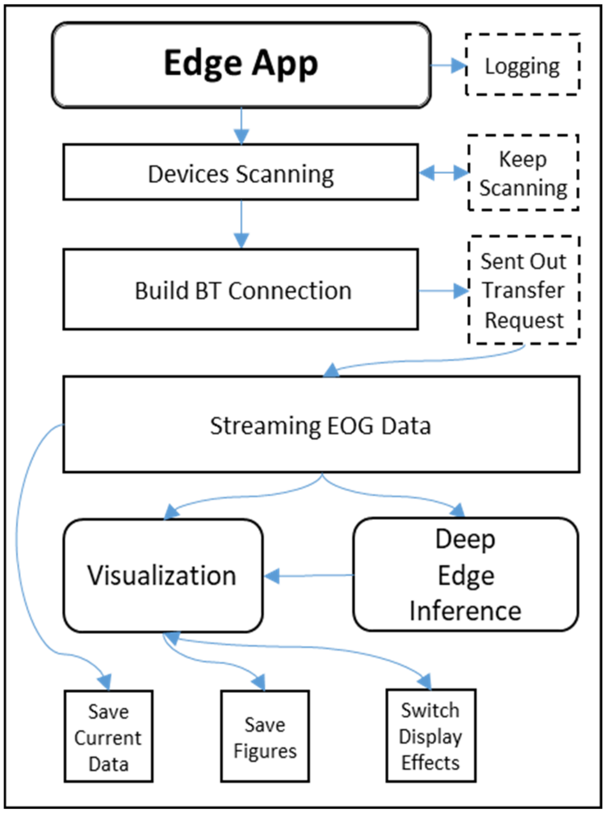

Fig. 3. The proposed edge computing platform that consists of wireless streaming (BT: Bluetooth), real-time brain visual dynamics visualization, real-time deep inference on the edge, and real-time demonstration of decoding results.

run the deep learning tasks instead of uploading them to the cloud, which yields multiple advantages. (1) It can avoid large latency and provide real-time data analysis ability. Unlike traditional cloud computing where every task should be performed on central servers, edge computing schedules tasks on the peripheral side. (2) Less data transmission to the cloud minimizes the chance of data disclosure and privacy leakage. Otherwise, large amounts of data need to be encrypted for transmission to the cloud if the edge device cannot handle the sensible tasks locally.

Edge computing platform design and develop requires optimization on multiple design dimensions. To achieve efficient edge deployment, we have firstly leveraged CNN for efficient data learning, considering CNN uses shared convolutional filters that have relatively low computation requirements, compare to the fully connected neural network. Secondly, we have transformed the learned CNN to a more lightweight deep learning package, which can be efficiently executed on the smart phone [44]. Further, we directly feed the EOG signals to CNN without complicated data processing steps and let the powerful CNN to analyze the raw data for deep pattern abstraction.

One thing to further note is the TensorFlow Lite library. It is an open-source machine learning library specifically optimized for mobile and edge computing applications. It can make the edge inference lightweight, low latency, and optimal for power consumption. We have designed and trained the deep learning architecture firstly, and afterwards used TensorFlow Lite to generate the software component that is then integrated to our mobile APP.

E. Wireless Data Streaming

The eyeSay application builds wireless data streaming upon Bluetooth Low Energy (BLE) module provided by Android. With BLE, the application greatly saves the battery consumption by automatically switching to power-saving mode depending on current streaming status.

The mobile APP we have developed keeps scanning for available devices until they are paired. After establishing the connection, the EOG data is then continuously streamed to the mobile APP. The data is segmented appropriately to accommodate the BLE transmission characteristics. The real-time data received is then visualized on the phone and fed into the edge inference module for eye-written word recognition.

F. Visualization

To provide real-time feedback to the users, the EOG signals received on the phone are visualized on the top half of the screen. The visualizer module applies smooth animation for the flow of data in a comfortable refreshing rate. In addition, it also supports zooming, highlighting, screen interactions, various animating effects and other visualization enhancement, allowing the caregiver to understand the scenarios. As the only part users would directly interact with, the design principle of this module is clear and easy-to-use.

Furthermore, we put another visualization module that can immediately demonstrate the edge inference results on the lower screen. In such a way, both the patient and caregiver can get the decoded eye-written words in real-time. The overall edge computing platform, with the wireless transmission, edge interference, and visualization functions, is expected to greatly advance the voice-free communication application for ALS patients.

III. Results

In this section, we demonstrate the comprehensive evaluation of the system we have proposed and built.

A. Experimental Setup

The application is designed for devices running Android 9.0 and later. A GPU is not necessary. We have tested it on the Samsung Galaxy Note 10+ (SM-975U) and Samsung S8+ (SM-G955F). The application would require functional permissions including Bluetooth, as well as location and external file access. The first two permissions are required by BLE API and the last one is used by edge device module for saving screenshot or received data.

To demonstrate the effectiveness of the proposed eyeSay system, we have applied an eye-writing database, which includes 4,500 eye-writing trials: 150 words and 5 repeats from 6 subjects. It is called Japanese Katakana database [35]. It covers common human names, actions and some adjective

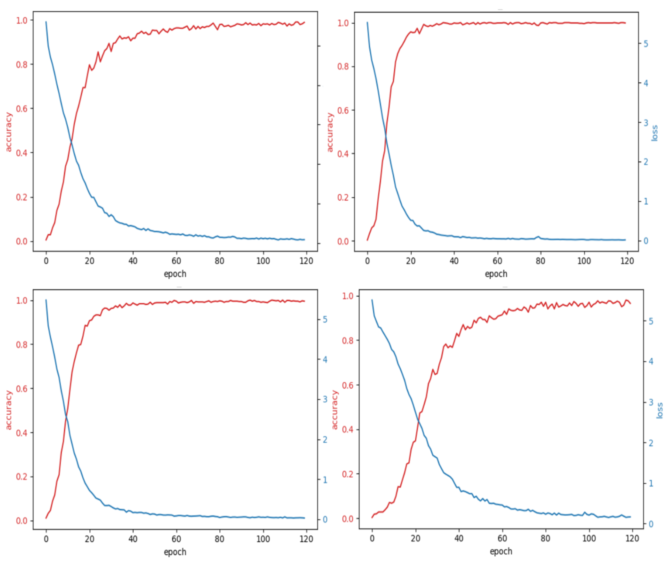

Fig. 4. The deep learning process for difference cases, indicating the effective convergence trend.

nouns, which are suitable for voice-controlling the phone calls as well as simple phone operations.

One thing worth noting is that the system can be easily generalized to other languages, such as English word recognition, by adapting current deep learning model to the new eye-writing dynamics. Another thing to mention is that, to test the wireless streaming function, we use a simulation device, called sender, to transmit the EOG data continuously to the phone, called receiver, which executes visualization and inference. A monitor can be used in future as the sender. But the overall experimental setup is effective enough now to demonstrate the proposed functions of the system.

B. Deep Model Learning

Our deep CNN model uses a gradient-based method, Adam, to automatically adjust the model parameters [45]. Adam is a fast first-order gradient and moment optimizer, which needs very little tuning and can robustly converge for either stationary or non-stationary problem.

We applied leave-one-trial strategy for each subject in the evaluation. Fig. 4 gives the training loss and accuracy for difference cases, indicating the effectiveness of the learning process. To be visualized in the next Edge Inference section, both intra-subject and inter-subject signal variability is high, making EOG decoding very challenging. However, the learning curves in Fig. 4 all converge well and thus demonstrate the deep CNN model has learned effective patterns from the EOG signals.

C. Wireless Data Streaming

In the system, the sender transmits data to the edge receiver, e.g. a smart phone, for both visualization and decoding. Once the connection between the sender and the receiver is built, the sender calls the data fragmentation to and sends them over BLE in sequence. Meanwhile, it indexes every single piece of data to avoid disorder. After the edge receiver gets the starting notification from the sender, it begins to visualize and decode the data.

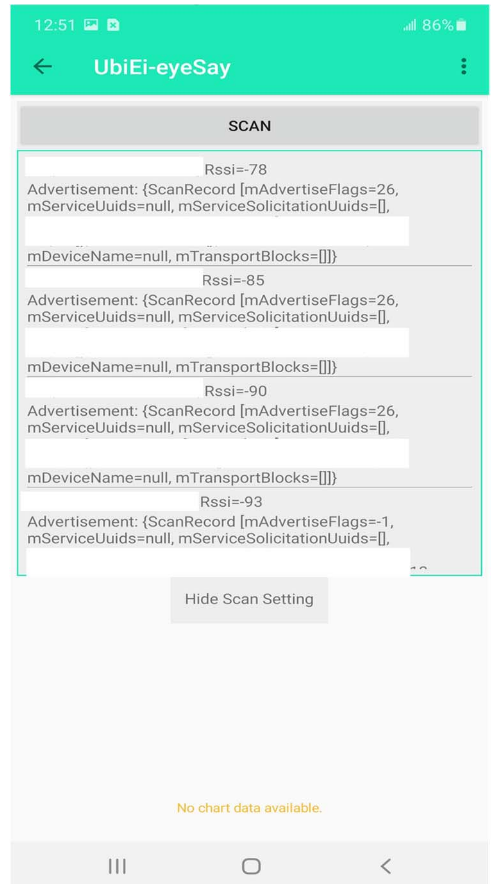

Fig. 5. The wireless streaming function on the edge platform, showing the edge APP is scanning for devices to connect.

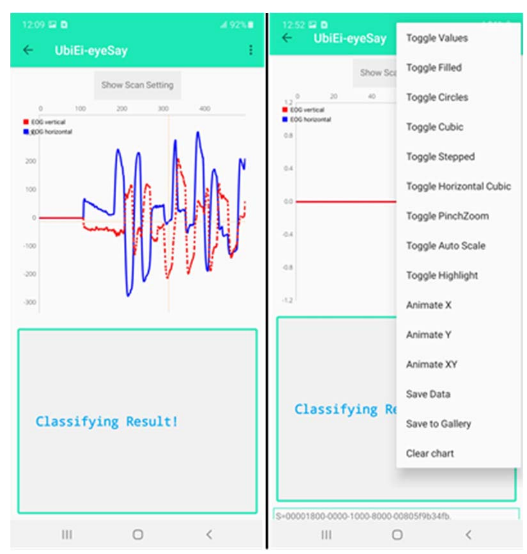

Fig. 6. The visualization function of the edge APP, showing the real-time EOG signal and providing many visualization adjustments options for the caregiver.

Once the connection is built, the sender would automatically sample EOG data at 50 Hz and send data packet by packet.

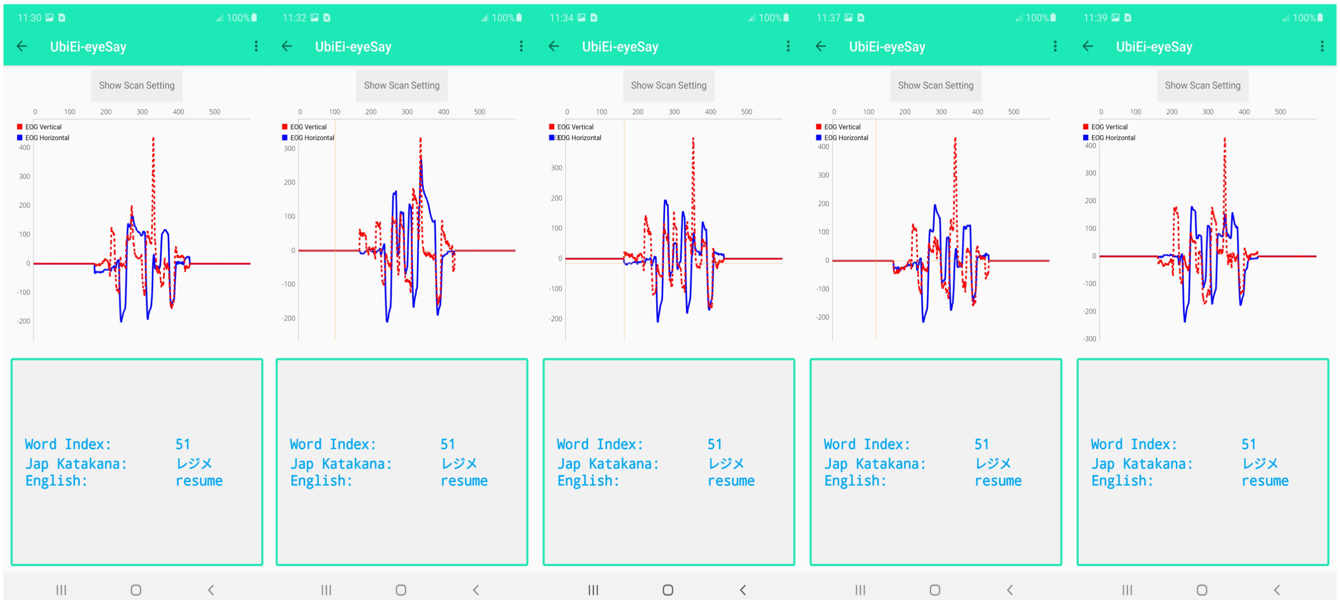

Fig. 7. The EOG signal and edge inference results for the same subject and same word, but different trials. There are two major observations: (1) there is a high intra-subject inter-trial variability that poses challenges to deep decoding; (2) the edge inference model has successfully decoded these signals with high variability.

Fig. 8. The EOG signal and edge inference results for the same trial and same word, but different subjects. There is a very high inter-subject variability, indicating the biological and behavioral differences among subjects. The decoding results also indicate very robust word recognition.

For the receiver's APP, there is an expandable drawer to set up scanning configurations, as shown in Fig. 5. It contains a list of nearby devices including their names, unique identities, physical addresses, advertisements, manufacturing information, and transmission power indicators. Clicking a device in the list would have current edge device to establish pairing.

D. Edge Visualization of Incoming Data

The visualizing chart is a 2D line chat of time and magnitude information, as shown in Fig. 6. Two-channel ECG is visualized simultaneously, corresponding to the horizontal and vertical EOG, respectively. Besides, the chart leverages animation to smoothly update itself with the new data coming in. Meanwhile the chart keeps its center focusing on the latest data point and automatically scale the axis for a better view. Furthermore, there are many other display effects to use, e.g., showing referencing values of the lines, filling the downside of curves, marking or highlighting data points, disabling gesture control, and imitating the incoming of dataset.

E. Edge Inference – Same Subject & Same Word

The edge device, i.e., the receiver, handles deep learning tasks through TensorFlow Lite, a light-weight framework optimized for energy efficiency. We have thoroughly illustrated the effectiveness of the edge inference function in Fig. 7 to 9.

On the main screen of the edge APP, the upper half visualizes the received EOG data, and the lower page displays the

Fig. 9. The EOG signal and edge inference results for different subjects, trials, and words. This comparison further demonstrates the variability among the EOG signals, which make the deep learning even more challenging. But the proposed deep CNN model on the edge can still robustly decode the eye-written words.

real-time inference. The deep learning classifier begins to work when the assembled input meets the dimension request of the model. Furthermore, the edge APP keeps updating the results with a refresh interval of 0.5 second till the data transmission is over. Experiment shows that eyeSay can successfully decode the long, continuous and complex EOG signal to Katakana words for each user.

In Fig. 7, we have illustrated the EOG signal and edge inference results for the same subject and same word, but different trials. There are two major observations: (1) there is a high intra-subject inter-trial variability that poses challenges to deep decoding; (2) the edge inference model has successfully decoded these signals even with high variability. When the same subject writes the same word for several times, there may still be significant variability, due to the high degree of freedom of the eye movements. But our proposed deep learning model can still robustly decode the highly complex and time-varying eye dynamics.

F. Edge Inference – Different Subjects & Same Word

In Fig. 8, we have further demonstrated the EOG signal and edge inference results for the same trial and same word, but different subjects. Obviously, there is a very high inter-subject variability, indicating the biological and behavioral differences among subjects. The decoding results also indicate very robust word recognition.

G. Edge Inference – Different Subjects & Different Words

Fig. 9 gives EOG signal and edge inference results for cases of different trials, different words, and different subjects. When comparing the EOG visualizations, we can find the huge inter-subject and inter-word difference. Though, the decoding results indicate very robust word recognition again.

Overall, our experimental results on 4,500 eye-writing trials and a detection accuracy of 90.47%, have clearly demonstrated the feasibility and effectiveness of the proposed eyeSay system.

H. Clinical Applications and Future Study

It is promising to apply the system for ALS patient applications to enable voice-free communications through visual dynamics decoding, and further, it is promising to apply the system in cognitive quantization and neurodegeneration disease evaluations. In future, it will be interesting to conduct more experiments and data to train the deep learning algorithm for robust data analytics in diverse scenarios. Besides, it will be promising to introduce recurrent learning algorithms [46] to the framework to further learn the temporal patterns.

IV. Conclusion

In this study, we have proposed eyeSay, an innovative system empowered both deep learning and edge computing for brain visual dynamics decoding, targeting the ALS patients. We have developed a multi-stage deep CNN model to dynamically decode the eye-generated EOG signals into meaningful words. We have further designed and developed an edge computing platform, to wirelessly stream, visualize, and decode eye EOG in real-time. Evaluated on 4,500 eye-written trials performed by multiple users, our novel system has demonstrated a recognition rate up to 90.47%. The system is demonstrated to be intelligent, effective, and convenient for enabling human-computer interaction for ALS patients. Besides, the novel system can be easily generalized to other languages. This research thus is expected to greatly advance the understanding of the neural system and dynamics decoding, by leveraging machine learning and edge computing innovations.

References

[1] Q. Zhang, D. Arney, J. M. Goldman, E. M. Isselbacher, and A. A. Armoundas, "Design implementation and evaluation of a mobile continuous blood oxygen saturation monitoring system," *Sensors*, vol. 20, no. 22, p. 6581, Nov. 2020.

[2] H. Thapliyal, "Internet of Things-based consumer electronics: Reviewing existing consumer electronic devices, systems, and platforms and exploring new research paradigms," *IEEE Consum. Electron. Mag.*, vol. 7, no. 1, pp. 66–67, Jan. 2018.

[3] P. Datta and B. Sharma, "A survey on IoT architectures, protocols, security and smart city based applications," in *Proc. 8th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jul. 2017, pp. 1–5.

[4] A. Mosenia, S. Sur-Kolay, A. Raghunathan, and N. K. Jha, "Wearable medical sensor-based system design: A survey," *IEEE Trans. Multi-Scale Comput. Syst.*, vol. 3, no. 2, pp. 124–138, Apr./Jun. 2017.

[5] Y. A. Qadri, A. Nauman, Y. B. Zikria, A. V. Vasilakos, and S. W. Kim, "The future of healthcare Internet of Things: A survey of emerging technologies," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1121–1167, 2nd Quart., 2020.

[6] V. Neuro. *ALS Incidence*. Accessed: 2022. [Online]. Available: https://alstreatment.com/amyotrophic-lateral-sclerosis-incidence/

[7] Q. Zhang *et al.*, "IEEE access special section editorial: Smart health sensing and computational intelligence: From big data to big impacts," *IEEE Access*, vol. 9, pp. 30452–30455, 2021.

[8] Q. Zhang, V. Piuri, E. A. Clancy, D. Zhou, T. Penzel, and W. W. Hu, "IEEE access special section editorial: Advanced information sensing and learning technologies for data-centric smart health applications," *IEEE Access*, vol. 9, pp. 30404–30407, 2021.

[9] T. Rajpathak, R. Kumar, and E. Schwartz, "Eye detection using morphological and color image processing," in *Proc. Florida Conf. Recent Adv. Robot.*, 2009, pp. 1–6.

[10] K. Arai and R. Mardiyanto, "Autonomous control of eye based electric wheel chair with obstacle avoidance and shortest path findings based on Dijkstra algorithm," *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 12, pp. 19–25, 2011.

[11] P. Xu, K. A. Ehinger, Y. Zhang, A. Finkelstein, S. R. Kulkarni, and J. Xiao, "TurkerGaze: Crowdsourcing saliency with webcam based eye tracking," 2015, *arXiv:1504.06755*.

[12] J. Ten Kate, E. Frietman, F. Stoel, and W. Willems, "Eye-controlled communication aids," *Med. Prog. Through Technol.*, vol. 8, no. 1, pp. 1–21, 1980.

[13] D. J. Ward and D. J. MacKay, "Fast hands-free writing by gaze direction," *Nature*, vol. 418, no. 6900, p. 838, 2002.

[14] V. Raudonis, R. Simutis, and G. Narvydas, "Discrete eye tracking for medical applications," in *Proc. 2nd Int. Symp. Appl. Sci. Biomed. Commun. Technol.*, Nov. 2009, pp. 1–6.

[15] Y. Ishiguro, A. Mujibiya, T. Miyaki, and J. Rekimoto, "Aided eyes: Eye activity sensing for daily life," in *Proc. 1st Augmented Hum. Int. Conf.*, 2010, pp. 1–7.

[16] R. G. Lupu, R. G. Bozomitu, F. Ungureanu, and V. Cehan, "Eye tracking based communication system for patient with major neoro-locomotor disabilites," in *Proc. 15th Int. Conf. Syst. Theory, Control Comput.*, Oct. 2011, pp. 1–5.

[17] K. Ogaki, K. M. Kitani, Y. Sugano, and Y. Sato, "Coupling eye-motion and ego-motion features for first-person activity recognition," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2012, pp. 1–7.

[18] M. Tonsen, J. Steil, Y. Sugano, and A. Bulling, "InvisibleEye: Mobile eye tracking using multiple low-resolution cameras and learning-based gaze estimation," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 3, pp. 1–21, 2017.

[19] N. Zdarsky, S. Treue, and M. Esghaei, "A deep learning-based approach to video-based eye tracking for human psychophysics," *Frontiers Human Neurosci.*, vol. 15, Jul. 2021.

[20] J. Xu, J. Min, and J. Hu, "Real-time eye tracking for the assessment of driver fatigue," *Healthcare Technol. Lett.*, vol. 5, no. 2, pp. 54–58, 2018.

[21] N. Neshov and A. Manolova, "Drowsiness monitoring in real-time based on supervised descent method," in *Proc. 9th IEEE Int. Conf. Intell. Data Acquisition Adv. Comput. Syst., Technol. Appl. (IDAACS)*, vol. 2, Sep. 2017, pp. 660–663.

[22] P. J. Murphy, A. L. Duncan, A. J. Glennie, and P. C. Knox, "The effect of scleral search coil lens wear on the eye," *Brit. J. Ophthalmol.*, vol. 85, no. 3, pp. 332–335, 2001.

[23] S. I. Tonni, T. A. Aka, M. M. Antik, K. A. Taher, M. Mahmud, and M. S. Kaiser, "Artificial intelligence based driver vigilance system for accident prevention," in *Proc. Int. Conf. Inf. Commun. Technol. Sustain. Develop. (ICICT4SD)*, Feb. 2021, pp. 412–416.

[24] E. M. Bearly and R. Chitra, "Automatic drowsiness detection for preventing road accidents," in *Proc. Smart Technol., Commun. Robot. (STCR)*, Oct. 2021, pp. 1–5.

[25] I. Perlman, "The electroretinogram: ERG," in *Webvision: The Organization of the Retina and Visual System [Internet]*. Salt Lake City, UT, USA: Univ. of Utah Health Sciences Center, 2007.

[26] K. T. Brown, "The electroretinogram: Its components and their origins," *Vis. Res.*, vol. 8, no. 6, pp. 633–677, 1968.

[27] R. Yin *et al.*, "Soft transparent graphene contact lens electrodes for conformal full-cornea recording of electroretinogram," *Nature Commun.*, vol. 9, no. 1, pp. 1–11, Dec. 2018.

[28] C. Guger, G. Krausz, B. Z. Allison, and G. Edlinger, "Comparison of dry and gel based electrodes for P300 brain–computer interfaces," *Frontiers Neurosci.*, vol. 6, p. 60, May 2012.

[29] M. Spüler, "A high-speed brain–computer interface (BCI) using dry EEG electrodes," *PLoS ONE*, vol. 12, no. 2, Feb. 2017, Art. no. e0172400.

[30] S. K. Paul, M. Z. Nine, M. Hasan, and M. A. Amin, "Cognitive task classificaiton from wireless EEG," in *Proc. Int. Conf. Brain Inform. Health*. London, U.K.: Springer, 2015, pp. 13–22.

[31] R. Barea, L. Boquete, M. Mazo, and E. Lopez, "System for assisted mobility using eye movements based on electrooculography," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 10, no. 4, pp. 209–218, Dec. 2002.

[32] J. Xiao, J. Qu, and Y. Li, "An electrooculogram-based interaction method and its music-on-demand application in a virtual reality environment," *IEEE Access*, vol. 7, pp. 22059–22070, 2019.

[33] F. Wang, X. Li, and J. Pan, "A human-machine interface based on an EOG and a gyroscope for humanoid robot control and its application to home services," *J. Healthcare Eng.*, vol. 2022, pp. 1–14, Mar. 2022.

[34] J. Heo, H. Yoon, and K. S. Park, "A novel wearable forehead EOG measurement system for human computer interfaces," *Sensors*, vol. 17, no. 7, p. 1485, 2017.

[35] F. Fang and T. Shinozaki, "Electrooculography-based continuous eye-writing recognition system for efficient assistive communication systems," *PLoS ONE*, vol. 13, no. 2, Feb. 2018, Art. no. e0192684.

[36] F. Fang, T. Shinozaki, Y. Horiuchi, S. Kuroiwa, S. Furui, and T. Musha, "Improving eye motion sequence recognition using electrooculography based on context-dependent HMM," *Comput. Intell. Neurosci.*, vol. 2016, pp. 1–9, 2016.

[37] X. Ding and Z. Lv, "Design and development of an EOG-based simplified Chinese eye-writing system," *Biomed. Signal Process. Control*, vol. 57, Mar. 2020, Art. no. 101767.

[38] W.-D. Chang, J.-H. Choi, and J. Shin, "Recognition of eye-written characters using deep neural network," *Appl. Sci.*, vol. 11, no. 22, p. 11036, Nov. 2021.

[39] D.-H. Kang and W.-D. Chang, "Recognition of eye-written characters with limited number of training data based on a Siamese network," *Electronics*, vol. 10, no. 23, p. 3009, Dec. 2021.

[40] F. D. Pérez-Reynoso, L. Rodríguez-Guerrero, J. C. Salgado-Ramírez, and R. Ortega-Palacios, "Human–machine interface: Multiclass classification by machine learning on 1D EOG signals for the control of an omnidirectional robot," *Sensors*, vol. 21, no. 17, p. 5882, 2021.

[41] J. Zou and Q. Zhang, "eyeSay: Eye electrooculography decoding with deep learning," presented at the 39th IEEE Int. Conf. Consum. Electron., 2020.

[42] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. pattern Recognit.*, Jun. 2015, pp. 1–9.

[43] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. ICML*, 2010, pp. 1–8.

[44] Google. *TensorFlow Lite*. Accessed: 2022. [Online]. Available: https://www.tensorflow.org/lite

[45] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[46] K. Cho *et al.*, "Learning phrase representations using RNN encoder–decoder for statistical machine translation," 2014, *arXiv:1406.1078*.