## 33.2 A Sub-1μJ/class Headset-Integrated Mind Imagery and Control SoC for VR/MR Applications with Teacher-Student CNN and General-Purpose Instruction Set Architecture

Zhiwei Zhong*, Yijie Wei*, Lance Christopher Go, Jie Gu

Northwestern University, Evanston, IL
*Equally Credited Authors (ECAs)

2024 IEEE International Solid-State Circuits Conference (ISSCC) | 979-8-3503-0620-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISSCC49657.2024.10454317

Virtual Reality (VR) and Mixed Reality (MR) systems, e.g., Meta Quest and Apple Vision Pro, have recently gained significant interest in consumer electronics, creating a new wave of developments in metaverse for gaming, social networking, workforce assistance, online shopping, etc. Strong technological innovations in AI computing and multi-modular human activity tracking and control have produced immersive virtual realistic user experiences. However, most existing VR headsets only rely on traditional joysticks or camera-based user gestures for input control and human tracking, missing an important source of information, namely, brain activity. Hence there is a growing interest in incorporating brain-machine interfaces (BMIs) into VR/MR systems for consumer and clinical applications [1]. As illustrated in Fig. 33.2.1, an existing VR/MR system integrated with EEG channels typically consists of a VR headset, a 16/32-channel EEG cap, a neural recording analog frontend, and a PC for signal classification. Major drawbacks of such systems include: (1) cumbersome wear and poor user appearance, (2) lack of *in situ* computing support for low-latency operation, (3) inability for real-time mind imagery control and feedback based on brain activity, (4) high power consumption due to AI classification. To overcome these challenges, this work introduces a mind imagery device integrated into existing VR headsets without extra wearing burden for mind-controlled BMI for a VR/MR system. The contributions of this work include: (1) an SoC supporting *in situ* mind imagery control for VR/MR systems, (2) seamless integration with existing VR headset and optimized selection of EEG channels to enhance user acceptance and experience, (3) a general-purpose instruction set architecture (ISA) with flexible dataflow, supporting a broad range of mind imagery operations, (4) a confusion-matrix-guided teacher-student CNN scheme to save power during AI operations, (5) sparsity enhancement on EEG signals to reduce energy. A 65nm SoC test chip is fabricated with *in situ* demonstrations on various mind imagery-based VR controls. While prior works address EEG-based seizure detection or similar biomedical applications [2-6], this work focuses on emerging BMI in a VR/MR environment. The digital core of the SoC achieves an energy consumption <1μJ/class for compute-intensive CNN operations thanks to the low-power features and system-level optimizations of the design.

Figure 33.2.2 shows EEG channel selection and integration into the Meta Quest 2 VR headset with a tradeoff between accuracy and user convenience. To support a variety of mind imagery tasks, 8 EEG channels T3, T5, O1, O2, T6, T4, PZ, and CZ are selected and subtly incorporated into the head-strap to maintain user aesthetics. Different mind tasks activate a subset of the eight selected channels, e.g. T3/T5/CZ/T4/T6 for mental imagery, T5/CZ for affect (e.g., emotion) monitoring, or O1/O2/PZ for steady-state visual evoked potential (SSVEP). The reduction of channels leads to a minor drop in the average accuracy for the three main tasks (from 90.4% to 85.2%) but significantly improves the user experience and usability. Commercial Hydro-link electrodes with saline solution are used to capture EEG signals via pre-cut holes in the headband. Figure 33.2.2 also shows the top-level diagram of the fully integrated SoC. Up to 16 programmable channels of AFE are used for signal acquisition and digitalization. Each channel of the AFE includes a two-stage chopper amplifier with 45-to-72dB gain and 0.05-to-400Hz bandwidth, a low-pass filter with a corner frequency at 60Hz, and an 8b SAR ADC operating from 128Hz to 10kHz. The digital core for integrated AI operations comprises an 8×10 Processing Element (PE) array, control logic, and associated memory banks. An instruction memory with a specially developed ISA provides global control to the chip's operation for supporting a range of mind imagery tasks. The real-time classified brain states and mind-control commands are transmitted to the VR headset via an external Bluetooth module for control of the VR scenes.

While most existing works only focus on a fixed dataflow [4] and CNN model [2,3], a highly flexible computing architecture is required to support a variety of mind imagery tasks. Figure 33.2.3 shows the specially developed general-purpose ISA for dataflow control, model configuration, channel selection, etc. An ultra-wide ISA command of 128b is used to supervise various computing tasks, e.g., IIR filter, Convolutional (Conv) layer, discrete Fourier transform (DFT), and fully connected (FC) layer with high hardware efficiency. To support ever-changing AI models, the configuration of each sub-task, such as the number of kernels, number of layers, branch target address (BTA), sparsity settings, etc., are also integrated into the ISA for efficient scheduling and execution of different tasks. Figure 33.2.3 also shows the detailed architecture of the digital neural processor. The 8×10 PE array can be flexibly turned on or gated off by rows or columns. CNN, FC, DFT, and IIR filtering operations can be specially performed by reusing the same PE array through different dataflows, e.g., weight-stationary for Conv layers, or output-stationary for FC layers and DFT. Instead of the conventional systolic array, which engages significant pipelined flip-flops, this design purposely removes most of the pipeline stages for power saving while still meeting the classification latency target of 5 to 10ms. Low-power features, e.g., sparsity enhancement, fine clock gating, and a teacher-student CNN scheme are also developed as described next.

Given the slow pacing of mental states, a teacher-student CNN scheme is developed to strike a balance between sensitivity, computational power, and accuracy, as depicted in Fig. 33.2.4. Offline-trained teacher-student models are downloaded into the chip for brain activity monitoring. The small-size student model is about 3× faster, with 14% lower accuracy but 70% less energy consumption than the teacher model. Pre-determined user-specific confusion matrices are stored on the chip to judiciously decide which model to be activated for classifications. As shown in Fig. 33.2.4, for the example of affect monitoring, while the initial classification is performed by the teacher CNN model for high accuracy, as user's mental state lasts, the small student CNN model is turned on for power saving. When a state transition is detected, the confusion matrix is checked to evaluate the possibility of true transition or false alarm. A rejection is issued if the confusion matrix shows a high possibility of false detection. The rejection is followed by the engagement of the teacher CNN for confirmation. Essentially, the confusion matrix is used to reduce the false classification rate from the student CNN leading to enhanced overall accuracy for the student model. Experiments show that 55% energy/class saving can be achieved through the teacher-student CNN scheme with an accuracy drop of only 2.3% in the affect detection case. Figure 33.2.4 also shows the proposed sparsity enhancement technique where small noisy signals are zeroed using comparators with a preset threshold. Direct sparse enhancement leads to a significant accuracy drop of over 15%. A special sparsity-aware training process that adds sparsity operation into the training process reduces the accuracy impacts. With the special training flow, a total CNN power saving of 12% is achieved with an accuracy impact of up to 4.6%.

The SoC is fabricated in 65nm CMOS with a total area of 7.5mm² and supports four mind imagery and affect monitoring/control tasks, including: (1) mental imagery-based VR interface control, i.e., users control the GUI operation through the imagination of pictures, (2) real-time affect state tracking and feedback control during VR gaming, (3) motor imagery, i.e., user's imagination of hand or leg motions, (4) SSVEP, i.e., user focuses on pictures flashing with various frequency as input to the system. Figure 33.2.5 shows demonstrations of mental imagery and affect based VR control. In the mental imagery control task, the user issues pre-trained "focus" action to pop up a selection menu and imagines photos of rainy days or surfing to make a selection of the menu items. *In situ* measurement on mental imagery and control shows the mind commands are successfully injected into the VR scenes with an accuracy of 79 to 83%. In affect tracking and control, a customized Endless Running game is built and the gaming difficulty is dynamically adjusted based on the CNN-classified user arousal level to enhance user engagement, e.g., increasing moving speed when a user's arousal level is low. Measurement results show an accuracy above 90% in tracking a gamer's affect states with successful affect-based pace adjustment.

Figure 33.2.6 shows more measurement results. Public SSVEP [8] and motor imagery [9] datasets are evaluated by the on-chip CNN, achieving 86% and 80% accuracy, with 2% drop from a baseline model in the SSVEP dataset [8]. The teacher-student CNN scheme achieves 1.97μJ/class for teacher CNN, 0.6μJ/class for student CNN and 0.89μJ/class for combined operation. Figure 33.2.6 makes a comparison with prior works on biomedical SoCs with integrated digital cores. Thanks to the low-power features, the digital core in this work achieves the state-of-the-art energy consumption at sub-1μJ/class for CNN operation. While prior works mainly focus on medical diagonosis, e.g., seizure detection, this work extends brain-machine-interface technology to consumer electronics for flourishing VR/MR systems. The chip micrograph is shown in Fig. 33.2.7.

*Acknowledgement:*
This work was supported by the National Science Foundation under CNS-1816870 and CCF-2208573.

*References:*

[1] M. Li, et al., "Neurophysiological and Subjective Analysis of VR Emotion Induction Paradigm," *IEEE Trans. Visualization and Computer Graphics*, vol. 28, no. 11, pp. 3832-3842, Nov. 2022.

[2] J. Liu, et al., "4.5 BioAIP: A Reconfigurable Biomedical AI Processor with Adaptive Learning for Versatile Intelligent Health Monitoring," *ISSCC*, pp. 62-63, 2021.

[3] C.-W., Tsai et al., "SciCNN: A 0-Shot-Retraining Patient-Independent Epilepsy-Tracking SoC," *ISSCC*, pp. 488-489, 2023.

[4] A. Chua, et al., "SOUL: An Energy-Efficient Unsupervised Online Learning Seizure Detection Classifier," *IEEE JSSC*, vol. 57, no. 8, pp. 2532-2544, Aug. 2022.

[5] M. Zhang, et al., "A One-Shot Learning, Online-Tuning, Closed-Loop Epilepsy Management SoC with 0.97μJ/Classification and 97.8% Vector-Based Sensitivity," *IEEE VLSI Circuits Symp.*, 2021.

[6] U. Shin, et al., "A 256-Channel 0.227μJ/class Versatile Brain Activity Classification and Closed-Loop Neuromodulation SoC with 0.004mm²-1.51 μW/channel Fast-Settling Highly Multiplexed Mixed-Signal Frontend," *ISSCC*, pp. 338-339, 2022.

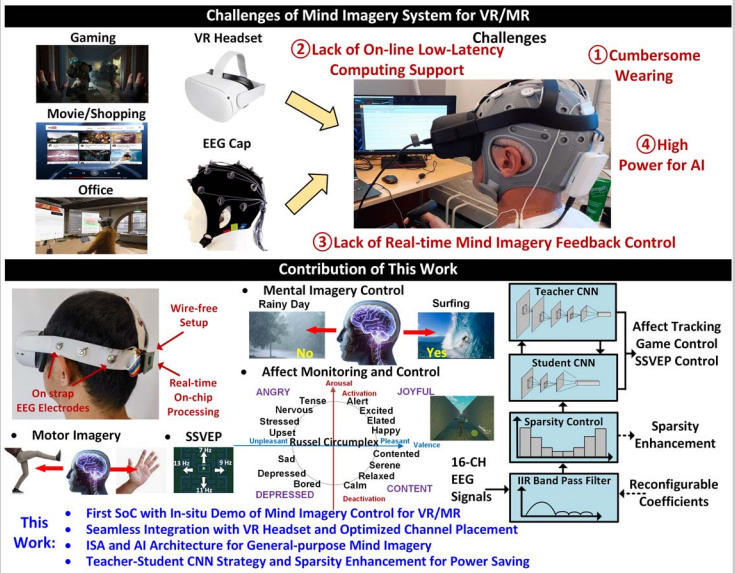



Figure 33.2.1: Challenges of existing mind-imagery systems for VR/MR applications(top). Proposed system solution and contribution of this work (bottom).

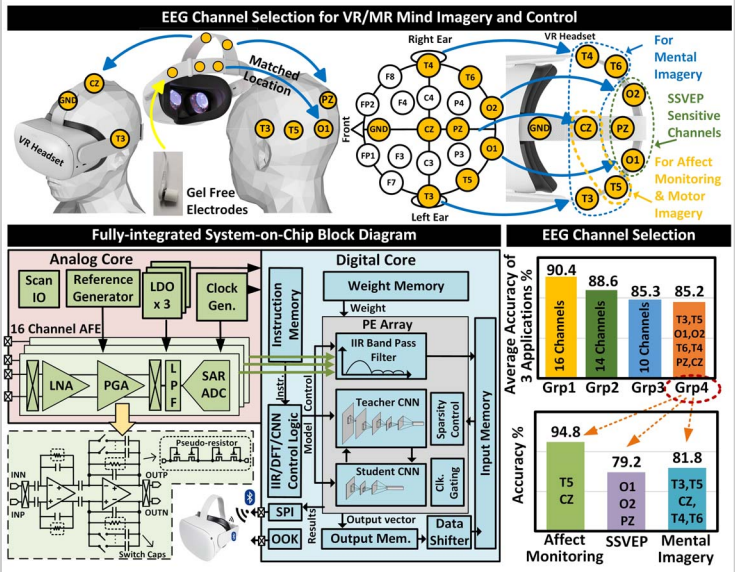



Figure 33.2.2: EEG channel selection for VR/MR mind imagery and control and proposed fully integrated SoC block diagram.

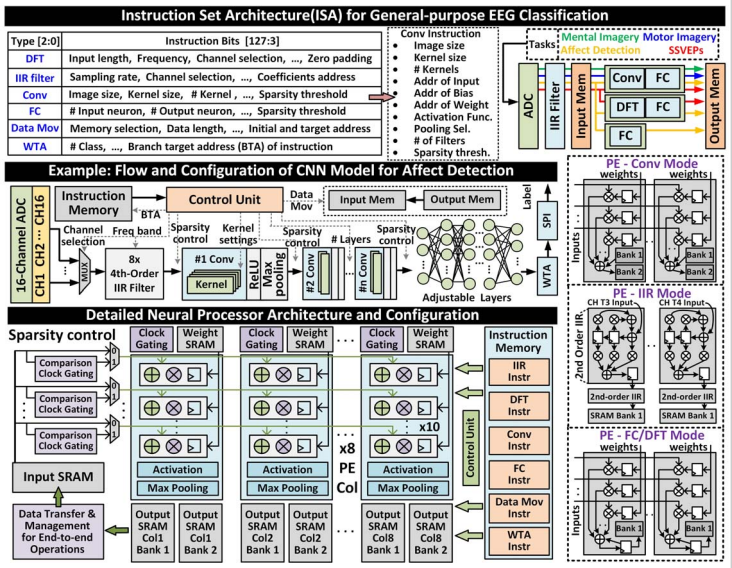



Figure 33.2.3: ISA for general-purpose mind imagery and control with an example flow of affect detection, and neural processor architecture and configurations.

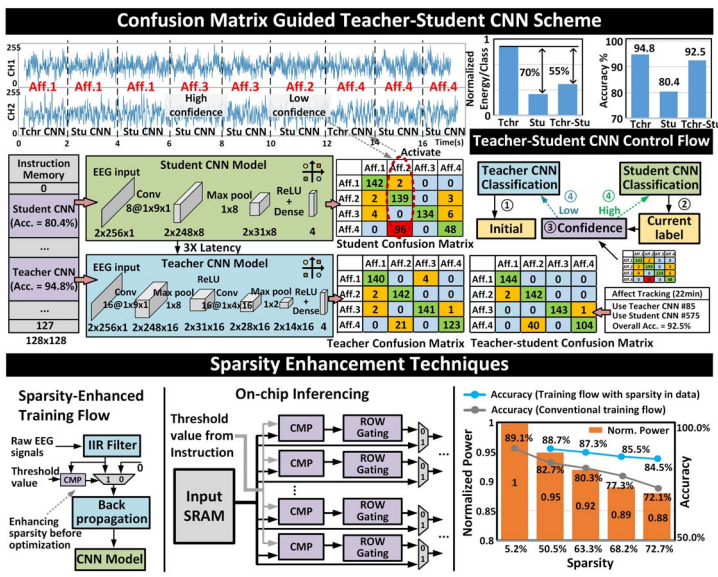



Figure 33.2.4: Confusion-matrix-guided teacher-student CNN scheme, and sparsity-enhancement technique.

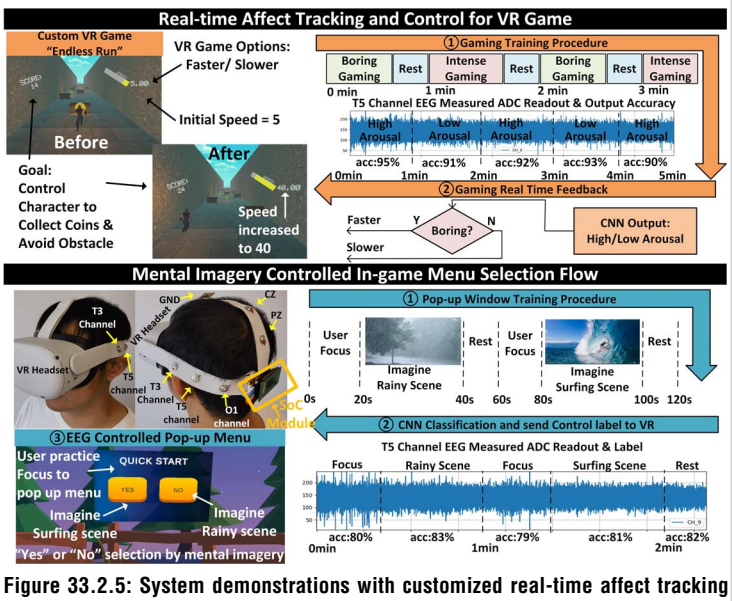
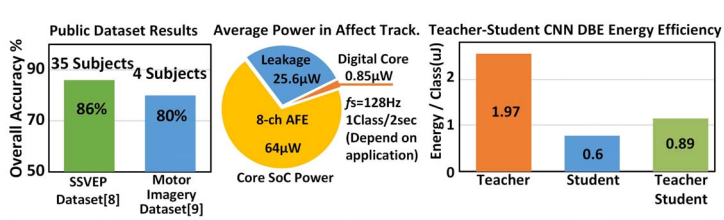



Figure 33.2.5: System demonstrations with customized real-time affect tracking based VR game control and mental imagery controlled in-game menu selection.




| | ISSCC'21[2] | ISSCC'23[3] | JSSC'22[4] | VLSI'21[5] | ISSCC'22[6] | CICC'20[7] | This Work |
|---|---|---|---|---|---|---|---|
| Process | 65nm | 40nm | 28nm | 40nm | 65 | 180nm | 65nm |
| SoC | NO | YES | NO | YES | YES | YES | YES |
| Supply Voltage(V) | 0.75 | 1(AFE)/ 0.9(DBE) | 0.5 | 1.1(AFE)/ 0.7(DBE) | 1.2 | 1 | 1.0(AFE)/0.8(DBE) |
| # Channel | 2 | 22 | 8 | 16 | 256 | 4 | 16 |
| Area(mm²) | 1.74 | 6 | 4 | 11.2 | 8 | 16 | 7.5 |
| Application Space & task | Biomedical: ECG/Seizure/ EMG | Biomedical: EEG Seizure | Biomedical: EEG Seizure | Biomedical: EEG Seizure | Biomedical: EEG Seizure + Stimulation | Biomedical: Autism spectrum disorder | VR/MR, EEG Affect Tracking/ Imagery control/ SSVEP |
| Classification Methods | CNN/MLP | SciCNN | LR | GTCA SVM | NeuralTree | DNN | CNN,IIR,FC,DFT |
| Energy (uJ/Class) | 4.36/2.06/5.25* | 28.33** | 1.5nJ* | 0.97** | 0.227** | 10.13** | 0.89(teacher-student)* |
| Verification (# subjects) | Bonn (N/A) | CHBMIT (24) | CHBMIT (24) | CHB-MIT (24) | CHB-MIT(24) iEEG(6) | DEAP(32) SEED(15) | THU-SSVEP(35) Motor Imagery(4) |
| in-situ Demostration | NO | YES | N/A | N/A | YES | N/A | Affect tracking & imagery control |
| AFE Gain(dB) | NO | N/A | NO | 36-54 | 40.7-57.9 | 50-64 | 45-72 |
| AFE Power/Ch | | | | N/A | 1.51uW | 1.63uW | 8uW |

*Digital backend only **Analog frontend + digital backend

Figure 33.2.6: Measurement results and comparison table.

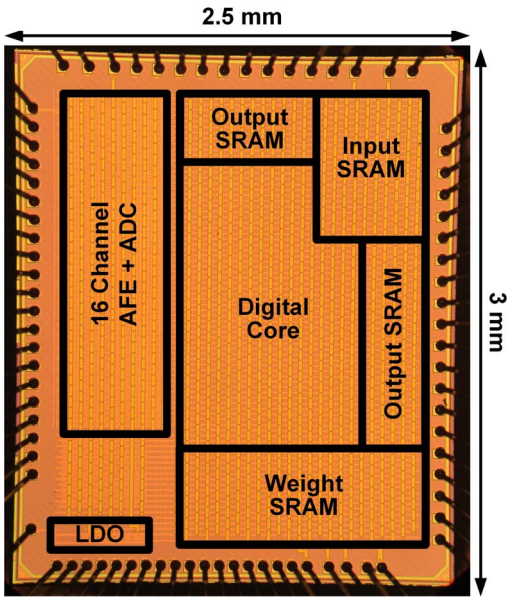



**Figure 33.2.7: Die photo.**

**Additional References:**
[7] A. R. Aslam, et al.,"A 10.13μJ/classification 2-channel Deep Neural Network-based SoC for Emotion Detection of Autistic Children," *IEEE CICC*, 2020.
[8] Y. Wang, et al.,"A Benchmark Dataset for SSVEP-Based Brain-Computer Interfaces," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 25, no. 10, pp. 1746-1752, 2017.
[9] H. Cho, et al., "EEG Datasets for Motor Imagery Brain-Computer Interface," *GigaScience*, vol. 6, no. 7, 2017.




979-8-3503-0620-0/24/$31.00 ©2024 IEEE