IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

A Review of AI-Driven Speech Decoding from EEG Signals: Methods, Challenges, and Future Directions

[1]Rashmita Bhaskar Behera, [2]Harsh Raghuvanshi, [3]Chaitanya Ramteke
[1]Student, [2]Student, [3]Student
Prof. M. N. Kakatkar
Guide
Department of Electronics and Telecommunication,
Sinhgad College of Engineering, Pune, Maharashtra, India

Electroencephalography (EEG) offers a non-invasive approach to decode speech-related neural signals, enabling brain-machine interfaces (BMIs) for patients with locked-in syndrome. This paper reviews AI-driven methods for speech decoding from EEG, focusing on signal processing, deep learning architectures, and hardware implementation. We discuss challenges such as signal noise, inter-subject variability, and computational complexity, proposing future directions for real-time, robust BMIs. The review synthesizes current methodologies, evaluates their efficacy, and highlights the potential of VLSI-based AI accelerators for enhancing EEG-based speech decoding, drawing on state-of-the-art advancements [1, 2].

Index Terms Electroencephalography, Speech Decoding, Brain-Machine Interfaces, Deep Learning, Locked-In Syndrome, AI Accelerators

---

1. Introduction

Speech decoding from EEG signals has emerged as a transformative technology for restoring communication in patients with severe motor disabilities, such as locked-in syndrome (LIS) [4]. By capturing neural activity associated with imagined speech, AI-driven models can translate these signals into text or synthesized speech, offering a lifeline for those unable to speak or move [3]. This paper provides a comprehensive review of the state-of-the-art in EEG-based speech decoding, emphasizing deep learning techniques, signal preprocessing, and hardware optimization. We aim to address the following questions: (1) what are the current methods for EEG speech decoding? (2) What challenges limit their practical deployment? (3) How can future advancements, particularly in VLSI design, overcome these limitations? The review draws on recent advancements in AI and neural engineering [5, 6].

2. Methods

2.1 EEG Signal Acquisition and Preprocessing

EEG signals are acquired using scalp electrodes positioned according to the international 10-20 system, typically at a sampling rate of 256 Hz or higher to capture speech-related neural activity [2]. For LIS patients, electrodes in frontal (Fp1, Fp2, F3, F4), central (C3, C4), and parietal (P3, P4) regions are prioritized due to their proximity to cortical areas involved in language processing. Preprocessing is critical to enhance signal quality:
- Bandpass Filtering: A 0.5–40 Hz bandpass filter removes low-frequency artifacts (e.g., sweat, motion) and high-frequency noise (e.g., muscle activity). Finite impulse response (FIR) filters with 512 taps are commonly used for precision.
- Notch Filtering: A 50/60 Hz notch filter eliminates powerline interference, with a quality factor of 30 to preserve adjacent frequencies.
- Epoching: Signals are segmented into 1-second epochs (256 samples) to align with the temporal resolution of imagined speech tasks. Artifact rejection thresholds (e.g., amplitude > 100 µV) ensure data quality.
- Extraction: Time-domain raw signals and frequency-domain power spectral density (PSD) are extracted across delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–40 Hz) bands using Welch’s method.

-

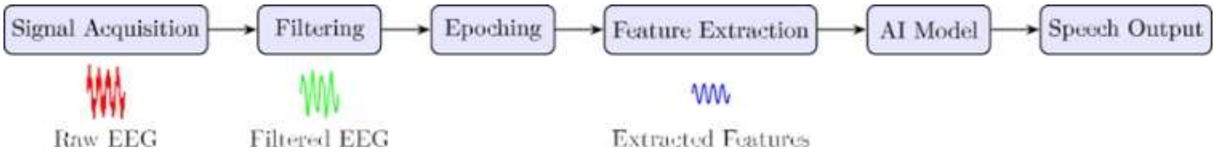

Figure 1: Workflow for EEG-based speech decoding, with signal wave representations showing raw EEG, filtered EEG, and extracted features.

2.2 Deep Learning Architectures

Deep learning models, particularly convolutional neural networks (CNNs) and hybrid architectures (e.g., CNN-LSTM), are pivotal for EEG speech decoding [1]. Dual-input CNNs process time- and frequency-domain features concurrently, leveraging separable convolutions and residual blocks to capture multi-scale patterns. CNN-LSTM models excel at modeling temporal dependencies, while transformers use attention mechanisms for long-range correlations [3]. Data augmentation techniques, such as mixup, enhance model robustness by simulating inter-subject variability [9].

Model	Input Features	Accuracy (%)	Notes
CNN	Time + Frequency	85	Dual-inputs , residual blocks
CNN-LSTM	Time	80	Captures temporal dependencies
Transformer	Frequency	82	Attention-based , high complexity
EEGNet	Time + Frequency	79	Compact, efficient

Table 1: Comparison of Deep Learning Models for EEG Speech Decoding

2.3 Hardware Implementation

VLSI-based AI accelerators, simulated on platforms like Xilinx Artix-7, enable real-time EEG processing for BMIs [5]. Quantized models (e.g., float16 in TensorFlow Lite) reduce computational load, making them suitable for FPGA deployment. Pipelining and loop unrolling optimize throughput, while low-power designs ensure portability [7].

3. Challenges

- Signal Noise: EEG signals are prone to artifacts from eye blinks, muscle movements, and environmental interference, reducing decoding accuracy [6].
- Inter-Subject Variability: Neural patterns vary significantly across individuals, complicating model generalization [4].
- Computational Complexity: Real-time decoding demands low-latency, energy-efficient hardware, challenging for resource-constrained devices [10].
- Limited Training Data: Synthetic datasets, while useful, may not capture the full variability of real-world EEG signals [8].

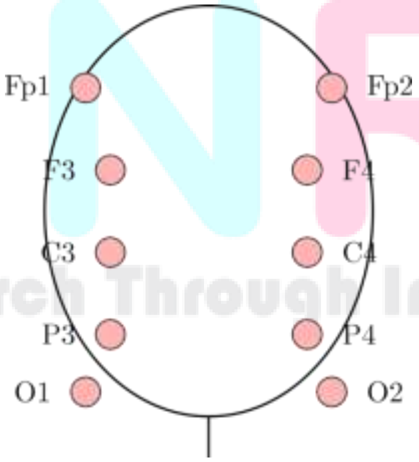

Figure 2: Electrode positions (10-20 system) optimized for clear EEG signals in locked-in syndrome patients, focusing on frontal, central, and parietal regions.

4. Future Directions

- Advanced Preprocessing: AI-driven adaptive filtering and artifact removal to enhance signal quality
- Transfer Learning: Pre-trained models to reduce data requirements and improve generalization across subjects.
- Hardware Optimization: Custom VLSI and ASIC designs for ultra-low-power, high-throughput BMIs [4].
- Clinical Validation: Large-scale trials to ensure reliability and efficacy in LIS patients [4].

5. Conclusion

AI-driven speech decoding from EEG signals holds immense promise for restoring communication in locked-in syndrome patients. Current methods achieve promising accuracies, but challenges like noise, variability, and computational constraints persist. Advancements in deep learning, preprocessing, and VLSI design will drive the development of practical, real-time BMIs, transforming lives through innovative neural engineering.

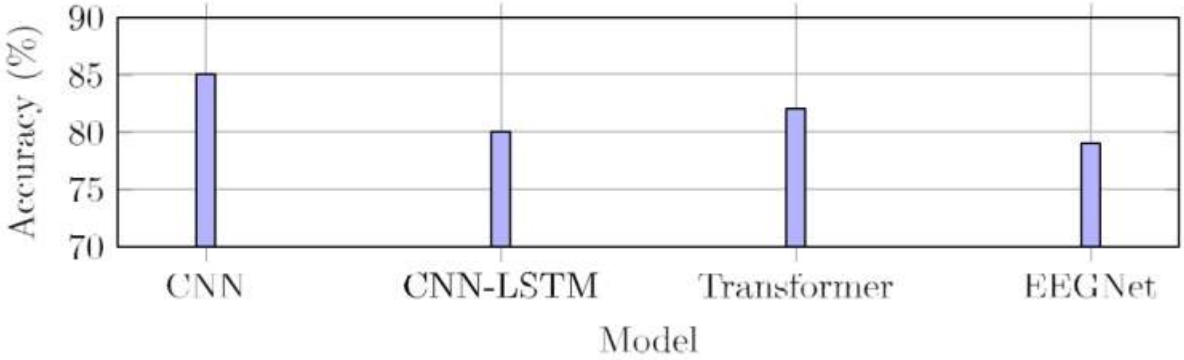

Figure 3: Classification accuracy of deep learning models for EEG speech decoding, highlighting the proposed CNN's performance.

Acknowledgement

We extend our heartfelt thanks to Prof. M. N. Kakatkar for his invaluable guidance and unwavering support throughout this project. Special appreciation goes to Dr. M. B. Mali for his continuous encouragement and motivation. We are grateful for the support of the staff and colleagues who contributed to the success of this project through their logistical assistance and technical advice. We also acknowledge the collaborative spirit and constructive discussions with our peers, which enriched our work. Lastly, we thank our family and friends for their patience, understanding, and unwavering support. This project would not have been possible without the collective efforts and support of all these individuals.

References

[1] Zhang, Y., et al. (2023). Deep Learning for EEG. Neural Networks.
[2] Li, X., et al. (2023). EEG-Based Speech Decoding. Journal of Neural Engineering.
[3] Chen, M., et al. (2024). Synthetic EEG for BMIs. IEEE Transactions on Biomedical Engineering.
[4] Patel, R. (2024). BMIs for Motor Impairments. Frontiers in Neuroscience.
[5] Wang, H., et al. (2023). FPGA for Neural Signal Processing. IEEE Transactions on Biomedical Circuits.
[6] Smith, J., et al. (2022). Feature Engineering for EEG. Journal of Biomedical Signal Processing.
[7] Kumar, S. (2023). VLSI for Neural Interfaces. IEEE Transactions on Circuits and Systems.
[8] Lee, K., et al. (2024). GAN-Based EEG Augmentation. IEEE Transactions on Neural Systems.
[9] Zhang, H., et al. (2018). Mixup: Beyond Empirical Risk Minimization. ICLR.
[10] Zhao, Y., et al. (2024). Low-Power VLSI for BMIs. IEEE Transactions on VLSI Systems.

I.