Optimizing Multichannel EEG Data: An Investigation of Current EEG Data Compression Methods

Andrews Damoah
adamoah@terpmail.umd.edu
University of Maryland - College Park
College Park, MD, USA

Teng Liang
tl40355@pausd.us
Palo Alto High School
Palo Alto, CA, USA

ABSTRACT

In this paper, we utilized a systematic literature review scheme to understand the current methods utilized to compress multichannel Electroencephalography (EEG) signals and how these techniques could be applied to the novel EEGEyeNet dataset. Our review will shed light on the current trends within the EEG data compression field and simplify the explanation of these techniques along with how to utilize them. By compiling a comprehensive list of the most recent and relevant research on this topic, we hope to provide a solid foundation for understanding the most up-to-date EEG data compression standards, their capabilities, and how these techniques compare to each other in terms of performance. We hope to expand the knowledge and accessibility of EEG data compression methods to broaden their utilization in EEG analysis.

CCS CONCEPTS

• **Information systems → Data compression**.

KEYWORDS

Review, Data Compression, Computational Efficiency, Compressive Sensing, Wavelet Transform, Huffman Coding, EEG, Time Series, Spatiotemporal Data, Lossy, Lossless, Clustering.

ACM Reference Format:
Andrews Damoah and Teng Liang. 2024. Optimizing Multichannel EEG Data: An Investigation of Current EEG Data Compression Methods. In *Proceedings of KDD Undergraduate Consortium (KDD-UC '24)*. ACM, New York, NY, USA, 6 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Electroencephalography (EEG) has been an extremely popular method for extracting and interpreting neurological information due to its non-invasive procedure. Additionally, its high temporal resolution makes it especially useful for real-time brain analytics often used in medical diagnostics. However, high temporal resolution, combined with multiple channels, can make the storage, transmission, and processing of EEG signals difficult, especially when working with large datasets such as EEGEyeNet, a high-quality EEG and eye-tracking dataset storing hours of recordings across a vast number of participants [19]. As such, improving the computational efficiency of EEG signal processing requires a deep understanding of the compression techniques applicable to EEG.

To address this, our paper provides a systematic review of the current literature, specifically focusing on techniques to compress multichannel EEG data, highlighting the recent trends within this domain. By compiling these findings, we aim to facilitate a deeper understanding of the current state of EEG data compression, evaluate the various techniques within the space, and highlight the potential applications of these methods for other datasets, specifically EEGEyeNet.

1.1 Research Questions

In this paper, we address the following research questions to better understand and evaluate the various trends within the current EEG data compression research field:

(1) What are the most effective and widely adopted techniques for compressing multichannel EEG data, according to the latest research?
(2) How do these EEG data compression techniques compare in terms of efficiency, accuracy, computational cost, and applicability to hardware implementations?
(3) How can these identified methods be effectively applied to the EEGEyeNet dataset to optimize performance and maintain data integrity? How can they address challenges related to big data acquisition and processing in healthcare applications?

By addressing these research questions, this paper aims to explain and evaluate the current methods used to compress EEG data, highlight the most promising techniques for large-scale, multichannel EEG datasets like EEGEyeNet, and equip the data science

Abbreviation	Definition
CR	Compression Ratio
PRD	Percentage Root Mean Square Difference
EEG	Electroencephalography
ECG	Electrocardiography
CS	Compressed Sensing
DWT	Discrete Wavelet Transform
DCT	Discrete Cosine Transform
SPIHT	Set Partitioning in Hierarchical Trees
GOP	Group of Pictures

Table 1: List of Frequently Used Acronyms

KDD-UC '24, August 25–29, 2024, Barcelona, Spain

ACM ISBN 978-1-4503-XXXX-X/23/08…$15.00
https://doi.org/XXXXXXX.XXXXXXX

community with resource-conscious solutions for EEG signal analysis. For clarity, Table 1 contains commonly used acronyms in this paper and their definitions.

2 RELATED WORK

Many previous review papers present key insights into the trends in EEG data compression techniques [3, 6, 7, 11, 20, 23, 29–36, 40, 41]. While these literature reviews provide a solid foundation for established techniques, rapid advancements in the field make it imperative to reevaluate the most up-to-date EEG compression techniques to ensure a deeper understanding of current trends in the domain.

Existing literature on EEG data compression techniques tends to be narrow in the compression techniques they evaluate, often examining a specific algorithm or class of algorithms. This can potentially make it difficult to comprehensively evaluate the performance of all methods and accurately weigh the benefits and tradeoffs of each. Additionally, some literature reviews broaden their analysis to other types of physiological signals, such as Electrocardiography (ECG) signals. While EEG and ECG signals are similar, they both possess intrinsic patterns that may influence which compression techniques are most effective.

The paper "Trends in Compressive Sensing for EEG Signal Processing Applications" by Gurve et al. [12] explores the integration of Compressive Sensing (CS) with neural engineering, specifically focusing on EEG signals for Brain-Computer Interfaces (BCIs). CS offers fast and energy-saving solutions for handling large volumes of neurological data. Key points from the review include the necessity of CS in EEG due to the growing challenges in big data acquisition and processing in healthcare and the potential of EEG signals for BCIs. The paper also examines current practices, scientific opportunities, and challenges related to CS in BCIs, such as major CS reconstruction algorithms, sparse bases, and measurement matrices used in processing EEG signals. Additionally, the review provides an overview of the reconstruction-free CS approach, aiming to improve EEG signal processing efficiency without full reconstruction, and discusses opportunities and challenges in integrating the CS framework into BCI applications.

The paper "Compressive Sensing of Electroencephalogram: A Review" by de Oliveira et al. [10] explores the application of Compressive Sensing (CS) to EEG signals. The review provides an overview of the current state of CS applications for EEG, highlighting the benefits such as efficient data acquisition and reduced power consumption. The authors discuss challenges related to implementing CS for EEG, including selecting appropriate measurement matrices and reconstruction algorithms, and explore opportunities for improving EEG signal processing efficiency through CS. This paper highlights the potential of CS in EEG-based applications, particularly in the context of BCIs and neurofeedback.

The paper "Compressed Sensing Approach for Physiological Signals: A Review" by Lal et al. [22] reviews the application of Compressed Sensing (CS) to various physiological signals, including EEG, ECG, EMG, and EDA. The authors discuss the advantages and disadvantages of using CS for these signals, evaluating its suitability for hardware implementation. Performance metrics such as Compression Ratio (CR), Signal-to-Noise Ratio (SNR), Percentage Root Mean Square Difference (PRD), and processing time are emphasized for assessing CS performance. The paper also explores current practices, challenges, and opportunities related to CS in healthcare applications, highlighting its potential to address the increasing volume of physiological data, transmission bandwidth limitations, and power consumption in telemonitoring.

Our paper aims to take a more current and holistic approach to analyzing contemporary data compression techniques specific to EEG. By incorporating insights from recent advancements, such as those presented by Gurve et al., de Oliveira et al., and Lal et al., we aim to better understand the different techniques and accurately compare their efficiency, accuracy, and computational costs. Furthermore, we investigate how these identified methods can be effectively applied to the EEGEyeNet dataset to optimize performance and maintain data integrity. By doing so, we aim to deliver the most up-to-date analysis of the most recent EEG data compression trends.

3 METHODS

3.1 Keywords

We conducted a systematic review using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method to identify relevant EEG data compression techniques. The search focused on research papers in the Google Scholar database.

We used the following keywords for the search: *("EEG" OR "Electroencephalography" OR "EEG Signal") AND ("Data Compression" OR "Compression") AND ("Multichannel" OR "Spatial") AND ("Lossy" OR "Lossless") AND ("Neural Network Predictor" OR "Autoencoder" OR*

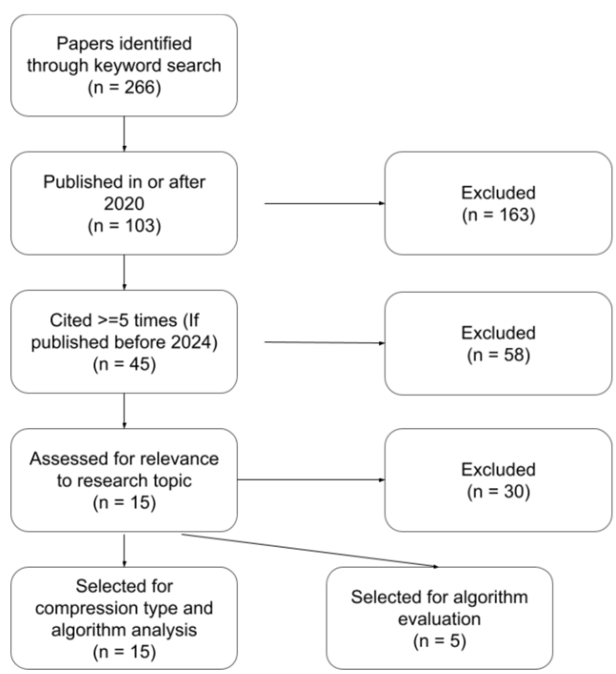

Figure 1: Selection process for the papers

"Wavelet Transform" OR "Channel Clustering" OR "Deep Learning") AND ("Storage" OR "Transmission" OR "Mobile" OR "Review").

Filtered by titles that contained: *("EEG" OR "Electroencephalography") AND ("Data Compression" OR "Compression")*

The search strategy aimed to identify papers relevant to studies and reviews in the context of EEG data compression techniques. Figure 1 visually represents the search process, illustrating the number of papers found at each step and the number of papers excluded based on predefined inclusion and exclusion criteria. We narrowed down fifteen papers for analyzing method frequency and five papers for comparing the performance of specific EEG compression algorithms.

3.2 Paper Selection Criteria

To maintain the relevance and quality of the compression techniques we highlight, the research papers we selected were chosen based on the following criteria:

- **EEG Focus:** The compression methods presented or discussed in the selected papers must be specifically applied to and tested on multichannel EEG datasets.
- **Publication Timeframe:** To ensure the selected papers present the most up-to-date methods, we only review research published in 2020 or later.
- **Relevance and Impact:** To ensure that the reviewed papers propose the most effective methods and are in line with current standards, we select papers that are frequently cited relative to their publication date, indicating their impact and relevance in the field.
- **Technique Diversity:** Selected papers as a whole must cover a range of compression techniques, including both lossy and lossless methods, and recent advancements such as compressive sensing and deep learning approaches.

By adhering to these criteria, we aim to provide a comprehensive and current evaluation of EEG data compression techniques, focusing on those that demonstrate high impact and relevance in the field.

4 RESULTS

4.1 Types of Compression Algorithms

Table 2 reveals the distribution of data compression algorithm types, including lossy and lossless, presented in the fifteen papers we selected [1, 2, 8, 13–18, 21, 22, 37–39, 42]. However, it is also important to note that some papers present both lossy and lossless compression, either separately or as a hybrid algorithm.

4.1.1 *Lossy Algorithms.*

Lossy algorithms are of data compression in which some information is permanently removed during the compression process. This often allows for much higher rates of compression but also usually results in some error between the original and reconstructed data.

Our review revealed that lossy compression algorithms are the most common method for EEG data compression. This is likely due to the high compression rates, which can significantly condense the often large and high-dimensional EEG datasets. Even so, the deviations present in the reconstructed data can have a significant impact on signal accuracy and can pose issues in medical diagnostics.

4.1.2 *Lossless Algorithms.*

Lossless compression algorithms are techniques whereby no information is lost during compression. This means that the reconstructed data will always contain the same information as the original, though this often comes at a cost to the compression rate.

Lossless algorithms are seemingly much more underutilized than lossy algorithms. They typically yield much lower levels of compression, meaning the size difference between the compressed and original signals is a lot smaller. However, it is important to note that since the reconstructed data from lossless algorithms is the same as the original, there are no accuracy trade-offs to consider. This is particularly important for medical applications, as there is no compromise to diagnosis accuracy when using this type of compression.

4.2 Compression Algorithms

In addition to presenting the relative frequency between the two types of compression methods present in the fifteen selected papers [1, 2, 8, 13–18, 21, 22, 37–39, 42]. We also report the individual algorithms that most frequently appear within our review of the literature. A breakdown of the top five can be found in Table 3.

4.2.1 *Compressed Sensing.*

Compressed, or Compressive Sensing (CS), is one of the most popular techniques for the compression of EEG signals. This algorithm leverages the sparsity of signals to reconstruct the input signal from only a few samples, vastly reducing the amount of stored information needed to represent a full signal [1]. CS also has the advantage that signals with perfect sparsity requirements can be reconstructed losslessly. Despite compression being efficient, the reconstruction stage of CS is still computationally intensive. Additionally, CS is also quite susceptible to noise in most practical applications [9].

4.2.2 *Signal Transforms.*

Signal transforms work by decomposing a signal into some alternative domain. The most common are

Type	Paper Count
Lossy	12
Lossless	5

Table 2: Most common types of EEG data compression that were discussed in research papers (some papers may include both lossy and lossless compressions and are counted as both).

Algorithm	Paper Count
Compressed Sensing	4
Signal Transformations	3
Clustering	3
Coding Algorithms	3
Autoencoders	2

Table 3: Breakdown of popular EEG compression Algorithms

Discrete Cosine Transform (DCT), where the signal is decomposed into cosine frequencies, and Discrete Wavelet Transform (DWT), where the signal is decomposed into temporally dependent frequencies. [38]. The coefficients of the decomposed signals can then be extracted and stored. Contrary to CS, DCT, and DWT have complex compression stages but are much easier to reconstruct [1]..

4.2.3 ***Clustering.*** Clustering algorithms, in general, operate by separating similar samples into groups [14–16]. This takes advantage of the redundancy and repetition present in EEG signals, allowing the data to be effectively represented by their clusters, significantly reducing the overall size.

4.2.4 ***Coding Algorithms.*** Coding algorithms aim to compress data by encoding the optimal bit representations for parts of the input, usually based on frequency, and are usually paired with other compression techniques [14–16]. This method is particularly effective for high-dimensionality EEG data with frequent signal repetition, as the repeated sequences can be encoded to have significantly lower bit representations.

4.2.5 ***Autoencoders.*** Autoencoders are a neural network with two main structures: the encoder, which converts the given input into a latent space, and the decoder, which reconstructs the original data from the latent representation. Since autoencoders are based on neural networks, they greatly benefit from the ability to adapt to the data they are training on, allowing them to better filter out the noise and redundant information from the compressed data [18]. However, autoencoders can be bulky and hard to implement on weaker devices. The training overhead may also be a concern depending on the situation.

4.3 EEG Compression Evaluation

To better understand how effective specific algorithms are at compressing EEG data, we selected five papers to compile and compare their specific techniques and performance, shown in Table 4 [2, 15, 16, 21, 42]. The results collected from the five selected papers show the algorithms are evaluated using Compression Ratio (CR) and Percentage Root Mean Square Difference (PRD). CR measures the compression level as a ratio between the original and compressed data sizes. PRD measures the relative difference between the original and reconstructed signals. Visualizations for the performance metrics can be seen in Figure 2. The x-axis represents the different algorithms, and the y-axis represents CR and PRD respectively. Higher CR and lower PRD values indicate better performance.

4.3.1 ***Fractal Compression.*** In the fractal compression technique, EEG signals are separated into blocks and matched with another block, typically much smaller, that contains similar signals at some scale and offset factor. The original signal can then be reconstructed by applying those factors to the smaller block.

This technique can achieve a CR as high as 160 with very low PRD [2]. Additionally, the CR and PRD values are heavily dependent on the block size, making this technique more adaptable for a variety of applications.

4.3.2 ***VLSI.*** This technique integrates an algorithm for EEG data compression into a Very Large Scale Integrated (VLSI) circuit [42].

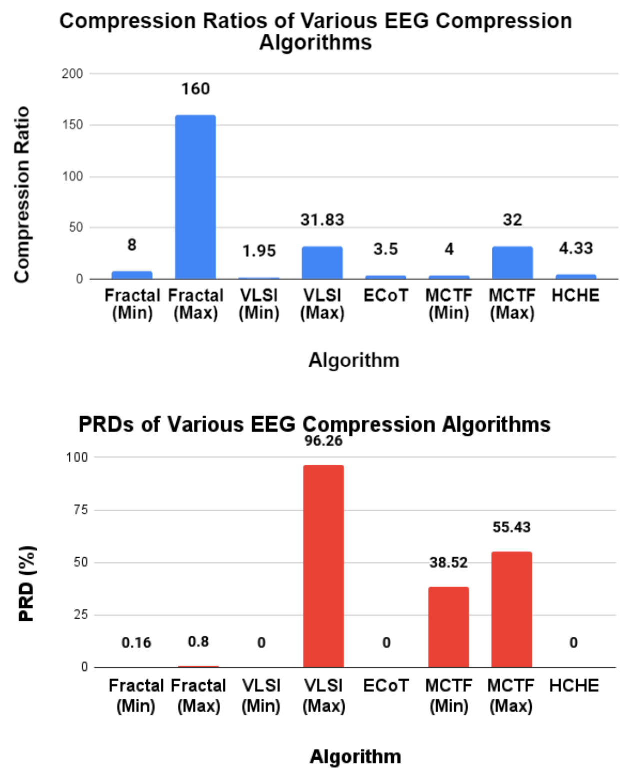

Figure 2: Visualization of EEG compression algorithm performance. Algorithms with ranging values are denoted by their min and max. PRD means Percent Root Mean Square Difference.

The algorithm first decomposes the signal via 2D DWT. Then Set Partitioning in Hierarchical Trees (SPIHT) is utilized to encode the wavelet coefficients as bitstreams, which act as memory locations for retrieving the wavelet coefficients from the circuit.

The algorithm's CR is mostly dependent on the number of discarded bits from the stream with higher CR requiring more discarded bits, but resulting in higher PRD. Alternatively, lossless compression is possible by discarding zero bits, although at a fairly low

Technique	CR	PRD(%)
Fractal Compression	8-160	0.16-0.8
VLSI	1.95-31.83	0-96.26
ECoT	3.5	0
MCTF	4-32	38.52-55.43
HCHE	4.33	0

Table 4: Breakdown of the performance of techniques used in 5 of the reviewed papers. CR means the Compression Ratio. PRD means Percent Root Mean Square Difference.

CR, about 1.95. Even so, since this algorithm is optimized for specific hardware, the enhanced resource efficiency makes this method particularly useful for low-memory EEG processing devices.

4.3.3 *ECoT*.

In the Efficient Compression Technique (ECoT), EEG signals are first clustered into groups using Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [15]. Then delta encoding is used to save the difference between the indices of the EEG data within each cluster. Finally, Huffman encoding is used to further compress the final size of the list of delta encoding differences.

While the CR of ECoT is not particularly high, it results in a lossless compression. Additionally, ECoT also benefits from having fairly fast compression and decompression times, improving the overall speed of data transmission.

4.3.4 *MCTF*.

In the Motion Compensated Temporal Filtering (MCTF) algorithm, each channel of the EEG data is temporally segmented and arranged into a 2D Group of Pictures (GOPs) [21]. Then motion compensating and temporal filtering are used to generate a set of low-pass segments, high-pass segments, and motion vectors. Next, 2D DWT is used to decompose the low-pass and high-pass segments. Finally, SPIHT is used to encode the spatiotemporal coefficients from DWT.

MCTF achieved a CR of between 4 and 32 when tested using combinations of GOP and segment sizes. Typically, larger GOP sizes result in lower PRD levels, while higher segment sizes lead to increased PRD.

4.3.5 *HCHE*.

The combined Hierarchical Clustering and Huffman Encoding (HCHE) technique first utilizes agglomerative hierarchical clustering to separate each sample into individual clusters, which are then merged hierarchically [16]. Then Huffman encoding is used to encode each cluster by its signal frequencies.

HCHE has one of the highest CR, when compared to the other lossless compression techniques, at about 4.33. This makes HCHE a highly efficient method for lossless EEG data compression.

5 DISCUSSION

Our comprehensive review of the current EEG data compression field highlights the most popular methods over the past four years, as well as providing a solid foundation for understanding and evaluating these techniques. While providing a general overview is important, we also aim to examine which methodologies could potentially apply to the EEGEyeNet dataset.

Lossy compression algorithms' ability to exhibit high levels of compression is especially helpful for a large dataset, such as EEGEyeNet. However, it is also important to recognize the effect of reconstruction error on gaze predictions from the decompressed EEG signals.

Similarly, algorithms like discrete wavelet transform and clustering could potentially be useful for exploiting redundancy in large EEG datasets. Although autoencoders may not be the most popular method, they could still be useful for the EEGEyeNet dataset, considering the large sample count would directly benefit its ability to interpret and encode the EEG signals.

Fractal compression is seemingly one of the best-performing EEG data compression algorithms that has been evaluated. The high compression ratio combined with very low reconstruction error is optimal for balancing compression strength with reconstruction accuracy. However, it is important to note that fractal compression can be quite computationally intensive at low block sizes, potentially slowing down processing on a large dataset [2]. The hierarchical clustering and Huffman encoding can exhibit comparatively high CR to other lossless algorithms. This is particularly useful for the EEGEyeNet dataset, as high data integrity can be maintained even after compression.

Limitations: While our systematic review provides insights into the current EEG data compression field, there are still limitations to acknowledge. Namely, due to time and selection criteria, our review is not exhaustive, and thus other valuable resources may exist. Additionally, as an evolving field, the information presented in this paper can quickly become outdated, underscoring the need for consistent and up-to-date reviews of current trends.

Future Work: In the future, investigations on the applications of data compression techniques for other types of data on EEG may be valuable, as some of the algorithms discussed, such as SPIHT and fractal compression, were originally designed for image compression [2, 42]. Examining other modalities may reveal key insights on applying new techniques to improve EEG data compression [4, 5, 24–28, 43–45]. Alternatively, as we have discussed the potential for these algorithms to be applied to additional EEG datasets, such as EEGEyeNet, conducting experiments to determine the viability of these techniques empirically will be vital.

We hope our findings elucidate a deeper understanding of where EEG data compression is today and better equip the data science community to navigate this realm, expanding the accessibility and advancement of EEG signal processing as a whole.

6 CONCLUSION

This paper presents a systematic review of trends in the current EEG data compression space, illustrating the most up-to-date techniques to provide a solid understanding of signal compression methods for EEG data.

We find lossy algorithms to be the most common types of compression used in recent years, with CS being the most common algorithm in recent literature. To enable a more robust evaluation, we also examine and report the performance of specific algorithms for better comparison of their benefits and tradeoffs, observing fractal compression to be the best option for high compression and HCHE to be the best option for low reconstruction error.

Finally, we analyze the different algorithms to determine their potential applications for new EEG datasets, namely EEGEyeNet. Due to the dataset's large sample size and synchronization with eye-tracking recordings, algorithms with high compression rates and low reconstruction error, such as fractal compression, are likely to be most effective.

By addressing our initial research questions, we've conducted a comprehensive evaluation of current methods in EEG data compression, laying a solid foundation for further exploration in the field.

REFERENCES

[1] Roghayeh Aghazadeh, Javad Frounchi, Fabio Montagna, and Simone Benatti. 2020. Scalable and energy efficient seizure detection based on direct use of compressively-sensed EEG data on an ultra low power multi-core architecture. *Computers in Biology and Medicine* 125 (2020), 104004.

[2] Kahlaa K Al-Nassrawy, Dhiah Al-Shammary, and Ali Kadhum Idrees. 2020. High performance fractal compression for EEG health network traffic. *Procedia Computer Science* 167 (2020), 1240–1249.

[3] Riyam Faisal Alwash, Ali Kadhum Idrees, and Salah Al-Obaidi. 2023. A Methodological Review on EEG Data Reduction in Edge/Fog computing-based IoMT networks. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, 1–7.

[4] Sizhe An, Ganapati Bhat, Suat Gumussoy, and Umit Ogras. 2023. Transfer learning for human activity recognition using representational analysis of neural networks. *ACM Transactions on Computing for Healthcare* 4, 1 (2023), 1–21.

[5] Sizhe An, Yigit Tuncel, Toygun Basaklar, and Umit Y Ogras. 2023. A survey of embedded machine learning for smart and sustainable healthcare applications. In *Embedded Machine Learning for Cyber-Physical, IoT, and Edge Computing: Use Cases and Emerging Challenges*. Springer, 127–150.

[6] Darren Craven, Brian McGinley, Liam Kilmartin, Martin Glavin, and Edward Jones. 2015. Compressed Sensing for Bioelectric Signals: A Review. *IEEE Journal of Biomedical and Health Informatics* 19, 2 (2015), 529–540. https://doi.org/10.1109/JBHI.2014.2327194

[7] Phuong Thi Dao, Xue Jun Li, and Hung Ngoc Do. 2015. Lossy compression techniques for EEG signals. In *2015 International Conference on Advanced Technologies for Communications (ATC)*. 154–159. https://doi.org/10.1109/ATC.2015.7388309

[8] Samiran Das and Chirag Kyal. 2021. Efficient multichannel EEG compression by optimal tensor truncation. *Biomedical Signal Processing and Control* 68 (2021), 102749.

[9] Mark A. Davenport, Jason N. Laska, John R. Treichler, and Richard G. Baraniuk. 2012. The Pros and Cons of Compressive Sensing for Wideband Signal Acquisition: Noise Folding versus Dynamic Range. *IEEE Transactions on Signal Processing* 60, 9 (2012), 4628–4642. https://doi.org/10.1109/TSP.2012.2201149

[10] Mateus M de Oliveira, Mahdi Khosravy, Henrique LM Monteiro, Thales W Cabral, Felipe M Dias, Marcelo AA Lima, Leandro R Manso Silva, and Carlos A Duque. 2020. Compressive sensing of electroencephalogram: a review. *Compressive sensing in healthcare* (2020), 247–268.

[11] Guangyao Dou, Zheng Zhou, and Xiaodong Qu. 2022. Time majority voting, a PC-based EEG classifier for non-expert users. In *International Conference on Human-Computer Interaction*. Springer, 415–428.

[12] Dharmendra Gurve, Denis Delisle-Rodriguez, Teodiano Bastos-Filho, and Sridhar Krishnan. 2020. Trends in compressive sensing for EEG signal processing applications. *Sensors* 20, 13 (2020), 3703.

[13] Fereidoon Hashemi Noshahr, Morteza Nabavi, and Mohamad Sawan. 2020. Multi-channel neural recording implants: A review. *sensors* 20, 3 (2020), 904.

[14] Ali Kadhum Idrees, Sara Kadhum Idrees, Raphael Couturier, and Tara Ali-Yahiya. 2022. An edge-fog computing-enabled lossless EEG data compression with epileptic seizure detection in IoMT networks. *IEEE Internet of Things Journal* 9, 15 (2022), 13327–13337.

[15] Ali Kadhum Idrees and Marwa Saieed Khlief. 2023. Efficient compression technique for reducing transmitted EEG data without loss in IoMT networks based on fog computing. *The Journal of Supercomputing* 79, 8 (2023), 9047–9072.

[16] Sara Kadhum Idrees and Ali Kadhum Idrees. 2022. New fog computing enabled lossless EEG data compression scheme in IoT networks. *Journal of Ambient Intelligence and Humanized Computing* 13, 6 (2022), 3257–3270.

[17] Uthayakumar Jayasankar, Vengattaraman Thirumal, and Dhavachelvan Ponnurangam. 2021. A survey on data compression techniques: From the perspective of data quality, coding schemes, data type and applications. *Journal of King Saud University-Computer and Information Sciences* 33, 2 (2021), 119–140.

[18] Tiejia Jiang, Jiahua Zhu, Dinghan Hu, Weidong Gao, Feng Gao, and Jiuwen Cao. 2022. Early seizure detection in childhood focal epilepsy with electroencephalogram feature fusion on deep autoencoder learning and channel correlations. *Multidimensional Systems and Signal Processing* 33, 4 (2022), 1273–1293.

[19] Ard Kastrati, Martyna Beata Płomecka, Damián Pascual, Lukas Wolf, Victor Gillioz, Roger Wattenhofer, and Nicolas Langer. 2021. EEGEyeNet: a simultaneous electroencephalography and eye-tracking dataset and benchmark for eye movement prediction. *arXiv preprint arXiv:2111.05100* (2021).

[20] Matthew L Key, Tural Mehtiyev, and Xiaodong Qu. 2024. Advancing EEG-based gaze prediction using depthwise separable convolution and enhanced pre-processing. In *International Conference on Human-Computer Interaction*. Springer, 3–17.

[21] Beenish Khalid, Muhammad Majid, Imran Fareed Nizami, Syed Muhammad Anwar, and Majdi Alnowami. 2020. EEG compression using motion compensated temporal filtering and wavelet based subband coding. *IEEE Access* 8 (2020), 102502–102511.

[22] Bharat Lal, Raffaele Gravina, Fanny Spagnolo, and Pasquale Corsonello. 2023. Compressed sensing approach for physiological signals: A review. *IEEE Sensors Journal* (2023).

[23] Weigeng Li, Neng Zhou, and Xiaodong Qu. 2024. Enhancing Eye-Tracking Performance Through Multi-task Learning Transformer. In *International Conference on Human-Computer Interaction*. Springer, 31–46.

[24] Yingzhou Lu, Tianyi Chen, Nan Hao, Capucine Van Rechem, Jintai Chen, and Tianfan Fu. 2024. Uncertainty quantification and interpretability for clinical trial approval prediction. *Health Data Science* 4 (2024), 0126.

[25] Yingzhou Lu, Kosaku Sato, and Jialu Wang. 2023. Deep Learning based Multi-Label Image Classification of Protest Activities. *arXiv preprint arXiv:2301.04212* (2023).

[26] Yingzhou Lu, Minjie Shen, Huazheng Wang, Xiao Wang, Capucine van Rechem, and Wenqi Wei. 2023. Machine learning for synthetic data generation: a review. *arXiv preprint arXiv:2302.04062* (2023).

[27] Xiaobo Ma. 2022. *Traffic performance evaluation using statistical and machine learning methods*. Ph. D. Dissertation. The University of Arizona.

[28] Xiaobo Ma, Abolfazl Karimpour, and Yao-Jan Wu. 2024. Data-driven transfer learning framework for estimating on-ramp and off-ramp traffic flows. *Journal of Intelligent Transportation Systems* (2024), 1–14.

[29] Nathan Koome Murungi, Michael Vinh Pham, Xufeng Caesar Dai, and Xiaodong Qu. 2023. Empowering Computer Science Students in Electroencephalography (EEG) Analysis: A Review of Machine Learning Algorithms for EEG Datasets. (2023).

[30] Yansheng Qiu, Ziyuan Zhao, Hongdou Yao, Delin Chen, and Zheng Wang. 2023. Modal-aware visual prompting for incomplete multi-modal brain tumor segmentation. In *Proceedings of the 31st ACM International Conference on Multimedia*. 3228–3239.

[31] Xiaodong Qu. 2022. *Time Continuity Voting for Electroencephalography (EEG) Classification*. Ph. D. Dissertation. Brandeis University.

[32] Xiaodong Qu and Timothy J Hickey. 2022. Eeg4home: A human-in-the-loop machine learning model for eeg-based bci. In *International Conference on Human-Computer Interaction*. Springer, 162–172.

[33] Xiaodong Qu, Peiyan Liu, Zhaonan Li, and Timothy Hickey. 2020. Multi-class time continuity voting for EEG classification. In *Brain Function Assessment in Learning: Second International Conference, BFAL 2020, Heraklion, Crete, Greece, October 9–11, 2020, Proceedings 2*. Springer, 24–33.

[34] Xiaodong Qu, Saran Liukasemsarn, Jingxuan Tu, Amy Higgins, Timothy J Hickey, and Mei-Hua Hall. 2020. Identifying clinically and functionally distinct groups among healthy controls and first episode psychosis patients by clustering on EEG patterns. *Frontiers in psychiatry* 11 (2020), 541659.

[35] Xiaodong Qu, Qingtian Mei, Peiyan Liu, and Timothy Hickey. 2020. Using EEG to distinguish between writing and typing for the same cognitive task. In *Brain Function Assessment in Learning: Second International Conference, BFAL 2020, Heraklion, Crete, Greece, October 9–11, 2020, Proceedings 2*. Springer, 66–74.

[36] Laxmi Shaw, Daleef Rahman, and Aurobinda Routray. 2018. Highly Efficient Compression Algorithms for Multichannel EEG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26, 5 (2018), 957–968. https://doi.org/10.1109/TNSRE.2018.2826559

[37] Geevarghese Titus and MS Sudhakar. 2020. A simple but efficient EEG data compression algorithm for neuromorphic applications. *IETE Journal of Research* 66, 3 (2020), 303–314.

[38] Min Wang, Xuefei Yin, Yanming Zhu, and Jiankun Hu. 2022. Representation learning and pattern recognition in cognitive biometrics: a survey. *Sensors* 22, 14 (2022), 5111.

[39] Kang Yang, Sungwon Kim, and Joel B Harley. 2023. Guidelines for effective unsupervised guided wave compression and denoising in long-term guided wave structural health monitoring. *Structural Health Monitoring* 22, 4 (2023), 2516–2530.

[40] Yiming Yao, Peisheng Qian, Ziyuan Zhao, and Zeng Zeng. 2022. Residual channel attention network for brain glioma segmentation. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2132–2135.

[41] Long Yi and Xiaodong Qu. 2022. Attention-based CNN capturing EEG recording's average voltage and local change. In *International Conference on Human-Computer Interaction*. Springer, 448–459.

[42] Haotian Yu, Yaguang Yang, Daibo Zhang, Qiliang Zhang, and Zhiqiang Li. 2024. VLSI design for adjustable compression rate in lossless/lossy compression of EEG signal. *Microelectronics Journal* 148 (2024), 106193.

[43] Zhengming Zhang, Renran Tian, and Zhengming Ding. 2023. Trep: Transformer-based evidential prediction for pedestrian intention with uncertainty. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37. 3534–3542.

[44] Zhengming Zhang, Renran Tian, Vincent G Duffy, and Lingxi Li. 2024. The comfort of the soft-safety driver alerts: Measurements and evaluation. *International Journal of Human–Computer Interaction* 40, 4 (2024), 904–914.

[45] Zhengming Zhang, Renran Tian, Rini Sherony, Joshua Domeyer, and Zhengming Ding. 2022. Attention-based interrelation modeling for explainable automated driving. *IEEE Transactions on Intelligent Vehicles* 8, 2 (2022), 1564–1573.