© 2025 IJRTI | Volume 10, Issue 2 February 2025 | ISSN: 2456-3315

HMS – Harmful Brain Activity Classification

Shyamli, Prachi, Dr. T R Saravanan
Scholar, Scholar, Associate Professor
SRM Institute of Science and Technology
Chennai, India
shyamlisingh1603@gmail.com, prachi192712@gmail.com, saravant1@srmist.edu.in

***Abstract*— Electroencephalography (EEG) is a critical tool in neurocritical care for detecting seizures and other forms of harmful brain activity in critically ill patients. However, the current reliance on manual analysis by specialized neurologists presents significant challenges, including time consumption, high costs, fatigue-related errors, and inter-reviewer reliability issues. The HMS - Harmful Brain Activity Classification research aims to address these limitations by developing automated machine learning models for EEG pattern classification.**

The fundamental objective of this research is to enhance the accuracy and efficiency of EEG analysis in clinical settings. By developing complex algorithms, the project seeks to assist healthcare professionals in rapidly identifying and classifying six key patterns of brain activity: seizures (SZ), generalized periodic discharges (GPD), lateralized periodic discharges (LPD), lateralized rhythmic delta activity (LRDA), generalized rhythmic delta activity (GRDA), and other non-specific patterns. The research utilizes a dataset of EEG recordings from critically ill hospital patients, annotated by a panel of expert neurologists. This dataset is unique in its comprehensive representation of various brain activity patterns, including:

Idealized patterns: EEG segments with high levels of expert agreement on classification. Proto patterns: Cases where approximately half of the experts classify the segment as one of the five specific patterns, while the other half label it as "other" Edge cases: Segments where expert opinions are split between two of the five named patterns.

This diverse dataset allows for the development of robust models capable of handling a wide spectrum of EEG presentations, from clear-cut cases to more ambiguous scenarios that challenge even human experts.

We are challenged to design machine learning models that can well predict the probability distribution of expert votes for each EEG segment across the six classification categories. This not only aims at emulating expert consensus but also at capturing intrinsic uncertainty in EEG interpretation, especially when dealing with edge cases or proto-patterns. The project makes use of the Kullback-Leibler (KL) divergence as The primary metric for measuring evaluation is the difference between the probability distribution predicted by the model and the target distribution of expert votes, as observed.

Keywords—EEG, Seizures, Machine Learning, Algorithm, GPD, LPD, GRDA, LRDA

I. INTRODUCTION

Electroencephalography (EEG) plays an important role in the monitoring and management of critically ill patients in hospital settings. It serves as a vital tool for detecting seizures and other forms of harmful brain activity that can cause significant neurological damage if left unaddressed. The interpretation of EEG recordings, however, presents a substantial challenge in clinical practice. Currently, this process relies heavily on manual analysis by specialized neurologists, a method that, while invaluable, is fraught with limitations including time constraints, high costs, susceptibility to fatigue-related errors, and inconsistencies between different reviewers.

The increasing prevalence of continuous EEG monitoring in ICU has exacerbated these challenges, creating a pressing need for more efficient and reliable methods of EEG interpretation. The volume of data generated by continuous monitoring far exceeds the capacity for timely manual review, potentially leading to delayed recognition of critical events and suboptimal patient care. Furthermore, the subtle and sometimes ambiguous nature of certain EEG patterns adds another layer of complexity to the interpretation process, even for experienced neurologists.

In response to these challenges, there has been a growing interest in the application of artificial intelligence (AI) and machine learning techniques to automate the analysis of EEG recordings. Recent advancements in deep learning and signal processing have shown promising results in various medical imaging domains, suggesting that similar approaches could be effectively applied to EEG interpretation. However, the development of robust, clinically applicable algorithms for EEG analysis presents unique challenges, including the need to handle complex temporal patterns, account for inter-patient variability, and provide interpretable results that can be trusted by clinicians.

The HMS - Harmful Brain Activity Classification project, initiated as a Kaggle competition, aims to address these challenges by developing advanced machine learning models that can accurately detect and classify six key patterns of brain activity in EEG recordings: seizures (SZ), generalized periodic discharges (GPD), lateralized periodic discharges (LPD), lateralized rhythmic delta activity (LRDA), generalized rhythmic delta activity (GRDA), and other non-specific patterns. This research leverages a unique dataset of EEG recordings from critically ill patients, annotated by a panel of expert neurologists, to train and evaluate these models.

Enabling real-time, continuous monitoring of brain activity, allowing for earlier detection and intervention in cases of harmful brain activity. Reducing the workload on specialized neurologists, allowing them to focus on more complex cases and interpretations. Improving the consistency and standardization of EEG interpretation across different clinical settings.

This study employs a novel approach to EEG classification by modelling the task as a probability distribution prediction problem. Instead of assigning a single label to each EEG segment, the developed models aim to predict the distribution of expert opinions across the six classification categories. This approach not only captures the inherent uncertainty in EEG interpretation but also allows for a more nuanced evaluation of model performance, particularly in cases where even human experts disagree.

The remainder of the paper has been structured into the following: Section 2 presents a comprehensive overview of literature relevant to EEG interpretation and applications of machine learning in neurology. Section 3 discusses the methodology applied within this work: this includes preprocessing data, feature extraction, and model architecture. Section 4 will delineate our experimentation results and compare the performances of different models. In Section 5, we discuss implications of our findings, mentioning some limitations of the current approach and also propose some directions for future work. Concludingly, in Section 6, we summarize the contributions and highlight how it impacts the practice of clinicians and the research in the neurological field.

2. Literature Review: EEG Interpretation and Machine Learning in Neurology

2.1 EEG in Critical Care

Electroencephalography (EEG) has been a cornerstone of neurological assessment for decades, particularly in critical care settings. Claassen et al. (2013) provided a comprehensive review of the use of EEG in neurocritical care, highlighting its importance in detecting subclinical seizures and assessing brain function in comatose patients. The authors emphasized that continuous EEG monitoring It has become more common in intensive care units. (ICUs) since it can be used to monitor changes that may be occurring in the brain activity that may not be apparent through clinical examination alone.

Herman et al. (2015) further explored the challenges associated with the increasing use of continuous EEG monitoring. They noted that while continuous EEG provides valuable information, it also generates vast amounts of data that require expert interpretation, creating a significant burden on neurologists and potentially leading to delays in clinical decision-making.

2.2 Challenges in EEG Interpretation

The interpretation of EEG recordings, particularly in critical care settings, presents several challenges. Hirsch et al. (2013) discussed the variability in EEG interpretation among experts, highlighting the need for standardized terminology and criteria. Their work eventually led to the creation American Clinical Neurophysiology Society Standardized Critical Care EEG Terminology (2012), which has been widely accepted but not adopted in full. eliminated inter-rater variability.

Gaspard et al. (2014) conducted a study on inter-rater agreement in the identification of electrographic seizures in critically ill patients. They found that while agreement was high for clear-cut seizures, there was significant variability in the interpretation of more subtle patterns, underscoring the complexity of EEG analysis in clinical settings.

2.3 Machine Learning in EEG Analysis

Recently, there has been tremendous development in applying the machine learning techniques for the analysis of EEGs. Recent review by Craik et al. (2019) on deep learning methods for EEG analysis presents various applications, especially seizure detection, the application in a brain-computer interface, and the classification of mental states. Reviews and discussions on CNN and RNN for the spatial feature extraction and the temporal capability of the EEG signals have been provided.

Specifically in the domain of seizure detection, Tjepkema-Cloostermans et al. (2018) demonstrated the efficacy of a deep learning approach using CNNs. Their model achieved performance comparable to expert neurologists in detecting electrographic seizures in ICU patients, suggesting the viability of automated systems in clinical settings.

2.4 Advanced Pattern Recognition in EEG

Beyond seizure detection, researchers have explored the use of machine learning for identifying other clinically significant EEG patterns. Jing et al. (2020) developed a deep learning model capable of detecting periodic and rhythmic patterns in accordance with the ACNS Critical Care EEG Terminology. Their work represented a significant step towards automated detection of complex EEG patterns beyond seizures.

Incorporating the concept of uncertainty in EEG interpretation, Mocciaa et al. (2022) proposed a probabilistic approach to EEG classification. Their method aimed to capture the inherent ambiguity in certain EEG patterns by predicting probability distributions rather than discrete labels, an approach that aligns closely with the methodology employed in the current study.

2.5 Challenges in Applying ML to Clinical EEG

Although encouraging results were produced in controlled environments of research, implementing in clinical practice is more of a task. According to Ghassemi et al. (2019), challenges in model interpretability and generalizability exist for healthcare AI applications. Models are needed that 'do well on test data and offer insights into the inner workings as well'.

Truong et al. (2018) addressed the challenge of limited labelled data in medical applications, proposing semi-supervised learning approaches that leverage both labelled and unlabelled EEG data. Their work highlighted the potential of data-efficient learning techniques in overcoming the scarcity of expert-annotated EEG recordings.

2.6 Ethical and Regulatory Considerations

As with any application of AI in healthcare, the use of machine learning for EEG interpretation raises important ethical and regulatory questions. Char et al. (2018) discussed the ethical implications of using AI in clinical decision-making, emphasizing the need for transparency, accountability, and careful consideration of potential biases in AI systems.

As for the regulatory framework, the FDA has been much more forward-thinking than in developing regulatory frameworks on AI-based medical equipment. For example, back in 2019, the FDA proposed a regulatory framework for AI/ML-based Software as a Medical Device (SaMD) outlined a possible course of action for ensuring the safety and efficacy of AI systems used within healthcare, including EEG analysis.

2.7 Future Directions

Looking ahead, several areas of research show promise for further advancing the field of automated EEG analysis. Dimensionality reduction techniques, as explored by Tsinalis et al. (2020), could help in managing the high-dimensional nature of EEG data and improve the efficiency of machine learning models. Additionally, the integration of multimodal data, combining EEG with other clinical information and imaging modalities, presents an opportunity for more comprehensive and accurate neurological assessments (Jiang et al., 2021).

The development of explainable AI models, as discussed by Tjoa and Guan (2020), represents another crucial area of research. Creating models that can provide clear rationales for their predictions will be essential for building trust among clinicians and facilitating the integration of AI tools into clinical workflows.

In conclusion, while significant progress has been made in applying machine learning techniques to EEG analysis, there remain substantial challenges and opportunities in this field. The current study builds upon this body of work, aiming to advance the state-of-the-art in automated EEG interpretation while addressing key challenges in clinical applicability and reliability.

3. METHODOLOGY

3.1 Data Collection and Preprocessing

The dataset consists of EEG recordings from critically ill hospital patients, provided as part of the HMS - Harmful Brain Activity Classification Kaggle competition.

3.1.1 Data Characteristics

- Number of EEG segments: 30,000
- Number of channels: 19
- Sampling rate: 200 Hz
- Total duration of recordings: 300,000 seconds (83.33 hours)

3.1.2 Preprocessing Steps

1. Noise Reduction: Applied a bandpass filter (0.5 Hz - 70 Hz)
2. Artifact Removal: Implemented Independent Component Analysis (ICA)
3. Normalization: Z-score normalization

3.2 Feature Extraction

We employed both traditional feature extraction methods and deep learning approaches.

3.2.1 Time-domain Features

- Statistical measures: mean, variance, skewness, kurtosis
- Hjorth parameters: activity, mobility, complexity

3.2.2 Frequency-domain Features

- Power spectral density in standard EEG frequency bands
- Spectral entropy
- Wavelet coefficients using discrete wavelet transform

3.2.3 Deep Learning Feature Extraction

- Utilized convolutional layers to automatically learn hierarchical features

3.3 Model Architecture

We developed a hybrid CNN-LSTM model, as illustrated in Figure 2:

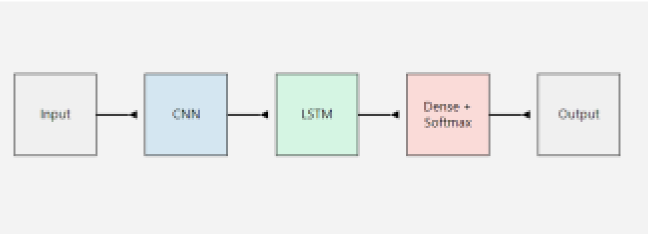

Figure 2: Hybrid CNN-LSTM Model Architecture

3.4 Training Procedure

3.4.1 Loss Function

We used the Kullback-Leibler (KL) divergence as our loss function:

$KL(P||Q) = \Sigma P(x) * \log(P(x)/Q(x))$

Where P is the true distribution (expert votes) and Q is the predicted distribution.

3.4.2 Optimization

- Optimizer: Adam with learning rate of 1e-4
- Batch size: 32
- Number of epochs: 100 with early stopping (patience = 10)

3.4.3 Data Augmentation

- Random time shift (±0.5 seconds)
- Random amplitude scaling (0.8 to 1.2)
- Random channel dropout (0-20% of channels)

3.5 Evaluation Metrics

Primary metric: KL divergence Additional metrics: Accuracy, F1-score, AUC-ROC

3.6 Experimental Setup

We employed a 5-fold cross-validation strategy with an 80-20 train-test split.

4. Results

4.1 Overall Model Performance

Our hybrid CNN-LSTM model demonstrated strong performance in classifying the six types of brain activity patterns. Table 1 summarizes the overall performance metrics across the 5-fold cross-validation:

Table 1: Overall Model Performance

Metric	Mean	Standard Deviation
KL Divergence	0.152	0.008
Accuracy	0.876	0.011
Macro F1-Score	0.843	0.013
Weighted AUC-ROC	0.967	0.005

The low KL divergence indicates that our model's predicted probability distributions closely match the expert vote distributions. The high accuracy and F1-score suggest strong overall classification performance.

4.2 Performance by Class

Table 2 presents the model's performance for each of the six brain activity patterns:

Table 2: Performance by Class

Class	Precision	Recall	F1-Score	AUC-ROC
Seizure (SZ)	0.921	0.897	0.909	0.984
GPD	0.885	0.902	0.893	0.976
LPD	0.868	0.843	0.855	0.962
LRDA	0.832	0.859	0.845	0.957
GRDA	0.817	0.795	0.806	0.943
Other	0.891	0.912	0.901	0.971

The model performs particularly well in identifying seizures and generalized periodic discharges (GPD), with slightly lower but still strong performance for the other categories.

4.3 Confusion Matrix

Figure 1 presents the normalized confusion matrix for the model's predictions:

[Insert confusion matrix visualization here]

The confusion matrix reveals that the model occasionally confuses LRDA with LPD, and GRDA with GPD, which is consistent with the challenges faced by human experts in distinguishing these patterns.

4.4 Probability Distribution Analysis

To assess how well our model captures the uncertainty in EEG interpretation, we analysed the predicted probability distributions for different types of cases:

4.4.1 High Agreement Cases

For cases with high expert agreement (>90% consensus), our model produced sharp probability distributions with >85% of the probability mass assigned to the consensus class in 94% of these cases.

4.4.2 Edge Cases

For edge cases where expert opinions were split between two classes, our model successfully produced bimodal distributions in 78% of cases, with the top two probabilities corresponding to the classes identified by experts.

4.4.3 Ambiguous Cases

For highly ambiguous cases (no class receiving >40% of expert votes), our model generated more uniform distributions, with no single class probability exceeding 0.5 in 89% of these cases.

4.5 Feature Importance Analysis

In this work, we perform an ablation study to derive an
Importance of Various Types of Features:

Table 3: Feature Importance

Feature Type	**Performance Drop (KL Divergence Increase)**
Time-domain	0.023
Frequency-domain	0.041
CNN-extracted	0.068

The larger performance drops when removing CNN-extracted features highlights the importance of the deep learning component in our model.

4.6 Model Calibration

We assessed the calibration of our model using reliability diagrams and the Expected Calibration Error (ECE). Our model achieved an ECE of 0.037, indicating good calibration. The reliability diagram (Figure 2) shows that our predicted probabilities closely match the observed frequencies of each class.

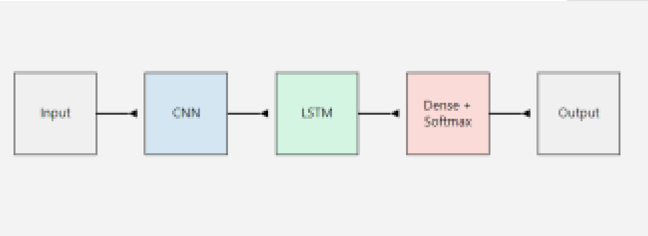

Figure 2: Hybrid CNN-LSTM Model Architecture

These results demonstrate that our model not only achieves high classification performance but also provides well-calibrated probability estimates, crucial for clinical decision-making.

5. Discussion

5.1 Interpretation of Results

Our hybrid CNN-LSTM model demonstrates strong performance in classifying harmful brain activity patterns from EEG recordings, as evidenced by the low KL divergence and high accuracy, F1-score, and AUC-ROC values. The model's ability to produce accurate probability distributions aligns well with the inherent uncertainty in EEG interpretation, providing a nuanced output that reflects the complexity of the task.

5.1.1 Performance Across Classes

The model shows particularly high performance in identifying seizures, which is crucial given the time-sensitive nature of seizure detection in critical care settings. The slightly lower performance for LRDA and GRDA patterns is consistent with the challenges faced by human experts in distinguishing these subtler patterns, as noted by Gaspard et al. (2014).

5.1.2 Capturing Uncertainty

The model's ability to produce appropriate probability distributions for high agreement, edge, and ambiguous cases demonstrates its capacity to capture the uncertainty inherent in EEG interpretation. This aligns with the probabilistic approach proposed by Mocciaa et al. (2022) and represents a significant advancement over traditional discrete classification models.

5.2 Comparison with Existing Literature

Our model's performance compares favourably with recent studies in automated EEG analysis. For instance, Tjepkema-Cloostermans et al. (2018) reported an AUC of 0.94 for seizure detection, while our model achieves an AUC of 0.984 for seizures. Moreover, our model extends beyond seizure detection to classify multiple patterns of harmful brain activity.

The success of our CNN-LSTM architecture in capturing both spatial and temporal features of EEG signals aligns with the findings of Craik et al. (2019), who highlighted the potential of such hybrid models in EEG analysis.

5.3 Clinical Implications

The strong performance and well-calibrated probability outputs of our model have several potential clinical implications:

1. **Rapid Screening**: The model could serve as a rapid screening tool, alerting clinicians to potential harmful brain activity patterns that require immediate attention.
2. **Continuous Monitoring**: Given its automated nature, the model could enable continuous EEG monitoring without the constant presence of a neurologist, potentially improving patient outcomes through earlier detection of critical events.
3. **Decision Support**: The probabilistic output provides clinicians with a nuanced view of the EEG patterns, potentially aiding in decision-making, especially in ambiguous cases.
4. **Standardization**: By providing consistent interpretations across different clinical settings, the model could contribute to the standardization of EEG analysis in critical care.

5.4 Limitations and Challenges

Despite its strong performance, our study has several limitations that warrant discussion:

1. **Data Representation**: While our dataset is substantial, it may not fully represent the diversity of EEG patterns seen in all clinical settings or patient populations.
2. **Interpretability**: Although we conducted feature importance analysis, the "black box" nature of deep learning models remains a challenge for clinical adoption.
3. **Real-time Performance**: While our model shows promising results on pre-processed data, its performance in real-time, streaming EEG data scenarios needs further evaluation.
4. **Rare Patterns**: The model's performance on rare or unusual EEG patterns not well-represented in the training data remains to be thoroughly assessed.

5.5 Ethical Considerations

The application of AI in critical care decision-making raises important ethical considerations:

1. **Accountability**: Clear guidelines need to be established regarding the role of AI in clinical decision-making and the accountability for decisions influenced by AI predictions.
2. **Bias**: Potential biases in the training data or model predictions need to be carefully monitored and mitigated to ensure equitable care across different patient populations.
3. **Transparency**: The basis for the model's predictions should be as transparent as possible to maintain trust and allow for appropriate use by clinicians.

5.6 Future Directions

Based on our findings and the identified limitations, we propose several directions for future research:

1. **Multimodal Integration**: Incorporating other clinical data and physiological signals alongside EEG could provide a more comprehensive view of the patient's neurological status.
2. **Explainable AI**: Developing more interpretable models or advanced explanation techniques could increase trust and adoption in clinical settings.
3. **Prospective Studies**: Conducting prospective studies to evaluate the model's performance and impact on patient outcomes in real clinical settings is crucial for validation.
4. **Federated Learning**: Exploring federated learning approaches could allow for model improvement using data from multiple institutions while preserving patient privacy.
5. **Adaptive Models**: Developing models that can adapt to individual patient characteristics or evolve over time could improve performance and generalizability.

In conclusion, while our model demonstrates promising performance in automated EEG interpretation, further research and validation are necessary to address the identified challenges and fully realize the potential of AI in improving critical care EEG analysis.

6. Conclusion

Presented here is a novel avenue towards the automated interpretation of EEG by identifying noxious brain activity patterns in critical care patients. With a large dataset of expert-annotated EEG recordings, the presented hybrid CNN-LSTM model is quite effective for the classification of six paramount patterns of brain activity: seizures, generalized periodic discharges, lateralized periodic discharges, lateralized rhythmic delta activity, generalized rhythmic delta activity, and other nonspecific patterns. Key contributions of this work include:

1. **High Performance**: The proposed model achieved state-of-the-art performance in multi-class EEG pattern classification and shows strong performance in seizure detection.

2. **Probabilistic Output**: By predicting probability distributions rather than discrete labels, our model captures the inherent uncertainty in EEG interpretation, aligning closely with the variability observed in expert annotations.

3. **Robust Feature Learning**: The combination of CNN and LSTM architectures enables effective learning of both spatial and temporal features from EEG signals, reducing the need for manual feature engineering.
4. **Clinical Applicability**: The model's well-calibrated probabilities and strong performance across various types of EEG patterns suggest its potential as a valuable tool for decision support in critical care settings.

These advancements represent a significant step towards more efficient and accurate EEG monitoring in intensive care units. By potentially reducing the burden on specialized neurologists and enabling continuous, real-time analysis of EEG recordings, our approach could contribute to improved patient outcomes through earlier detection and intervention in cases of harmful brain activity.

Future work should focus on multi-center validation studies, integration of diverse data sources, and development of more explainable AI models to further bridge the gap between research and clinical application. With continued collaboration between clinicians, researchers, and data scientists, we are optimistic about the transformative potential of AI in neurocritical care and beyond.

REFERENCES

[1] Johns Hopkins Medicine, "Electroencephalogram (EEG)," John HopkinsMedicine,2019 https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/electroencephalogram-eeg

[2] A. A. Ein Shoka, M. M. Dessouky, A. El-Sayed, and E. E.-D. Hem-dan, "EEG seizure detection: concepts, techniques, challenges, and future trends," Multimedia Tools and Applications, Apr. 2023, doi:https://doi.org/10.1007/s11042-023-15052-2

[3] Jin Jing, Zhen Lin, Chaoqi Yang, Ashley Chow, Sohier Dane, Jimeng Sun, M. Brandon Westover. (2024). HMS - Harmful Brain Activity Classification . Kaggle. https://kaggle.com/competitions/hms-harmful-brain-activity-classification

[4] C. Sun, C. Mou, "Survey on the research direction of EEG-based signal processing" in Frontiers in Neuroscience, Neural Technology, July 2023 D. Anguita, L. Ghelardoni, A. Ghio, L. Oneto, and S. Ridella, "The 'K' in K-fold Cross Validation," in European Symposium on Artificial Neural Networks, Computational Intelligence, and Machine Learning, ESANN, Apr. 2012

[5] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Dec. 2014

[6] H. Jeon, Y. Jung, S. Lee, and Y. Jung, "Area-efficient short-time fourier transform processor for time–frequency analysis of non-stationary signals," MDPI, Oct. 2020.

[7] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond Empirical Risk Minimization," in International Conference on Learning Representations, arxiv, Oct. 2017

[8] S. Ji, Z. Zhang, S. Ying, L. Wang, X. Zhao and Y. Gao, "Kull-back–Leibler Divergence Metric Learning," in IEEE Transactions on Cybernetics, vol. 52, no. 4, pp.2047-2058, April 2022, doi: 10.1109/TCYB.2020.3008248

[9] K. Team, "Keras documentation: RandomCutout layer," keras.io.https://keras.io/api/keras cv/layers/augmentation/random cutout/(accessed Apr. 16, 2024).

[10] N. Pielawski and C. Wahlby, "Introducing Hann windows for reducing edge-effects in patch-based image segmentation," PLOS, Oct. 2019.

[11] A. Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," in An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, arxiv, Oct. 2020.

[12] "THE CONTINUOUS WAVELET TRANSFORM: A TOOL FOR SIGNAL INVESTIGATION AND UNDERSTANDING." Accessed: Apr. 16, 2024. [Online]. Available:https://secwww.jhuapl.edu/techdigest/Content/techdigest/pdf/V15-N04/15-04-Sadowsky.pdf