

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/189989/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Hariharan, Vigneshwar, Reghuvaran, Chithra, John, Arlene, Pham, Nhat (Nick), Rana, Omer , John, Deepu and Iyer, Ganesh Neelakanta 2026. EEG-Fuseformer: A transformer-driven feature fusion framework for seizure onset prediction. Presented at: 2026 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Centre Prouvé, Nancy, France, 25-28 May 2026. IEEE Instrumentation and Measurement Technology Conference Proceedings. IEEE, 10.1109/i2mtc66907.2026.11694865

Publishers page: <https://doi.org/10.1109/i2mtc66907.2026.11694865>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



EEG-FuseFormer: A Transformer-Driven Feature Fusion Framework for Seizure Onset Prediction

Vigneshwar Hariharan^{1*}, Chithra Reghuvaran^{2*}, Arlene John³, Nhat Pham⁴,
Omer Rana⁴, Deepu John², Ganesh Neelakanta Iyer¹

¹National University of Singapore, ²University College Dublin, ³University of Twente, ⁴Cardiff University

Abstract—Epilepsy is one of the most common neurological disorders globally, characterized by recurring seizures and significantly impacting the quality of life. Despite advancements in diagnostic techniques, the mitigation of risks faced by epilepsy patients remains challenging due to the unpredictability of seizure events. An accurate forecast of seizure onset helps to reduce risks in epilepsy patients. In this paper, we propose EEG-FuseFormer, a transformer-based feature fusion framework for seizure-onset prediction that combines intermediate features extracted from Convolutional Neural Networks-Long Short-Term Memory (CNN-LSTM) and ResNet-18 networks. The CNN-LSTM architecture captures both spatial and temporal features directly from the raw signal, whereas the ResNet-18 extracts features from the Short-Time Fourier Transform (STFT) representation of the EEG signals. Fusion is carried out using a transformer encoder, and the final prediction is generated using fully connected dense layers. The CHB-MIT dataset was used to validate the proposed model. The results show that the proposed model achieves a mean recall of 98.85% and outperforms most of the state-of-the-art methods. This study evaluates the ability of the proposed feature fusion model to generalize in cross-patient testing scenarios. Fine-tuning pre-trained models on limited target patient data (target adaptation) within the cross-patient validation framework results in higher recall, precision, and F1-score metrics in comparison to the conventional cross-patient validation approach. Finally, the runtime-based computational complexity of the model is assessed across diverse hardware platforms to highlight the performance-complexity trade-off.

Index Terms—Feature Fusion, 1D-Convolutional Neural Networks (1D-CNN), Long Short-Term Memory (LSTM), Seizure Prediction, Electroencephalography, Short-Time Fourier Transform (STFT).

I. INTRODUCTION

Epilepsy is a brain condition that is characterized by recurring seizures, affecting nearly 50 million people around the world, making it one of the most common neurological disorders in the world [1]. An automated patient-specific seizure onset prediction approach could reduce the serious consequences of seizure by detecting the onset and raising an alarm. Electroencephalograms (EEG), which measure brain electrical activity, are commonly employed to diagnose seizure onset prediction [2]. In epileptic patients, brain activity typically consists of four phases: pre-ictal (just before a seizure), ictal (during a seizure), postictal (immediately after a seizure), and interictal (normal activity) [2]. The detection of pre-ictal and interictal states is the main task in seizure onset prediction.

EEG signals are inherently complex and exhibit irregular variations over time. Moreover, pre-ictal and interictal EEG patterns differ significantly across individuals, making seizure prediction highly challenging [3]. Feature extraction from decomposed EEG signals across time, frequency, and time-frequency domains is a key process in many seizure onset prediction studies [4]. These extracted features are then utilized to develop automated seizure prediction models. With the rapid advancement of deep learning, Convolutional Neural Networks (CNN) [3], [5], [6], Long Short-Term Memory (LSTM) [7], [8], CNN-Bidirectional-LSTM [9], Residual network (ResNet)-LSTM [10] architectures among others have become the most popular methods for seizure prediction in the literature. A single EEG feature cannot comprehensively represent the EEG signals, leading to insufficient accuracy and reliability in the detection of epileptic seizures [11]. To address this, various studies fuse multiple features to further enhance detection performance [12], [13]. Tsiouris *et al.* [7] fused both time and frequency domain features to select the optimal feature set for classification, to further improve classification accuracy. Zhang *et al.* in [11] proposed a Convolutional Neural Network-Gated Recurrent Unit-Attention mechanism (CNN-GRU-AM) model, where EEG signals are first decomposed using wavelet transform. Time-frequency domain and nonlinear features are then extracted, followed by further feature extraction and classification using the CNN-GRU-AM model.

However, simply combining various features can significantly increase runtime and introduce irrelevant features that may impact the final results. Hence, we propose a transformer-driven feature fusion model, EEG-FuseFormer, for seizure onset prediction. The proposed model uses deep learning networks to extract spatial and temporal dynamic features from EEG data. A two-channel feature fusion model based on CNN-Bi-LSTM was proposed in [14]. In this approach, CNN was used to extract spatial features, Bi-LSTM captures temporal features, and an attention module assigns weights to the features from each module. The proposed EEG-FuseFormer aims to further enhance seizure prediction performance by incorporating a transformer encoder that can carry out self-attention to fuse the features from the ResNet-18 and 1D-CNN-LSTM architectures.

Furthermore, most studies on seizure onset prediction in the literature have concentrated on patient-specific models aimed at predicting seizures for individual patients [6], [7], [10], [15]. Seizure symptoms vary among patients, from brief

* Vigneshwar Hariharan and Chithra Reghuvaran are co-first authors.

interruptions in activity to severe convulsions [2]. The real-life use of seizure prediction methods would require the deployment of cross-patient evaluation methods, such as leave-one-patient-out validation schemes [2]. The ability of the proposed EEG-FuseFormer model to generalize in cross-patient testing scenarios is evaluated. The most common approach in the literature for addressing cross-patient evaluation challenges is the application of the adaptation technique [16], which allows fine-tuning of pre-trained models using a limited amount of data from target patient. This paper evaluates the performance of the proposed model with and without adaptation under the cross-patient testing framework. The results show a significant improvement in recall, precision, and F1-score metrics with target adaptation in comparison to the conventional cross-patient validation approaches. Lastly, the runtime-based computational complexity of the proposed model is evaluated across diverse hardware platforms, including GPUs such as NVIDIA L4 and NVIDIA A100 SXM4 as well as embedded devices like the NVIDIA Jetson Nano, to illustrate the performance-complexity trade-off.

The remainder of this paper is organized as follows: Section II details the proposed EEG-FuseFormer Feature Fusion model and the cross-patient testing framework. Section III outlines the experimental setup and presents the results. Finally, Section IV concludes the paper.

II. METHODOLOGY

A. Datasets

The CHB-MIT Scalp EEG dataset [17] was used in this paper, which comprises of continuous scalp EEG recordings from 23 pediatric subjects with intractable seizures. The recordings were sampled at 256Hz and used the international 10-20 system for recordings using bipolar channels. To ensure consistency across all subjects, 18 channels were selected: FP1-F7, F7-T7, T7-P7, P7-O1, FP1-F3, F3-C3, C3-P3, P3-O1, FP2-F4, F4-C4, C4-P4, P4-O2, FP2-F8, F8-T8, T8-P8-O, P8-O2, FZ-CZ, and CZ-PZ.

Seizure activity in this work is categorized according to the proximity of the signal sample to the seizure (ictal) event. The pre-ictal period refers to the 1 hour preceding a seizure, post-ictal period refers to the 30 minutes immediately following a seizure event [15]. The inter-ictal period refers to periods outside of the two previously defined intervals and the ictal interval. Due to the high irregularity of post-ictal brain activity and inter-patient variability, post-ictal segments were excluded from analysis in this work. Additionally, to ensure relevance of this study to real-world applications, cases with greater than 12 seizures were omitted from analysis since such patients would require medical intervention or continuous monitoring, making seizure prediction less valuable. Hence a total of 13 patients were selected for analysis as listed in Table I.

B. Pre-processing

To eliminate noise sources and to focus on frequencies relevant to epilepsy analysis, the raw EEG signals were filtered using a high-pass filter with a cutoff frequency of 1 Hz and a

TABLE I: Subject details from the CHB-MIT Dataset

Case	Age (years)	No. of channels	No. of Seizures	Recording Duration (hh:mm:ss)
chb01	11	18	7	40:33:08
chb02	11	18	3	35:15:59
chb03	14	18	7	38:00:06
chb05	7	18	5	39:00:10
chb07	14.5	18	3	67:03:08
chb09	10	18	4	67:52:18
chb13	3	18	12	33:00:00
chb14	9	18	8	26:00:00
chb18	18	18	6	35:38:05
chb19	19	18	3	29:55:46
chb20	6	18	8	27:36:06
chb21	13	18	4	32:49:49
chb23	6	18	7	26:33:00

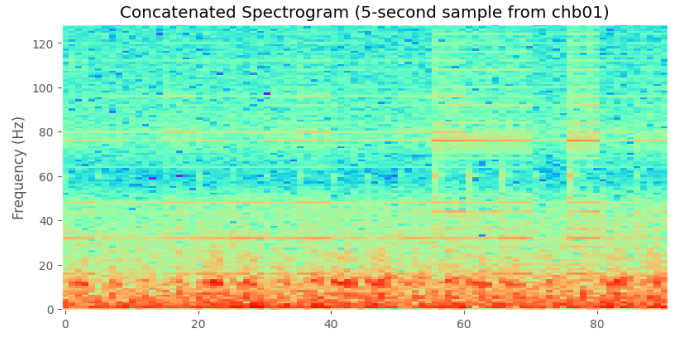


Fig. 1: Sample Concatenated Spectrogram from chb01

notch filter to remove power-line interference between 50 to 60 Hz. The data were then converted to a dataframe and interictal and preictal events are labeled according to the seizure phase. The accurate classification of pre-ictal events is considered as early detection of seizure onset in this paper. The post-ictal periods were discarded from the analysis. The class window size is 5 seconds [7]. Since there exists a class imbalance problem with fewer pre-ictal samples, the use of a sliding window approach was employed to artificially augment the number of samples by creating overlapping segments [5] during the training phase. The stride length of the sliding window selected was unique to each patient and was determined based on the label distribution. Lastly, following the aforementioned over-sampling, random under-sampling was applied to further balance cases where severe imbalance was not resolved simply using data augmentation. To ensure a leakage-free evaluation, we adopted dataset split before windowing and augmentation.

C. Feature engineering

Short-Time Fourier Transform (STFT) was used to extract time-frequency features from a short-duration segment of EEG signal by computing its Fourier Transform. This results in the creation of spectrograms that describe the power distributions of frequency over a time segment. STFT was applied to 5-second segments of EEG to extract a spectrogram representation of the sample. These were used as image inputs to Convolutional Neural Networks (CNNs) in order to extract features or classify signals. Lastly, the spectrograms from the

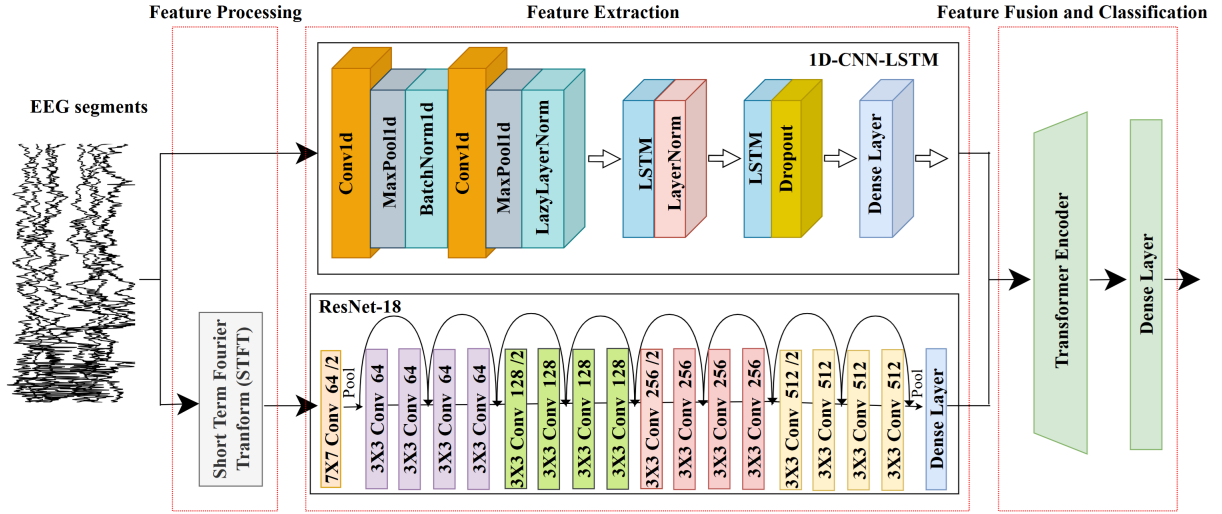


Fig. 2: Proposed EEG-FuseFormer Feature-Fusion model

eighteen channels were concatenated to generate the input representation for these networks. Fig. 1 shows the concatenated spectrogram of a 5-second sample from chb01.

D. EEG-FuseFormer model

In this paper, we propose a fusion model based on ResNet-18 and 1D-CNN-LSTM for feature extraction, feature selection, and classification of epileptic EEG signals, as illustrated in Fig. 2. The model consists of two parallel channels that process EEG signals simultaneously. Channel 1 operates on the filtered time-domain signal and employs a 1D-CNN-LSTM architecture to capture both spatial and temporal features. Channel 2 performs frequency-domain feature engineering using an STFT block and applies ResNet-18 with ReLU activation to extract spatial features. These features are projected into a common vector dimension of size 64 using a linear layer before being connected to a multi-head transformer encoder and dense layers for final classification. The classification of interictal and preictal stages is performed using a softmax function.

1) *1D-CNN-LSTM*: The proposed model contains two convolutional blocks that act on the raw signal, each consisting of a batch normalization layer, a Rectified Linear Unit (ReLU) activation function and a max pooling layer. The first convolutional block uses a 1D convolutional layer with 18 convolutional kernels of size 7 and max-pooling with a kernel size of 5. The second block uses 128 kernels with a kernel size of 5 and a max-pooling kernel of size 7. The output of these blocks is layer-normalized and fed into a two-layer LSTM model with 128 and 64 hidden units, respectively. A ReLU activation function is used in all hidden layers. Lastly, to prevent overfitting, batch normalization was applied before convolutional layers, while layer normalization was applied before recurrent layers [18]. A dropout probability of 0.3 was also used along with L2-regularization with a coefficient of 0.00001 in the 1D-CNN-LSTM model. Across the models,

we employed the categorical cross entropy loss function in conjunction with the Adam optimizer to reduce the need for extensive hyperparameter tuning.

2) *ResNet-18*: ResNets improve upon the traditional CNN architecture due to the addition of residual/skip connections, which enable deeper architectures without performance degradation [19]. ResNet-18 consists of 5 blocks as shown in Fig. 2. The input layer of ResNet-18 architecture is modified in this work to accept a concatenated spectrogram input of size 128×90 . The extracted feature from ResNet-18 is converted into a 512-dimensional vector using a pooling operation followed by a flatten layer.

3) *Transformer Encoder*: The transformer encoder integrates features from both ResNet-18 and 1D-CNN-LSTM architectures. The encoder used here consists of two layers, each containing two sub-layers: (1) a multi-head self-attention mechanism with 8 heads, and (2) a fully connected feed-forward network [20]–[22]. For the transformer encoder, all layers output dimension is 128, which is then passed through two dense layers with input sizes of 128 and 64 for the final classification.

E. Cross-patient Testing Framework

The evaluation of seizure prediction models can be conducted using several testing frameworks, as shown in Fig. 3 [2]. Patient-specific modelling involves using data from a single patient to train a specific model for that patient. However, this model does not generalize to seizure onset prediction for unseen patients. To address this, we employed cross-subject modelling, where labeled data from $(N_1, N_2, \dots, N_{I-1})$ patients were used for training, and data from N_I patient was reserved for testing. The study investigates adaptation method to improve the performance of cross-subjects modelling approach. The target adaptation helps to finetune pre-trained models on a small sample of data from the N_I patient data for 3-5 epochs before testing. Target adaptation allows good

TABLE II: Patient-specific results for 1D-CNN-LSTM, ResNet-18 models and EEG-FuseFormer. P, R and FAR denote Precision, Recall and False Alarm Rate respectively.

Case	1D-CNN-LSTM				ResNet-18				EEG-FuseFormer			
	P(%)	R(%)	F1-Score(%)	FAR(/h)	P(%)	R(%)	F1-Score(%)	FAR(/h)	P(%)	R(%)	F1-Score(%)	FAR(/h)
chb01	91	99	94	0.26	97	98	97	0.2	97	99	98	0.1
chb02	94	98	95	0.25	99	99	98	0.19	99	98	99	0.13
chb03	95	100	97	0.14	96	99	99	0.14	98	99	98	0.06
chb05	81	90	85	0.24	93	95	94	0.2	94	98	96	0.1
chb07	94	99	96	0.23	98	100	98	0.18	99	100	100	0.089
chb09	87	96	91	0.22	96	97	96	0.17	96	99	98	0.109
chb13	88	96	91	0.24	97	99	98	0.1	97	100	100	0.06
chb14	72	80	75	0.27	91	94	92	0.18	95	97	94	0.14
chb18	98	100	99	0.13	98	99	99	0.09	99	100	99	0.03
chb19	100	100	99	0.14	98	99	99	0.06	99	100	100	0.06
chb20	95	99	97	0.16	98	98	98	0.09	98	100	100	0.04
chb21	84	97	89	0.24	95	93	94	0.20	95	96	95	0.13
chb23	75	61	67	0.50	96	98	97	0.32	97	99	96	0.15

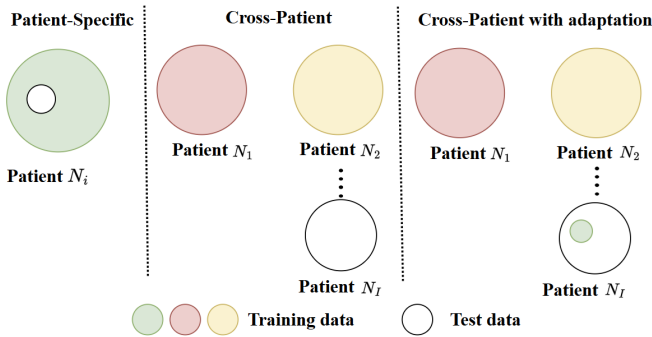


Fig. 3: Visualization of the Testing Frameworks

classification results in cross-subjects modelling approach. These approaches aim to learn a set of relevant features from the general population that are subsequently adapted to the clinical data from the target patients.

III. RESULTS AND DISCUSSION

The section discusses the results of proposed models when evaluated on the thirteen chosen cases in the CHB-MIT dataset. Pytorch and Torchvision were used to implement the proposed architecture. The models were trained on a 60-15-25 training-validation-test split ratio.

Most of the works reported in the literature report patient-specific results, due to the individualized nature of seizures (i.e., models were trained and tested on the same patient's data). In order to ensure a consistent comparison, the models were first tested using the same patient-specific setup. Table II shows the results on individual cases with 1D-CNN-LSTM, ResNet-18 and EEG-FuseFormer models. Precision, Recall, F1-Score and False Alarm Rate (FAR) are the performance metrics employed for the evaluation. Precision quantifies the proportion of predicted positive events that correspond to true positive outcomes. Recall denotes the percentage of actual seizure events that are correctly identified by the predictive model [3]. FAR quantifies how many inter-ictal samples are misclassified as pre-ictal per hour of inter-ictal activity [3]. The results in Table II show that the 1D-CNN-LSTM model

achieves a mean recall of 93.4% and a mean F1-score of 90.38%. In comparison, ResNet-18 demonstrates a superior performance achieving a mean recall of 97.53% and a mean F1-Score of 96.8%. The EEG-FuseFormer has the highest predictive performance, with a mean recall of 98.85% and a mean F1-Score of 97.92%. Additionally, EEG-FuseFormer shows a low FAR(/h) of 0.12, whereas the mean FAR(/h) is at 0.27 for 1D-CNN-LSTM.

Table III shows the results of the proposed feature fusion model and its comparison with the state-of-the-art models. The results indicate that feature fusion models by integrating multiple features outperform standalone models like 1D-CNN-LSTM and ResNet-18. The proposed fusion model outperforms the state-of-the-art methods of Truong *et al.* [5], Zhang *et al.* [3], Jemal *et al.* [23], Wu *et al.* [8], Ma *et al.* [14], Lee *et al.* [10] and Zhang *et al.* [11]. Additionally, the models in [7] require expert knowledge on the appropriate use of features due to the elaborate extraction of features such as entropy characteristics, peak-to-peak voltage, etc. The proposed model achieves comparable performance with Tsiouris *et al.* [7] but with significantly lower feature engineering efforts.

A. Cross-patient Testing Framework

Table IV shows the results of the proposed models for the cross-patient testing framework with and without adaptation. In the cross-patient testing, data from 12 patients were used for model training, while the remaining patient data was used for target adaptation and testing. Here, chb023 was the target patient for adaptation and testing. Observing the results in Table IV, it is apparent that a major limitation of these models is their patient-specificity. However, after finetuning the training model with the target adaptation technique, we can improve the model performance. This is evident from the significant improvement in the recall values, which increased from 40 – 60% across different models. Furthermore, Table IV clearly demonstrates the superior performance of the EEG-FuseFormer model compared to CNN-LSTM and ResNet-18. The proposed model achieves a 22% improvement over 1D-CNN-LSTM and a 4% improvement over ResNet-18 within the cross-patient testing framework with target adaptation.

TABLE III: Comparison of the proposed EEG-FuseFormer model with the state-of-the-arts. NA: Not Available

Ref	Features	Classifier	Precictal du- ration (min)	Recall (%)	FAR(h)
Truong <i>et al.</i> [5]	STFT	CNN	5	81.20	0.16
Tsiouris <i>et al.</i> [7]	Time-domain: Statistical moments, Standard Deviation, Peak-to-peak voltage Frequency-domain: Fast Fourier Transform, Power Spectral Density, Discrete Wavelet Transform	LSTM	60	99	0.06
Zhang <i>et al.</i> [3]	Spatial pattern statistics	CNN	30	92.2	0.12
Jemal <i>et al.</i> [23]	Raw Data	CNN inspired by Filter Bank Common Spatial Pattern (FBCSP)	30	96.1	0.04
Wu <i>et al.</i> [8]	Raw Data	LSTM	60	92.17	0.29
Ma <i>et al.</i> [14]	Raw Data	CNN-Bi-LSTM	NA	94.84	NA
Lee <i>et al.</i> [10]	STFT	ResNet-LSTM	30	92.71	0.05
Zhang <i>et al.</i> [11]	Time-frequency domain with non-linear fusion	CNN-GRU-AM	NA	95.47	NA
This Work	Raw Data	1D-CNN-LSTM	60	93.46	0.23
This Work	STFT	ResNet-18	60	97.53	0.15
This Work	Raw Data and STFT fusion	EEG-FuseFormer	60	98.85	0.09

B. Computational Complexity Analysis

Table V shows the computational complexity analysis of the three models CNN-LSTM, ResNet-18, and EEG-FuseFormer based on runtime performance. Our evaluation utilizes multiple GPU platforms, from NVIDIA L4 and NVIDIA A100 SXM4 to embedded devices like the Jetson Nano. To assess computational complexity, we compared the CNN-LSTM, ResNet-18, and EEG-FuseFormer architectures in terms of both runtime and the number of trainable parameters. Table V shows the mean runtime for patient-specific testing of the subjects listed in Table I. To evaluate runtime on a memory-constrained device such as the Jetson Nano, the data were divided into 10-hour segments and processed sequentially to estimate computational complexity. In real-time operation, this segmentation will not be necessary, as seizure-detection data will be streamed continuously, processed on the fly, and discarded after prediction.

Based on the results in Table II and Table V, EEG-FuseFormer delivers the highest predictive performance by leveraging multi-scale feature fusion and attention mechanisms. However, this gain comes with increased architectural complexity and runtime overhead. In contrast, CNN-LSTM model achieves competitive performance with relatively low complexity, making it suitable for resource-constrained deployment. Thus, selecting an appropriate model depends on whether the target application is cloud-based or edge deployment. Overall, the results highlight a clear performance and complexity trade-off across CNN-LSTM, ResNet-18, and EEG-FuseFormer.

IV. CONCLUSION

The paper investigates the use of feature fusion deep learning models for seizure onset prediction. 1D-CNN-LSTM and ResNet-18 architectures were explored that leverage both time- and frequency-domain features from EEG signals. ResNet-18 appeared to be the best performing model with a mean sensitivity of 97.9%. The paper proposed EEG-FuseFormer, a novel feature-fusion framework combining the outputs of both

TABLE IV: Cross-patient testing with and without adaptation technique.

Ref	Model	Recall (%)	Precision (%)	F1- score(%)
Cross-patient testing without adaptation				
Jemal <i>et al.</i> [2]	CNN	-	-	55.34
This work	1D-CNN-LSTM	54	40	46
This work	ResNet-18	59	56	57
This work	EEG-FuseFormer	59	59	58
Cross-patient testing with adaptation				
Jemal <i>et al.</i> [2]	CNN	-	-	66.45
Deng <i>et al.</i> [14]	CNN	-	-	75
This work	1D-CNN-LSTM	76	67	66
This work	ResNet-18	89	84	86
This work	EEG-FuseFormer	93	82	88

TABLE V: Computational Complexity in terms of runtime and the number of trainable parameters

Model	Runtime (seconds)			Trainable Parameters
	NVIDIA L4	NVIDIA A100 SXM4	Jetson Nano	
CNN-LSTM	0.50K	0.20K	1.58K	0.12M
ResNet-18	3.28K	1.23K	8.18K	11.17M
EEG-FuseFormer	22.47K	11.29K	56.17K	12.40M

the ResNet-18 model and the 1D-CNN-LSTM model, exploiting the self-attention framework of transformer encoders. The proposed fusion model was compared to existing state-of-the-art models and demonstrated improved performance. To better assess the ability of the proposed model to generalize to new patients, a cross-patient modelling approach was employed. The adaptation techniques enhanced the performance of the cross-patient testing framework by over 40%. The predictive performance and complexity analysis shows that there should be a trade-off across CNN-LSTM, ResNet-18, and EEG-FuseFormer in selection of different seizure prediction target applications. The results of the complexity analysis show that

selecting CNN-LSTM, ResNet-18, or EEG-FuseFormer models involves navigating a trade-off between performance and complexity based on the target seizure prediction application.

V. ACKNOWLEDGEMENTS

This project received funding from the European Union's Horizon Europe Excellent Science programme under the Marie Skłodowska-Curie Actions Grant Agreement [Grant Agreement No 101081457] and in part from CÚRAM, Science Foundation Ireland Research Centre under grant number 13/RC/2073_P2, as well as from Research Ireland-WIN Research Alliance Award. We also acknowledge the support and computational facilities provided by the Irish Centre for High-End Computing (ICHEC).

REFERENCES

- [1] World Health Organization, "Epilepsy," World Health Organization, 2024. [Online]. Available: <https://www.who.int/news-room/factsheets/detail/epilepsy>
- [2] I. Jemal, L. Abou-Abbas, K. Henni, A. Mitiche, and N. Mezghani, "Domain adaptation for EEG-based, cross-subject epileptic seizure prediction," *Frontiers in Neuroinformatics*, vol. 18, 2024.
- [3] Y. Zhang, Y. Guo, P. Yang, W. Chen, and B. Lo, "Epilepsy seizure prediction on EEG using common spatial pattern and convolutional neural network," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 2, pp. 465–474, 2020.
- [4] S. Jahan, F. Nowsheen, M. M. Antik, M. S. Rahman, M. S. Kaiser, A. S. M. S. Hosen, and I.-H. Ra, "Ai-based epileptic seizure detection and prediction in internet of healthcare things: A systematic review," *IEEE Access*, vol. 11, pp. 30 690–30 725, 2023.
- [5] N. D. Truong, A. D. Nguyen, L. Kuhlmann, M. R. Bonyadi, J. Yang, S. Ippolito, and O. Kavehei, "Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram," *Neural Networks*, vol. 105, pp. 104–111, 09 2018.
- [6] T. Dissanayake, T. Fernando, S. Denman, S. Sridharan, and C. Fookes, "Deep learning for patient-independent epileptic seizure prediction using scalp EEG signals," *IEEE Sensors Journal*, vol. 21, no. 7, pp. 9377–9388, 2021.
- [7] K. M. Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A Long Short-Term Memory deep learning network for the prediction of epileptic seizures using EEG signals," *Computers in Biology and Medicine*, vol. 99, pp. 24–37, 08 2018.
- [8] X. Wu, Z. Yang, T. Zhang, L. Zhang, and L. Qiao, "An end-to-end seizure prediction approach using long short-term memory network," *Frontiers in Human Neuroscience*, vol. 17, 05 2023.
- [9] H. Daoud and M. Bayoumi, "Efficient epileptic seizure prediction based on deep learning," *IEEE Transactions on Biomedical Circuits and Systems*, pp. 1–1, 2019.
- [10] D. Lee, B. Kim, T. Kim, I. Joe, J. Chong, K. Min, and K. Jung, "A ResNet-LSTM hybrid model for predicting epileptic seizures using a pretrained model with supervised contrastive learning," *Scientific Reports*, vol. 14, p. 1319, 01 2024.
- [11] J. Zhang, S. Zheng, W. Chen, G. Du, Q. Fu, and H. Jiang, "A scheme combining feature fusion and hybrid deep learning models for epileptic seizure detection and prediction," *Scientific Reports*, vol. 14, 07 2024.
- [12] A. John, H. Wang, B. Cardiff, K. K. Parhi, and D. John, "Multimodal fusion for robust respiratory rate estimation in wearable sensing," *Information Fusion*, vol. 123, p. 103253, 2025.
- [13] A. John, K. K. Parhi, B. Cardiff, and D. John, "Mlfusion: Multilevel data fusion using cnns for atrial fibrillation detection," *Computers in Biology and Medicine*, vol. 189, p. 109901, 2025.
- [14] Y. Ma, Z. Huang, J. Su, H. Shi, D. Wang, S. Jia, and W. Li, "A multi-channel feature fusion CNN-Bi-LSTM epilepsy EEG classification and prediction model based on attention mechanism," *IEEE Access*, vol. 11, pp. 62 855–62 864, 2023.
- [15] R. V. Godoy, T. J. S. Reis, P. H. Polegato, G. J. G. Lahr, R. L. Saute, F. N. Nakano, H. R. Machado, A. C. Sakamoto, M. Becker, and G. A. P. Caurin, "EEG-based epileptic seizure prediction using temporal multi-channel transformers," 2022.
- [16] S. Shafieezadeh, G. Marco Duma, M. Pozza, and A. Testolin, "A systematic review of cross-patient approaches for eeg epileptic seizure prediction," *Journal of neural engineering*, vol. 21, no. 6, 2024.
- [17] J. Gutttag, "CHB-MIT scalp EEG database (version 1.0.0)," *PhysioNet*, 2010.
- [18] J. L. Ba, J. R. Kiros, and G. E. Hinton, "Layer normalization," *arXiv:1607.06450 [cs, stat]*, 07 2016. [Online]. Available: <https://arxiv.org/abs/1607.06450>
- [19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *arXiv.org*, 12 2015. [Online]. Available: <https://arxiv.org/abs/1512.03385>
- [20] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," 2023. [Online]. Available: <https://arxiv.org/abs/1706.03762>
- [21] L. J. P. W. H. H. P. J. and W. F., "STGATE: Spatial-temporal graph attention network with a transformer encoder for EEG-based emotion recognition," *Frontiers in Human Neuroscience*, vol. 17, no. 1169949, 2023.
- [22] A. John, B. Cardiff, and D. John, *Deep Learning and Signal-Processing Methods for Multisensor Data Fusion. Applications to Ambulatory Health Monitoring*, 1st ed. Springer Cham, 2026.
- [23] I. Jemal, N. Mezghani, L. Abou-Abbas, and A. Mitiche, "An interpretable deep learning classifier for epileptic seizure prediction using eeg data," *IEEE Access*, vol. 10, pp. 60 141–60 150, 2022.