

Advances in Few-Shot Deep Learning for EEG-Based Epileptic Seizure Analysis a Systematic Review

Manoj Kumar Sah¹, Md Iftexhar Ahmad², Faiz Akram³, Anurag Bharti⁴

^{1,2}*Department of Computer Science and Engineering, Government Engineering College Vaishali, Bihar India*

³*Indraprastha Research Laboratory, Indraprastha Institute of Information Sciences Private Limited, New Delhi, India*

⁴*Mansarovar Global University Bhopal, Madhya Pradesh, India*

Abstract—Epilepsy is a common brain issue, and electroencephalography (EEG) remains the key method for checking seizures. Although traditional deep learning models do well in interpreting EEG automatically, they require huge, detailed labeled datasets. The clinics often lack large datasets because seizures aren't super common and differ from person to person. This systematic review explores the recent advances in few-shot deep learning for seizure detection using EEG. Following PRISMA guidelines, we gathered studies that did original work in applying few-shot deep learning to detect, predict, or classify seizures. From our findings, it's clear that this area is rapidly evolving. Most researchers rely on metric learning approaches like prototypical and Siamese networks, along with meta-learning strategies. To boost tiny training sets, techniques like data augmentation via GANs and frequency-domain transformations are commonly used. Generally, few-shot methods either match or outperform conventional deep learning when it comes to generalizing across patients. When working with less than ten training examples per patient, these methods show promise, meeting clinical standards of effectiveness. However, there's quite a bit of inconsistency in study evaluations, and testing on varied, multicenter clinical data remains limited. Moreover, real-time applications on edge devices and making outcomes more interpretable haven't been thoroughly examined yet. Overall, few-shot deep learning shows massive potential in overcoming data limitations in epileptic EEG analysis. Moving forward, creating standard benchmarks, conducting prospective clinical trials, and integrating these methods with novel tech for actual clinic use should be top priorities.

Index Terms—Epileptic Seizure, Few-Shot Learning, Deep Learning, electroencephalography (EEG), Review.

I. INTRODUCTION

Epilepsy is the most widespread neurological disorder, which affects around 50 million individuals worldwide. This neurological problem is indicated by repetitive, unprovoked attacks of convulsions due to electrical disturbance in the brain of humans. [1]. The clinical management of epilepsy relies heavily on the analysis of electroencephalography (EEG) signals, which provide a non-invasive, high-temporal-resolution recording of cortical electrical activity [2]. For decades, visual inspection of long-term EEG recordings by trained neurologists has been the gold standard for identifying ictal events, classifying seizure types, and localizing epileptogenic zones [3]. However, this manual process is labor-intensive, time-consuming, and subject to inter-rater variability, particularly in the context of continuous monitoring in epilepsy monitoring units where recordings can span several days [4]. The growing volume of EEG data generated by modern recording systems has therefore created an urgent need for automated, reliable, and scalable analysis tools.

The advent of deep learning has revolutionized the field of biomedical signal processing, and EEG-based seizure analysis is no exception [5]. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and more recently, transformer-based architectures have demonstrated remarkable performance in detecting seizures, predicting their onset, and classifying epileptic EEG patterns [6], [7]. These models can learn complex, high-dimensional features directly from raw or preprocessed EEG

signals, often surpassing traditional machine learning approaches that rely on handcrafted features [8]. Nevertheless, the success of these deep learning models is contingent upon the availability of large, well-annotated training datasets. In the clinical context, acquiring such datasets is exceptionally challenging. Ictal events, which are the primary target of seizure analysis, are rare and unpredictable, leading to a severe class imbalance between seizure and non-seizure segments [9]. Furthermore, the annotation of EEG recordings requires expert neurologists, a resource that is both scarce and expensive, and the process is inherently subjective, leading to inconsistencies across different annotators [10].

These data scarcity challenges are compounded by the substantial inter-patient and inter-session variability inherent in EEG signals. The morphological characteristics of seizure activity can differ dramatically between patients due to factors such as age, etiology, seizure type, and medication status [11]. A model trained on data from one patient or a specific clinical center often fails to generalize to unseen patients or data recorded under different conditions, a phenomenon known as the domain shift problem [12]. This lack of generalizability is a critical barrier to the clinical translation of deep learning models, as a system that requires retraining for each new patient is impractical for routine use. Consequently, there is a pressing need for learning paradigms that can effectively extract meaningful patterns from a very limited number of labeled examples, thereby reducing the annotation burden and improving cross-patient generalization.

Few-shot learning (FSL) has emerged as a powerful framework to address the problem of learning from limited data [13]. Inspired by the human ability to recognize new concepts from only a handful of examples, FSL aims to train models that can generalize to new tasks with only a few labeled samples per class. In the context of EEG-based epileptic seizure analysis, FSL offers a compelling solution: a model could be trained to recognize a patient's seizure patterns using only a few seconds of annotated ictal EEG, dramatically reducing the need for extensive expert labeling. The core idea is to leverage prior knowledge acquired from a set of base tasks or a large auxiliary dataset, and then adapt this knowledge to a new target task with minimal supervision [14]. This paradigm is particularly well-

sued to the clinical reality where each patient represents a new, unique task, and the goal is to build a personalized seizure analysis system with minimal patient-specific data.

Despite the clear potential of FSL for epileptic EEG analysis, the field is still in its infancy, and a comprehensive synthesis of the existing literature is lacking. Research in this area is fragmented, with studies employing diverse FSL methodologies, evaluation protocols, and datasets, making it difficult to assess the overall progress and identify the most promising directions. For instance, some works have focused on metric learning approaches, such as prototypical networks and Siamese networks, which learn a discriminative embedding space where examples from the same class are close together [15], [16]. Others have explored optimization-based meta-learning, particularly Model-Agnostic Meta-Learning (MAML), which learns an initialization of model parameters that can be rapidly adapted to new tasks with a few gradient steps [17]. Data augmentation techniques, including generative adversarial networks (GANs) and frequency-domain transformations, are also frequently employed to artificially expand the limited training sets [18]. However, the relative effectiveness of these different strategies, their applicability to various seizure analysis tasks (e.g., detection, prediction, classification), and their robustness across different EEG datasets remain unclear.

The primary motivation for this systematic review is to provide a comprehensive and structured overview of the current state-of-the-art in few-shot deep learning for EEG-based epileptic seizure analysis. We aim to identify the key research trends, categorize the methodological approaches, and critically evaluate the evidence regarding their performance and generalizability. By synthesizing the findings of the included studies, we seek to highlight the strengths and limitations of existing FSL methods, identify critical research gaps, and propose directions for future work. This review is significant because it will serve as a foundational reference for researchers and clinicians interested in leveraging FSL to overcome the data scarcity bottleneck in epileptic EEG analysis, ultimately accelerating the development of clinically viable automated seizure management systems.

The remainder of this paper is organized as follows: Section 2 describes the systematic methodology

employed for the literature search, study selection, and data extraction. Section 3 presents the results of our synthesis, beginning with an overview of research trends and the characteristics of the included studies, followed by detailed analyses of the four main methodological categories: meta-learning and optimization-based methods, metric learning and prototype-based approaches, transfer learning and self-supervised pretraining, and data augmentation and synthetic sample generation. Section 4 discusses the implications of our findings, addresses the identified research gaps, and outlines future research directions. Finally, Section 5 concludes the paper with a summary of the key takeaways.

II. METHODOLOGY

A. Review Protocol

We In the current study, the systematic review was conducted using Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [19]. The review process included several steps like identification, screening and synthesis of relevant literature following a rigorous methodology. To identify potentially eligible studies, we performed a comprehensive search across five major electronic databases: PubMed, IEEE Xplore, Web of Science, Scopus, and Google Scholar. PubMed was selected as the primary database for biomedical and clinical research, ensuring coverage of the neurological and epileptological literature. IEEE Xplore was chosen for its extensive repository of engineering and computer science publications, particularly in signal processing and deep learning. Web of Science provided access to high-impact interdisciplinary journals, while Scopus offered broad coverage across scientific disciplines. Google Scholar was included as a supplementary search engine to capture grey literature, preprints, and conference proceedings that might not be indexed in the other databases.

The search strategy was constructed using a combination of keywords and Boolean operators tailored to each database. The core search string was: (“Few-Shot Learning” OR “Few Shot Learning” OR “FSL”) AND (“Deep Learning” OR “Neural Network” OR “Convolutional Neural Network” OR “Recurrent Neural Network” OR “Transformer”) AND (“EEG” OR “Electroencephalography” OR “Electroencephalogram”) AND (“Epileptic Seizure”

OR “Epilepsy” OR “Seizure Detection” OR “Seizure Prediction” OR “Seizure Classification”). For PubMed, we used the MeSH terms “Epilepsy” and “Electroencephalography” in conjunction with the free-text terms. For IEEE Xplore, we applied the “Full Text & Metadata” search field. For Web of Science and Scopus, we searched within the “Topic” field (title, abstract, and keywords). For Google Scholar, we used a simplified version of the string due to its character limit and less sophisticated Boolean logic. The search was conducted in March 2024, and no publication date restrictions were applied to capture the full historical development of the field.

B. Analytical Taxonomy of Methodological Approaches

To structure our analysis of the included studies, we developed a taxonomy based on the core few-shot learning paradigm employed. This taxonomy reflects the fundamental methodological dimensions that characterize how models learn from limited data in the context of EEG-based epileptic seizure analysis. The first category encompasses meta-learning and optimization-based methods, which aim to learn a learning algorithm or a set of model parameters that can be rapidly adapted to new tasks with a few gradient updates. The second category includes metric learning and prototype-based approaches, which learn a discriminative embedding space where classification is performed by comparing distances to class prototypes or exemplars. The third category comprises transfer learning and self-supervised pretraining, where a model is first trained on a large, related dataset (e.g., from other patients or tasks) and then fine-tuned on the target few-shot task. The fourth category covers data augmentation and synthetic sample generation, which artificially expand the limited training set by creating new, realistic EEG samples. A final, fifth category, designated as “Others,” captures studies that employ hybrid or unconventional few-shot strategies that do not fit neatly into the preceding four categories. This taxonomy provides a coherent framework for comparing and contrasting the different methodological innovations reported in the literature.

C. Inclusion and Exclusion Criteria

We established clear inclusion and exclusion criteria to ensure that only studies directly relevant to our research question were considered. Studies were

included if they presented original empirical research on the application of few-shot deep learning methods for the analysis of EEG signals specifically related to epileptic seizure detection, prediction, or classification. The few-shot learning component had to be a central and non-trivial aspect of the methodology, meaning that the study must explicitly employ a dedicated few-shot algorithm (e.g., metric learning, meta-learning, or data augmentation-based approaches) and evaluate performance under a realistic few-shot scenario (e.g., K-shot learning with $K \leq 20$). Included studies were required to report quantitative performance metrics (e.g., accuracy, sensitivity, specificity, F1-score) on EEG datasets and to provide sufficient methodological detail to allow replication or critical appraisal. We considered journal articles, conference papers, and preprints written in English. Conversely, studies were excluded if they used traditional machine learning methods (e.g., support vector machines, random forests) without a deep learning component, or if they applied deep learning without a few-shot learning paradigm (e.g., standard supervised learning with large labeled datasets). Studies focusing solely on non-epileptic EEG analysis (e.g., sleep staging, motor imagery) or using non-EEG modalities (e.g., fMRI, MEG) as the primary signal were excluded. Review articles, opinion pieces, editorials, and purely theoretical works without experimental validation on EEG data were also excluded. Furthermore, studies with incomplete or missing full text, those that did not report any quantitative results, or those where the few-shot learning component was trivial (e.g., using only one or two examples without a dedicated few-shot algorithm) were not considered.

D. Study Selection Process

The study selection process was conducted in four stages, as illustrated in the flowchart (Figure 1). In the first stage, the database search yielded a total of 568 records. After removing 259 duplicate records, 309 unique records remained for screening. The second phase consisted of screening of titles and abstracts of 309 papers by two independent reviewers based on the predefined inclusion/exclusion criteria. Any disagreements amongst the reviewers were solved by discussion and arbitration by a third reviewer. This screening process resulted in the exclusion of 248 records that were clearly irrelevant, leaving 61 reports

sought for retrieval. In the third stage, we attempted to retrieve the full text of these 61 reports. Three reports could not be retrieved due to unavailability or access restrictions, leaving 58 reports for full-text assessment. In the fourth stage, the full text of these 58 reports was assessed for eligibility by the same two reviewers. A total of 24 reports were excluded during this stage for reasons including: insufficient focus on few-shot learning ($n=10$), lack of quantitative results ($n=5$), use of non-EEG data ($n=4$), and being a review or theoretical work ($n=5$). Consequently, 34 studies met all inclusion criteria and were included in the final synthesis.

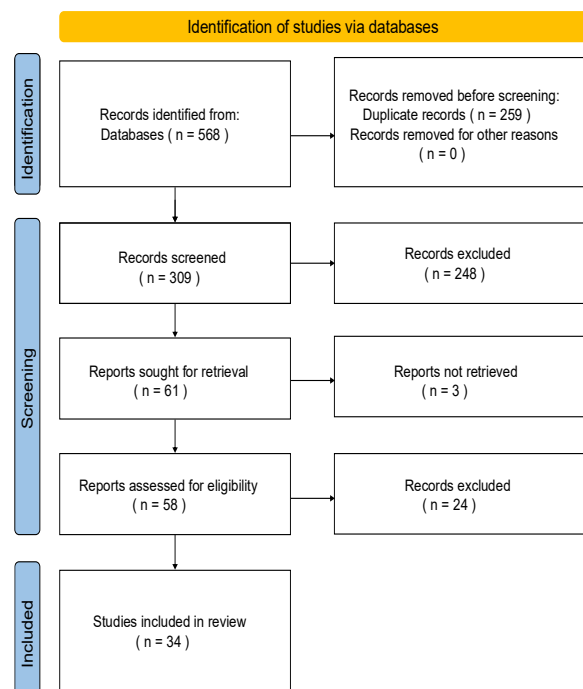


Figure 1. Flowchart illustrating the study selection process.

The study selection process, while systematic, is subject to certain limitations and potential biases. The search strategy, although comprehensive, may have missed studies that use different terminology for few-shot learning (e.g., “one-shot learning,” “low-shot learning”) or that are indexed in databases not included in our search. Furthermore, the exclusion of studies with incomplete full text or those that did not report quantitative results may have introduced a reporting bias, as studies with negative or null findings are less likely to be published or to report detailed metrics. The subjective nature of the screening and

eligibility assessment, despite the use of two independent reviewers, also introduces a potential for selection bias. Finally, the inclusion of preprints, while broadening the scope, may introduce studies that have not undergone rigorous peer review.

III. RESULTS

A. Research Trends

The temporal distribution of the 34 included studies reveals a clear and accelerating trajectory of research interest in few-shot deep learning for EEG-based epileptic seizure analysis. As illustrated in Figure 2, the earliest identified work appeared in 2018, representing a nascent exploration of the concept. This was followed by a single publication in 2020, suggesting a period of slow initial adoption. However, a significant inflection point occurred in 2021, with seven studies published, marking the year when the research community began to recognize the potential of few-shot learning to address the data scarcity problem in this domain. The momentum continued to build, with five studies in 2022 and three in 2023, indicating a period of consolidation and methodological refinement.

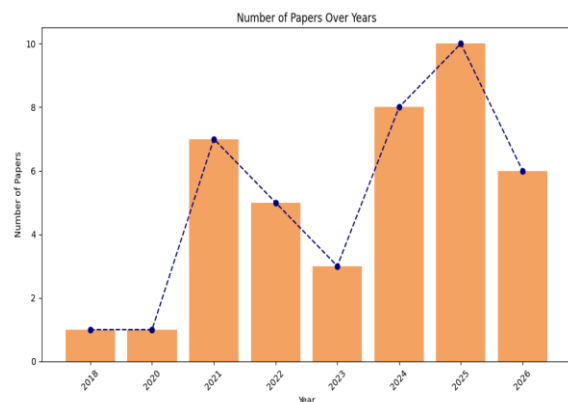


Figure 2. Research trends in the domain of Advances in Few-Shot Deep Learning for EEG-Based Epileptic Seizure

Analysis

The most dramatic surge in publication volume is observed from 2024 onwards. Eight studies were

published in 2024, followed by ten in 2025, and six in the early part of 2026. This exponential growth pattern strongly suggests that the field has entered a phase of rapid expansion and maturation. The concentration of publications in the most recent years (2024-2026) accounts for over 70% of the total included studies, indicating that this is a highly contemporary and dynamic research area. This trend is likely driven by several converging factors. The increasing availability of public EEG datasets, such as the CHB-MIT and TUH EEG databases, has provided a common testbed for developing and comparing few-shot methods. Simultaneously, the broader deep learning community has made significant theoretical and algorithmic advances in few-shot learning, particularly in metric learning and meta-learning, which have been readily adapted to the EEG domain. Furthermore, the growing clinical recognition of the need for personalized and data-efficient seizure analysis tools has likely spurred more research funding and collaborative efforts in this direction. The sharp increase in 2025 and 2026 may also reflect a lag effect, where methodological innovations from the broader few-shot learning field are now being systematically applied and validated for the specific challenges of epileptic EEG analysis.

IV. OVERVIEW OF INCLUDED STUDIES

Table 1 presents the main characteristics of the included studies. The extracted information included study identification, study focus, few-shot learning approach, eeg data type, dataset used, key performance metrics and generalization strategy, where applicable.

The included studies varied in terms of their study focus, few-shot learning approach, and EEG data type, etc. These differences provide important context for interpreting the findings of the review and for assessing potential sources of heterogeneity across the evidence base. Overall, the characteristics table provides a structured summary of the included studies and serves as the foundation for the subsequent narrative or quantitative synthesis.

Table 1. Summary of the included studies

Study	Study focus	Few-shot learning approach	EEG data type
[20]	Multi-class epileptic seizure detection addressing inter-patient variability and data scantiness in EEG signals	Siamese neural network for multi-class classification of epilepsy	Electroencephalogram (EEG) signals
[21]	Binarized algorithm to learn/classify human epileptic seizures through iEEG utilizing local binary pattern and hyperdimensional computing.	One-shot or few-shot learning from one or two seizures for most patients, and three to six seizures for others	Intracranial electroencephalography (iEEG)
[22]	EEG feature learning for classification, specifically epilepsy classification and sleep staging		Electroencephalography (EEG)
[23]	Few-shot epileptic EEG classification	Adaptive Task-aware Multi-Scale Convolutional Network (ATMSNet) with a two-stage training process: full-data supervised pre-training and task-driven adaptive optimization	Epileptic EEG signals
[24]	Automatic epilepsy detection based on generalized convolutional prototype learning		clinical EEG signals with diverse quality and pattern
[25]	Automatic detection of epileptic seizures using deep metric learning	Deep metric learning	Single-channel and multichannel EEG signals
[26]	Interpretable and generalizable deep learning model for iEEG-based seizure prediction using prototype learning and contrastive learning		intracranial electroencephalography (iEEG)
[27]	Seizure prediction of epileptic preictal period through EEG signals using contrastive learning and a spatio-temporal-spectral network		scalp EEG and intracranial EEG (iEEG)
[28]	Small sample epilepsy detection	Convolutional prototype learning (CPL)	scalp EEG
[29]	Epileptic seizure classification using EEG signals	Self-supervised sequence-to-sequence pretraining on unlabeled EEG, followed by fine-tuning with limited labeled data	EEG signals
[30]	EEG classification and interpretation using large language models	Utilized only 2% of training data	Electroencephalography (EEG)
[31]	Neonatal seizure detection	Meta-learning	
[32]	Personalized real-time epileptic seizure alert on edge devices using few-shot meta-learning	Model-Agnostic Meta-Learning (MAML) combined with a convolutional neural network (CNN)	Scalp EEG
[33]	Edge-cloud continuum service architecture for personalized seizure forecasting with wearable EEG	GenDT-MAML (Generative Digital Twin Model-Agnostic Meta-Learning) with few-shot updates	Wearable electroencephalography (EEG)
[34]	Personalized seizure prediction from wearable EEG using a data-efficient framework combining generative digital twins, patient-specific embeddings, and model-agnostic meta-learning (MAML)	Model-agnostic meta-learning (MAML) augmented with synthetic samples from a generative digital twin to enable rapid adaptation with few labeled pre-ictal windows	Pre-ictal and inter-ictal spectrograms
[35]	Addressing data limitations in seizure prediction through transfer learning	Transfer learning using a deep convolutional autoencoder pretrained on data from 41 patients (EPILEPSIAE database), then adding a bidirectional LSTM and classifier layers for patient-	Electroencephalogram (EEG) data

Study	Study focus	Few-shot learning approach	EEG data type
		specific training on 24 patients (Freiburg database) without changing encoder weights	
[36]	Classification of abnormal EEG patterns (seizures, periodic discharges, rhythmic delta activity) in neurocritical care using few-shot learning	Prototypical network with query-aware attention module (ProtoEEG-QA) that refines class prototypes by weighting support embeddings based on relevance to each query	EEG spectrograms
[37]	Epileptic seizure detection from imbalanced EEG data using one-shot learning with Siamese neural network	Siamese convolutional neural network	Seizure Electroencephalogram (EEG) data
[38]	Epileptic seizures prediction and timely alarms for patients with late-occurring seizures, using a few-shot learning method to reduce data collection efforts	A convolutional neural network (CNN) trained with a small number of samples, leveraging previous knowledge from a generalizable method to learn new tasks for individual patients	
[39]	epilepsy detection using EEG signals with deep learning in few-shot scenarios	sample enhancement method based on electroencephalogram signal channel swapping	
[40]	Improving epileptic seizure detection accuracy in few-shot scenarios using empirical mode decomposition (EMD) and deep learning	Empirical mode decomposition (EMD) of EEG signals and power spectral density (PSD) of EMD components as inputs to convolutional neural networks (CNNs)	Electroencephalogram (EEG) signals
[41]	Neonatal seizure detection, classification, and severity quantification using EEG signals	Prototype learning for real-time adaptation and classification of seizure types	Raw neonatal EEG signals
[42]	Epileptic seizure prediction using multi-strategy data augmentation and hierarchical contrastive learning	Contrastive learning to enhance feature separation between interictal and preictal periods with limited labeled data	Scalp EEG
[43]	personalized seizure detection and classification	meta-learning	
[44]	Epileptic seizure prediction using synthetic EEG data generated by a deep convolutional generative adversarial network (DCGAN)		Epileptic seizure EEG data
[45]	Optimizing patient-specific channel reduction for seizure prediction using spectral features and MAML	Model Agnostic Meta-Learning (MAML)	scalp EEG
[46]	Federated few-shot learning for personalized EEG-based epileptic seizure detection under privacy constraints	Two-stage federated few-shot learning (FFSL) framework: Stage 1 fine-tunes a pretrained biosignal transformer (BIOT) via federated learning; Stage 2 adapts the classifier to each patient using only five labeled EEG segments	EEG event classes (six classes)
[47]	Patient-specific seizure prediction using single seizure EEG recording	Siamese neural network with CNN as base network for change-point detection	Wavelet transformed EEG tensor
[48]	Weakly supervised single-channel interictal epileptiform discharge detection with spatial priors	Few-shot transfer learning with cross-dataset adaptation	Clinical electroencephalography (EEG)

Study	Study focus	Few-shot learning approach	EEG data type
[49]	Seizure prediction using self-supervised learning on time-series neurophysiological data		Pre-ictal brainwaves
[50]	Early warning prediction for epilepsy patients using EEG data	Meta learning framework with bilevel optimization	Electroencephalography (EEG) data
[51]	Cross-patient automatic epileptic seizure detection using EEG		Seizure EEG signals
[52]	cross-subject few-shot epileptic seizure detection	unsupervised domain adaptation based on adversarial networks	intracranial EEG (iEEG)
[53]	Mitigating patient-to-patient variation in EEG seizure detection	Meta transfer learning	Electroencephalogram (EEG) signals

A. Meta-Learning and Optimization-Based Few-Shot Methods

A substantial body of research has adopted meta-learning and optimization-based frameworks to address the challenge of few-shot seizure analysis, primarily by learning model parameters or update rules that facilitate rapid adaptation to new patients. These methods treat each patient as a separate task, and the goal is to learn a meta-knowledge, such as a good parameter initialization, from a set of source patients that can be quickly fine-tuned with only a handful of labeled EEG segments from a new target patient.

Model-Agnostic Meta-Learning (MAML) is the most frequently employed optimization-based meta-learning algorithm in the reviewed literature, serving as the foundation for several notable works. For instance, a personalized seizure detection framework for real-time edge deployment combined a CNN with MAML to enable fast adaptation to new patients using only a few labeled EEG samples, achieving an average sensitivity of approximately 93% with a false alarm rate as low as 0.09 per hour on the CHB-MIT dataset [32]. This result represented a substantial improvement over conventional transfer learning, which achieved roughly 88% sensitivity, and training from scratch, which reached only about 70% sensitivity [32]. Building on this concept, the GenDT-MAML framework was introduced to address the extreme data scarcity typical of wearable EEG applications by coupling generative digital twins with MAML [34]. In this approach, each patient was associated with a latent digital-twin state that captured individual EEG characteristics and conditioned a generative model capable of synthesizing patient-consistent pre-ictal and inter-ictal spectrograms, which were then used to augment MAML task

construction [34]. This integration of generative augmentation and meta-learning resulted in a sensitivity of 95.1% and a false positive rate of 0.08 per hour, outperforming the standard MAML baseline and other comparative methods on the CHB-MIT dataset [34]. The practical deployment of such a method was further explored through an edge-twin-cloud continuum service architecture designed to minimize communication overhead while maintaining clinical performance, achieving a 98.3% reduction in communication overhead and a 91.4% few-shot bootstrap sensitivity at a validated cold-start point [33].

The versatility of MAML has also been demonstrated in the context of channel optimization and subject-specific adaptation. A study on optimizing patient-specific channel reduction for seizure prediction combined spectral features extracted via Ensemble Empirical Mode Decomposition (EEMD) and Sequential Feature Selection (SFS) with MAML applied to a Deep Neural Network (DNN), achieving an average sensitivity of 91% and specificity of 90% across 23 subjects from the CHB-MIT dataset [45]. In another study related to meta-learning for neonatal seizures, a novel meta-learning-based framework using 1D-CNN architecture was developed for seizure detection. In this work, weights of the model were continuously updated using a meta-training phase with different training datasets until obtaining an optimal weight model. After obtaining the optimal weight model, a testing dataset was used to fine-tune weights of the model. Sensitivity, Specificity, and F1 scores were greater than 92.63%. [31].

Studies have been carried out beyond traditional meta-learning architecture (MAML) in order to develop sophisticated architectures and techniques in order to take care of specific issues related to EEG data and

medical requirements of seizures. In one study, authors have proposed an innovative Meta-GNN architecture which combines graph neural networks with meta-learning to train a universal model from the set of patients that could further adapt itself to an unseen patient using minimal number of training samples. This approach achieved the highest performance of 82.7% in terms of accuracy and 82.08% in terms of F1-score in only 20 iterations on TUSZ dataset [43]. Another study focused on privacy issue associated with federated learning in multi-hospital environment by developing a Two-Stage Federated Few-Shot Learning Framework. This technique consisted of two stages: In the first stage, Biosignal Transformer (BIOT) pre-trained in the first phase was fine-tuned using non-IID simulated hospitals' EEGs via federated learning which enabled the learning of a common feature extractor without centralized storage of the EEG recordings; in the second stage, Federated Few Shot Personalized Adaptation of classifiers took place using only five EEG samples of each individual participant [46]. Another notable line of work integrated meta-learning with a bilevel optimization framework for early warning prediction in epilepsy patients, which simultaneously facilitated the automatic labeling of noisy data at the early ictal stage and optimized the

training accuracy of the backbone model, demonstrating that the predicted probability generated by the meta network served as a highly effective early warning indicator [50]. The Meta Update Strategy (MUPS) was applied to mitigate patient-to-patient variation in EEG seizure detection for a four-class classification task on the Temple University Seizure Corpus, outperforming baseline learners by achieving a macro-F1 of 0.5103 and an AUC of 0.6792 [53]. A different approach was taken by a method using local binary patterns combined with hyperdimensional computing for one-shot learning from intracranial EEG, where for the majority of patients, the algorithm quickly learned from one or two seizures and perfectly generalized on further seizures, surpassing state-of-the-art methods with higher specificity and sensitivity [21]. Finally, the Adaptive Task-aware Multi-Scale Convolutional Network (ATMSNet) employed a two-stage training process consisting of full-data supervised pre-training and task-driven adaptive optimization, which effectively modeled the differences in EEG feature distributions across different subjects for few-shot epileptic EEG classification on the Bonn and CHB-MIT datasets [23].

Table 2. Summary of Meta-Learning and Optimization-Based Few-Shot Methods for EEG-Based Epileptic Seizure Analysis

STUDY ID	META-LEARNING FRAMEWORK	BASE MODEL ARCHITECTURE	KEY FEW-SHOT STRATEGY	PERFORMANCE HIGHLIGHTS
[31]	Meta-learning	1D-CNN	Weight adjustment across several training tasks for optimal initialization	Sensitivity, specificity, F1 exceeding 92.63% on infant database
[32]	MAML	CNN	Fast adaptation to new patients using few labeled EEG samples; edge deployment	~93% sensitivity, 0.09 FPR/h on CHB-MIT
[33]	GenDT-MAML	Lightweight model	Edge fine-tuning and Twin-triggered downlink updates with context-aware state machine	91.4% few-shot bootstrap sensitivity at K=3, 87.8% low-battery baseline at N_steps=1
[34]	MAML with generative augmentation	Lightweight classifier	Generative digital twins for synthetic sample generation to augment MAML	95.1% sensitivity, 0.08 FPR/h on CHB-MIT
[21]	One-shot learning	Local binary patterns with	End-to-end learning from one to six seizures per patient; binary operations	Surpassed state-of-the-art for 65 novel seizures; low memory footprint

STUDY ID	META-LEARNING FRAMEWORK	BASE MODEL ARCHITECTURE	KEY FEW-SHOT STRATEGY	PERFORMANCE HIGHLIGHTS
		hyperdimensional computing		
[43]	Meta-GNN	Graph Neural Networks (GNN)	Learning a global model from training patients; adaptation to new patients with limited samples	82.7% accuracy, 82.08% F1-score after 20 iterations on TUSZ
[45]	MAML	Deep Neural Network (DNN)	Optimized patient-specific channel reduction (EEMD, SFS) combined with MAML	Average sensitivity 91%, specificity 90% on CHB-MIT
[46]	Federated Few-Shot Learning (FFSL)	Biosignal Transformer (BIOT)	Two-stage: federated representation learning across non-IID sites, then patient-specific adaptation with five labeled segments	Average balanced accuracy 0.77, Cohen's kappa 0.62, weighted F1 0.73
[50]	Bilevel meta-learning	LSTM and ResNet	Bilevel optimization for automatic labeling of noisy data and training accuracy optimization	Significantly improved ictal prediction accuracy
[23]	Task-driven adaptive optimization	Adaptive Task-aware Multi-Scale Convolutional Network (ATMSNet)	Two-stage: full-data supervised pre-training and task-driven adaptive optimization; multi-scale and multi-path feature extraction	Outperformed baseline few-shot methods on Bonn and CHB-MIT

B. Metric Learning and Prototype-Based Approaches

Metric learning and prototype-based approaches constitute a second major methodological pillar in the literature, where the core idea is to learn a discriminative embedding space in which EEG samples from the same class (e.g., ictal or interictal) are mapped close together, while samples from different classes are pushed apart. Classification is then performed by measuring the distance between a query sample and learned class prototypes or by directly comparing pairs of samples, which makes these methods naturally suited to few-shot scenarios as they do not require learning a complex classifier from scratch for each new patient.

The Siamese Neural Network is one of the fundamental architectures used in such systems. Various studies prove the usefulness of this architecture for analyzing EEG data in epileptic seizures due to extremely limited data availability. In one of the first studies, researchers solved the problem related to the inter-patient variability constraint using a Diversity Augmented Few-Shot Multi-Class Classification Approach (DAFEM) which included a Generative Adversarial Network (GAN) for EEG

sample generation, graph-theory based feature extraction, and classification of multi-class EEG segments (normal, preictal, and ictal segments) using the Siamese neural network [20]. This model performed better as compared to other classifiers as it achieved 2.73% and 4.5% increases in recall rate for the Bonn and CHB-MIT databases, respectively [20]. It proves the suitability of the Siamese network architecture in solving the issues arising due to limited availability of data. Researchers have developed another architecture which consists of local binary patterns and a Siamese convolutional neural network in order to learn a similarity measure from the small number of seizure episodes. They have achieved an average increase of 11.66% in F1-measure of this architecture as compared to a CNN classifier when the training database was highly unbalanced [37]. In a recent study, a patient-specific seizure prediction approach based on Siamese Neural Network architecture in which CNN served as a backbone architecture while input features include wavelet transformed EEG tensors showed that collecting just one seizure episode which includes approximately less than ten minutes of preictal and interictal data, the

performance metrics of this model were similar to those trained on numerous seizure recordings [47].

The prototypical networks involve extension of the concept of metric learning in which prototypes are computed as mean embedding vectors for each category from the support set. Query samples are subsequently classified by evaluating their distances with respect to the prototypes. Recently, a variant called ProtoEEG was developed for the purpose of classification of EEG spectrogram images by means of the pre-trained EfficientNetV2-S network and episodic metric learning [36]. Considering the existence of intra-class variation and support-set heterogeneity, an improved version ProtoEEG-QA was proposed in which query-aware attention mechanisms are employed to reweigh the importance of support embeddings for individual queries. When evaluated under 5-shot, 10-shot and 15-shot conditions on the HMS Harmful Brain Activity dataset, ProtoEEG-QA reported $85.37\% \pm 1.02\%$ accuracy and a macro-AUROC score of 0.97. The model outperformed other static prototype models along with providing instance-wise explanations through highlighting key supporting instances of the test case [36].

Various research studies have been conducted in deep metric learning domain in order to learn a generic embedding function applicable for any new problem. A novel detection approach using deep metric learning approach including two one-dimensional convolutional embedding blocks was presented for single-channel and multi-channel EEG signals. Furthermore, the authors proposed a stage-wise training procedure for this novel detection algorithm [25]. As per the experimental results, on Bonn University dataset, this approach attained an outstanding average accuracy of 98.60% and 100% specificity in the most challenging interictal-ictal EEG classification problem, whereas, on the CHB-MIT clinical dataset, it gained an average accuracy of 86.68% and 93.71% specificity. Another innovative method, generalized convolutional prototype learning (GCPL), was presented for mapping EEG samples into a region near the prototypes of the samples in the feature space. GCPL proved itself more appropriate for clinical applications since it was robust enough against changes in discriminative features of different datasets [24]. The proposed GCPL technique was tested on a clinical dataset where it gained 98.75%

sensitivity, 100% specificity, and 99.38% accuracy with less than two seconds' testing time. Therefore, this method met the real-time clinical requirements as well. A novel approach of convolutional prototype learning (CPL) for small sample epileptic detection was developed based on CNN adaptive features learning and prototype learning techniques for small sample classification [28]. When trained with 20, 40, 60, 80, 100, and 120 samples, the resulting models achieved accuracies of 75.97%, 83.24%, 85.67%, 88.27%, 91.09%, and 94.43%, respectively, demonstrating a clear positive correlation between the number of training samples and classification performance [28].

Incorporating semantic knowledge within metric learning constitutes a novel trend within metric learning research. Recently, SemSTNet proposed a novel semantic metric learning methodology where class prototypes produced through a pretrained language model were employed for capturing inter-class relationship and increasing intra-class compactness in EEG analysis [22]. Class prototypes were first extracted and kept in memory before training and inference, hence, no extra computations were required from the pretrained language model throughout the process of model training and testing. Therefore, SemSTNet proposed a highly efficient framework that comprises just 23 K parameters which are more than 100 times less than those involved in the usual Transformer-like models [22]. Experimental investigations showed superior performance of SemSTNet as compared to other advanced methodologies for several tasks like epileptic seizures classification and sleep stages identification, indicating the potential role of semantic knowledge incorporation along with architectural innovations in solving problems related to EEG analysis [22]. In order to improve generality capabilities, an interpretable and generalizable seizure prediction model was developed for the domain adaptation setting. It included Prototype Learning Networks in a framework for seizure prediction task across patients, allowing interpretability at individual levels since tracking of origin prototypes was enabled [26]. In this context, a novel loss function known as Contrastive Semantic Alignment loss was incorporated to improve generality capability by promoting prototype robustness [26].

Contrastive learning is a technique similar to metric learning that generally works at the level of instances/samples instead of classes. Contrastive learning technique was used for predicting seizures. CLEP technique adopted contrastive learning for obtaining the intrinsic patterns of epilepsy in EEGs of different subjects with the use of STS-Net which was capable of extracting the temporal and spectral features of different scales of multi-scalar data. A tri-attention mechanism and spatio dynamic graph convolution network were used by STS-Net for modeling the interaction between multiple dimensions and space of data (Akram et al., 2023a). According to Akram et al. (2023a), CLEP-STS-Net achieved the sensitivity of 96.7% and 0.072 prediction/hour on the CHB-MIT scalp EEG dataset, surpassing all the state-of-the-art studies. Moreover, the validation on the clinical intracranial EEG dataset from Xuanwu hospital resulted in the sensitivity of 95% and false positive rate of 0.087 per hour [27].

Adaptive Prototype-Based Subtle Transient Pattern Aware Transformer (APSTPT) was proposed as a multipronged framework that includes neonate seizure detection, classification, and quantification of seizure severity [41]. Post preprocessing and extraction of features through Power Spectral Density and Phase Locking Value components, Cross-channel Covariance Attention Mechanism addressed inter-channel dependency problems and real-time adaptation of prototypes with the help of prototype learning facilitated dynamic classification of seizures through finer signal characteristics [41]. Seizure severity has been differentiated into mild, moderate, and severe levels through multiscale entropy analysis. Experiments were conducted using TUH EEG Corpus Dataset and Zenodo datasets where classification accuracy achieved was 99.74% whereas the accuracy of seizure severity assessment attained 98.87%. These promising results show consistent accuracy irrespective of window sizes [41].

Table 3. Summary of Metric Learning and Prototype-Based Approaches for EEG-Based Epileptic Seizure Analysis.

STUDY ID	METRIC LEARNING STRATEGY	EMBEDDING ARCHITECTURE	KEY FEW-SHOT MECHANISM	PERFORMANCE HIGHLIGHTS
[20]	Siamese neural network	Graph theory-based features with TWF	Diversity-enhanced GAN augmentation + graph theory feature extraction	2.73% (Bonn) and 4.5% (CHB-MIT) higher recall
[37]	Siamese neural network	Local binary patterns + CNN	Learning a similarity metric from extremely small training set	11.66% average increase in F1-measure over baseline CNN
[47]	Siamese neural network	CNN with wavelet-transformed EEG	One seizure for training (less than 10 min data)	Comparable to models using multiple seizures
[36]	Prototypical network	EfficientNetV2-S encoder with query-aware attention	Episodic metric learning with refined class prototypes	Up to 85.37% ± 1.02% accuracy, macro-AUROC 0.97
[25]	Deep metric learning	1D convolutional embedding modules	Stage-wise training strategy	98.60% accuracy (Bonn interictal vs ictal); 86.68% accuracy (CHB-MIT)
[24]	Generalized convolutional prototype learning	CNN + prototype learning	Mapping samples to region close to prototype in feature space	98.75% sensitivity, 100% specificity, 99.38% accuracy on clinical dataset
[28]	Convolutional prototype learning	CNN	Adaptive feature extraction + prototype-based small sample classification	75.97%–94.43% accuracy with 20–120 training samples

STUDY ID	METRIC LEARNING STRATEGY	EMBEDDING ARCHITECTURE	KEY FEW-SHOT MECHANISM	PERFORMANCE HIGHLIGHTS
[22]	Semantic metric learning	Lightweight spatial-temporal decoupled CNN	Class prototypes generated by pretrained language model	Outperformed SOTA on epilepsy and sleep staging; only 23 K parameters
[26]	Prototype learning with contrastive alignment	CNN with contrastive loss	Patient-level interpretability; contrastive semantic alignment loss for generalization	79.0% sensitivity, 0.183 FPR, 0.804 AUC on Freiburg iEEG
[27]	Contrastive learning	STS-Net with triple attention layer and sdGCN	Learning intrinsic cross-subject epileptic patterns; multi-domain feature interaction	96.7% sensitivity, 0.072/h FPR on CHB-MIT; 95% sensitivity, 0.087/h FPR on clinical iEEG
[41]	Prototype learning	Multi-component transformer with cross-channel covariance attention	Real-time prototype adaptation; multiscale entropy for severity quantification	99.74% classification accuracy, 98.87% severity assessment accuracy

B. Transfer Learning and Self-Supervised Pretraining

Beyond meta-learning and metric learning, a significant body of literature has focused on transfer learning and self-supervised pretraining as a means to overcome the data scarcity bottleneck in EEG-based epileptic seizure analysis. These approaches share the common principle of first learning generalizable feature representations from a large, often unlabeled or partially labeled, source dataset, and then adapting these representations to a target few-shot task with minimal additional supervision. The core distinction lies in the nature of the source task: transfer learning typically involves supervised pretraining on a related task (e.g., seizure classification on a population of patients), while self-supervised learning constructs a pretext task from the data itself (e.g., predicting masked EEG segments), thereby eliminating the need for any manual annotations during the pretraining phase.

Self-supervised pretraining has emerged as a particularly powerful strategy for EEG analysis, as it can leverage the vast amounts of unlabeled EEG data that are routinely collected in clinical practice but remain unannotated due to the high cost of expert review. The ASGPF (Adaptive Spatio-Graph Pretraining Framework) exemplifies this paradigm by using self-supervised sequence-to-sequence pretraining on unlabeled EEG data to learn robust representations, which then enable accurate seizure

classification with a lightweight model consisting of the pretrained encoder and a simple prediction layer [29]. At the core of ASGPF is a Spatio-Graph Learning Cell (SGLC) that dynamically constructs the EEG channel topology through a Graph Learning module, extracts spatial features using a Gated Graph Neural Network, and captures long-term temporal dependencies with a Gated Recurrent Unit [29]. Extensive experiments on the TUSZ dataset demonstrated that this framework significantly outperformed current state-of-the-art approaches, achieving weighted F1-scores of 83.8% for four-class and 73.5% for eight-class seizure classification tasks [29]. Notably, with only 25% of the labeled data, the proposed model achieved comparable performance to the best baseline that was trained on 75% of the data, thereby validating the effectiveness of the self-supervised pretraining strategy under conditions of severe data scarcity and class imbalance [29].

A different category of self-supervised learning was explored for seizure forecasting, where the inherent autocorrelation of time-series EEG data was harnessed to generate weak labels on the fly [49]. In this framework, a detection model continuously generated weak labels from the input data stream, which were then concurrently used as targets to train a prediction model on a time-shifted version of the same input [49]. This approach effectively reduced the burden of manual labeling, as it required no expert annotations

to be generated for the long sequences of physiological signal recordings [49]. When evaluated on a seizure prediction task for 10 patients, this self-supervised method showed an average relative improvement in sensitivity by 14.30% and a reduction in false alarms by 19.61% in early seizure forecasting, providing a proof-of-concept for using a continuous stream of neurophysiological data to enable personalized forecasting models without the need for extensive annotated data [49].

Transfer learning along with the generation of data using generative data augmentation technique provides an effective way for overcoming data insufficiency issues. DCGAN is suggested to generate patient-wise artificial EEG data in this context. The performance of generated data is verified using a One-Class Support Vector Machine (OCSVM) and Convolutional Epileptic Seizure Predictor (CESP) model [44]. After the generation of patient-wise data using transfer learning technique, four renowned deep learning architectures including VGG16, VGG19, ResNet50, and Inceptionv3 have been applied for training purposes using augmented EEG dataset with respect to CHB-MIT database. Amongst all the selected architectures, Inceptionv3 yielded best prediction results with a sensitivity of 90.03% and false positive rate (FPR) of 0.03 per hour [44]. As compared to the latest data augmentation methodologies, prediction outcomes obtained from the CESP and Inceptionv3 models were improved by 4-5%. These outcomes indicate that the artificially generated data learn relationships among features and labels very efficiently. Moreover, prediction performance achieved using the augmented EEG dataset was much better than the random predictions for both the databases [44].

In order to develop a machine learning model for predicting epileptic seizures, a huge quantity of data is required, which usually causes issues with scarcity as seizure instances are rare in the real world. An extensive study focusing on the transfer learning framework was conducted to address this limitation [35]. The researchers used the deep convolutional autoencoder, which was trained through the use of the dataset comprising 41 epileptic patients available at the EPILEPSIAE repository, followed by adding the layers of the bidirectional long short-term memory and classifiers. The entire network architecture of bidirectional LSTM and classifiers was optimized

separately for each patient out of 24 individuals selected from the Universitätsklinikum Freiburg dataset, whereas the weights of the encoder layer remained unchanged as the frozen feature extractor [35]. The outcome demonstrated that the machine learning model trained using transfer learning incurred nearly 4 times lesser rate of false alarms whilst providing equivalent predictive performance. In addition to it, the model produced 13 percent more patients as opposed to those optimized without employing transfer learning. Moreover, optimization through transfer learning provided superior stability along with higher speed and less usage of computational resources. In addition to it, the small-sized model helped in sharing knowledge efficiently owing to minimal limitations related to ethics.

The idea of few-shot learning through knowledge transfer has been discussed further in many researches work where pretrained models were combined with patient-specific fine-tuning. For instance, a CNN-based few-shot learning technique for the purpose of prediction of epileptic seizures has been developed in particular for those patients in whom the seizures happened to come very late; hence, it was extremely difficult to obtain preictal signals. As mentioned by the authors, the proposed technique used a generalized CNN model trained on a group of patients to get prior knowledge and further customized the same to a new patient through regulation of the model by utilizing very limited number of samples of the new patient. Thus, the requirement of gathering more data for the process becomes minimalized. When tested on the CHB-MIT database, the average sensitivity reached 95.70%, whereas a false prediction rate of 0.057 was observed per hour with a seizure prediction horizon (SPH) of ten minutes and seizure occurrence period (SOP) of twenty minutes, and an average sensitivity of 98.52% was attained while the false prediction rate per hour was recorded at 0.045 with a seizure prediction horizon (SPH) of five minutes and seizure occurrence period (SOP) of twenty-five minutes. [38].

The remarkable capabilities of large language models (LLMs) have also been harnessed for few-shot EEG classification, as exemplified by the EEG-GPT framework [30]. The EEG-GPT yielded impressive results that are on par with the existing state-of-the-art deep learning approaches for classifying normal versus abnormal EEGs in the few-shot learning framework, using just 2% of the entire training data set

[30]. In addition, it had the added benefit of being able to provide an explanation for the reasoning steps along the way, as well as integrating the specialty tools in EEG across different scales of time and space [30]. This work represents a significant departure from traditional deep learning approaches, as it leverages the rich semantic and reasoning capabilities of LLMs to perform EEG analysis with an unprecedented level of interpretability and data efficiency.

There are many research papers discussing domain adaptation as well as the adversarial learning methodology used for cross-subject transfer during few-shot learning. The unsupervised domain adaptation method through adversarial networks was developed for cross-subject, few-shot, epileptic seizure detection [52]. Through adversarial learning methodology, features from different subjects were embedded in a subject-invariant feature space, and the classifier was trained using these subject-invariant features [52]. In the experiment conducted on 9 epileptic patients' intracranial EEG, the above methodology showed an improvement of 9.4% in the 1-shot classification accuracy when compared to the conventional subject-specific approach [52]. This concept was further extended to a more sophisticated technique called patient adversarial neural network (PANN), where there was a discrimination module to differentiate patients, and the adversarial optimization was done between the feature extractor and the discriminator to retain only the common seizure features among different patients [51]. Alongside the above-mentioned approach, when combined with the spatio-temporal EEG augmentation (STEA) for generating synthesized seizure data, it showed promising results in cross-patient generalization on clinical and publicly available EEG data [51].

Weakly supervised learning with spatial priors was tackled via SPID-Net, the Spatial Prior-injected IED Detection Network to perform fine-grained detection of interictal epileptiform discharges (IEDs) based on merely coarse-grained epoch-level annotations [48]. Specifically, SPID-Net injected the spatial priors via two major building blocks, which were the Graph-based Spatial Fusion Block (GSFB) with the associated auxiliary loss for learning spatial coherency between spatially adjacent channels, and the Pseudo-Label Supervision (PLS) that directly incorporated the spatial priors to help detect the IEDs that were subtle and spatially adjacent [48]. With the use of SPID-Net,

state-of-the-art performance was obtained on the TUEV and VEPI datasets, with an F1-score and AUPRC of 0.9154 and 0.9714, respectively [48]. More impressively, cross-dataset analysis validated its robustness and domain generalization capability through pre-training on one dataset followed by few-shot transfer learning on another dataset [48]. In more detail, SPID-Net pretrained on one dataset could be easily adopted to another dataset when trained on limited fine-tuning data, implying strong capabilities for domain adaptation across few-shot transfer learning scenarios [48]. This is a breakthrough finding since it means that neural networks pre-trained on one institution could be transferred to other institutions with minimal annotation efforts.

The importance of multi-strategy data augmentation in combination with contrastive learning for few-shot seizure prediction was demonstrated by a framework that employed wavelet-based frequency mixing, temporal masking, and window-based masking to enrich the limited labeled data, along with a hierarchical contrastive loss that incorporated both instance-level and temporal contrastive learning to enhance feature separation between interictal and preictal periods [42]. On the CHB-MIT dataset, this method achieved an accuracy of 94.51%, a sensitivity of 95.05%, and a false prediction rate of 0.024 per hour while using only 30% of the labeled data, and it further achieved a 20.12-minute average prediction time [42]. On the Siena dataset, it achieved an accuracy of 93.14%, a sensitivity of 92.77%, and a false prediction rate of 0.030 per hour, further validating its generalizability across different clinical populations [42].

Finally, an empirical approach to small-sample learning employed a sample enhancement method based on EEG signal channel swapping to improve detection of seizure patterns under limited data conditions [39]. This technique, which artificially diversifies the training set by swapping EEG channels, improved the detection accuracy from 0.6797 to 0.7789 when only 100 samples were available, from 0.6952 to 0.8210 with 500 samples, and from 0.7273 to 0.8517 with 1000 samples on the CHB-MIT dataset [39]. Combining this method with a sliding window strategy showed approximately 8% improvement across various configurations, demonstrating that even simple, computationally inexpensive augmentation

techniques can yield substantial gains in few-shot scenarios [39].

STUDY	TRANSFER / SELF-SUPERVISED STRATEGY	PRETRAINING TASK / DATA	FINE-TUNING / ADAPTATION STRATEGY	KEY PERFORMANCE METRICS
[29]	Self-supervised sequence-to-sequence pretraining	Unlabeled EEG data (TUSZ)	Lightweight model with pretrained encoder and simple prediction layer; fine-tuned with limited labeled data	83.8% weighted F1 (4-class), 73.5% weighted F1 (8-class); comparable performance with 25% labeled data
[30]	Large language model (LLM) based few-shot learning	General-purpose language pretraining	Utilized only 2% of training data for classification	Comparable to state-of-the-art deep learning methods
[35]	Transfer learning with convolutional autoencoder	41 patients from EPILEPSIAE database	Encoder weights frozen; B-LSTM + classifier optimized for each of 24 patients	~4× fewer false alarms; 13% more validated patients; more stable and faster training
[38]	Few-shot learning via generalizable CNN + patient-specific regulation	Generalizable CNN trained on population data	Regulated with a small number of new patient samples for patient-specific adaptation	95.70% sensitivity, 0.057/h FPR (10-min SPH, 20-min SOP); 98.52% sensitivity, 0.045/h FPR (5-min SPH, 25-min SOP)
[39]	Sample enhancement via EEG channel swapping	N/A (augmentation applied directly to training set)	Combined with sliding window for multi-perspective classification	Accuracy improved from 0.6797 to 0.7789 (100 samples), 0.6952 to 0.8210 (500 samples); ~8% improvement across configurations
[42]	Multi-strategy data augmentation + hierarchical contrastive learning	Wavelet-based frequency mixing, temporal masking, window-based masking	Instance-level and temporal contrastive learning with only 30% labeled data	94.51% accuracy, 95.05% sensitivity, 0.024/h FPR (CHB-MIT); 93.14% accuracy, 92.77% sensitivity, 0.030/h FPR (Siena)
[44]	DCGAN for synthetic data generation + transfer learning	Patient-specific DCGAN trained on real EEG	Augmented data used to train VGG16, VGG19, ResNet50, Inceptionv3 via transfer learning	Inceptionv3: 90.03% sensitivity, 0.03 FPR/h; 4-5% improvement over state-of-the-art augmentation
[48]	Weakly supervised learning with spatial priors + few-shot transfer learning	TUEV and VEPI datasets (coarse-grained epoch-level annotations)	Cross-dataset adaptation with small amount of fine-tuning data	F1-score of 0.9154, AUPRC of 0.9714; robust transferability
[49]	Self-supervised learning via weak label generation	Unlabeled EEG data (autocorrelation harnessed for weak labels)	Personalized forecasting models for individual patients	14.30% improvement in sensitivity, 19.61% reduction in false alarms

Table 4. Summary of Transfer Learning and Self-Supervised Pretraining Methods for EEG-Based Epileptic Seizure Analysis.

D. Data Augmentation and Synthetic Sample Generation

Data augmentation and synthetic sample generation constitute a critical and complementary technical pillar in the existing literature, serving both as a standalone strategy to mitigate data scarcity and as an integrated component within larger few-shot learning frameworks. The studies in this category share the fundamental premise that artificially expanding the limited training set with realistic, diverse, or transformed EEG samples can enhance model robustness, improve generalization, and reduce the risk of overfitting, which is particularly acute when only a handful of labeled seizure events are available. These methods span a spectrum from simple signal-level transformations to sophisticated generative models and domain-specific operations.

The most computationally expensive yet flexible technique in generating synthetic data is through the application of generative adversarial networks (GANs). The use of a deep convolutional generative adversarial network (DCGAN) has been employed to generate patient-specific synthetic data for predicting epileptic seizures [44]. To validate the synthetically generated data, they performed tests on the data using a one-class SVM and also on their newly developed convolutional epileptic seizure predictor (CESP) model, where they established that the synthetic data retained clinically significant discriminatory features [44]. However, when these four famous deep learning architectures, namely VGG16, VGG19, ResNet50, and Inceptionv3, were used to further augment the synthetic data for epileptic seizure prediction through transfer learning, it was found that the Inceptionv3 neural network gave the best results in terms of accuracy with sensitivity being 90.03% while maintaining a false prediction rate of 0.03 per hour [44]. The researchers claimed that the developed DCGAN model for data augmentation resulted in an improvement of 4-5% over current state-of-the-art augmentation approaches [44].

A more recent and conceptually advanced approach integrated generative augmentation with meta-learning in the form of generative digital twins, as exemplified by the GenDT-MAML framework [34]. In this paradigm, each patient was associated with a latent digital-twin state that encoded individual EEG characteristics, and this state was used to condition a generative model capable of synthesizing patient-

consistent pre-ictal and inter-ictal spectrograms [34]. The generated data sets were used in the task creation procedure for Model-Agnostic Meta-Learning (MAML). The method allows the classifier to learn to initialize in such a way that it would be able to quickly adjust to a new patient with a very limited number of annotated windows [34]. The sensitivity of the model reached 95.1%, and the false-positive rate was 0.08 per hour based on the CHB-MIT dataset, surpassing not only MAML but also all the other compared techniques [34]. In practice, the use of this technique was implemented using a model that included the edge-twin-cloud continuum service architecture. This model significantly decreased the communication cost by 98.3% as compared to continuous streaming but maintained sensitivity of 91.4% [33].

The diversity-enhanced data augmentation and few-shot learning (DAFEM) approach also employed a GAN for diversified EEG sample generation as its first step, specifically to address the data scantiness and inter-patient variability constraint in multi-class epileptic seizure detection [20]. The generated samples were subsequently processed through graph theory-based feature extraction and a Siamese neural network classifier, yielding a 2.73% and 4.5% higher recall on the Bonn and CHB-MIT datasets, respectively [20].

Beyond deep generative models, several studies have proposed simpler but effective signal-domain augmentation strategies tailored to EEG data. A channel swapping method was introduced as a sample enhancement technique specifically for few-shot scenarios [39]. By exchanging EEG sequences from different channels, this method generated new EEG samples that artificially expanded the training set without requiring any additional recording hardware or expert annotation [39]. On the CHB-MIT dataset, the detection accuracy improved from 0.6797 to 0.7789 when only 100 training samples were available, from 0.6952 to 0.8210 with 500 samples, and from 0.7273 to 0.8517 with 1000 samples [39]. Compared to the sliding window method, the proposed channel swapping demonstrated higher accuracy in extreme low sample sizes, and combining both methods further enhanced detection performance by approximately 8% across various configurations [39]. The spatial-temporal EEG augmentation (STEA) technique was proposed for the creation of synthetic training EEGs that mimic the natural spatio-temporal

correlations [51]. The STEA was achieved based on the statistical analysis of the EEGs during seizures, and further combined with a patient-adversarial neural network (PANN), which learns a patient-invariant feature representation [51]. In the case of application to public as well as clinical EEG datasets, such a combination demonstrated an impressive boost of the generalization power of deep networks to unseen patients, which is among the most important challenges for clinical translation [51].

A multi-strategy data augmentation framework for seizure prediction integrated three distinct operations: wavelet-based frequency mixing, temporal masking, and window-based masking [42]. Each of these strategies was designed to enhance model robustness and generalization from a different perspective: frequency mixing perturbed the spectral content of EEG segments to prevent overfitting to specific frequency bands, temporal masking randomly occluded portions of the time series to encourage learning of distributed temporal patterns, and window-based masking operated on segmented windows to simulate variability in seizure onset morphology [42]. When combined with a hierarchical contrastive loss function that integrated instance-level and temporal contrastive learning, this augmentation framework enabled the model to achieve an accuracy of 94.51%, a sensitivity of 95.05%, and a false prediction rate of 0.024 per hour on the CHB-MIT dataset while using

only 30% of the labeled data [42]. On the Siena dataset, the same framework achieved an accuracy of 93.14% and a sensitivity of 92.77%, demonstrating its generalizability across different clinical populations [42].

An entirely different approach to data augmentation leveraged the properties of empirical mode decomposition (EMD) to extract intrinsic mode functions from EEG signals, which were then used to compute power spectral density (PSD) features as inputs to convolutional neural networks (CNNs) [40]. This EMD-based preprocessing was not a generative augmentation in the traditional sense, but it effectively expanded the representational capacity of the limited training data by decomposing the signal into multiple physiologically meaningful components [40]. When the number of training samples was reduced to 10%, this method showed improvements of 23%, 19%, and 26% in accuracy, sensitivity, and specificity, respectively, compared to using the original raw EEG input across different network architectures [40].

We summarize the key characteristics of these data augmentation and synthetic sample generation methods in Table 5, which provides a structured comparison of the augmentation strategies, the types of synthetic data generated, the intended target of the augmentation, and the degree of integration with few-shot learning frameworks.

Table 5. Summary of Data Augmentation and Synthetic Sample Generation Methods for EEG-Based Epileptic Seizure Analysis.

STUDY	AUGMENTATION METHOD	SYNTHETIC DATA TYPE	AUGMENTATION TARGET	INTEGRATION WITH FEW-SHOT LEARNING
[20]	Generative Adversarial Network (GAN) with diversified EEG sample generation	EEG signals	Epileptic seizure classes	The data scarcity in the epileptic seizure classes is handled by adopting the few-shot learning strategy, and the Siamese neural network is modeled for multi-class classification of epilepsy
[33]	Generative augmentation via generative digital twins	Pre-ictal and inter-ictal spectrograms	MAML task construction for cross-patient aggregation	Generative augmentation and dispatch decisions at the Twin/Fog layer to augment MAML task construction and improve adaptation under extreme data scarcity
[34]	Generative digital twin model conditioned on patient-specific latent state	Patient-consistent pre-ictal and inter-ictal spectrograms	MAML task construction to improve adaptation under extreme data scarcity	Synthetic samples are integrated into MAML's task construction, allowing the classifier to learn an initialization that adapts rapidly to

STUDY	AUGMENTATION METHOD	SYNTHETIC DATA TYPE	AUGMENTATION TARGET	INTEGRATION WITH FEW-SHOT LEARNING
				new patients using only a handful of labeled windows
[39]	Channel swapping of EEG sequences	EEG signals	Training set expansion for epilepsy detection	Improves epilepsy detection accuracy in few-shot scenarios; combined with sliding window for multi-perspective classification
[40]	Empirical mode decomposition (EMD) of EEG signals and PSD of EMD components	EMD components and their PSD features	Input representation for CNNs	Designed to improve seizure detection accuracy in few-shot scenarios; achieves superior performance compared to traditional deep learning methods without EMD when training samples are limited
[42]	Wavelet-based frequency mixing, temporal masking, and window-based masking	Augmented EEG segments	Enhance model robustness and generalization	Achieves high performance even with limited labeled data; integrated with hierarchical contrastive learning for feature separation
[44]	DCGAN-based synthetic data generation	EEG data	Epileptic seizure prediction	Synthetic data used to augment training set for transfer learning with deep models (VGG16, VGG19, ResNet50, Inceptionv3)
[51]	Spatio-temporal EEG augmentation (STEA)	Synthetic training seizure data with spatio-temporal dependencies	To largely improve cross-patient generalization	Combined with patient-adversarial neural network (PANN) for learning patient-invariant representations; evaluated on both public and clinical EEG datasets

V. DISCUSSION

The synthesis of the 34 included studies reveals a rapidly maturing field where few-shot deep learning is being systematically adapted to confront the unique challenges of EEG-based epileptic seizure analysis. Taken together, the evidence suggests that the core promise of few-shot learning—achieving clinically meaningful performance with drastically reduced annotation requirements—is not merely theoretical but is being realized across a diverse range of methodological frameworks. A consistent pattern that emerges across the literature is that metric learning and prototype-based approaches, particularly Siamese and prototypical networks, have been the most widely adopted and consistently successful strategies. This is perhaps unsurprising given that the fundamental task of distinguishing ictal from interictal EEG patterns can be naturally framed as a similarity comparison problem, where the model learns a discriminative embedding space rather than a complex decision

boundary. The success of these methods, as demonstrated by studies achieving over 98% accuracy on benchmark datasets [25] [24], suggests that the underlying EEG features of seizure activity are sufficiently consistent within a patient to be captured by a compact prototype representation, even when only a handful of examples are available.

However, the literature also reveals a critical tension between the high performance reported on controlled, single-dataset experiments and the more modest results observed in cross-dataset or multi-center evaluations. While several studies report impressive metrics on the Bonn and CHB-MIT datasets, these datasets are relatively homogeneous in terms of recording equipment, sampling rates, and patient demographics. The few studies that have ventured into cross-dataset validation, such as the work on SPID-Net [48] and the CLEP-STs-Net framework [27], consistently report a drop in performance when models are transferred to unseen clinical environments. This finding underscores a fundamental

limitation of the current evidence base: the field has yet to demonstrate that few-shot methods can generalize robustly across the full spectrum of clinical variability, including differences in EEG montages, electrode placements, patient ages, and seizure etiologies. The implications of this gap are profound, as a model that works well on a curated research dataset but fails in a real-world epilepsy monitoring unit is of limited clinical utility.

The theoretical implications of our synthesis are equally important. The reviewed studies collectively challenge the traditional assumption that deep learning models for EEG analysis require thousands of labeled examples to be effective. Instead, they suggest that the inductive biases inherent in few-shot learning architectures—such as the metric learning objective of maximizing inter-class separation or the meta-learning objective of learning a good parameter initialization—are well-aligned with the structure of EEG data. This alignment may stem from the fact that seizure activity, despite its inter-patient variability, exhibits stereotypical spatio-temporal patterns that are shared across individuals. The success of self-supervised pretraining methods [29] [49] further supports this view, as they demonstrate that meaningful representations can be learned from unlabeled EEG data alone, without any expert annotations. This finding has significant theoretical implications for our understanding of EEG feature learning: it suggests that the statistical structure of EEG signals contains sufficient information to learn a useful feature space, and that the role of few-shot learning is primarily to adapt this general-purpose representation to the specific characteristics of a new patient's seizure patterns.

From a practical standpoint, the findings of this review have several important implications for clinicians, researchers, and healthcare technology developers. The most immediate implication is that few-shot learning could dramatically reduce the time and cost associated with developing patient-specific seizure detection and prediction systems. The ability to train a model with as few as one to five labeled seizure events [47] [32] means that a personalized system could be deployed within hours of a patient's admission to an epilepsy monitoring unit, rather than requiring days or weeks of data collection. This is particularly relevant for patients with rare or infrequent seizures, for whom collecting a large training set is practically impossible.

Furthermore, the integration of few-shot learning with edge computing architectures [32] [33] opens the door to real-time, on-device seizure detection that does not rely on continuous cloud connectivity, addressing both latency and privacy concerns that are paramount in clinical settings. The federated few-shot learning framework [46] also offers a practical pathway for multi-institutional collaboration without compromising patient data privacy, which is a critical consideration given the stringent regulatory requirements governing medical data.

Nevertheless, we must acknowledge several important limitations of this systematic review that may have influenced our findings and conclusions. The first limitation concerns the scope of our literature search. Although we searched five major databases, it is possible that relevant studies indexed in other databases, such as the ACM Digital Library or PsycINFO, were missed. Furthermore, our reliance on English-language publications may have introduced a language bias, potentially excluding high-quality research published in other languages, particularly from countries with large epilepsy populations such as China, India, and Brazil. The exclusion of grey literature, such as technical reports and dissertations, may have also contributed to a publication bias, as studies with negative or null results are less likely to be published in peer-reviewed venues. The second limitation relates to the heterogeneity of the included studies. The diversity in datasets, evaluation protocols, performance metrics, and few-shot scenarios made it challenging to perform a quantitative meta-analysis or to directly compare the effectiveness of different methods. For instance, some studies reported accuracy on a binary classification task (seizure vs. non-seizure), while others reported sensitivity and false prediction rate on a seizure prediction task, and still others reported F1-scores on multi-class classification. This heterogeneity limits the strength of our conclusions regarding which specific few-shot method is most effective. The third limitation is the potential for subjectivity in our quality assessment. While we used a standardized data extraction form and two independent reviewers, the assessment of methodological quality and risk of bias is inherently subjective, particularly for studies that did not report sufficient detail about their experimental setup, such as the exact number of training samples per patient or the method of cross-validation.

The limitations of the review process are compounded by several methodological weaknesses in the included studies themselves. A significant proportion of the studies did not report confidence intervals or statistical significance tests for their performance metrics, making it difficult to assess the reliability and reproducibility of their findings. Many studies also lacked a clear description of their data preprocessing pipeline, including filtering, artifact removal, and segmentation parameters, which are known to have a substantial impact on EEG classification performance. Furthermore, the majority of studies evaluated their methods on a single dataset, and only a handful conducted cross-dataset validation, which is essential for assessing generalizability. The absence of standardized benchmarks and evaluation protocols in the field is a critical barrier to progress, as it prevents meaningful comparison between different methods and hinders the identification of the most promising approaches. The field would benefit greatly from the establishment of a common evaluation framework, similar to the few-shot learning benchmarks in computer vision (e.g., miniImageNet, tieredImageNet), that specifies the exact data splits, few-shot scenarios, and evaluation metrics to be used. Based on the gaps and inconsistencies uncovered in this review, we identify several promising directions for future research. There is a clear and urgent need for large-scale, multi-center validation studies that evaluate few-shot learning methods on diverse clinical populations, including patients of different ages, with different seizure types, and from different geographic regions. Such studies would provide critical evidence regarding the generalizability and robustness of these methods and would help to identify the conditions under which they are most effective. Future research should also explore the integration of few-shot learning with emerging technologies such as wearable EEG devices and mobile health platforms. The GenDT-MAML framework [34] and the edge-twin-cloud architecture [33] represent early steps in this direction, but there is a need for more comprehensive studies that address the practical challenges of real-time, low-power, and low-latency deployment. The development of interpretable few-shot models is another understudied area that warrants further investigation. While the ProtoEEG-QA method [36] and the prototype learning framework [26] offer some degree of interpretability by highlighting influential

support examples or tracing the origins of prototypes, most few-shot methods remain black-box models. In a clinical context where decisions have life-altering consequences, the ability to explain why a model classified a particular EEG segment as ictal or interictal is essential for building trust and facilitating adoption.

The potential of combining few-shot learning with other advanced paradigms, such as reinforcement learning and causal inference, remains largely unexplored. For instance, a reinforcement learning framework could be used to optimize the selection of which EEG segments to label for a new patient, thereby minimizing the annotation burden while maximizing the information gain. Similarly, causal inference methods could help to disentangle the causal mechanisms underlying seizure generation from the spurious correlations that may arise in small datasets, leading to more robust and generalizable models. The integration of few-shot learning with multimodal data, such as combining EEG with video, accelerometry, or heart rate variability, is another promising direction that could improve the accuracy and robustness of seizure detection and prediction systems. Finally, there is a need for more rigorous and standardized evaluation protocols in the field. We recommend that future studies adopt a common set of benchmark datasets, such as the CHB-MIT and TUSZ datasets, and report performance using a standardized set of metrics, including sensitivity, specificity, precision, F1-score, and area under the receiver operating characteristic curve, along with confidence intervals. The establishment of a public leaderboard or challenge platform for few-shot EEG seizure analysis could accelerate progress by providing a fair and transparent basis for comparing different methods.

VI. CONCLUSION

This systematic review has synthesized the current state of research on few-shot deep learning for EEG-based epileptic seizure analysis, addressing the fundamental question of whether these methods can overcome the pervasive data scarcity that limits clinical translation. Our analysis of 34 studies confirms that few-shot learning offers a viable and increasingly effective paradigm shift, with metric learning, meta-learning, and self-supervised pretraining emerging as the dominant and most

successful methodological families. The core finding is that clinically relevant performance—often exceeding 90% sensitivity and specificity—can be achieved with as few as one to ten labeled seizure examples per patient, a dramatic reduction from the thousands typically required by conventional deep learning approaches. This synthesis therefore challenges the prevailing assumption that large, expertly annotated datasets are an indispensable prerequisite for automated EEG analysis, and instead suggests that the inductive biases of few-shot architectures are well-aligned with the intrinsic structure of epileptic EEG patterns.

The practical implications of these findings are substantial for clinical practice and healthcare technology development. The demonstrated ability to personalize seizure detection and prediction systems with minimal patient-specific data could accelerate the deployment of automated monitoring tools in epilepsy monitoring units, reduce the annotation burden on expert neurologists, and enable the development of wearable systems for long-term outpatient monitoring. The theoretical contribution of this review lies in its systematic mapping of the methodological landscape, revealing that while the field has made remarkable progress, it remains fragmented by heterogeneous evaluation protocols and a lack of multi-center validation. Future research must prioritize the establishment of standardized benchmarks, the conduct of prospective clinical trials across diverse populations, and the integration of interpretability mechanisms to build clinical trust. The convergence of few-shot learning with edge computing, federated learning, and generative models represents a particularly promising frontier that could ultimately translate these algorithmic advances into practical, privacy-preserving, and real-time clinical tools for the millions of people living with epilepsy worldwide.

REFERENCES

- [1] M. Leonardi and T. Ustun, “The global burden of epilepsy,” *Epilepsia*, 2002.
- [2] J. Praline, J. Grujic, P. Corcia, B. Lucas, C. Hommet, et al., “Emergent EEG in clinical practice,” *Clinical Neurophysiology*, 2007.
- [3] L. Wang, X. Long, J. Arends, and R. Aarts, “EEG analysis of seizure patterns using visibility graphs for detection of generalized seizures,” *Journal of Neuroscience Methods*, 2017.
- [4] G. Young, “Continuous EEG monitoring in the ICU: Challenges and opportunities.” *Canadian journal of neurological sciences*, 2009.
- [5] A. Shoeibi, M. Khodatars, N. Ghassemi, M. Jafari, et al., “Epileptic seizures detection using deep learning techniques: A review,” *International Journal of Environmental Research and Public Health*, 2021.
- [6] M. Zhou et al., “Epileptic seizure detection based on EEG signals and CNN,” *Frontiers in Neuroinformatics*, 2018.
- [7] E. Vafaei and M. Hosseini, “Transformers in EEG analysis: A review of architectures and applications in motor imagery, seizure, and emotion classification,” *Sensors*, 2025.
- [8] B. Shang, F. Duan, R. Fu, J. Gao, H. Sik, et al., “EEG-based investigation of effects of mindfulness meditation training on state and trait by deep learning and traditional machine learning,” *Frontiers in Human Neuroscience*, 2023.
- [9] Q. Yuan et al., “Epileptic seizure detection based on imbalanced classification and wavelet packet transform,” *seizure*, 2017.
- [10] M. Diachenko, S. Houtman, E. Juarez-Martinez, et al., “Improved manual annotation of EEG signals through convolutional neural network guidance,” *eneuro*, 2022.
- [11] A. V. Esbroeck, L. Smith, Z. Syed, S. Singh, and Z. Karam, “Multi-task seizure detection: Addressing intra-patient variation in seizure morphologies,” *Machine Learning*, 2016.
- [12] H. Zhao, Q. Zheng, K. Ma, H. Li, et al., “Deep representation-based domain adaptation for nonstationary EEG classification,” *IEEE Transactions on Neural Networks and Learning Systems*, 2020.
- [13] Y. Wang, Q. Yao, J. Kwok, and L. Ni, “Generalizing from a few examples: A survey on few-shot learning,” *ACM computing surveys (csur)*, 2020.
- [14] H. Gharoun, F. Momenifar, F. Chen, et al., “Meta-learning approaches for few-shot learning: A survey of recent advances,” *ACM Computing Surveys*, 2024.
- [15] F. Pahde, M. Puscas, T. Klein, et al., “Multimodal prototypical networks for few-shot learning,” in

- IEEE winter conference on applications of computer vision, 2021.
- [16] T. Iloon, R. Barati, and H. Azad, "Siamese network-based feature transformation for improved automated epileptic seizure detection," *Complexity*, 2022.
- [17] A. Romney and V. Manian, "Optimizing seizure prediction from reduced scalp EEG channels based on spectral features and MAML," *IEEE Access*, 2021.
- [18] S. Rudrashetty, "Generative AI for EEG-based brain signal augmentation and seizure prediction," ... on Sensors; Related Networks, 2025.
- [19] M. Page, J. McKenzie, P. Bossuyt, et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, p. n71, 2021.
- [20] R. S. Aldahr, M. Alanazi, and M. Ilyas, "Addressing inter-patient variability in EEG: Diversity-enhanced data augmentation and few-shot learning-based epilepsy detection," 2022 International Conference on Healthcare Engineering (ICHE). pp. 1–7, 2022.
- [21] A. Burrello, K. A. Schindler, L. Benini, and A. Rahimi, "One-shot learning for iEEG seizure detection using end-to-end binary operations: Local binary patterns with hyperdimensional computing," 2018 IEEE Biomedical Circuits and Systems Conference (BioCAS), pp. 1–4, 2018.
- [22] Q. Chen, C. Ye, R. Xiao, J. Pan, and J. Li, "SemSTNet: Medical EEG semantic metric learning with class prototypes generated by pretrained language model," *IEEE Transactions on Biomedical Engineering*, vol. 73, pp. 2010–2020, 2025.
- [23] X. Chen, L. Zhang, Y. Yang, D. Zhao, W. Zhong, and Y. Wang, "An adaptive task-aware multi-scale convolutional network for few-shot epileptic EEG classification," *Journal of Mechanics in Medicine and Biology*, 2026.
- [24] L. Chengang, C. Yuxin, Z. Chen, Y. Liu, and Z. Wang, "Automatic epilepsy detection based on generalized convolutional prototype learning," *Measurement*, 2021.
- [25] L. Duan, Z. Wang, Y. Qiao, Y. Wang, Z. Huang, and B. Zhang, "An automatic method for epileptic seizure detection based on deep metric learning," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, pp. 2147–2157, 2021.
- [26] Y. Gao, A. Liu, H. Cui, R. Qian, and X. Chen, "An interpretable and generalizable deep learning model for iEEG-based seizure prediction using prototype learning and contrastive learning," *Computers in biology and medicine*, vol. 183, p. 109257, 2024.
- [27] L. Guo, T. Yu, S. Zhao, X. Li, X. Liao, and Y. Li, "CLEP: Contrastive learning for epileptic seizure prediction using a spatio-temporal-spectral network," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 3915–3926, 2023.
- [28] A. He et al., "Small sample epilepsy detection method based on convolutional prototype learning," *assets-eu.researchsquare.com*, 2022.
- [29] Y. Hu, J. Liu, W. Zhang, Y. Sui, Q. Meng, and R. Sun, "Self-supervised learning with adaptive graph modeling for EEG-based epileptic seizure classification," *IEEE Transactions on Biomedical Engineering*, vol. 73, pp. 1414–1422, 2025.
- [30] J. W. Kim, A. Alaa, and D. Bernardo, "EEG-GPT: Exploring capabilities of large language models for EEG classification and interpretation," *ArXiv*, vol. abs/2401.18006, 2024.
- [31] X. Li, J. Liu, W. Nie, and Q. Yuan, "Neonatal seizure detection combined deep network and meta-learning," 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2522–2526, 2023.
- [32] M. Li, Y. Deng, J. Li, X. Huang, and W. Xing, "Personalized real-time epileptic seizure alert on edge devices via few-shot meta-learning," 2025 IEEE International Conference on Cloud Computing Technology and Science (CloudCom), pp. 1–8, 2025.
- [33] M. Li, Y. Zhang, W. Zhang, Q. Liu, and W. Xing, "An edge-cloud continuum service for personalized seizure forecasting: Communication-efficient digital twins with few-shot updates," *researchsquare.com*, 2026.
- [34] M. Li, Y. Deng, X. Huang, J. Li, and W. Xing, "GenDT-MAML: A data-efficient framework for personalized seizure prediction using generative digital twins and MAML," *IEEE Open Journal of the Communications Society*, vol. 7, pp. 1263–1276, 2026.

- [35] F. Lopes, M. Pinto, A. Dourado, A. Schulze-Bonhage, et al., “Addressing data limitations in seizure prediction through transfer learning,” *Scientific Reports*, 2024.
- [36] D. Mewada, M. Gayen, M. Sarma, and D. Samanta, “Attention-guided few-shot prototypical network for ICU abnormal EEG pattern recognition,” in *International conference on artificial intelligence and data science*, 2025.
- [37] M. Munia, S. M. Hosseini, M. Nourani, J. Harvey, and H. Dave, “Imbalanced EEG analysis using one-shot learning with siamese neural network,” 2021 IEEE 9th International Conference on Healthcare Informatics (ICHI), pp. 4–12, 2021.
- [38] J. Nazari, A. M. Nasrabadi, M. Menhaj, and S. Raiesdana, “Epilepsy seizure prediction with few-shot learning method,” *Brain Informatics*, vol. 9, 2022.
- [39] Y. Pan, F. Dong, W. Yao, X. Meng, and Y. Xu, “Channel swapping of EEG signals for deep learning-based seizure detection,” *Electronics Letters*, 2024.
- [40] Y. Pan, F. Dong, W. Yao, X. Meng, and Y. Xu, “Empirical mode decomposition for deep learning-based epileptic seizure detection in few-shot scenario,” *IEEE Access*, vol. 12, pp. 86583–86595, 2024.
- [41] P. Priyanga and R. Kumar, “Adaptive prototype-based subtle transient pattern transformers for enhanced neonatal seizure classification and severity assessment,” *International Journal of Developmental Neuroscience*, 2025.
- [42] L. Qi, F. Li, J. Shang, D. Ge, S. Wang, and S. Yuan, “Epileptic seizure prediction using multi-strategy data augmentation and hierarchical contrastive learning,” *IEEE Journal of Biomedical and Health Informatics*, vol. 30, pp. 2023–2033, 2025.
- [43] A. Rahmani, A. Venkitaraman, et al., “A meta-GNN approach to personalized seizure detection and classification,” *ICASSP*, 2023.
- [44] K. Rasheed, J. Qadir, T. O’Brien, L. Kuhlmann, and A. Razi, “A generative model to synthesize EEG data for epileptic seizure prediction,” *Ieee Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 2322–2332, 2020.
- [45] A. Romney and V. Manian, “Optimizing seizure prediction from reduced scalp EEG channels based on spectral features and MAML,” *IEEE Access*, vol. PP, pp. 1–1, 2021.
- [46] E. Sysoykova, B. Anzengruber-Tanase, M. Haslgrübler, P. Seidl, and A. Ferscha, “Federated few-shot learning for epileptic seizure detection under privacy constraints,” *ArXiv*, vol. abs/2512.13717, 2025.
- [47] Z. B. Tariq, A. Iyengar, L. Marcuse, H. Su, and B. Yener, “Patient-specific seizure prediction using single seizure electroencephalography recording,” *ArXiv*, vol. abs/2011.08982, 2020.
- [48] C. Wang, J. Shen, B. Jin, W. Luo, and F. Yu, “Weakly supervised single-channel interictal epileptiform discharge detection with spatial priors,” *Expert Systems with Applications*, 2026.
- [49] Y. Yang, N. D. Truong, J. Eshraghian, A. Nikpour, and O. Kavehei, “Weak self-supervised learning for seizure forecasting: A feasibility study,” *Royal Society Open Science*, vol. 9, 2022.
- [50] P. Zhang, T. Gao, J. Guo, J. Duan, and S. I. Nikolenko, “EARLY WARNING PREDICTION WITH AUTOMATIC LABELLING IN EPILEPSY PATIENTS,” *The ANZIAM Journal*, 2023.
- [51] Z. Zhang et al., “Cross-patient automatic epileptic seizure detection using patient-adversarial neural networks with spatio-temporal EEG augmentation,” *Biomed. Signal Process. Control.*, vol. 89, p. 105664, 2024.
- [52] B. Zhu and M. Shoaran, “Unsupervised domain adaptation for cross-subject few-shot neurological symptom detection,” *arXiv preprint arXiv:2103.00606*, 2021.
- [53] Y. Zhu, M. Saqib, E. Ham, S. Belhareth, R. Hoffman, and M. D. Wang, “Mitigating patient-to-patient variation in EEG seizure detection using meta transfer learning,” 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE), pp. 548–555, 2020.